Ryohei Nakatsu
Matthias Rauterberg
Paolo Ciancarini
Editors

Handbook of Digital Games and Entertainment Technologies

SpringerReference

Handbook of Digital Games and Entertainment Technologies

Ryohei Nakatsu • Matthias Rauterberg
Paolo Ciancarini
Editors

Handbook of Digital Games and Entertainment Technologies

With 373 Figures and 36 Tables

Springer Reference

Editors
Ryohei Nakatsu
Design School
Kyoto University
Kyoto, Japan

Matthias Rauterberg
Industrial Design
Eindhoven University of Technology
Eindhoven, The Netherlands

Paolo Ciancarini
Dipto. Scienze dell'Informazione
Università di Bologna
Bologna, Italy

ISBN 978-981-4560-49-8 ISBN 978-981-4560-50-4 (eBook)
ISBN 978-981-4560-51-1 (print and electronic bundle)
DOI 10.1007/978-981-4560-50-4

Library of Congress Control Number: 2016940980

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer Science+Business Media Singapore Pte Ltd.

Preface

We are very proud of presenting to the reader this Springer *Handbook of Digital Games and Entertainment Technologies*.

The Handbook covers all aspects of designing and building the most advanced interactive systems and devices for playing and entertaining, such as human-machine interfaces, networks and robots, artificial intelligence, and interactive television, and includes interdisciplinary studies on serious games, digital art, edutainment, entertainment ethics and sociology, and many more. The scope of each part spans from basic theories to enabling technologies, and from advanced applications to psychological and sociological reflections on those.

Entertainment is an essential part of our everyday activities. When we are children we play with our friends and listen to stories told by our relatives: these experiences are the basis of our ability to communicate, discuss, and negotiate with others. Johan Huizinga showed in his book *Homo Ludens* that playing is the basis of our culture. In the course of the human history, however, somehow entertainment has been thought as a marginal activity of lesser importance with respect to other activities such as education, work, medicine, etc.

Thanks to the development of digital information and communication technologies (ICT), recently a plethora of new and interactive entertainment systems and products have emerged: from lean-back consumption to lean-forward interaction. Not only the younger generations but also more mature generations enjoy playing video games, communicating via social networks, and using new enhanced entertainment media like interactive television or immersive virtual reality systems. These new systems and products are blurring the distinction between work and play just as the psychologist Mihaly Csikszentmihalyi indicated in his "Flow Theory."

The entertainment market is huge; the companies offering products in the areas of playing consoles, smart toys, online games, digital music, interactive TVs, movies, robots, etc., are economically very relevant. However, until recently entertainment was not considered as a major research topic in academia. In the first decade of the current twenty-first century, some pioneering researchers including us working in the area of entertainment met several times in specific conferences and agreed to define a new research area called "entertainment computing." In this new area, we wanted to get together various types of interdisciplinary research. We asked the International Federation on Information Processing (IFIP) to setup a new technical committee focusing on entertainment computing. In 2002, our proposal was accepted and we formed a new group called Specialist Group on Entertainment Computing that in 2006 was upgraded to Technical Committee on Entertainment Computing (IFIP TC14). Members of TC14 have been working in various areas of entertainment computing and have been promoting academic activities in this area.

Today, the design of digital games exploiting entertainment technologies has been recognized as an important and attractive topic in academic research. There are many people both in academia and industry who want to know the most recent topics and developments. Therefore we accepted the invitation by Springer to edit this Handbook. We hope that this work will contribute to a prospering development of entertainment computing both in academia and industry.

The aim of this Handbook is to serve as a key reference work as it provides the readers with a holistic picture of this interdisciplinary field covering technical issues, aesthetic/design theories, and sociological investigations. The Handbook consists of invited contributions from top class scholars and researchers from several topic areas. Each author was assigned the task to recall the foundations of a specific subject in the field of entertainment computing, to survey the current state of the art in the same field, and finally to sketch the most advanced entertainment applications related to that field.

The parts and their editors are the following:

1. Artificial Intelligence and Games (part editor Paolo Ciancarini): Artificial intelligence is a fundamental enabling technology for improving the playing experience in several types of games. This part includes four chapters dealing with algorithms and technologies for solving games, especially based on machine learning from large sets of playing data.
2. Brain-Computer Interfaces and Games (part editor Anton Nijholt): The direct exploitation of the brain activities of players is a radically new way to interact with entertainment products. This part includes six chapters describing how special devices allow to play in novel ways and how they influence the design of new videogames.
3. Entertainment Games (part editor Junichi Hoshino): Digital games are the core of entertainment computing. This part includes four chapters on different types of videogames exploiting a variety of entertainment computing technologies. Especially a survey on digital game industry would give readers the latest and deep insight into this fast moving area.
4. Interactive Storytelling (part editors Marc Cavazza and Michael Young): Storytelling is a very ancient activity; interactive storytelling is based on software which supports a narrative whose storyline is not predetermined. Interactive storytelling fulfills an old dream: the ability of the listener to “enter” the story she is told. This part includes five chapters which display a very interdisciplinary panorama on this subject.
5. Networking in Games (part editor Marco Roccetti): The global availability of the Internet and the widespread diffusion of powerful smartphones and personal

computers allow millions of people to play anytime everywhere, alone or in very large parties. Entertainment systems need advanced network technologies which connect devices with very different capabilities. This part includes five chapters on the main issues in networking for entertainment.

6. Serious Games (part editor Alessandro De Gloria): Serious games are one of the most promising areas in bridging the gap between enjoyful play and professional use through gamification. This part includes three chapters on serious gaming regarding science, technology, engineering, and mathematics (STEM); corporate identity; and ethics, privacy, and trust.
7. Art and Entertainment (part editors Ryohei Nakatsu and Naoko Tosa): Entertainment computing is a discipline whose aim is to combine technology with other areas such as art, culture, etc. Digital arts are novel forms of expression that we are learning to appreciate. This part includes seven chapters showing various examples on how entertainment computing handle art and culture.
8. Edutainment (part editor Wolfgang Mueller): The combination of edutainment and entertainment technologies – sometimes called "gamification" – offers new possibilities to educators and learners. This part includes two chapters.
9. Entertainment Robots (part editors Hooman Samani and Elham Saadatian): Robots are just starting to coexist with humans in several fields. Playful robotic devices offer new challenges in human-machine interactions and enable new kinds of user experiences that need to be studied with special care. This part includes four chapters.
10. Interactive Television and Online Experiences (part editor Marianna Obrist): Digital technologies enable new ways of interacting with old media: interactive TV is one major example, where the viewer is allowed to participate in the TV experience. This part includes five chapters.
11. Social and Ethical Issues (part editor Matthias Rauterberg): Because entertainment products have not only a technical and economical impact but also an enormous societal impact, this part addresses all related topics. This part includes six chapters on social and ethical aspects regarding positive and negative effects, in particular addiction, emerging media technology, and unconscious emotions.

This Handbook is a work in progress (a living reference work in Springer terms). This means that the authors and the part editors are allowed to update the online version of the papers even before the next edition of the Handbook. We are already planning a new edition, to include the new developments and topics that the exciting field of entertainment computing will study in the next future. Therefore, we will invite additional chapters from recognized experts in such fields.

The editors wish to thank all those who contributed to this Handbook, especially all part editors who have collected valuable chapter papers and reviewed them to guarantee the high quality of this Handbook.

This Handbook can be cited as follows:

Ryohei Nakatsu, Matthias Rauterberg, and Paolo Ciancarini (eds.), Handbook of Digital Games and Entertainment Technologies, Springer Singapore, 2016. DOI 10.1007/978-981-4560-52-8. ISBN (online) 978-981-4560-50-4

Ryohei Nakatsu
Matthias Rauterberg
Paolo Ciancarini

Contents

Volume 1

About the Editors

Paolo Ciancarini is Professor of Computer Science at the University of Bologna since 1992.

He got a Ph.D. in Informatics at the University of Pisa in 1988.

In Bologna, he lectures on Software Engineering and Software Architecture, and is member of the Faculty of the PhD School in Computer Science.

He currently is the President of the Italian Association of University Professors in Computer Science.

He is also the Vice-Director of CINI (National Inter-University Consortium for Informatics), a consortium of 43 universities engaged in national and international research projects.

In the period November 2011–June 2013, he has served in the national panel (ANVUR GEV01) for the evaluation of computer science research in Italian universities and research centers (VQR 2004–2010).

In the period March 2012–December 2013, he has been a member of the Italian ICT Delegation at the European Union for the 7th ICT Framework Program.

His research interests include: coordination languages and models, software architectures and infrastructures, advanced Web technologies, and software engineering for computer games.

He has been involved as a site leader in several projects funded by the European Commission and by the Italian Government.

He is the author of over 120 scientific papers and books.

He is married, has two children, and is a passionate chess player and book collector.

Prof. Dr. Matthias Rauterberg Eindhoven University of Technology (Netherlands), received a B.S. in Psychology (1978) at the University of Marburg (Germany), a B.A. in Philosophy (1981), a B.S. in Computer Science (1983), an M.S. in Psychology (1981), an M.S. in Computer Science (1986) at the University of Hamburg (Germany), and a Ph.D. in Computer Science/Mathematics (1995) at the University of Zurich (Switzerland). He was a senior lecturer for "usability engineering" in computer science and industrial engineering at the Swiss Federal Institute of Technology (ETH) in Zurich, where later he was heading the Man-Machine Interaction research group (MMI) (1989–1998).

Since 1998, he is Fulltime Professor for "Interactive Systems Design" first at IPO – Centre for User System Interaction Research, and later at the Department of Industrial Design at the Eindhoven University of Technology (TU/e, The Netherlands). From 1999 till 2002, he was director of IPO. He was director of the graduate program at the Department of Industrial Design of the TU/e (2012–2014). He was the head of the Designed Intelligence research group (2006–2015). He was the Swiss representative in the IFIP TC13 on "Human Computer Interaction" (1994–2002) and the chairman of the IFIP WG13.1 on "HCI and Education" (1998–2004). He is now the Dutch representative in the IFIP TC14 on "Entertainment Computing" and the founding vice-chair of this TC14 (2006–2012). Since 2012, he is the IFIP TC14 chair (2013–2015). He was appointed as visiting professor at Kwansei Gakuin University (Japan) (2004–2007); he is senior honorary research fellow of Taicang University Science and Technology Park (since 2012) and guest professor at Jiangnan University (Wuxi, China) (2011–2015) and at East China University of Science and Technology (Shanghai, China) (2013–2016).

He received the German GI-HCI Award for the best Ph.D. in 1997 and the Swiss Technology Award for the BUILD-IT system in 1998. In 2004, he was nominated as member of the "Cream of Science" in the Netherlands (the 200 top-level Dutch researchers) and among the 10 top-level TU/e scientists. Since 2007, he is holder of the IFIP Silver Core Award.

He has over 400 publications in international journals, conference proceedings, books, etc. He acted also as editor and member of the editorial board of several leading international journals. Since 2009, he is co-editor-in-chief of the journal *Entertainment Computing* (Elsevier). He acts regularly as reviewer for national and international funding bodies, individual selection and departmental assessments committees, and large-scale European funding schemas. He was appointed as member of one of the few expert and evaluation panels for the most esteemed European grant from the European Research Council (2010–2014).

Ryohei Nakatsu received the B.S., M.S., and Ph.D. degrees in Electronic Engineering from Kyoto University in 1969, 1971, and 1982, respectively. After joining NTT in 1971, he mainly worked on speech recognition technology. In 1994, he joined ATR (Advanced Telecommunications Research Institute) as Director of ATR Media Integration and Communications Research Laboratories. In 2002, he became Professor at School of Science and Technology, Kwansei Gakuin University. Since March 2008 until December 2014, he was Professor at National University of Singapore (NUS) and was Director of Interactive and Digital Media Institute (IDMI) at NUS. In December 2014, he retired from NUS and came back to Japan. Now he is Adjunct Professor of Kyoto University, Kyoto/Japan, and Visiting Professor of Seika University, Kyoto/Japan. Also he has established two start-up companies and now he is serving as CEO of Hexogon Japan and Executive Director of NT & Associates.

His research interests include interactive media, entertainment technologies, and communication robot/agent.

In 1978, he received Young Engineer Award from the Institute of Electronics, Information and Communication Engineers Japan (IEICE-J), in 1996 the best paper award from the IEEE International Conference on Multimedia, in 1999, 2000, and 2001, Telecom System Award from Telecommunication System Foundation and the best paper award from Virtual Reality Society of Japan, and in 2000 the best paper award from Japanese Society for Artificial Intelligence. Also he received in 2010 IEEE Kansai Section medal, in 2011 IEEE Virtual Reality Service Award, and in 2012 IFIP TC14 Contribution Award.

He is a fellow of the IEEE since 2001 and a life fellow since 2014. Also he is a fellow of the Institute of Electronics, Information and Communication Engineers Japan (IEICE-J) since 2001 and Virtual Reality Society of Japan since 2012. Also he is a honorary member of Japanese Society for Artificial Intelligence. He is a member of various academic societies such as IEEE, IEICE-J, Japanese Society for Artificial Intelligence, and others. He was a chair of IFIP Technical Committee on Entertainment Computing (TC14) since 2006 until 2012 and now is an honorary member of IFIP TC14.

Contributors

Nadjib Achir L2TI – Institut Galilée, University Paris 13, Sorbone Paris Cité, Villetaneuse, France

Craig A. Anderson Center for the Study of Violence, Department of Psychology, Iowa State University, Ames, IA, USA

Lynne Baillie Department of Mathematical and Computer Science, Heriot-Watt University, Edinburgh, UK

Francesco Bellotti University of Genoa, Genoa, Italy

Riccardo Berta University of Genoa, Genoa, Italy

Magdalena Bielenia-Grajewska Intercultural Communication and Neurolinguistics Laboratory, Department of Translation Studies, Institute of English, Faculty of Languages, University of Gdansk, Gdańsk, Poland

Yngvi Björnsson School of Computer Science, Reykjavik University, Menntavegur, Reykjavík, Iceland

Mayumi Bono National Institute of Informatics, National Center of Sciences, Tokyo, Japan

Khaled Boussetta Urbanet, CITI Insa Lyon / INRIA Grenoble Rhône-Alpes, CITI lab, Villeurbanne, France

Ulrich Brandstätter Interface Culture, University of Art and Design Linz, Linz, Austria

Armir Bujari Department of Mathematics, University of Padua, Padua, Italy

Aine Carroll National Rehabilitation Hospital, Dun Laoghaire, Republic of Ireland

Marc Cavazza School of Electronics and Digital Arts, University of Kent, Canterbury, UK

Pablo Cesar CWI: Centrum Wiskunde and Informatica, Amsterdam, The Netherlands

Huang-Ming Chang Department of Industrial Design, Eindhoven University of Technology, Eindhoven, The Netherlands

Fred Charles School of Computing, Teesside University, Middlesbrough, UK

Adrian David Cheok Imagineering Institute, Nusajaya, Malaysia

City University London, London, UK

Chun-Hsiang Chuang Brain Research Center, National Chiao Tung University, Hsinchu, Taiwan

Damien Coyle Intelligent Systems Research Centre, Ulster University, Derry, Northern Ireland, UK

Michael J. Darnell User Experience Center America – Visual Displays, Samsung Research America, Mountain View, CA, USA

Alex Davies Creative Robotics Lab, NIEA, UNSW Art and Design, Sydney, NSW, Australia

Richard Davison School of Computing Science, Newcastle University, Newcastle, UK

Gabriele Ferri School of Informatics and Computing, Indiana University, Bloomington, IN, USA

Doron Friedman The Advanced Reality Lab, The Interdisciplinary Center, Herzliya, Israel

Elisabeth V. C. Friedrich Department of Cognitive Science, University of California San Diego, La Jolla, CA, USA

Marco Furini University of Modena and Reggio Emilia, Modena, Italy

Jhonatan Garcia Intelligent Systems Research Centre, Ulster University, Derry, Northern Ireland, UK

Xiaoyu Ge Artificial Intelligence Group, Research School of Computer Science, The Australian National University, ANU College of Engineering and Computer Science, Canberra, Australia

David Geerts CUO, iMinds/KU Leuven, Leuven, Belgium

Irini Giannopulu Virtual Reality Prism, IHU-A-Brain and Spine Institute (ICM), UPMC, Groupe Hospitalier Pitié-Salpêtrière, Paris, France

Jeffrey H. Goldstein Institute for Cultural Inquiry, Utrecht University, Utrecht, The Netherlands

Mark D. Griffiths International Gaming Research Unit, Psychology Division, Nottingham Trent University, Nottingham, UK

Christopher L. Groves Center for the Study of Violence, Department of Psychology, Iowa State University, Ames, IA, USA

Hayrettin Gurkok Department EWI Research Group, Human Media Interaction (HMI), Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, Enschede, The Netherlands

Junichi Hoshino Graduate School of Systems and Information Engineering¸ Entertainment Computing Laboratory, University of Tsukuba, Tsukuba-shi, Ibaraki, Japan

Kiyoshi Hoshino University of Tsukuba, Tsukuba, Ibaraki, Japan

Hiroshi Ishiguro Osaka University, Osaka, Japan

Leonid Ivonin Department of Industrial Design, Eindhoven University of Technology, Eindhoven, The Netherlands

Arnav Jhala University of California Santa Cruz, Santa Cruz, CA, USA

Tzyy-Ping Jung Swartz Center for Computational Neuroscience, University of California San Diego, San Diego, CA, USA

Kasun Karunanayaka Imagineering Institute, Nusajaya, Malaysia

Stephan Keller Knowledge Technologies Institute, Graz University of Technology, Graz, Austria

Peter Khooshabeh Cognitive Sciences Branch, U.S. Army Research Laboratory, Aberdeen Proving Ground, Aberdeen, MD, USA

Akihiro Kishimoto IBM Research, Ireland Research Lab, Dublin, Ireland

Tatsuya Kitamura Department of Intelligence and Informatics, Faculty of Intelligence and Informatics, Konan University, Kobe, Hyogo, Japan

Noam Knoller Interface Studies Group, Amsterdam School for Cultural Analysis (ASCA), University of Amsterdam, Amsterdam, The Netherlands

Hartmut Koenitz Department of Entertainment and Media Studies, University of Georgia, Athens, GA, USA

Jeffrey Koh Creative Robotics Lab, NIEA, UNSW Art and Design, Sydney, NSW, Australia

Brent J. Lance Translational Neuroscience Branch, U.S. Army Research Laboratory, Aberdeen Proving Ground, Aberdeen, MD, USA

Anatole Lécuyer INRIA, Centre de Recherche Rennes Bretagne-Atlantique, Campus de Beaulieu, Rennes Cedex, France

Augustin Lefebvre Sorbonne Nouvelle University Paris 3, Paris, France

David Levy Imagineering Institute, Nusajaya, Malaysia

Retro Computers Ltd, Luton, UK

Chen Li Intelligent Systems Research Centre, Ulster University, Derry, Northern Ireland, UK

Chin-Teng Lin Brain Research Center, National Chiao Tung University, Hsinchu, Taiwan

Fabien Lotte INRIA, Centre de Recherche Bordeaux Sud-Ouest, Talence Cedex, France

Guillaume Loup LIUM, Université du Maine, Laval Cedex 9, France

Emilie Loup-Escande CRP-CPO (EA7273), Université de Picardie Jules Verne, Amiens, France

Shao-Wei Lu Brain Research Center, National Chiao Tung University, Hsinchu, Taiwan

Dario Maggiorini Department of Computer Science, University of Milan, Milan, Italy

Perla Maiolino Goldsmiths University of London, London, UK

Amar Marathe Translational Neuroscience Branch, U.S. Army Research Laboratory, Aberdeen Proving Ground, Aberdeen, MD, USA

Alke Martens Institute for Computer Science and Electrical Engineering, University of Rostock, Rostock, Germany

Fulvio Mastrogiovanni Department of Informatics, Bioengineering, Robotics, and Systems Engineering, University of Genoa, Genoa, Italy

Rod McCall Environmental Research and Innovation, Luxembourg Institute of Science and Technology, Esch-sur-Alzette, Luxembourg

Karl. A. McCreadie Intelligent Systems Research Centre, Ulster University, Derry, Northern Ireland, UK

Kaleb McDowell Translational Neuroscience Branch, U.S. Army Research Laboratory, Aberdeen Proving Ground, Aberdeen, MD, USA

Jacinta McElligott National Rehabilitation Hospital, Dun Laoghaire, Republic of Ireland

Laurent Mignonneau Interface Culture, University of Art and Design Linz, Linz, Austria

Youichiro Miyake Square Enix Co., Ltd., Tokyo, Japan

Graham Morgan School of Computing Science, Newcastle University, Newcastle, UK

Hiroshi Mori Graduate school of Engineering, Utsunomiya University, Utsunomiya-City, Tochigi, Japan

Yukihiro Morisawa Saitama Institute of Technology, Saitama, Japan

Martin Mueller University of Alberta, Edmonton, AB, Canada

Wolfgang Müller Media Education and Visualization Group (MEVIS), University of Education Weingarten, Weingarten, Germany

Gernot Müller-Putz Institute for Knowledge Discovery, Graz University of Technology, Graz, Austria

Akiyo Nadamoto Department of Intelligence and Informatics, Faculty of Intelligence and Informatics, Konan University, Kobe, Hyogo, Japan

Ryohei Nakatsu Design School, Kyoto University, Kyoto, Japan

Anton Nijholt Department EWI Research Group, Human Media Interaction (HMI), Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, Enschede, The Netherlands

Hiroki Nishino NUS Graduate School for Integrative Sciences and Engineering, National University of Singapore and Graduate School of Media Design, Keio University, Singapore, Singapore

Claudio E. Palazzi Department of Mathematics, University of Padua, Padua, Italy

Viktoria Pammer-Schindler Know-Center GmbH, Graz, Austria

Johanna Pirker Institute of Information Systems and Computer Media, Graz University of Technology, Graz, Austria

Mannes Poel Department EWI Research Group, Human Media Interaction (HMI), Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, Enschede, The Netherlands

Halley M. Pontes International Gaming Research Unit, Psychology Division, Nottingham Trent University, Nottingham, UK

Julie Porteous School of Computing, Teesside University, Middlesbrough, UK

Christian Quadri Department of Computer Science, University of Milan, Milan, Italy

Siddharth Ramakrishnan Neuroscience Program, Department of Biology, University of Puget Sound, Tacoma, WA, USA

Matthias Rauterberg Industrial Design, Eindhoven University of Technology, Eindhoven, The Netherlands

Jochen Renz Artificial Intelligence Group, Research School of Computer Science, The Australian National University, ANU College of Engineering and Computer Science, Canberra, Australia

Mark Rice Institute for Infocomm Research, A*STAR, Singapore, Singapore

Laura Anna Ripamonti Department of Computer Science, University of Milan, Milan, Italy

Elham Saadatian School of Electronics and Computer Science, Interaction, Complexity Group, University of Southampton, Southampton, UK

Paul Sajda Department of Biomedical Engineering, Columbia University, New York, NY, USA

Jose Saldana Department of Electrical Engineering and Communications EINA, Aragon Institute of Engineering Research (I3A), University of Zaragoza, Zaragoza, Spain

Hooman Samani Department of Electrical Engineering, College of Electrical Engineering and Computer Science, National Taipei University, NTUP, Taipei, Taiwan

Reinhold Scherer Institute for Knowledge Discovery, Graz University of Technology, Graz, Austria

Stephan Schiffel School of Computer Science, Reykjavik University, Menntavegur, Reykjavík, Iceland

Christa Sommerer Interface Culture, University of Art and Design Linz, Linz, Austria

Mark Springett Middlesex University, London, UK

Jacqueline Stow National Rehabilitation Hospital, Dun Laoghaire, Republic of Ireland

Mirko Suznjevic Department of Telecommunications, Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, Croatia

Alistair D. Swale Screen and Media Studies, School of Arts, Faculty of Arts and Social Sciences, University of Waikato, Hamilton, New Zealand

Tsjalling Swierstra Department of Philosophy, Faculty of Arts and Social Sciences, Maastricht University, Maastricht, MD, The Netherlands

Naoko Tosa Academic Center for Computing and Media Studies, Kyoto University, Kyoto, Japan

Jon Touryan Translational Neuroscience Branch, U.S. Army Research Laboratory, Aberdeen Proving Ground, Aberdeen, MD, USA

Lars L. Tremper Institute of Sport Science, Technische Universitaet Darmstadt, Darmstadt, Germany

Tomohiro Umetani Department of Intelligence and Informatics, Faculty of Intelligence and Informatics, Konan University, Kobe, Hyogo, Japan

Gary Ushaw School of Computing Science, Newcastle University, Newcastle, UK

Erik van der Spek Technische Universiteit Eindhoven, Eindhoven, The Netherlands

Victoria Vesna Department of Design Media Arts, Art | Sci Center, University of California Los Angeles, Los Angeles, CA, USA

Program in Empowerment Informatics, School of Integrative and Global Majors, University of Tsukuba, Tsukuba, Japan

Yu-Kai Wang Brain Research Center, National Chiao Tung University, Hsinchu, Taiwan

Josef Wiemeyer Institute of Sport Science, Technische Universitaet Darmstadt, Darmstadt, Germany

Karin Wilding Knowledge Technologies Institute, Graz University of Technology, Graz, Austria

Mark H. M. Winands Department of Data Science and Knowledge Engineering, Maastricht University, Maastricht, The Netherlands

Thomas Winkler University of Luebeck, Lübeck, Germany

Wang Xuan Rolls-Royce, Singapore, Singapore

R. Michael Young North Carolina State University, Raleigh, NC, USA

Part I

AI and Games

Game Solvers

1

Akihiro Kishimoto and Martin Mueller

Contents

A. Kishimoto (✉)
IBM Research, Ireland Research Lab, Dublin, Ireland
e-mail: akihirok@ie.ibm.com

M. Mueller
University of Alberta, Edmonton, AB, Canada
e-mail: mmueller@ualberta.ca

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_35

Abstract

Games have simple, fixed rules as well as clear results such as win, draw, or loss. However, developing algorithms for solving games has been a difficult challenge in Artificial Intelligence, because of the combinatorial complexity that the algorithms must tackle.

This chapter presents an overview of successful approaches and results accomplished thus far on game solving. Conducting tree search is a standard way to solve games and game positions. Remarkable progress has been made in developing efficient search algorithms over the last few decades. The chapter describes several standard techniques including $\alpha\beta$ search, proof-number search, and endgame databases.

Keywords

AND/OR tree • Search • $\alpha\beta$ algorithm • Proof-number search • df-pn • Endgame databases

Introduction

Researchers have invested significant resources on research in *two-player zero-sum games with perfect information*. Many popular board games such as chess, checkers, and Go fall into this category, and these games have been used as test beds for testing algorithms in artificial intelligence (AI) research. In this type of zero-sum game, the two players' goals are strictly opposite: when one player wins, the opponent loses. Perfect information means that all information is available to both players. Game positions are typically represented by a board state and the turn to play. Depending on the game, extra information such as the history of the moves played so far may be needed to play and score the result according to the rules.

In a two-player zero-sum game with perfect information, if both players continue playing optimal moves from a position, the final outcome for that position, called the *game-theoretic value* or *value*, is either a win for the first player (i.e., a loss for the second player), or a loss for the first player (i.e., a win for the second player), or a draw (if allowed by the rules of the game). Any finite game can be *solved* in principle since the value of the starting position can be determined by following optimal moves of both players.

Allis (1994) defines three levels of solving a game:

1. *Ultra-weakly solved*. The game-theoretic value of the start position has been determined.
2. *Weakly solved*. A strategy from the start position has been determined to obtain the game-theoretic value of the start position under reasonable computing resources.
3. *Strongly solved*. The game-theoretic value and a strategy have been determined for all legal positions under reasonable computing resources.

There are often significant differences among these three levels in terms of difficulties of achieving the levels of solving. For example, the game of $n \times n$ Hex can be proven to be a win for the first player (Nash 1952). However, winning strategy is only known for $n \leq 10$ (Pawlewicz and Hayward 2014).

For weakly or strongly solving games, the availability of computing resources is restricted to only reasonable ones. In principle, as remarked in Allis (1994), given a large enough amount of time, CPU, and memory resources, games such as chess or Go could be weakly or strongly solved by performing $\alpha\beta$ search or retrograde analysis described later in this chapter. In practice, many games are far too large for a brute force approach, and therefore the development of game solvers that work efficiently under the available resources has been an ongoing challenge.

This chapter gives an overview of the most popular computational approaches for finding strategies for game positions of interest, that is, for at least weakly solving them. Search algorithms are the core of these approaches. In practice, high-performance game solvers combine game-independent search algorithms with game-specific knowledge. While both game-independent and game-specific approaches are necessary to significantly improve the performance of the solvers, the chapter mainly deals with game-independent search algorithms due to their applicability to many games and even to other domains.

Terminology and Definitions on AND/OR Tree and Minimax Tree

Assume that a player p tries to prove a win for a position where p is to play. Then, p must prove that *at least one* of the legal moves leads to a win. However, if it is the opponent's turn to play, p must be able to win against *all* the opponent's moves. This check can be performed recursively, leading to the concept of an AND/OR tree search. The definitions and terminology for AND/OR tree search introduced in this section follow (Kishimoto et al. 2012).

An AND/OR tree is a rooted, finite tree consisting of two types of nodes: *OR* and *AND* nodes. OR nodes correspond to positions where the first player is to play and AND nodes to positions where the second player moves next. A directed edge representing a legal move is drawn from node n to node m if that move played at position n leads to position m.

All nodes except the *root* node have a parent. In this chapter, players are assumed to move alternately. Therefore, each child of an OR node is an AND node, and each child of an AND node is an OR node. In addition, the root is assumed to be an OR node with no loss of generality, but it can be an AND node as well in practice.

Each node in an AND/OR tree has three possible types of values: *win*, *loss*, or *unknown*. As in Kishimoto et al. (2012), the phrase "a node is x" is short for "a node has value x." A node of value win/loss indicates that its corresponding position is a sure win/loss for the first player, respectively. For the sake of simplicity, the value of draw is defined to be the value of loss if possible game outcomes are not explicitly defined. Several techniques for dealing with draws are surveyed in

Kishimoto et al. (2012). A node of value *unknown* indicates that the game-theoretic value of its corresponding position has not yet been proven. To determine its game-theoretic value, such a node must be examined further. *Expanding* a node is the procedure of generating all children of the node, which represent legal moves, and connecting the node to these children by directed edges.

A node with no children is called a *terminal* node. A terminal node is either a win or a loss, as determined by the rules of the game. An *internal* node is a node that has at least one child. A *leaf* node is an unexpanded node with unknown value. A leaf node must be expanded to determine whether it is internal or terminal.

AND/OR tree search aims to *solve* an AND/OR tree, i.e., determine whether the root is a win or a loss. The value of an internal node is calculated from the values of its children. If at least one child of an internal OR node n is a win, then n is also a win. At position n, the first player can play a move that leads to that child and win against the second player. If all children of n are losses, n is a loss since all the legal moves of the first player at position n lead to losing positions. Otherwise, n is unknown. Similarly, an internal AND node n is a loss if at least one of its children is a loss, a win if all its children are wins, and unknown otherwise.

A node that is a win is also called a *proven* node, while a node that has been determined to be a loss is a *disproven* node. A *proof* is a computed win, while a *disproof* is a computed loss.

When a node is computed to be a win, a subtree of an AND/OR tree contains a winning strategy for the first player. Such a subtree is called a *proof* tree and guarantees that node is a win. A proof tree T with root node r is constructed as follows:

1. T contains r.
2. For each internal OR node of T, T contains at least one child.
3. For each internal AND node of T, T contains all children.
4. All terminal nodes in T are wins.

A *disproof tree*, which contains a winning strategy for the second player, is defined in an analogous way, by swapping AND and OR in the definition above, and requiring all terminal nodes to be losses.

Figure 1 illustrates an example of an AND/OR tree. OR nodes are shown by squares and AND nodes are shown by circles. Values win, loss, and unknown are shown by W, L, and U, respectively. Nodes D, F, H, and K are terminal nodes and nodes G, J, and L are leaf nodes. The other nodes are internal nodes for which values are calculated by propagating back the values of the leaf and terminal nodes. For example, node C is a loss because node F is a loss. Irrespective of the value of G, the second player can win against the first player by selecting a move that leads to F. Node I is unknown because node K is a win and node L is unknown. The second player still has a chance to win against the first player by examining L. By following this back-propagation procedure, the value of the root node A is determined to be a win. A proof tree of A is shown with bold lines. Note that for weakly solving the root, AND/OR tree search can ignore nodes that are not part of the

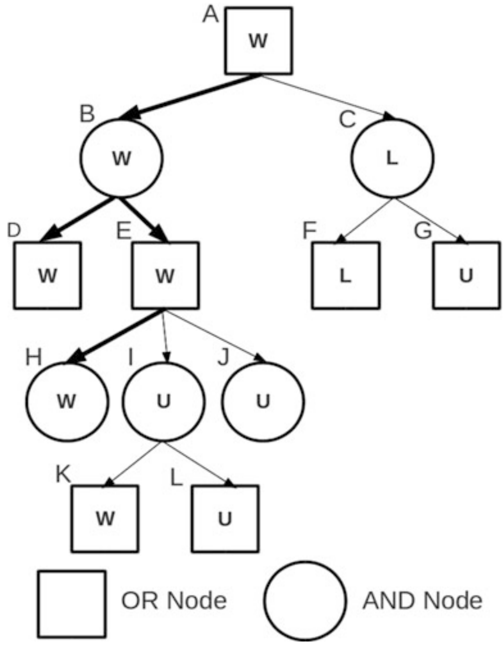


Fig. 1 Example of AND/OR tree

constructed proof tree. For example, in Fig. 1, there is no need to examine the nodes that are not along the bold lines. A high-performance AND/OR tree search algorithm focuses on finding a proof tree as quickly as possible. Assume the standard search space for trees with depth d and branching factor (the number of legal moves at each internal node) b. Also, assume there is only one proof tree in this search space and all terminal nodes are located at depth d. In the worst case, search examines all b^d terminal nodes to find a proof. In contrast, the proof tree contains only $\left\lfloor b^{\frac{d}{2}} \right\rfloor$ terminal nodes.

There may be many proof trees, but finding one is sufficient to solve the root.

In many games, more than one sequence of moves lead to the same position (e.g., Hex and Othello) and a move sequence may lead to a repeated position (e.g., chess and checkers). In other words, the search space of such games can be represented by a directed acyclic graph (DAG) or a directed cyclic graph (DCG). The notion of AND/OR trees, proof, and disproof trees can be generalized for such graphs.

Minimax trees are a generalization of AND/OR trees. Instead of Boolean values, numerical scores are assigned to leaf and terminal nodes. An OR node in such a minimax tree is called a *Max node,* and an AND node is called a *Min node*. The scores are assigned by calling an *evaluation function* that approximates the chance of the first player winning. A larger score indicates that a position is more favorable for the first player. As in AND/OR trees, the score at each internal node of a minimax tree is calculated from the leaf nodes in a bottom-up manner. At an

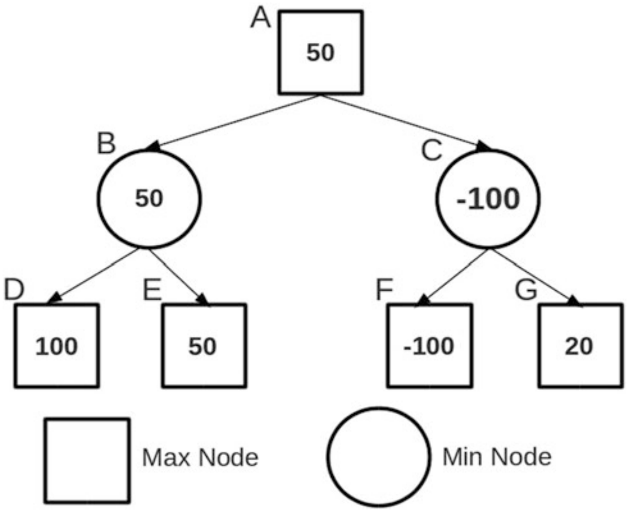


Fig. 2 Example of minimax tree

internal Max node, the first player aims to maximize its advantage by taking the maximum score of all children of that node. At an internal Min node, the second player aims to minimize the advantage of the first player by calculating the minimum score of the children.

Figure 2 illustrates an example of a minimax tree where Max and Min nodes are represented by squares and circles, respectively, and evaluation scores are written inside the squares and circles. The score of the root node *A* becomes 50 by propagating back the scores of leaf nodes *D*, *E*, *F*, and *G*.

The *solution tree* which contains a strategy in the minimax framework is defined in a similar way to the proof tree. For details, see e.g., (de Bruin et al. 1994).

Algorithms for Game Solvers

This section describes general approaches to solve games or game positions. *Forward search* explores a tree from the root until it reaches terminal nodes. *Depth-first search (DFS)* and *best-first search (BFS)* are standard search methods commonly used in many applications including game solvers. While DFS requires only a small amount of memory, it suffers from a combinatorial explosion of its time complexity when the search space is large. In contrast, while BFS tends to explore much smaller search space than DFS, BFS suffers from a combinatorial explosion of its space complexity caused by storing the explored search space in memory. However, game research has revealed that the issue on BFS large memory requirement can be resolved by preserving only important portions of the search space. In addition, BFS can often be enhanced further by incorporating ideas behind DFS.

Backward search is the other approach to search for a solution and deals with the scenario where all terminal nodes can be enumerated with the available computing resources. Backward search starts with terminal nodes and determines the values of positions toward the root. It can be combined with forward search (e.g., Gasser 1996; Schaeffer et al. 2007). The following sections introduce standard forward and backward search algorithms that can be used for game solvers.

The $\alpha\beta$ Algorithm

The $\alpha\beta$ algorithm (Knuth and Moore 1975) is a depth-first forward search algorithm commonly used in many two-player game-playing programs. $\alpha\beta$ is used to determine the next move to play but can be applied to solve games, for example, by assigning a score of ∞ to terminal positions that are wins for the first player, $-\infty$ to terminal positions that are not wins, and other heuristic values to undecided positions. Basic $\alpha\beta$ examines a minimax tree in a depth-first manner with a fixed depth d to compute the best score of the root node. If d is set deep enough, $\alpha\beta$ returns the winner at the root by returning the score of either ∞ or -∞.

$\alpha\beta$ preserves a lower bound α and a upper bound β on the score of a minimax tree. During performing search, the scores of α and β are updated and used for pruning subtrees that are irrelevant for calculating the score at the root. By incorporating good move ordering such as (Schaeffer 1989), $\alpha\beta$ can reduce the search space to examine by increasing the frequency of pruning subtrees. This enables $\alpha\beta$ to search much deeper and contributes to significantly improving the performance of $\alpha\beta$-based game solvers.

Many variants and enhancements to $\alpha\beta$ have been developed over decades (see the literature review such as (Marsland 1986). *Iterative deepening (ID)* (Slate and Atkin 1977) is a standard enhancement to $\alpha\beta$. ID carries out a series of shallower depth-first search before performing direct search to depth d. That is, ID first performs depth-first search from the root with $d = 1$. Then, if ID finds no solution, it performs depth-first search again from the root with $d = 2$, then $d = 3, 4, \cdots,$ and so on. This procedure is repeated until ID either finds a solution, proves that there is no solution, or exhausts resources. Intuitively, because of extra overhead of reexamining previously examined nodes, $\alpha\beta$ combined with ID looks less efficient than basic $\alpha\beta$ that performs direct search to depth d. However, $\alpha\beta$ with ID is empirically more efficient, because previous search results can be used to improve move ordering. As a result, the cost paid for shallower search becomes a small price in order to significantly increase the chance of pruning subtrees for deeper search. For example, the best move calculated in previous shallower search has a high probability that is also the best in deeper search. Examining this move first therefore reduces a large amount of work.

ID is enhanced further by using a *transposition table (TT)* (Greenblatt et al. 1967; Slate and Atkin 1977), a large cache which stores search results of previously examined nodes such as scores, flags indicating whether these scores are exact, lower bounds or upper bounds, search depths, best moves for shallower search, etc. The transposition table is usually constructed as a hash table and takes an advantage of the fact that the search space of many games are a graph where more than one path can lead to the same node, a so-called *transposition*. The transposition table prevents ID from examining the subtree again by merely retrieving and returning a score saved in the TT, when ID encounters a transposition and verifies that the cached score can be reused. In addition, even if the cached score does not result in eliminating the examination of subtrees, such as the case where a cached node has not been explored deep enough, the best move information in the TT can be used to improve move ordering of $\alpha\beta$, significantly reducing search effort.

Proof-Number Search Variants

An essential drawback of $\alpha\beta$ is that search is limited by fixed depth that causes the minimax tree to grow exponentially with the search depth. The drawback can be alleviated by introducing enhancements such as fractional depth and search extensions (e.g., Campbell et al. 2002; Tsuruoka et al. 2002). However, because they cure the problem of exponential tree growth only partially, $\alpha\beta$-based solvers are still unable to solve positions that require deep search. For example, it is difficult to adjust $\alpha\beta$ to solve positions that depend on narrow but deep lines of play, as occur in Go-Moku and checkmating puzzles in chess-like games (Kishimoto et al. 2012). This section describes proof-number search variants that address this problem.

Basic Proof-Number Search

Proof-Number Search (PNS) (Allis et al. 1994) is a best-first forward search algorithm. PNS calculates the proof and disproof numbers that estimate the difficulty of solving nodes. Based on the proof and disproof numbers, PNS aims to examine nodes in simplest-first order. As long as a node is considered to be easy because of a low proof or disproof number, PNS keeps exploring its subtree without any bound on the search depth. This characteristic enables PNS to find narrow but deep proofs or disproofs efficiently.

Formally, the proof number of a node is defined as the minimum number of leaf nodes in its subtree that must be proven to prove that the node is a win, while the disproof number is the minimum number of such leaf nodes that must be disproven to prove that the node is a loss. The smaller the proof/disproof number is, the easier PNS assumes that it is to prove that a node is a win/loss.

Let n be a node with children $n_1, \cdots, n_k$. One proven child suffices to prove a win at an OR node, while all children must be proven to show a win at an AND node

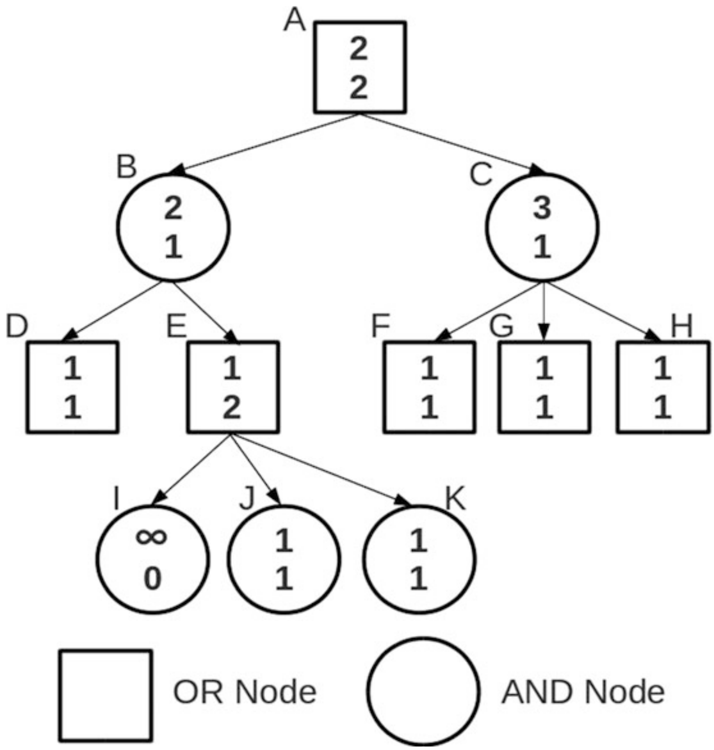


Fig. 3 Example of proof and disproof numbers

(and vice versa for disproof). The proof number **pn**(n) and the disproof number **dn** (n) of node n are therefore calculated as follows:

1. For a proven terminal node n, **pn**(n) $= 0$ and **dn**(n) $= \infty$.
2. For a disproven terminal node n, **pn**(n) $= \infty$ and **dn**(n) $= 0$.
3. For an unknown leaf node n, **pn**(n) $=$ **dn**(n) $= 1$.
4. For an internal OR node n that has children $c_1, \cdots, c_k$,
   pn(n) $= \min($**pn**$(c_1), \cdots,$ **pn**$(c_k))$, **dn**(n) $=$ **dn**$(c_1) + \cdots +$ **dn**(c_k).
5. For an internal AND node n that has children $c_1, \cdots, c_k$,
   pn(n) $=$ **pn**$(c_1) + \cdots +$ **pn**(c_k), **dn**(n) $= \min($**dn**$(c_1) + \cdots +$ **dn**$(c_k))$.

Figure 3 shows an example. The proof and disproof numbers of a node are shown inside that node. The proof number is shown above the disproof number. In this Figure, node I is a terminal node, a loss. Nodes D, F, G, H, J, and K are leaf nodes with proof and disproof numbers initialized to 1. The proof and disproof numbers of internal nodes are calculated by the rule described above. For example, $\mathbf{pn}(E) = \min(\mathbf{pn}(I), \mathbf{pn}(J), \mathbf{pn}(K)) = \min(\infty, 1, 1) = 1$ and $\mathbf{dn}(E) = \mathbf{dn}(I) + \mathbf{dn}(J) + \mathbf{dn}(K) = 0 + 1 + 1 = 2$.

PNS maintains a proof and a disproof number for each node. In the beginning, the AND/OR tree of PNS consists only of the root node and its proof/disproof numbers are initialized to 1. Then, until either the value of the root is determined, or resources are exhausted, PNS repeats the following four steps:

1. Starting from the root, one path in the tree is traversed until PNS finds a leaf node called a *most-promising node (MPN)* (aka *most-proving node*). To find a MPN, PNS selects an AND child with the smallest proof number among all children at internal OR nodes, and an OR child with the smallest disproof number at internal AND nodes. Ties are broken arbitrarily. In practice, game dependent knowledge can sometimes be used here.
2. The MPN is expanded by generating all its children and adding new edges from the MPN to them. The MPN becomes an internal node and the children are new leaf nodes.
3. If the MPN turns out to be a terminal node, the proof and disproof numbers of the MPN are set according to the rules of the game. Otherwise, the proof and disproof numbers of the new leaf nodes are initialized to 1.
4. The proof and disproof numbers of the affected nodes are recomputed along the path from the MPN back to the root.

Figure 4 illustrates an example of the procedure of PNS. Starting from the root node A, PNS traverses path $A \to B \to D$ and finds MPN D. Then, PNS expands D and generates three children of which proof and disproof numbers are initialized to 1. Next, it updates the proof and disproof numbers of D, B, and then A.

Depth-First Proof-Number Search

One inefficiency of basic PNS is that it always propagates back updated proof and disproof numbers from a MPN to the root even if the child with the smallest (dis) proof number remains the same. For example, assume that a MPN is located 100 levels down in the tree from the root. To expand only one leaf node (i.e., MPN), basic PNS must traverse back and forth 100 nodes along the path from the root to the MPN. The depth-first proof-number (df-pn) search algorithm (Nagai 2002) reduces the frequency of reexamining internal nodes. If the search space is a tree, Nagai proves that df-pn is equivalent to PNS in the sense that both algorithms can always select a MPN. This section briefly describes the idea behind df-pn. See (Kishimoto et al. 2012) as well as (Nagai 2002) for precise, detailed descriptions.

The node selection scheme of df-pn is still identical to PNS, which makes df-pn explore the search space in a best-first manner. However, df-pn uses two thresholds to explore the search space in a depth-first manner as well: one for the proof number and the other for the disproof number. If both proof and disproof numbers of a node

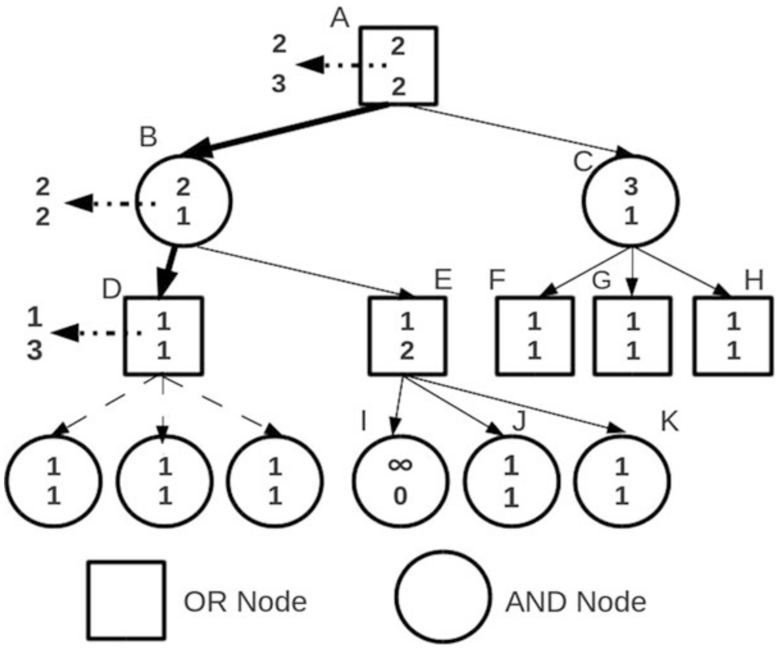


Fig. 4 Example of PNS procedure

n are smaller than the thresholds of proof and disproof numbers, respectively, df-pn continues examining n's subtree without backtracking to n's parent.

Let **thpn**(n) and **thdn**(n) be the thresholds of the proof and disproof numbers at node n, respectively. For example, in Fig. 5, the root OR node A has three children B, C, and D where **pn**(B) = 4, **pn**(C) = 8 and **pn**(D) = 10. B remains on a path to a MPN until **pn**(B) exceeds **pn**(C) = 8, the second smallest proof number among all children. Therefore, the proof number threshold for B, **thpn**(B) = 9, and search can stay in this subtree without updating exact proof numbers until the threshold is reached. When **pn**(B) $\geq 9 =$ **pn**(C) + 1, the MPN switches to a node below C in the tree.

Thresholds of the disproof number at AND nodes are handled analogously with a disproof threshold. Df-pn can remain in the subtree of the child c_1 with smallest disproof number as long as **dn**(c_1) is better than the disproof number of the second best child. For example, in Fig. 5, df-pn sets **thdn**(E) = **dn**(F) + 1 = 4 + 1 = 5 to be able to switch as soon as the disproof number of F becomes strictly smaller than E.

Thresholds for proof numbers of children at AND nodes and disproof numbers of children at OR nodes are set as illustrated below for AND node B in Fig. 5. Assume that **thpn**(B) = 9, and B has two children E and F with **pn**(E) = 3 and **pn**(F) = 1. E is selected for expansion since **dn**(E) < **dn**(F), and search can stay in its subtree until **pn**(E) + **pn**(F) reaches the threshold for the parent,

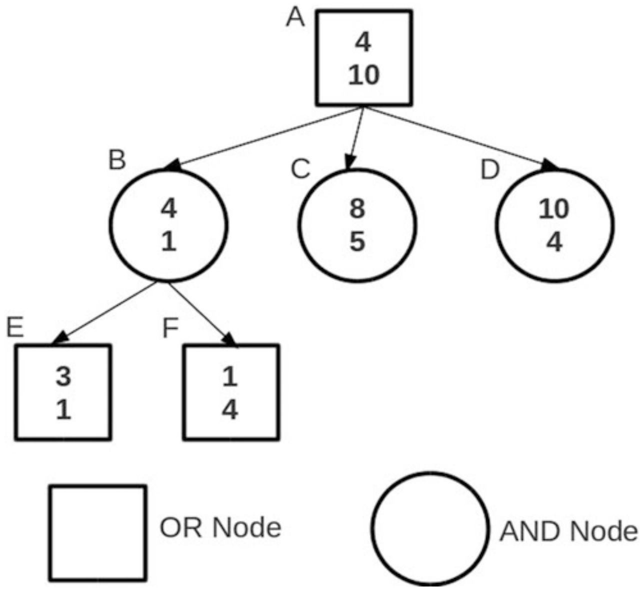


Fig. 5 Illustration of df-pn behavior

thpn(B) = 9. Therefore, for searching E, its threshold **thpn**(E) is set to **thpn**(B) – **pn**(F) = 9–1 = 8.

Reduction of Memory Requirement

The baseline PNS and df-pn algorithms store all nodes that they expand in main memory. For difficult problems, they quickly run out of space. More practical algorithms use a fixed-size transposition table (TT) with a replacement strategy (Breuker 1998; Seo et al. 2001; Nagai 2002). This solves the memory problem but adds overhead from re-expanding nodes that have been overwritten in the table.

The *SmallTreeGC* algorithm (Nagai 1999) is an effective replacement strategy used in many high-performance solvers (Nagai 2002; Kishimoto and Müller 2005; Kishimoto 2010; Kaneko 2010). The TT entry for each node stores the size of the subtree rooted at that node. SmallTreeGC assumes that a small subtree can be reconstructed easily with a small amount of effort even if that subtree is not available in the TT. When the TT becomes full, SmallTreeGC discards a fixed fraction R of the TT entries, starting with those of smallest subtree size. The value of R is determined empirically.

According to Kishimoto et al. (2012), Pawlewicz uses the multiple-probe replacement strategy of Beal and Smith (1996) in their Hex df-pn implementation. Their multiple-probing TT replacement strategy implementation is popular with $\alpha\beta$ chess programs such as Fruit and Stockfish. The technique probes four consecutive entries in a single hash table, and overwrites a TT entry with smallest subtree size among the four.

PN^2 and its variants are another approach which reduces the memory requirement of PNS (Allis 1994; Breuker 1998; Winands et al. 2004). The main idea is to use two levels of PNS. The first level, PN_1, stores a tree starting from the root node. At a leaf node of PN_1, PN^2 invokes the second level of PNS, PN_2, with limited number of nodes expanded and a separate memory allocation. After PN_2 completes search, this PN_2 tree is discarded and only the proof and disproof numbers of the root, corresponding to a PN_1 leaf node, are passed back to PN_1.

A hybrid approach was used to solve the game of checkers (Schaeffer et al. 2007). As in PN^2, a disk-based first-level search called the *front-end manager* was based on PNS. It invoked a memory-only second level search called the *back-end prover*. This back-end used df-pn with a TT and SmallTreeGC.

PNS Variants in Directed Acyclic and Cyclic Graphs

There are three problems to address when PNS and df-pn searches in DAG or DCG. The overestimation problem results from double-counting proof and disproof numbers of nodes in a DAG that can be reached along multiple paths (Allis et al. 1994). The Graph-History Interaction (GHI) problem (Palay 1983; Campbell 1985), which can occur both in $\alpha\beta$ and PNS variants, originates from incorrect handling of cycles in DCG. The infinite loop problem (Kishimoto and Müller 2003; 2008) refers to the phenomenon that df-pn may loop forever without expanding any new nodes in DCG. Current solutions to these problems are summarized in Kishimoto et al. (2012).

Search Enhancements

Although PNS variants are already powerful without domain knowledge, high-performance game solvers incorporate many search enhancements to be able to solve difficult game positions. This section describes some of the well-known methods. See Kishimoto et al. (2012) for comprehensive survey.

The proof and disproof numbers of the leaf node are always set to 1 in their original definition. However, in practice, some leaf nodes are easier to (dis)prove than others. With *heuristic initialization*, leaf nodes are heuristically initialized with proof and disproof numbers (Allis 1994). One example for this is using the number of legal moves at a leaf node, which can be calculated with little overhead in some games (Allis et al. 1994). Another popular approach is using domain-specific heuristic functions (Nagai 2002). One important remark is that the thresholds of df-pn must be increased to reduce the overhead of internal node re-expansions when heuristic initialization is combined with df-pn, as in df-pn$^+$ (Nagai 2002; Kishimoto and Müller 2005).

If more than one promising sibling exist and the search space does not fit into the TT, df-pn sometimes suffers from thrashing the TT by fast switches between subtrees. The $1 + \varepsilon$ trick (Pawlewicz and Lew 2007) increases the threshold when df-pn selects the best child. This enables df-pn to stay longer in the subtree of one child, without frequently switching to other promising siblings.

Given a proven node n, tree *simulation* (Kawano 1996) performs a quick proof check for an unknown node m that looks similar to n. The similarity of positions is usually defined in a game-specific way. Unlike normal search such as df-pn that generates all legal moves, simulation is restricted to generate moves borrowed from n's proof tree at each OR node and checks if a proof tree of m can be constructed in a similar way to n's proof tree. If simulation succeeds, m is proven as well. Otherwise, m's value remains unknown and normal search is performed. Considering that any search must examine all nodes in m's proof tree even in an ideal case, simulation makes almost the smallest effort in proving m, therefore, requires much less effort than normal search.

Schaeffer et al. (2007)'s PNS-based solver estimates the game-theoretic value of a node with high confidence by using scores from their $\alpha\beta$-search-based game-playing program. Let s be such a score of node n, and th be a threshold for an iterative PNS algorithm. If $s \geq th$, n is regarded as a terminal node with value *likely win*. Similarly, if $s \leq -th$, n is considered to be a *likely loss*. As in iterative deepening, the algorithm starts with a small value of th and gradually increases th after constructing each heuristic proof tree. When the root is solved with $th = \infty$, a true proof tree is constructed, where all leaf nodes are true values, and the proof is complete.

Endgame Databases

In many *converging* games (Allis 1994) such as chess and checkers, the number of pieces on the board decreases as the game progresses. This indicates that it is often feasible to enumerate all *endgame positions* that occur close to the end of the game. In this case, the game-theoretic values of these endgame positions can be precomputed and saved as *endgame databases* that map a position to its corresponding game-theoretic value. If the endgame databases contain the game-theoretic values of all the legal positions for a game, that game is strongly solved. As an example, the database for the game of Awari calculated by Romein and Bal (2003) contains the exact scores of all legal Awari positions.

Retrograde analysis (Bellman 1965; Thompson 1986) is a standard backward search algorithm to systematically construct endgame databases, starting from terminal nodes and progressing toward the root. Retrograde analysis has been successfully used in many domains (Lake et al. 1994; Schaeffer et al. 2003; Romein and Bal 2003). Additionally, endgame databases contribute to significantly improving the ability of game solvers that employ forward search (Schaeffer et al. 2007), because forward search can obtain the game-theoretic value of a node without deeper search.

Assume that the game-theoretic value of a game is either a win, a loss, or a draw, and a repeated position is defined to be a draw by the rule of that game. Then, one way to implement retrograde analysis is summarized as follows:

1. Let S be a set containing all positions, that may include unreachable ones from the root.
2. Initialize: assign the game-theoretic value determined by the rules of the game to all terminal nodes and a value of *unknown* to all other positions.
3. For each node $n \in S$ with value unknown, check if n's value can be determined from its children. For example, OR node n is a win if n has at least one child in S that is a win.
4. Repeat step 3 until the number of nodes with value unknown no longer decreases.
5. Mark the nodes of value unknown as draws since these nodes are either unreachable from the root, or no win or loss can be forced by either player.

When constructing a database for a difficult game in practice, retrograde analysis requires large CPU and storage resources. Implementations typically use parallel machines as well as large amounts of both main memory and hard disk space. There are several approaches that achieve efficient parallelism and reduce disk I/O. For example, as in paging of the operating system, retrograde analysis determines which part of the databases should be preserved in the main memory to alleviate the overhead of disk I/O (Lake et al. 1994; Romein and Bal 2003). Schaeffer et al. (2003) not only compress databases space-efficiently but also decompress the databases in real-time so that forward search can use them both time- and memory-efficiently. Techniques to reduce the main memory requirement at the cost of a small computational overhead are presented in Lincke (2002); Romein and Bal (2003).

To initiate parallelism, Romein and Bal (2003) present an asynchronous distributed-memory parallel algorithm that overlaps database computation and processor communication via network. In contrast, Lake et al. (1994) split the search space into a set of small slices that can be solved easily and independently. The correctness of databases must be verified, since both software and hardware errors including disk and network errors may occur during a large-scale computation (Schaeffer et al. 2003; Romein and Bal 2003).

Other Approaches

This section briefly describes other approaches related to forward search.

Threat-Based Approaches

Threats are moves to which a player must reply directly to win, or to avoid an immediate loss. If threats exist in a position, threat-space search (Allis 1994; Allis et al. 1996) considers only threats and safely ignores other legal moves. Threat-space search can significantly reduce the search space to explore, because the number of threats is usually much smaller than that of the legal moves. The idea behind threat-space search is used in solvers such as Henderson et al. (2009) and Kishimoto and Müller (2005).

Different levels of threat sequences, which are a generalization of direct threats, can be detected by using *pass* moves. Intuitively, moves are limited by using the information on how many moves a player needs to make in a row so that that player has a forced win. Examples of such methods are λ-search (Thomsen 2000) and generalized threats search (Cazenave 2002). These techniques can be applied to both $\alpha\beta$ and df-pn (Nagai 2002; Yoshizoe et al. 2007).

Early Win/Loss Detection

Game-specific features can sometimes be used to statically detect whether a player is winning or losing. Such static win/loss detection can significantly reduce the size of a proof tree, as well as the search space explored by forward search. Examples include virtual connections in the game of Hex (Anshelevich 2002; Henderson et al. 2009) and the detection of eye space and potential eye space in the life and death problem in Go (Wolf 1994; Kishimoto and Müller 2005).

Detecting a dominance relationship between positions can also contribute to recognizing wins or losses early. An important example is the checkmating problem in shogi (Japanese chess) (Seo 1999). For example, assume that forward search proves that node n is a win and then encounters unproven node m that dominates n. Then, m is also a win because the first player can copy the winning strategy from n.

Monte Carlo Tree Search Solver

Monte Carlo Tree Search (MCTS) is a forward search algorithm that has achieved remarkable success in playing the game of Go and many other games where accurate evaluation functions are difficult to develop (see the chapter on computer Go, written by Yoshizoe and Müller). MCTS combines tree search with Monte Carlo simulation that is used to evaluate a leaf node. In addition to Monte Carlo simulation results, if game-theoretic values of wins or losses are available, the MCTS-Solver propagates back these values (Winands et al. 2008). This modification enables MCTS-Solver to solve positions.

Probability Propagation

Instead of using proof and disproof numbers as the criteria to explore an AND/OR tree, probability propagation (PP) performs best-first forward search based on a probability of a first player win for each node in an AND/OR tree (Pearl 1984; Stern et al. 2007). Enhancements similar to the ones for PNS can be incorporated into PP, including heuristic initialization (Stern et al. 2007), transposition tables, and two-level search (Saffidine and Cazenave 2013).

Results Accomplished on Solving Games

The techniques described thus far contributed to solving many nontrivial, popular games that people actually play, such as Awari (Romein and Bal 2003), checkers (Schaeffer et al. 2007), Connect-Four (see http://tromp.github.io/c4/c4.html and (Allis 1988), and Go-Moku (Allis et al. 1996).

Researchers use games with smaller boards as a research test bed to develop new algorithms. Examples of games solved in this way are 5 × 5 Go (van der Werf et al. 2003), 6 × 6 Othello (see http://www.tothello.com/), 10 × 10 Hex (Pawlewicz and Hayward 2014), and 5 × 6 Amazons (Song and Müller 2014).

In chess, endgame databases are constructed for most positions with 3–7 pieces (see http://chessok.com/?page_id=27966). Puzzle or endgame solvers based on PNS variants have been developed to achieve state-of-the-art performance in many games, such as chess (Breuker 1998), checkers (Schaeffer et al. 2007), Hex (Arneson et al. 2011), life and death in Go (Kishimoto and Muller 2005), and tsume-shogi (Seo et al. 2001; Nagai 2002; Kishimoto 2010). In particular, research on tsume-shogi has achieved remarkable success by testing with many difficult tsume-shogi problems created by human experts. State-of-the-art solvers based on df-pn can solve all the existing hard problems with solution sequences longer than 300 steps, including Microcosmos with 1525 steps.

Sophisticated game solvers combine many of the techniques discussed. For example, the checkers solver of Schaeffer et al. (2007) incorporates PNS, df-pn, $\alpha\beta$, endgame databases, and many other game-independent and game-specific enhancements.

Conclusions

This chapter presented an overview of the techniques for weakly or strongly solving games or game positions. Game positions can be solved by forward search, backward search, or a combination. As standard algorithms to perform forward search, the chapter introduced $\alpha\beta$ and PNS variants. Backward search for building endgame databases uses retrograde analysis. Additionally, the chapter gave an overview of other approaches including threat-space search, MCTS and PP, followed by a summary of the results accomplished thus far on solving games.

Recommended Reading

L.V. Allis, *A Knowledge-Based Approach of Connect Four: The Game is Over, White to Move Wins*. Master's thesis, Vrije Universiteit Amsterdam, Amsterdam, 1988. Report No. IR-163

L.V. Allis, *Searching for Solutions in Games and Artificial Intelligence*. PhD thesis, University of Limburg, Maastricht, 1994
L.V. Allis, M. van der Meulen, H.J. van den Herik, Proof-number search. Artif. Intell. **66**(1), 91–124 (1994)
L.V. Allis, M.P.H. Huntjes, H.J. van den Herik, Go-moku solved by new search techniques. Comput. Intell. **12**(1), 7–23 (1996)
V. Anshelevich, A hierarchical approach to computer Hex. Artif. Intell. **134**(1–2), 101–120 (2002)
B. Arneson, R.B. Hayward, P. Henderson, Solving Hex: beyond humans, in *Computers and Games 2010*, ed. by H.J. van den Herik, H. Iida, A. Plaat. Lecture Notes in Computer Science (LNCS), vol. 6515 (Springer, Berlin, 2011), pp. 1–10
D. Beal, M.C. Smith, Multiple probes of transposition tables. ICCA J. **19**(4), 227–233 (1996)
R. Bellman, On the application of dynamic programming to the determination of optimal play in chess and checkers. Proc. Natl. Acad. Sci. U. S. A. **53**, 244247 (1965)
D.M. Breuker, *Memory Versus Search in Games*. PhD thesis, Universiteit Maastricht, Maastricht, 1998
M. Campbell, The graph-history interaction: on ignoring position history, in *Proceedings of the 1985 ACM Annual Conference on the Range of Computing: Mid-80's Perspective*, (ACM, New York, 1985). pp. 278–280
M. Campbell, A.J. Hoane Jr., F. Hsu, Deep Blue. Artif. Intell. **134**(1–2), 57–83 (2002)
T. Cazenave, A generalized threats search algorithm, in *Computers and Games 2002*, ed. by J. Schaeffer, M. Müller, Y. Björnsson. Lecture Notes in Computer Science (LNCS) (Springer, Heidelberg, 2002), pp. 75–87
A. de Bruin, W. Pijls, A. Plaat, Solution trees as a basis for game tree search. ICCA J. **17**(4), 207–219 (1994)
R. Gasser, Solving Nine Men's Morris, in *Games of No Chance*, ed. by R.J. Nowakowski. MSRI Publications, vol. 29 (Cambridge University Press, Cambridge, 1996), pp. 101–113
R. Greenblatt, D. Eastlake, S. Croker, The Greenblatt chess program, in *Proceedings of the Fall Joint Computer Conference*, 1967, pp. 801–810. Reprinted (1988) in *Computer Chess Compendium*, ed. by D.N.L. Levy (Batsford, London), pp. 56–66
P. Henderson, B. Arneson, R.B. Hayward, Solving 8 $\times$ 8 Hex, in *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI'09)*, ed. by C. Boutilier (AAAI Press, Pasadena, 2009), pp. 505–510
T. Kaneko, Parallel depth first proof number search, in *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence, (AAAI' 10)*, ed. by M. Fox, D. Poole (AAAI Press, Menlo Park, 2010), pp. 95–100
Y. Kawano, Using similar positions to search game trees, in *Games of No Chance*, ed. by R.J. Nowakowski. MSRI Publications, vol. 29 (Cambridge University Press, Cambridge, 1996), pp. 193–202
A. Kishimoto, Dealing with infinite loops, underestimation, and overestimation of depth-first proof-number search, in *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence, (AAAI' 10)*, ed. by M. Fox, D. Poole (AAAI Press, 2010)
A. Kishimoto, M. Müller, About the completeness of depth-first proof-number search, in *Computers and Games 2008*, ed. by H.J. van den Herik, X. Xu, Z. Ma, M.H.M. Winands. Lecture Notes in Computer Science, vol. 5131 (Springer, Heidelberg, 2008), pp. 146–156
A. Kishimoto, M. Müller, Df-pn in Go: an application to the one-eye problem, in *Advances in Computer Games 10 (ACG'03): Many Games, Many Challenges*, ed. by H.J. van den Herik, H. Iida, E.A. Heinz (Kluwer, Boston, 2003), pp. 125–141
A. Kishimoto, M. Müller, Search versus knowledge for solving life and death problems in Go, in *Proceedings of the 20th National Conference on Artificial Intelligence (AAAI'05)*, ed. by M.M. Veloso, S. Kambhampati (AAAI Press/MIT Press, Menlo Park, 2005), pp. 1374–1379
A. Kishimoto, M. Winands, M. Müller, J.-T. Saito, Game-tree search using proof numbers: the first twenty years. ICGA J. **35**(3), 131–156 (2012)

D.E. Knuth, R.W. Moore, An analysis of alpha-beta pruning. Artif. Intell. **6**(4), 293–326 (1975)
R. Lake, J. Schaeffer, P. Lu, Solving large retrograde analysis problems using a network of workstations. Advances in Computer Games **7**, 135–162 (1994)
T. Lincke, *Exploring the Computational Limits of Large Exhaustive Search Problems*. PhD thesis, ETH Zürich, 2002
T. Marsland, A review of game-tree pruning. ICCA J. **9**(1), 3–19 (1986)
A. Nagai, A new depth-first-search algorithm for AND/OR trees. Master's thesis, The University of Tokyo, Tokyo, 1999
A. Nagai, *Df-pn Algorithm for Searching AND/OR trees and its Applications*. PhD thesis, The University of Tokyo, 2002
J. Nash, Some games and machines for playing them. *Technical Report D-1164*, Rand Corp., 1952
A.J. Palay, *Searching with Probabilities*. PhD thesis, Carnagie Mellon University, 1983
J. Pawlewicz, R. Hayward, Scalable parallel depth first proof number search, in *Computers and Games (CG 2013)*, vol. 8427. *Lecture Notes in Computer Science* (Springer, 2014), pp. 138–150
J. Pawlewicz, L. Lew, Improving depth-first pn-search: 1+ε trick, in *Proceedings of the 5th Computers and Games Conference (CG'06)*, ed. by H.J. van den Herik, P. Ciancarini, H.H.L.M. Donkers. Lecture Notes in Computer Science, vol. 4630 (Springer, Heidelberg, 2007), pp. 160–170
J. Pearl, *Heuristics: Intelligent Search Strategies for Computer Problem Solving* (Addison-Wesley, Reading, 1984)
J.W. Romein, H. Bal, Solving the game of Awari using parallel retrograde analysis. IEEE Comput. **36**(10), 26–33 (2003)
A. Saffidine, T. Cazenave, Developments on product propagation. in *Computer Games 2013*, vol. 8427. *Lecture Notes in Computer Science*, 2013, pp. 100–109
J. Schaeffer, The history heuristic and alpha-beta search enhancements in practice. IEEE Trans. Pattern Anal. Mach. Intell. **11**(1), 1203–1212 (1989)
J. Schaeffer, Y. Björnsson, N. Burch, R. Lake, P. Lu, S. Sutphen, Building the checkers 10-piece endgame databases, in *Advances in Computer Games 10*: Many Games, ed. by H.J. van den Herik, H. Iida, E.A. Heinz. (Kluwer, Boston, 2003), pp. 193–210
J. Schaeffer, N. Burch, Y. Björnsson, A. Kishimoto, M. Müller, R. Lake, P. Lu, S. Sutphen, Checkers is solved. Science **317**(5844), 1518–1522 (2007)
M. Seo, On effective utilization of dominance relations in tsume-shogi solving algorithms, in *Proceedings of the 8th Game Programming Workshop*, 1999, pp. 137–144 (in Japanese)
M. Seo, H. Iida, J.W.H.M. Uiterwijk, The PN*-search algorithm: Application to tsume shogi. Artif. Intell. **129**(1–2), 253–277 (2001)
D.J. Slate, L.R. Atkin, Chapter 4. Chess 4.5 – Northwestern University chess program, in *Chess Skill in Man and Machine*, ed. by P.W. Frey (Springer, New York, 1977), pp. 82–118
J. Song, M. Müller, An enhanced solver for the game of Amazons, 2014, in *Accepted for IEEE Transactions on Computational Intelligence and AI in Games (TCIAIG),* 12 pp
D. Stern, R. Herbrich, T. Graepel, Learning to solve game trees, in *Proceedings of the 24th International Conference of Machine Learning (ICML)*, 2007, pp. 839–846
K. Thompson, Retrograde analysis of certain endgames. ICCA J. **9**(3), 131–139 (1986)
T. Thomsen, Lambda-search in game trees – with application to Go. ICGA J. **23**(4), 203–217 (2000)
Y. Tsuruoka, D. Yokoyama, T. Chikayama, Game-tree search algorithm based on realization probability. ICGA J. **25**(3), 132–144 (2002)
E.C.D. van der Werf, H.J. van den Herik, J.W.H.M. Uiterwijk, Solving Go on small boards. ICGA J. **26**(2), 92–107 (2003)
M.H.M. Winands, J.W.H.M. Uiterwijk, H.J. van den Herik, An effective two-level proof-number search algorithm. Theor. Comput. Sci **313**(3), 511–525 (2004)
M.H.M. Winands, Y. Björnsson, J.-T. Saito, Monte-Carlo tree search solver, in *Proceedings of the 6th Computers and Games Conference (CG'08)*, ed. by H.J. van den Herik, X. Xu, Z. Ma,

M.H.M. Winands. Lecture Notes in Computer Science, vol. 5131 (Springer, Berlin, 2008), pp. 25–36

T. Wolf, The program GoTools and its computer-generated tsume Go database, in *1st Game Programming Workshop in Japan (Hakone)*, 1994

K. Yoshizoe, A. Kishimoto, M. Müller, Lambda depth-first proof number search and its application to Go, in *Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI-07)*, ed. by M.M. Veloso (2007), Morgan Kaafmann, San Francisco, pp. 2404–2409

General Game Playing

2

Yngvi Björnsson and Stephan Schiffel

Contents

Y. Björnsson (✉) • S. Schiffel
School of Computer Science, Reykjavik University, Menntavegur, Reykjavík, Iceland
e-mail: yngvi@ru.is; stephans@ru.is

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_34

Abstract

Game playing is one of the oldest areas of investigation in artificial intelligence (AI) and has been at the forefront of AI research ever since the birth of the first computers, over half a century ago. The research focus was initially on developing general approaches for game playing, but gradually shifted towards building high-performance game-playing systems capable of matching wits with the strongest humans in the world in individual games. To renew interest in more general approaches to computer game playing, the AI community established the International General Game Playing Competition (IGGPC) in 2005, which has run annually ever since. General game playing (GGP) has in the decade since established itself as a fascinating research area, posing numerous interesting research challenges to a wide range of artificial intelligence subdisciplines. In here, we review the progress made in the field so far and highlight mainstay techniques used in contemporary state-of-the-art GGP agents.

Keywords

Artificial neural networks (ANN) • Competitions and events • Custom-made interpreters • Game description language (GDL) • Game specifications • Game-tree search in • Game-tree search, in GGP • GDL. *See* Game description language (GDL) • General game playing (GGP) • GGP. *See* General game playing (GGP) • Goblin and Ogre • KuhlPlayer • Learning heuristic evaluation functions • Learning simulation search control • Magline agent • Monte Carlo Tree Search in GGP • Reasoning

Introduction

In *General game playing (GGP)* the goal is to create intelligent agents that can automatically learn how to skillfully play a wide variety of games, given only the descriptions of the game rules. This requires that the agents learn diverse game-playing strategies without any game-specific knowledge being provided by their developers. A successful realization of this task poses interesting research challenges for artificial intelligence subdisciplines such as knowledge representation, agent-based reasoning, heuristic search, computational intelligence, and machine learning.

The two core components of any game-playing program are *search* and *evaluation*. The former provides the ability to think ahead, whereas the latter provides a mechanism for assessing the merits of the game positions that arise during the search. Domain-dependent knowledge plays an important part in both components, in particular for game-position evaluation, understandably, but also for providing effective search guidance. The main challenge faced by general game-playing systems, as opposed to game-playing systems for playing one specific game, is that the relevant game-specific knowledge required for expert-level play, whether for the search or the evaluation, must be effectively discovered during game play. This distinction provides unique challenges for GGP research.

In this chapter, we give an overview of the field of GGP. It has been an active research area over the last decade, or ever since the establishment of the International GGP Competition (IGGPC) (Genesereth and Love 2005) revitalized interest into general problem-solving approaches to computer game playing. The field has advanced steadily since and noteworthy research progress been made. However, given the large body of recent research into GGP systems, we out of necessity limit our overview for the most part to research results and techniques that have demonstrated themselves effective in competitive GGP agents. Main results are highlighted and references to more in-depth discussions given were relevant.

The chapter is structured as follows. In the next section, we show how games are specified in GGP and describe common approaches for representing and interpreting the game specifications. In the section that follows, we give an overview of the search techniques used in GGP agents for thinking ahead and highlight the unique challenges they face, in particular the need to autonomously learn relevant game-specific knowledge on the fly. The next three sections discuss in more detail how the agents learn search control, evaluation functions, and game structures, respectively. Finally, before concluding, we review competition results and provide references to useful online resources.

Game Specifications

GGP agents are capable of playing many different games. The rule set of a game to play is communicated to the agents in a formal language called *game description language (GDL)* and the agents then have a built-in mechanism to interpret and reason with the game description, enabling them to play a legitimate game.

Game Description Language

The *game description language (GDL)* (Genesereth et al. 2005; Love et al. 2008) is a declarative first-order logic-based language strongly resembling Datalog (Abiteboul et al. 1995), allowing function constants, negation, and recursion (in a restricted form). Its syntax, however, is adopted from the knowledge interface format (KIF) language (Genesereth et al. 1992).

The expressiveness of GDL allows for a wide range of deterministic, perfect information, single- or multiplayer games to be described. By definition, all games are simultaneous-move based; however, turn-based games can be modeled by having the players who do not have the turn return a special no-operation move (*noop*). Special relations are used to specify the setup of the initial game state, the rules for generating and executing legal moves in a given state, as well as rules for detecting and scoring terminal states. A game state is defined by the set of propositions that are true in that state (using the closed-world assumption, that is, the only true facts are those that are known to be so, others are assumed false). Figure 1 shows an example game written in GDL.

```
(role xplayer) (role oplayer)

(init (cell 1 1 b)) (init (cell 1 2 b)) (init (cell 1 3 b))
(init (cell 2 1 b)) (init (cell 2 2 b)) (init (cell 2 3 b))
(init (cell 3 1 b)) (init (cell 3 2 b)) (init (cell 3 3 b))
(init (control xplayer))

(<= (legal ?w (mark ?x ?y)) (true (cell ?x ?y b)) (true (control ?w)))
(<= (legal xplayer noop)    (true (control oplayer)))
(<= (legal oplayer noop)    (true (control xplayer)))

(<= (next (control xplayer)) (true (control oplayer)))
(<= (next (control oplayer)) (true (control xplayer)))
(<= (next (cell ?m ?n ?w))   (true (cell ?m ?n ?w)) (distinct ?w b))
(<= (next (cell ?m ?n x)) (does xplayer (mark ?m ?n)) (true (cell ?m ?n b)))
(<= (next (cell ?m ?n o)) (does oplayer (mark ?m ?n)) (true (cell ?m ?n b)))
(<= (next (cell ?m ?n b))
         (does ?w (mark ?j ?k)) (true(cell ?m ?n b)) (distinct ?m ?j))
(<= (next (cell ?m ?n b))
         (does ?w (mark ?j ?k)) (true(cell ?m ?n b)) (distinct ?n ?k))

(<= (line ?x) (true (cell ?m 1 ?x)) (true (cell ?m 2 ?x)) (true (cell ?m 3 ?x)))
(<= (line ?x) (true (cell 1 ?n ?x)) (true (cell 2 ?n ?x)) (true (cell 3 ?n ?x)))
(<= (line ?x) (true (cell 1 1 ?x))  (true (cell 2 2 ?x))  (true (cell 3 3 ?x)))
(<= open (true (cell ?m ?n b)))

(<= terminal    (line x))
(<= terminal    (line o))
(<= terminal    (not open))

(<= (goal  xplayer 100) (line x))
(<= (goal  xplayer 50)  (not (line x)) (not (line o)) (not open))
(<= (goal  xplayer 0)   (line o))
(<= (goal  oplayer 100) (line o))
(<= (goal  oplayer 50)  (not (line x)) (not (line o)) (not open))
(<= (goal  oplayer 0)   (line x))
```

Fig. 1 GDL description of *Tic-Tac-Toe.* The **role** relation indicates the players participating in the game (arbitrary many roles can be declared; however, once declared roles are fixed throughout the game). The **init** relation states the propositions that are true in the initial state. The **legal** and **next** relations determine and execute legal moves, respectively; the latter returns (collectively) all propositions that are true in the next state were the move provided in the **does** relation to be executed (**true** an **distinct** are general relations checking whether a proposition is true in the current game state and whether two values are distinct, respectively). The *line* and *open* are user-defined relations. The **terminal** relation checks if the game is over, in which case the **goal** relation gives the game outcome for each player in the range [0,100] (need not be zero-sum). The goal relation also applies to nonterminal states where it returns the intermediate score a player has already secured (e.g., number of tricks in a card game), although not relevant in this example

In addition to adhering to syntactic constraints, a valid GDL game description must be semantically *well-formed* (as described in the GDL specifications), that is: the game described must be finite, there must exist at least one move for each player in all nonterminal states (can be *noop*), goal values for all players must be monotonic (i.e., a player's goal value never decrease as the game progresses), and the game must be winnable.

Several extensions have been proposed to GDL. Most significant is GDL-II (GDL with imperfect information), which allows for incomplete information and nondeterministic games to be additionally described. This is achieved by introducing two new special relations, **sees** and **random**, as well as additional semantic constraints of what constitutes a well-formed game (Thielscher 2010, 2011).

Reasoning in GDL

GGP agents must accurately interpret the GDL game descriptions. The interpretation is typically encapsulated in a separate module, which in turn provides a generic interface for the agent to do state-space manipulations, for example, calls for generating and executing legal moves and for detecting end of game conditions. There are different ways of achieving this, but the three mainstream approaches are: (a) using a custom-made GDL interpreter; (b) translate the GDL descriptions into Prolog and then use an off-the-shelf Prolog interpreter; or, (c) translate GDL into an alternative representation that the agent knows how to interpret and reason with in an efficient manner.

Custom-Made Interpreters

Building a robust and efficient GDL interpreter from scratch is a highly involved task. Thankfully there exist several custom-made GDL interpreters written in popular imperative programming languages (Halderman et al. 2006; Schreiber 2013; Schultz et al. 2008); however, they still tend to be in their infancy. They are typically modeled after Prolog interpreters, for example, by using SLD-NF (Selective Linear Definite–Negation as Failure) resolution (Apt and van Emden 1982) for answering queries; however, they still lack many of the state-of-the-art optimization techniques commonly found in Prolog interpreters. There is some ongoing work on developing high-performance GDL interpreters (Schofield and Saffidine 2013; Waugh 2009); however, until that fully materializes, this approach still remains noncompetitive to the other two approaches efficiency wise. Nonetheless, as it allows for a convenient and straightforward integration a GDL reasoner into new GGP agents, this approach provides a good starting point for newcomers to the field.

Translation to Prolog

A common approach by established GGP agents is to translate the GDL game descriptions into Prolog programs and offload the interpretation to an (highly optimized) off-the-shelf Prolog engine. The translation of GDL to Prolog is for the most part straightforward as both languages are first-order-logic based and have a similar syntax. The semantics of GDL and Prolog do, however, differ somewhat and some precautions are necessary during the translation process to ensure that the correct semantic interpretation of GDL is preserved. For example, the ordering of clauses is inconsequential for the semantic in GDL (which is, as Datalog, fully declarative), whereas clause ordering is essential for determining a program's semantic in Prolog. In particular, one must make certain that negated clauses in generated Prolog implication rules are ordered such that their variable arguments are surely grounded before execution. In contrast to the first approach, the integration and interfacing of a Prolog engine into a GGP project can be a somewhat more involved task. For example, not all Prolog engines provide a convenient or efficient application-programming interface (API) to programs written in an imperative programming language. Another downside is that some of the publicly available Prolog engines are nonreentrant and as such cannot be safely used by host programs using thread-based parallelism.

Alternative Representations

An alternative approach is to translate the GDL game description into a representation other than Prolog that can be efficiently managed by the GGP agent. A few GGP agents do this, for example, by translating GDL into propositional nets (Cox et al. 2009; Schkufza et al. 2008) or binary decision diagrams (Bryant 1985; Kissmann and Edelkamp 2009). Whereas this can result in highly efficient reasoners, the main drawback is that such translations typically require the grounding of all possible GDL logic clauses, often leading to an exponential blowup of the size of the representation. Such an approach is thus feasible for only a limited subset of games (typically the simpler ones). GGP agents using this approach do thus also rely on one of the other two approaches as a fallback.

Discussion

General-purpose GDL reasoners are at least two orders of magnitude slower than their game-specific counterparts, although a clever representation such as propositional nets or binary decision diagrams can help bridge that gap for a subset of games. The current state-of-the-art in GDL reasoning is to use a high-performance propositional net when possible, but to fallback on one of the other two GDL reasoning methods if the game at hand proves too complex to be efficiently translated (space- or time-wise). A detailed comparison study of the efficiency and robustness of available GDL reasoners is given in Schiffel and Björnsson (2014).

Search

GGP agents use (at least) two levels of reasoning. First, the GDL reasoning for state-space manipulation and legitimate game play, and, second, the lookahead reasoning needed for achieving expert level of game play. The latter type of reasoning is – as in other game-playing programs – accomplished by using either traditional or simulation-based game-tree search algorithms.

Traditional Game-Tree Search in GGP

Early GGP agents used mostly traditional game-tree search methods, such as alphabeta (Knuth and Moore 1975) for 2-player games, max^n (Luckhart and Irani 1986) for multiplayer games, and variants of A* (Hart et al. 1968) or depth-first search for single-agent puzzles. The search was augmented with standard enhancements where applicable, such as iterative deepening, variable-depth exploration, move ordering, and transposition table. The most critical difference in design compared to game-specific agents lay in their heuristic evaluation functions for encapsulated game-specific knowledge, which had to be learned online during play (overviewed in a later section). GGP agents using this approach were dominant in the beginning of the IGGP competition and, for example, won the first 2 years the competition was held (Clune 2007; Kuhlmann et al. 2006; Schiffel and Thielscher 2007).

In addition to online learning of heuristic evaluation function, a few nonstandard approaches for playing general games were introduced, for example, using answer set programming (Muller et al. 2011; Thielscher 2009), constraint solvers (Clune 2008), and multi-valued variable planning (Kissmann and Edelkamp 2009) for solving single-agent games.

The complexity of these early agents, however, grew fast as they became increasingly specialized, typically using disparate search techniques for different type of games (i.e., single-, two-, and multiplayer games). Also, the traditional search techniques proved acutely reliant on having good heuristics to guide them, which is not always the case in GGP. For example, it typically takes some time for the agents to learn suitable heuristics for the game at hand, so early on in a match game there might not be much to build on; or, even worse, the heuristics once learned might have failed to capture the essence of the game and thus be nonuseful or even detrimental.

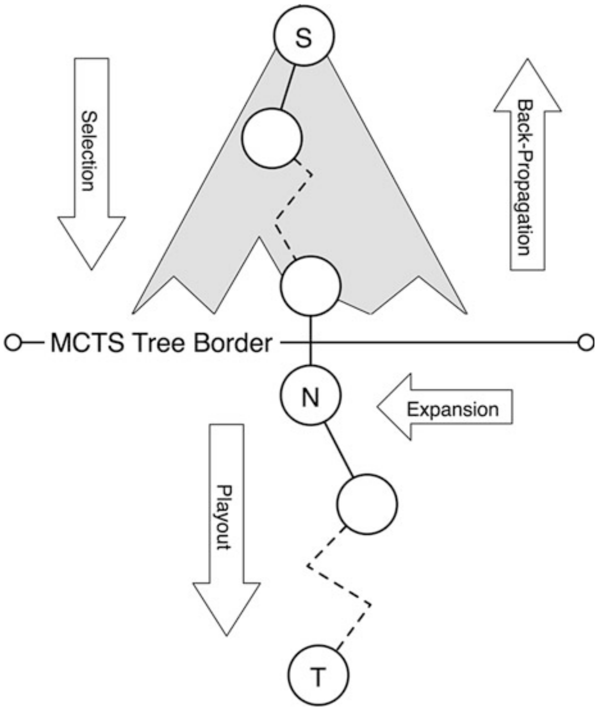
Monte Carlo Tree Search in GGP

A significant improvement in GGP came from the introduction of Monte Carlo tree search (MCTS) (Coulom 2006; Kocsis and Szepesvári 2006) into the field (Genesereth and Björnsson 2013). Although MCTS is not necessarily the ideal lookahead search method for every type of game, then overall it lends itself particularly well for general game playing. There are several reasons for this, but first and foremost is its relatively robust and effective reasoning across a wide range of games, even in the absence of well-informed heuristics. It is also applicable to single-, two-, and multiplayer games with only small-scale modifications, thus having the advantage of being the only search paradigm necessary in an all-embracing GGP agent. As a case in point, the most successful GGP agents since 2007 have all relied on MCTS as their primary lookahead search method.

Overview of MCST

Monte Carlo tree search (MCTS) continuously runs in-game simulations while deliberating on the best action to take in the current game state (position), that is, it repeatedly simulates games starting from that position and running to the end of the game. It does so until deliberation time is up, when it uses the game outcomes accumulated from the simulations to commit on the most promising action to play.

Fig. 2 MCTS overview: the four phases to a simulation



While running these Monte Carlo simulations, the search additionally gradually builds a game-tree in memory where it keeps track of the (averaged) outcome of played state-action pairs, $Q(s, a)$, where s is a game state and $a \in A(s)$ is one of the actions available in state s. The statistic kept in the game-tree is used by subsequent simulations to make progressively more informed action-selection choices (mimicking minimax-based action selection in the long run), thus the name Monte Carlo tree search.

Each simulation consists of four successive phases: *selection, playout, expansion,* and *back-propagation,* as depicted in Fig. 2. The *selection phase* is formed of the action-selection choices at the beginning of a simulation while still traversing down the in-memory game-tree. The action choices taken in this phase are typically guided by accumulated statistics kept in the tree. This phase is initially brief, but prolongs as the game-tree grows. The *expansion* phase commences when the simulation "falls off" the tree, but it controls how the game-tree is grown. A typical strategy is to expand the line of play just traversed by one level (or node) every few simulations, depending how fast the search is (care must be taken not to grow the tree too aggressively to limit memory usage). The *playout phase* then follows (the relative order of the expansion and playout phases is often inconsequential and can be interchanged). In there the simulation is played to the end of the game (or in some cases prematurely terminated) and the outcome of the game recorded. Because of the lack of accumulated state-specific statistics, it is typical for actions played in the playout phase to be less informed than in the selection phase, in the extreme sometimes even being played at random. It is common for MCTS to spend most of its search effort in the simulation phase, making it crucial for its effectiveness to be nonetheless guided by somewhat informed heuristics in that phase as well. In the *back-propagation* phase, the game outcome is backed up all the way to the root of the game-tree and the relevant statistics updated, such as $\mathsf{Q}(s, a)$.

One additional advantage of MCTS is that it is relatively straightforward to parallelize (Méhat and Cazenave 2011), and it is thus well poised to take advantage of contemporary multicore computer hardware.

MCTS in GGP

The MCTS algorithm has in recent years found wide-spread use in many computer game-playing domains. Its application in GGP is in many ways similar to how it is applied elsewhere, however, with the important distinction that search-control knowledge needs to be learned online during game play. Research into application of MCTS to GGP has thus to a large extend focused on online learning of search control for guiding the simulations, which we overview in the next section.

Using effective search control for simulation guidance is important for state-of-the-art MCTS-based GGP agents, which not only results in better quality information being returned from the simulations but also in increased simulation throughput (better informed simulations tends to be shorter). In state-of-the-art GGP agents, the simulation throughput is also increased by the use of highly parallel MCTS (Méhat and Cazenave 2011) and, as discussed in the previous section, the use of a fast GDL reasoner.

Learning Simulation Search Control

A robust and effective search-control mechanism for simulation guidance is a key component in any MCTS-based game-playing program. In other game-playing application domains where MCTS is successfully used, such as computer Go (Gelly et al. 2006), a typical approach is to manually define and encode game-specific domain knowledge that is useful for simulation control and then carefully tune its use, either by hand or by automated match play consisting of (possibly) thousands of games. Such an approach is, however, unworkable in GGP. In contrast, the search-control knowledge must be automatically discovered and tuned during online play consisting of (possibly) only a single match game – making the task at hand all the more challenging. In this section, we give an overview of search-control guidance techniques in GGP for simulation-based search.

Move-Average Sampling Technique

Move-average sampling technique (MAST) (Finnsson and Björnsson 2008) capitalizes on that some moves are on average better than others independent of the exact game state they are played in. In the back-propagation phase, while backing up the game outcome, the average outcome for each role-action pair encountered during the playout is maintained, that is $\mathsf{Q}_h(r, a)$ where a an action played by role r. These averages are then used to give preferences to some moves during the playout phase, thus biasing the simulations towards better lines of play. This is typically done in GGP either using the Gibbs (or Boltzmann) distribution or the ε-greedy action-selection strategy.

The Gibbs distribution chooses proportionally between good moves as:

$$\mathsf{P}(r, a) = \frac{e^{\mathsf{Q}_h(r,a)/\tau}}{\sum_{b=1}^{n} e^{\mathsf{Q}_h(r,b)/\tau}}$$

Where, $\mathsf{P}(r, a)$ is the probability that action a will be chosen by role r in the current playout state and $\mathsf{Q}_h(r, a)$ is the preference values given for action a. This results in actions with a high $\mathsf{Q}_h(r, a)$ value becoming more likely to be chosen. One can stretch or flatten the above distribution using the τ parameter ($\tau \rightarrow 0$ stretches the distribution, whereas higher values make it more uniform).

The ε-greedy action has a strong preference for the single best move, i.e., it chooses for each role the action with the highest preference score $\mathsf{Q}_h(r, a)$ most $(1 - \varepsilon)$ of the time, but a random action occasionally (ε of the time).

Despite its simplicity, this method is for many games capable of capturing relevant game strategies, for example, that it is beneficial to place a piece in the center in the game Tic-Tac-Toe or on an edge or corner in the game Othello. Furthermore, as the technique is trivial to implement and has low computational overhead, it is a good choice for a simple yet robust simulation search control in GGP agents.

Predicate-Average Sampling Technique

Predicate-average sampling technique (PAST) (Finnsson and Björnsson 2010) extends on MAST by adding state context. For every predicate-move pair that occurs during a simulation, an average $\mathsf{Q}_h(r, p, a)$ is maintained and updated in the back-propagation phase in a similar fashion to as in MAST. The parameters of $\mathsf{Q}_h(r, p, a)$ indicate the estimation for action a selected by role r when predicate p is present in the current state (an example predicate in Tic-Tac-Toe could for example be (cell 1 2 X)).

The move value PAST uses in the playout biasing to determine which move to select is obtained by comparing all $\mathsf{Q}_h(r, p, a)$ values possible in the current state S for each move and use the maximum pair value to represent the preference value for the move, i.e.:

$$\mathsf{Q}_h(r, a) = max_{p \in S} \mathsf{Q}_h(r, p, a)$$

If the sample count for a predicate/move pair is below a certain threshold, a default value of 50 (draw) is returned to avoid values with very high variance. Although this method is computationally more expensive than MAST, then it has in practice proven to be one of the overall most effective simulation search-control method for GGP.

Features-to-Action Sampling

The *Features-to-action sampling technique* (FAST) (Finnsson and Björnsson 2010) search control scheme, unlike MAST and PAST, imposes a game-specific meaning to the GDL game descriptions. That is, it tries to infer from the game descriptions high-level features such as piece types and board layout, which is then used to detect captures or strategically strong board cells. The feature discovery is done by matching templates to the GDL.

FAST uses the TD(λ) learning algorithm (Sutton 1988) to learn values for the detected features by having every simulation generate episode for the knowledge accumulation. For an episode with the path of $s_1 \rightarrow s_2 \rightarrow \ldots \rightarrow s_n$ the delta rule is applied to each state s_t:

$$\vec{\delta} = \vec{\delta} + \alpha \times \left[R_t^\lambda - V(s_t)\right] \times \nabla_{\vec{\theta}} V(s_t)$$

Where, R_t^λ is the λ-return (average of exponentially weighted n-step TD returns), $V(s)$ is the value function, and $\nabla_{\vec{\theta}} V(s)$ is its gradient. At the end of the episode, the difference of the players' goal values is used as the return. A weight vector $\vec{\theta}$ linearly weighs and combines the $\vec{\delta}$ with the detected features of the $\vec{f}(s)$ into the value function:

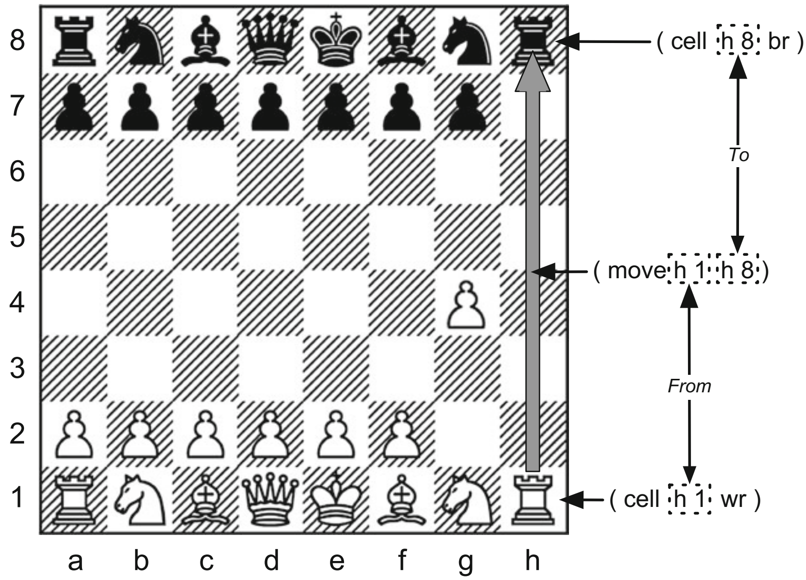
$$V(s) = \sum_{i=1}^{\left|\vec{f}\right|} \theta_i \times f_i(s)$$

Where, the features in $f_i(s)$ represent the number of pieces of a given type in state s. In the case of cell features, each feature is binary and indicates if a player's piece is placed on a given cell. Features are detected for all roles so a two-player game with n board cells would result $2 \times n$ detected features.

For biasing the action selection in the online simulations, the learned feature values are mapped into the $\mathrm{Q}_h(r, a)$ action-selection framework as shown in Figs. 3 and 4. The technique, although not applicable in all games, has overall proved quite effective.

N-Gram Selection Technique and Last Good Reply Policy

There has also been work on generalizing existing search-control methods to apply to GGP as well. The *N-Gram selection technique (NST)* (Tak et al. 2012) keeps track of



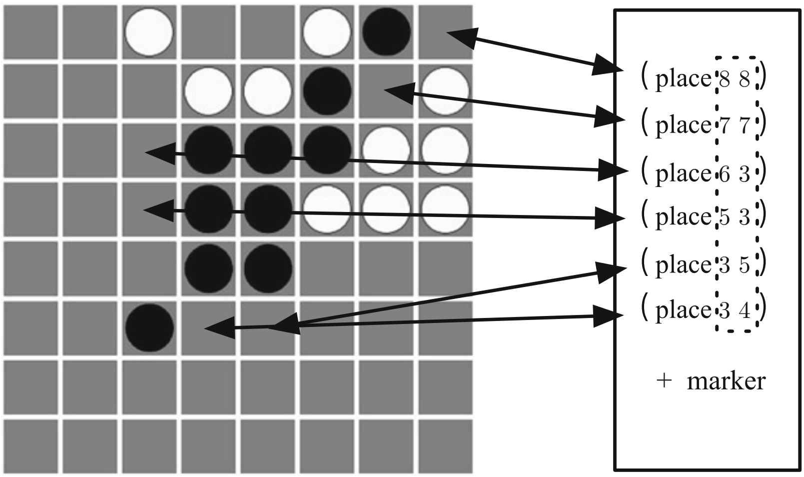
Feature	Value
...	
br	–4.98
...	
wr	5.26
...	

The values learnt for the piece type features found at the from and to positions are then fed to the formula for this move:

$$Q_h(\text{move h 1 h 8}) = -(2 * br + wr)$$
$$= -(2 * -4.98 + 5.26)$$
$$= 4.7$$

Fig. 3 FAST uses learned pieces values to give preference to moves that capture opponent pieces, in particular those where a valuable piece is being captured by a less valuable one. An example calculation for a capture move (move h 1 h 8) in chess in a state containing (cell h 1 wr) and (cell h 8 br) is shown here

the merits of move sequences, as opposed to only individual move merits. It is founded on N-Gram models, a structure common in statistical language processing. In NST, an N-Gram is defined as a consecutive move sequence of a length 1, 2, and 3. After a simulation ends the occurring N-Grams, from the point of view of each player, are extracted and their statistics updated, that is, for each N-Gram the average outcome of all simulations it has occurred in is incrementally modified and stored in a table. Figure 5 shows N-Grams generated from an example simulation (a corresponding N-Gram table would also be generated for the remaining role). When selecting a move to play in the playout phase (as well as in the selection phase if no other statistic is yet available), a preference score $Q_h(r, a)$ is given to each available move in a state by looking up in the table all matching N-Grams and average their score. The move to play is then chosen based on their individual preference scores, using either a Gibbs measure or an ε-greedy strategy.



Feature	Value
. . .	
3 4 white	0.7334
7 7 white	0.8382
. . .	
5 3 white	1.6626
6 3 white	1.6909
. . .	
3 5 white	1.2550
. . .	
8 8 white	4.1129
. . .	

The values learnt for the position with our marker are then fed to the formula for these moves :

$$Q_h(\text{place } 3\ 4) = c * [3\ 4\ \text{white}]) = 0.7334c$$

$$Q_h(\text{place } 3\ 5) = c * [3\ 5\ \text{white}]) = 1.2550c$$

$$Q_h(\text{place } 5\ 3) = c * [5\ 3\ \text{white}]) = 1.6626c$$

$$Q_h(\text{place } 6\ 3) = c * [6\ 3\ \text{white}]) = 1.6909c$$

$$Q_h(\text{place } 7\ 7) = c * [7\ 7\ \text{white}]) = 0.8382c$$

$$Q_h(\text{place } 8\ 8) = c * [8\ 8\ \text{white}]) = 4.1129c$$

Fig. 4 FAST calculations for a board feature in Othello; out of the available actions placing a piece is the upper right corner has the highest preference value

Last Good Reply Policy, a related technique that keeps tracks of best move replies as opposed to preference scores in association with N-Grams, was also generalized to be applicable in GGP (Tak et al. 2012).

Discussion

The learning of search control in GGP is still an active and open research, for example, how to mix different search-control strategies together and getting them to work in unison. Also, as we saw example of above, then specialized techniques applied in individual games can in many cases be generalized to also apply in GGP, for example, rapid-action-value estimates (RAVE) (Gelly and Silver 2011), which was initially developed for computer Go but had since then also be successfully adapted to also work in GGP.

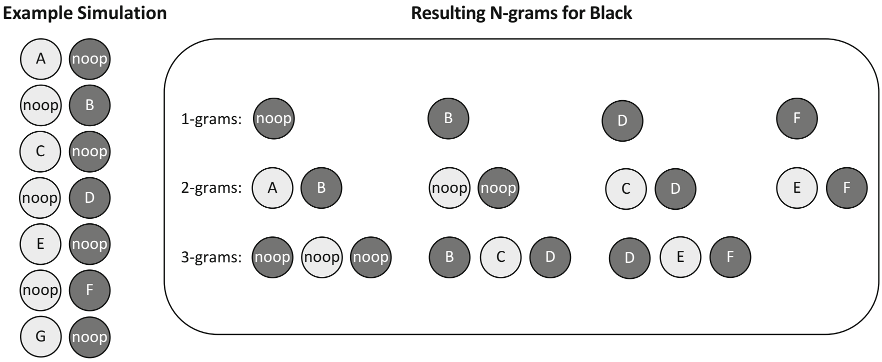


Fig. 5 N-Grams for role Black as a results of an example simulation: there are four 1-g, four 2-g, and three 3-g

Learning Heuristic Evaluation Functions

The first successful GGP agents were all based on traditional minimax-based game-tree search, augmented with an automatically learned heuristic evaluation function for encapsulating the domain-specific knowledge (Clune 2007; Kuhlmann et al. 2006; Schiffel and Thielscher 2007). The evaluation function's role is to assess the merits of the game states at the leaf nodes of the expanded game-tree, those merit values are then backed up the tree for determining the most promising action at the root. However, unlike game-specific game-playing programs that use a set of carefully hand-crafted game-specific features in their evaluation, GGP programs typically rely on a small set of generic features that apply to a wide range of games, such as mobility (the number of legal moves or approximate thereof), identifiable stable markers (e.g., pieces), goal proximity (number of steps to reach a goal), or partial subgoal fulfillment. The relevance and relative importance of the features are then automatically decided and tuned in real-time for the game at hand. Learning of game-independent heuristics is still an important research topic in GGP. We review the most significant work below.

ClunePlayer

CLUNEPLAYER's evaluation function is based on abstract features extracted from the GDL game descriptions. More specifically, sub-formulas are extracted from the *goal* rules (for features used in the payoff function) or the *legal* rules (for features used in the mobility function). Features are then constructed by imposing one of the following interpretations on these formulas:

- *Solution cardinality* counts the number of ground instances of the variables in a formula that fulfill the formula in the given state.

- *Symbol distance* uses a distance measure on the terms appearing in the formula, if such a distance measure could be identified in the game.
- *Partial solution* interpretation is applied to expressions that are conjunctions or disjunctions and counts how many of the sub-expressions are satisfied.

Random sample matches are then used to identify which of those features are stable, that is, their value does not fluctuate excessively when moving from one game state to the next. The intuition is that quantities that oscillate wildly do typically not provide as good a basis for an evaluation function as quantities that vary only incrementally. Nonstable features are considered irrelevant and not used in subsequent computations.

Given a set of stable features, then CLUNEPLAYER's heuristic evaluation function is then composed as follows:

- *Payoff* The payoff function estimates the future payoff (game outcome) of a role in a given state as a linear combination of the stable features, those weights are learned as a correlation of the feature values to actual payoffs in specially constructed artificial game states.
- *Mobility* The mobility function estimates the relative mobility of a role in a given state based on the extracted legality features. Interestingly, the actual number of legal moves in the state is not used to measure mobility; as such a measure could give misleading information in games with alternating moves or forced captures.
- *Termination* The termination function estimates how close a given state is to an end of the game (a terminal state). Termination is a linear function on the number of steps executed so far. The parameters of this linear function are learned from sample matches by least squares regression.

The above functions are then combined into a single heuristic evaluation of the game state in such a way that the mobility is more relevant at the start of the game and payoff has more influence towards the end of the game, where the progress of the game is estimated by the termination function. Both payoff and mobility are also regarded less important when their respective stability is low. This ensures that the value of the heuristics does not oscillate wildly.

CLUNEPLAYER was remarkably sophisticated for an early GGP agent and its evaluation function proved quite effective, resulting in it not only winning the first IGGP competition but also coming in close second the following 3 years, before retiring from competition.

FluxPlayer

FLUXPLAYER evaluation function is mainly based on the idea of partial subgoal fulfillment. The player generates an evaluation function from the *goal* formulas of the game. A state is evaluated against these *goal* conditions using fuzzy logic to compute a measure of partial goal fulfillment. The underlying idea is that in most environments goals can be reached step by step and states do not change wildly. Thus states in which more of the subgoals are fulfilled are probably closer to a goal state.

This fuzzy evaluation of the goal conditions is improved by using additional information about structures in the game that is obtained by game analysis. This includes information about counters (e.g., piece or point counters), boards, pieces, and ordering relations in the game description, among others. The game analysis is done partially by simulation of the game and partially by theorem proving using the game description (Haufe et al. 2012; Schiffel and Thielscher 2009). Knowledge of these structures is used in various ways to improve the evaluation function and the reasoning process. For example, knowledge of boards and pieces is used to compute the distance of pieces to their goal locations in racing games as part of the state evaluation. A more sophisticated approach for learning distance metrics was presented later in Michulke and Schiffel (2013).

Not unlike CLUNEPLAYER, FLUXPLAYER was a very successful early GGP agent with an effective evaluation function. Of all the players using evaluation functions, FLUXPLAYER seems to do the most extensive game analysis. The player was the winner of the second IGGP competition and scored top ranks for several more years.

KuhlPlayer

KuhlPlayer (Kuhlmann et al. 2006; Kuhlmann 2010) uses various methods to find relevant structures in a game, including successor relations, step counters, board, and pieces. This is done both by pattern-matching on the syntactic structure of the GDL rules and by playing random games. Numerous features are then generated from the detected structures, including:

- pieces, and their x and y coordinates (when on a two-dimensional board),
- Manhattan distance between each pair of pieces,
- number of pieces of each type,
- and various other quantities.

However, unlike the other knowledge-based agent, it does not combine all the different features into a single sophisticated evaluation function, but instead constructs many simplistic evaluation functions. Using massively parallel computer hardware, multiple game-tree searches are then run concurrently, each using its own distinct evaluation function. A majority vote from all searches is used to decide on the best move to play. The agent additionally contained a method to transfer knowledge from one game to similar games (Kuhlmann 2010). Knowledge transfer in a general game playing context had been previously studied in (Banerjee et al. 2006; Banerjee and Stone 2007). Knowledge transfer can greatly reduce the effort for game analysis and learning of evaluation functions. However, knowledge transfer requires computing the similarity of games, which is intractable in general. The agent participated in few of the early IGGP competitions, with a best place finishing of third in 2006.

Other Agents

The Magline agent, placing third and second in the 2009 and 2010 IGGP competitions respectively, used a game-independent feature learning algorithm named GIFL (Kirci et al. 2009) to learn patterns correlated with winning or losing based on differences in temporally related states. The learning is, however, applicable only in two-player alternating move games (for which the learning algorithm gave promising results).

Goblin and Ogre (Kaiser 2007a, b, c) are closely related agents that learn features for board games from playing random games. They learn features such as: board structure and piece types, occupied columns, Manhattan distances between pieces and their initial and goal locations. They learn several other more general features and patterns (e.g., mobility and subgoals). The features are then combined into a single heuristic value simply by summing their values (i.e., unweighted). Ogre placed second in the 2005 IGGP competition.

Artificial neural networks (ANN) have furthermore been used to learn heuristic evaluation functions for GGP agents; for example directly from (grounded) GDL *goal* definitions (Michulke 2012; Michulke and Thielscher 2009) or by evolving networks as in an agent participating in the IGGP 2006 competition (Reisinger et al. 2007).

Learning Structures

Apart from board game related structures such as boards and pieces, people have analyzed and tried to exploit more general structures of games, such as symmetries or the possibility of decomposing games.

Symmetries in games can come in various forms. For example, a game board setup can be symmetric, items or pieces in a game can be functionally identical (and thus symmetric), moves can lead to symmetric states, or the game can be symmetric with respect to switching the roles. Using the approach of Schiffel (2010), these symmetries in games are detected by transforming the game rules into a special graph and computing the automorphisms of this graph. Symmetries can be exploited by avoiding the expansion of symmetric states or symmetric actions, e.g., by using a transposition table and generating functions for symmetric states. The approach is limited to finding symmetries that are apparent from the game rules. However, it has the advantage of being somewhat independent on the size of the game's state space.

Some games can be decomposed (or factored) into several subgames whose actions and subgame states do not interact with each other. If possible this is very advantageous, because the state space of the composed game grows exponentially with the number of subgames. Thus, being able to search the subgames' state spaces instead of the states of the composed game results in considerable lower time and space complexities. In practice, subgames are never completely independent. They are at least connected through the goal and terminal conditions of the composed game. Thus, even if subgames mechanics are independent, which move is good in one subgame may depend on which states the other subgames are in.

So far, two different approaches to factoring games have been published. In Cox et al. (2009), the authors present an approach to factoring games using a representation of the game rules as a propositional network. This representation is already a graph structure, such that a game can be factored if the graphs of the subgames are only connected in specific ways at the goal, terminal and legal relations. In the presented form, the approach is only applicable for composed games in which the goal is to reach the goal in all of the subgames and where in every step an action can be chosen from any of the subgames. More complex relations between subgames goals or possible actions cannot be handled.

The approaches described in Gunther et al. (2009) for single-player games and in Zhao et al. (2009) for multiplayer games are more general. They allow for more complex relations between the subgames in the goal and terminal conditions of a game. Also, detection of subgames is done using syntactic and semantic analysis of the game rules instead of the propositional network. The problem with finding the best move in a subgame is solved by dividing the search process into two parts:

- **Subgame search** finds move sequences in each subgame that are significantly different with respect to the goal of the composed game – without needing a score for moves or subgame states.
- **Global game search** uses the move sequences from the subgame search to select a combination that is good with respect to the goal of the composed game.

This approach still retains the benefits of factoring games, except for pathological cases while being applicable to a wider range of composed games.

GGP Competitions and Events

To renew interest in general problem-solving approaches to computer game playing, the artificial intelligence community established the International GGP Competition (IGGPC) in 2005, which has run annually ever since. There is at present an active research community worldwide working on GGP. In here, we highlight the main ongoing GGP events.

Competitions

The IGGPC is the flagship competition. It is traditionally associated and collocated with one of the two main artificial intelligence conferences, AAAI or IJCAI. The *Computational Logic Group at Stanford University* is the main organizer of the competition. The exact format of the competition may vary slightly a year from year, although it is invariably conducted in two stages. In the preliminaries, which are open to everyone, a large number of disparate games is played on a remote server

Table 1 Winners of the IGGPC

Year	Agent	Authors
2014	SANCHO	Steve Draper and Andrew Rose
2013	TURBOTURTLE	Sam Schreiber
2012	CADIAPLAYER	Hilmar Finnsson and Yngvi Björnsson
2011	TURBOTURTLE	Sam Schreiber
2010	ARY	Jean Méhat
2009	ARY	Jean Méhat
2008	CADIAPLAYER	Yngvi Björnsson, Hilmar Finnsson, and Gylfi Gudmundsson
2007	CADIAPLAYER	Yngvi Björnsson and Hilmar Finnsson
2006	FLUXPLAYER	Stephan Schiffel and Michael Thielscher
2005	CLUNEPLAYER	Jim Clune

over an extended period of time and the (eight) overall best performing agents advance to the finals. In the finals, traditionally played on-site the collocated conference, the agents compete in a (double-elimination) playoff format. Each playoff match consists of a number of different games, often interesting variants of well-known board games (e.g., checkers on a cylindrical board).

Table 1 lists the winners of the IGGP competition over the years. In the first 2 years, 2005 and 2006, the competition was won by CLUNEPLAYER and FLUXPLAYER, respectively. Both agents emphasized knowledge-based heuristic state evaluation functions and used traditional game-tree search for thinking ahead. However, ever since CADIAPLAYER introduced MCTS to the field, winning consecutively in 2007 and 2008 (and in 2012), all winning agents have been MCTS based. For example, ARY, the winner in 2008 and 2009, was a highly parallel MCTS agent (Méhat and Cazenave 2011). The parallel search paradigm is now becoming increasingly mainstream in GGP competitions, with the prevalence of modern day multicore computer hardware. TURBOTURTLE has been a top contending agent in all recent competitions and winning in both 2011 and 2013. It is inarguably one of the strongest GGP agents today, along with the reigning IGPPC champion SANCHO. Both agents are MCTS based and use high-performance propositional networks for GDL reasoning (where applicable).

Occasional GGP tournaments are also scheduled now and then in association with various artificial intelligence events worldwide. For example, both the German and the Australian associations for artificial intelligence have organized GGP competitions as a part of their annual conferences. Finally, there are several online GGP servers that provide year-a-round twenty-four-seven playing opportunity for GGP agent, most notably the *Dresden GGP* and the *Tiltyard GGP* servers.

Workshops and Online Resources

The GGP research community also organizes a biennial workshop, *general intelligence in game-playing agents (GIGA)*, dedicated to general problem-solving approaches to computer game playing. New research results on GGP are published

Table 2 Useful online GGP resources

URL	Description
games.stanford.edu	This site, maintained by the Computational Logic Group at Stanford University, is the official home of the IGGP competition. The site also hosts a great deal of useful information on GGP, including educational material, a news forum, and an online GGP game server
general–game–playing.de	This site, maintained by the FLUXPLAYER team, also hosts a great deal of useful GGP related information, including educational material, starting-out software for aspiring GGP developers, and an online server (Dresden GGP server). The site also maintains a useful and (mostly) up-to-date bibliography of GGP publications
ggp.org	This site, maintained by the author of TURBOTURTLE, hosts various useful open-source GGP projects. The site also hosts an online server (Tileyard GGP server) and a comprehensive match archive for offline analysis

there as well as in mainstream artificial intelligence conferences, such as AAAI and IJCAI. There are also many useful online resources providing useful information and software on GGP. Table 2 lists a few such resources that are particularly comprehensive.

Conclusions

GGP has been an active research area over the last decade, and significant research progress been made toward general game-playing systems. There are, however, still many noteworthy research challenges to be addresses where the field has yet only scratched the surface. This includes more permanent type of learning, for example, where agents learn general game concepts by playing a wide variety of games over an extended period of time (months or even years) and then use that experience to expedite learning in newly encountered games. Another topic of interest where we foresee added research emphasis in the future is on general incomplete information and nondeterministic games. Applying general game-playing learning in such a demanding environment is a challenging task; however, promising preliminary work on the topic is already underway. Similarly, preliminary work is also underway on inductive learning in GGP, where game-playing systems – in addition to learning game-playing strategies – are also required to learn the rules of the games they play simply by observing others play.

Recommended Reading

S. Abiteboul, R. Hull, V. Vianu, *Foundations of Databases* (Addison-Wesley, Reading, 1995)

K.R. Apt, M.H. van Emden, Contributions to the theory of logic programming. J. ACM **29**(3), 841–862 (1982)

B. Banerjee, P. Stone, General game learning using knowledge transfer, in *The 20th International Joint Conference on Artificial Intelligence* Hyderabad, India (2007), pp. 672–677

B. Banerjee, G. Kuhlmann, P. Stone, Value function transfer for general game playing, in *ICML Workshop on Structural Knowledge Transfer for Machine Learning* Pittsburgh, Pennsylvania (2006)

R.E. Bryant, Symbolic manipulation of boolean functions using a graphical representation, in *Proceeding of the 22nd ACM/IEEE Design Automation Conference (DAC'85)*, Los Alamitos, USA June 1985 (IEEE Computer Society Press), pp. 688–694

J.E. Clune, Heuristic evaluation functions for general game playing, in *Proceedings of the Twenty-Second AAAI Conference on Artificial Intelligence, AAAI 2007,* Vancouver, 22–26 July 2007, pp. 1134–1139

J. Clune, Heuristic evaluation functions for general game playing. PhD thesis, University of California, Los Angeles, 2008

R. Coulom, Efficient selectivity and backup operators in monte-carlo tree search, in *Computers and Games*, ed. by H. Jaap van den Herik, P. Ciancarini, H.H.L.M. Donkers, Lecture Notes in Computer Science, vol. 4630, Berlin, (Springer, 2006), pp. 72–83

E. Cox, E. Schkufza, R. Madsen, M. Genesereth, Factoring general games using propositional automata, in *Proceedings of the IJCAI-09 Workshop on General Game Playing (GIGA'09)* Pasadena, California (2009), pp. 13–20

H. Finnsson, Y. Björnsson, Simulation-based approach to general game playing, in *Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence, AAAI 2008*, Chicago, 13–17 July 2008, ed. by D. Fox, C.P. Gomes, Boston, (AAAI Press, 2008), pp. 259–264

H. Finnsson, Y. Björnsson, Learning simulation control in general game-playing agents, in *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2010,* Atlanta, 11–15 July 2010, Boston, (AAAI Press, 2010)

S. Gelly, D. Silver, Monte-Carlo tree search and rapid action value estimation in computer Go. Artif. Intell. **175**(11), 1856–1875 (2011)

S. Gelly, Y. Wang, R. Munos, O. Teytaud, Modification of UCT with patterns in Monte-Carlo Go. Technical Report 6062, INRIA (2006)

M.R. Genesereth, Y. Björnsson, The international general game playing competition. AI Mag. 34(2), 107–111 (2013)

M. Genesereth, N. Love, General game playing: game description language specification. Technical Report March 15 2005, Stanford University (2005)

M.R. Genesereth, R.E. Fikes, R. Brachman, T. Gruber, P. Hayes, R. Letsinger, V. Lifschitz, R. Macgregor, J. Mccarthy, P. Norvig, R. Patil, L. Schubert, Knowledge interchange format version 3.0 reference manual, 14 Aug Stanford University, California 1992

M.R. Genesereth, N. Love, B. Pell, General game playing: overview of the AAAI competition. AI Mag. **26**(2), 62–72 (2005)

M. Gunther, S. Schiffel, M. Thielscher, Factoring general games, in *Proceedings of the IJCAI-09 Workshop on General Game Playing (GIGA'09)* Pasadena, California (2009), pp. 27–33

N. Halderman, M. Tung, J. Flatt, A. Willis-Woodward, Javaprover (2006), http://games.stanford.edu/resources/reference/java/java.html

P.E. Hart, N.J. Nilsson, B. Raphael, A formal basis for the heuristic determination of minimum cost paths. IEEE Trans. Syst. Cybern. **SSC-4**(2), 100–107 (1968)

S. Haufe, S. Schiffel, M. Thielscher, Automated verification of state sequence invariants in general game playing. Artif. Intell. **187–188**, 1–30 (2012)

D.M. Kaiser, Automatic feature extraction for autonomous general game playing agents, in *Proceedings of the Sixth International Joint Conference on Autonomous Agents and Multiagent Systems* Honolulu, Hawaii, (2007a)

D.M. Kaiser, The design and implementation of a successful general game playing agent, in *FLAIRS Conference* Honolulu, Hawaii, (AAAI Press, 2007b), pp. 110–115

D.M. Kaiser, The structure of games. PhD thesis, Florida International University, 2007c

M. Kirci, J. Schaeffer, N. Sturtevant, Feature learning using state differences, in *Proceedings of the IJCAI-09 Workshop on General Game Playing (GIGA'09)* Pasadena, California, (2009), pp. 35–42

P. Kissmann, S. Edelkamp, Instantiating general games, in *Proceedings of the IJCAI-09 Workshop on General Game Playing (GIGA'09)* Pasadena, California, (2009), pp. 43–50

D. Knuth, R. Moore, An analysis of alpha-beta pruning. Artif. Intell. **6**, 293–326 (1975)

L. Kocsis, C. Szepesvári, Bandit based Monte-Carlo planning, in *European Conference on Machine Learning (ECML)* Berlin, Germany, (2006), pp. 282–293

G. Kuhlmann, K. Dresner, P. Stone, Automatic heuristic construction in a complete general game player, in *21st AAAI,* Boston, USA, July 2006, pp. 1457–62

G.J. Kuhlmann, Automated Domain Analysis and Transfer Learning in General Game Playing. PhD thesis, University of Texas at Austin, 2010

N. Love, T. Hinrichs, D. Haley, E. Schkufza, M. Genesereth, General game playing: game description language specification. Technical Report LG–2006–01, Stanford Logic Group, Computer Science Department, Stanford University, 353 Serra Mall, Stanford, CA 94305, March 2008. Available at: games.stanford.edu

C. Luckhart, K. Irani, An algorithmic solution of N-person games. in *Proceedings of the 5th National Conference on Artificial Intelligence,* ed. by T.R.S. Kehler, vol. 1 (Morgan Kaufmann, Philadelphia, 1986), pp. 158–162

J. Méhat, T. Cazenave, A parallel general game player. Kunstl. Intell. **25**(1), 43–47 (2011)

D. Michulke, Evaluation Functions in General Game Playing. PhD thesis, Technische Universitt Dresden, 2012

D. Michulke, S. Schiffel, Admissible distance heuristics for general games, in *Agents and Artificial Intelligence,* vol. 358 (Springer, Berlin/Heidelberg, 2013), pp. 188–203

D. Michulke, M. Thielscher, Neural networks for state evaluation in general game playing, in *Proceedings of the European Conference on Machine Learning (ECML)* Bled, Slovenia, (2009), pp. 95–110

M. Muller, M. Schneider, M. Wegner, T. Schaub, Centurio, a general game player: parallel, java- and asp-based. Kunstl. Intell. **25**(1), 17–24 (2011)

J. Reisinger, E. Bahceci, I. Karpov, R. Miikkulainen, Coevolving strategies for general game playing, in *Proceedings of the IEEE Symposium on Computational Intelligence and Games* (IEEE, Piscataway, 2007), pp. 320–327

S. Schiffel, Symmetry detection in general game playing, in *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2010*, Atlanta, 11–15 July 2010, ed. by M. Fox, D. Poole, Boston, (AAAI Press, 2010)

Schiffel, Stephan, and Yngvi Bjornsson. "Efficiency of GDL reasoners." *Computational Intelligence and AI in Games, IEEE Transactions on* 6.4 (2014): 343-354.

S. Schiffel, M. Thielscher, Fluxplayer: a successful general game player, in *22nd AAAI,* Vancouver, British Columbia, (2007), pp. 1191–1196

S. Schiffel, M. Thielscher, Automated theorem proving for general game playing, in *Proceedings of IJCAI'09* Pasadena, California, (2009), pp. 911–916

E. Schkufza, N. Love, M.R. Genesereth, Propositional automata and cell automata: representational frameworks for discrete dynamic systems, in *Proceeding of the 21st Australasian Joint Conference on Artificial Intelligence* Auckland, New Zealand, (2008), pp. 56–66

M. Schofield, A. Saffidine, High speed forward chaining for general game playing, in *3rd Workshop on General Intelligence in Game-Playing Agents (GIGA)*, Beijing, Aug 2013, pp. 31–38

S. Schreiber, The general game playing base package (2013), http://code.google.com/p/ggp-base/

N. Schultz, N. Manthey, D. Müller, C++-reasoner (2008), http://www.general-game-playing.de/downloads.html

R.S. Sutton, Learning to predict by the methods of temporal differences. Mach. Learn. **3**, 9–44 (1988)

M.J.W. Tak, M.H.M. Winands, Y. Björnsson, N-grams and the last-good-reply policy applied in general game playing. IEEE Trans. Comput. Intell. AI Games **4**(2), 73–83 (2012)

M. Thielscher, Answer set programming for single-player games in general game playing, in *Proceedings of the International Conference on Logic Programming (ICLP)* Boston, (Springer, 2009)

M. Thielscher, A general game description language for incomplete information games, in *Proceedings of the AAAI Conference on Artificial Intelligence,* Atlanta, July 2010, Boston, (AAAI Press, 2010), pp. 994–999

M. Thielscher, GDL-II. Kunstl. Intell. **25**, 63–66 (2011)

K. Waugh, Faster state manipulation in general games using generated code, in *Proceeding of the IJCAI-09 Workshop on General Game Playing (GIGA'09),* Pasadena, California, (2009)

D. Zhao, S. Schiffel, M. Thielscher, Decomposition of multi-player games, in *Proceedings of the Australasian Joint Conference onArtificial Intelligence*, vol. 5866, Berlin, (Springer, 2009), pp. 475–484

Monte-Carlo Tree Search in Board Games

3

Mark H. M. Winands

Contents

M.H.M. Winands (✉)
Department of Data Science and Knowledge Engineering, Maastricht University, Maastricht, The Netherlands
e-mail: m.winands@maastrichtuniversity.nl

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_27

Abstract

Monte-Carlo Tree Search (MCTS) is a best-first search method guided by the results of Monte-Carlo simulations. It is based on randomized exploration of the search space. Using the results of previous explorations, the method gradually builds up a game tree in memory and successively becomes better at accurately estimating the values of the most promising moves. MCTS has substantially advanced the state of the art in board games such as Go, Amazons, Hex, Chinese Checkers, Kriegspiel, and Lines of Action.

This chapter gives an overview of popular and effective enhancements for board game playing MCTS agents. First, it starts by describing the structure of MCTS and giving pseudocode. It also addresses how to adjust MCTS to prove the game-theoretic value of a board position. Next, popular enhancements such as RAVE, progressive bias, progressive widening, and prior knowledge, which improve the simulation in the tree part of MCTS, are discussed in detail. Subsequently, enhancements such as MAST, N-Grams, and evaluation function-based strategies are explained for improving the simulation outside the tree. As modern computers have nowadays multiple cores, this chapter mentions techniques to parallelize MCTS in a straightforward but effective way. Finally, approaches to deal with imperfect information and stochasticity in an MCTS context are discussed as well.

Keywords

Minimax search techniques • Monte-Carlo Tree Search (MCTS) • MCTS • MCTS-solver • Tree and Root Parellization • Chance nodes • Multi-Armed Bandit (MAB) problem • Playout • Simulation strategy • Expansion strategy • Backpropagation • Final move selection strategies • Domain-independent • Rapid Action-Value Estimator (RAVE) • Progressive Bias (PB) • Implicit Minimax • Progressive Widening • Move-Average Sampling Technique (MAST) • N-Gram Selection Technique (NST) • Greedy strategy • Root parallelization • Tree parallelization • Determinization • Upper Confidence Bounds applied to Trees (UCT)

Introduction

For decades, *minimax* search techniques such as $\alpha\beta$, expectimax, or paranoid search have been the standard approach used by programs for playing board games such as chess, checkers, and backgammon (and many others). Over the years, many search enhancements have been proposed for this minimax framework that further enhances its effectiveness. This traditional adversarial search approach has, however, been less successful for other types of games, in particular where a large branching factor prevents a deep lookahead or the intricacies of the game hinder the construction of an effective board evaluation function. Go is an example of a game that has so far eluded this approach.

In recent years, a new paradigm for game-tree search has emerged, so-called Monte-Carlo Tree Search (MCTS) (Coulom 2007b; Kocsis and Szepesvári 2006). In the context of game playing, Monte-Carlo simulations were first used as a mechanism for dynamically evaluating the merits of leaf nodes of a traditional minimax-based search (Abramson 1990; Bouzy and Helmstetter 2003; Brügmann 1993), but under the new paradigm, MCTS has evolved into a full-fledged best-first search procedure that replaces traditional minimax-based search altogether.

In the past few years, MCTS has substantially advanced the state of the art in several deterministic game domains (Browne et al. 2012), in particular Go (Gelly et al. 2012), but other board games include Amazons (Lorentz 2008), Hex (Arneson et al. 2010), Lines of Action (Winands et al. 2010), and the ones of the General Game Playing competition (Björnsson and Finnsson 2009). MCTS has even increased the level of competitive agents in board games with challenging properties such as multiplayer (e.g., Chinese Checkers, Sturtevant 2008) and uncertainty (e.g., Kriegspiel (Ciancarini and Favini 2010), Lord of the Rings: The Confrontation (Cowling et al. 2012), and Scotland Yard (Nijssen and Winands 2012a)).

This chapter gives an overview of successful approaches for board game playing MCTS agents. Only properties are taken into account that are typical in board games such as turn-taking (sequential-move games), determinism or stochasticity, two or more players, and (im)perfect information. One-player games or simultaneous-move games are not considered.

This chapter is organized as follows. The structure of MCTS is described first in section "Monte-Carlo Tree Search." Additionally, section "MCTS-Solver" addresses the question how to adjust MCTS to prove the game-theoretic value of board positions. Next, popular enhancements to improve the selection and playout part of MCTS are provided in sections "Selection Strategies" and "Simulation Strategies," respectively. As modern computers have nowadays multiple cores, section "Parallelization" presents techniques to parallelize MCTS in a straightforward but effective way. Finally, section "Uncertainty" discusses handling imperfect information and stochasticity in the MCTS tree.

Monte-Carlo Tree Search

Monte-Carlo Tree Search (MCTS) (Coulom 2007b; Kocsis and Szepesvári 2006) is a best-first search method that does not require a positional evaluation function. It is based on a randomized exploration of the search space. Using the results of previous explorations, the algorithm gradually builds up a game tree in memory and successively becomes better at accurately estimating the values of the most promising moves.

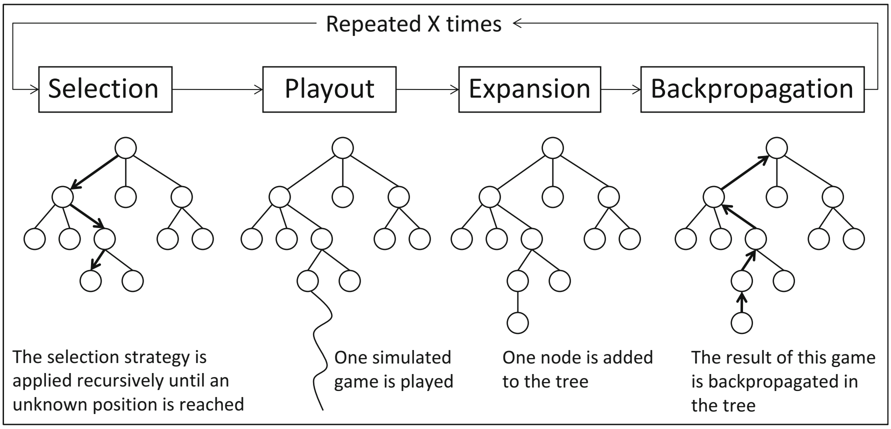


Fig. 1 Outline of Monte-Carlo tree search (adapted from Chaslot et al. 2008a; Winands et al. 2010)

The basic structure of MCTS is given in subsection "Structure of MCTS." Four strategic steps are discussed in subsection "The Four Strategic Steps." Next, the selection of the final move to be played in the actual game is described in subsection "Final Move Selection." Finally, pseudocode is provided in subsection "Pseudocode."

Structure of MCTS

In MCTS, each node i represents a given position (also called a state) of the game. A node contains at least the following two pieces of information: (1) the current value v_i of the position (usually the average result of the simulations passed through this node) and (2) the visit count of this position m_i. MCTS usually starts with a tree containing only the root node.

MCTS consists of four strategic steps, repeated as long as there is time left (Chaslot et al. 2008a). The steps, outlined in Fig. 1, are as follows: (1) In the *selection step*, the tree is traversed from the root node until a node E is reached, where a position is selected that is not added to the tree yet. (2) Next, during the *playout step*, moves are played in self-play until the end of the game is reached. (3) Subsequently, in the *expansion step*, children of E are added to the tree. (4) Finally, in the *backpropagation step*, the game result r is propagated back along the path from E to the root node, adding r to an incrementally computed average for each move along the way. In this chapter, *simulation* refers to the entire game played from the root until a terminal position is reached (i.e., the selection and the playout step).

The Four Strategic Steps

The four strategic steps of MCTS are discussed in detail below.

Selection

Selection chooses a child to be searched based on previous information. It controls the balance between exploitation and exploration. On the one hand, the task consists of selecting the move that leads to the best results so far (exploitation). On the other hand, the less promising moves still have to be tried, due to the uncertainty of the simulations (exploration).

This balancing of exploitation and exploration has been studied in particular with respect to the multiarmed bandit (MAB) problem (Robbins 1952). The MAB problem considers a gambling device and a player, whose objective is to maximize the reward from the device. At each time step, the player can select one of the arms of the device, which gives a reward. Usually this reward obeys a stochastic distribution. The selection step of MCTS can be viewed as a MAB problem for a given node. The problem is to select the next move to play, which will give an unpredictable result (the outcome of a single randomized game). Knowing the past results, the problem is to find the optimal move. However, the main difference with the MAB problem is that MCTS has to select moves at every level of the search tree.

Several *selection strategies* (Browne et al. 2012) have been suggested for MCTS such as BAST, EXP3, UCB1-tuned, but the most popular one is based on the UCB1 algorithm (Auer et al. 2002), called UCT (**U**pper **C**onfidence **B**ounds applied to **T**rees) (Kocsis and Szepesvári 2006). UCT works as follows. Let I be the set of nodes immediately reachable from the current node p. The selection strategy selects the child b of node p that satisfies Formula 1:

$$b \in \text{argmax}_{i \in I}\left(v_i + C \times \sqrt{\frac{\ln m_p}{m_i}}\right) \quad (1)$$

where v_i is the value of the node i, m_i is the visit count of i, and m_p is the visit count of p. C is a parameter constant, which can be tuned experimentally (e.g., $C = 0.4$). The value of v_i should lie in the range [0, 1]. In case a child has not been visited yet, the maximum value that a node can have by sampling (i.e., $v_{\max} = 1$) is assumed.

The aforementioned selection strategy is only applied in nodes with visit count higher than a certain threshold T (e.g., 5) (Chaslot et al. 2008a; Coulom 2007b). If the node has been visited less than this threshold, the next move is selected according to the *simulation strategy* discussed in the next strategic step. In this case, a well-tuned simulation strategy chooses moves more accurately than a selection strategy. However, when a node has been visited more, the selection strategy is more accurate.

Playout

The playout step begins when a position is entered that is not a part of the tree yet. Moves are selected in self-play until the end of the game. This task might consist of playing plain random moves or – better – semi-random moves chosen according to a *simulation strategy*. Smart simulation strategies have the potential to improve the level of play significantly (Gelly and Silver 2007). The main idea is to play interesting moves based on heuristics.

Designing an effective simulation strategy is a difficult task. If the strategy becomes too stochastic (e.g., if it selects moves nearly randomly), then the moves played are often weak, and the level of the MCTS program is reduced. In contrast, if the strategy becomes too deterministic (e.g., if the selected move for a given position is almost always the same), then the exploration of the search space is too selective, and the level of the MCTS program is reduced as well.

Next, when considering a simulation strategy, the trade-off between search and knowledge has to be balanced. The more informed the playout becomes, typically the slower it gets. On the one hand, this decreases the total number of simulations that can be run in the allotted time, but on the other hand, the result of each simulation is potentially more accurate. The former degrades the decision quality of MCTS, whereas the latter improves it, so the question is where the right balance lies. The trend seems to be in favor of fast simulation playouts where moves are chosen based on only computationally light knowledge, although recently, adding heuristic knowledge at the cost of slowing down the simulation playouts has proved beneficial in board games such as Chinese Checkers, Focus, and Lines of Action.

Expansion

Expansion is the procedure that decides whether nodes are added to the tree. A popular *expansion strategy* is to add one node per simulation (Coulom 2007b). The added leaf node L corresponds to the first position encountered during the traversal that was not already stored.

It is also possible to add all the children of a node to the tree. This strategy is possible when the branching factor is small. In contrast to this strategy, one may forbid any node expansion before a certain number T of simulations have been made through this node. This allows to save memory and reduces only slightly the level of play.

In general, the effect of these expansion strategies on the playing strength is small. The strategy of creating one node per simulation is sufficient in most cases.

Backpropagation

Backpropagation is the procedure that propagates the result r of a simulated game t back from the leaf node L, through the previously traversed nodes, all the way up to the root. If a game is won, the result of a player j is scored as a $r_{t,j} \leftarrow 1$, in the case of a loss as $r_{t,j} \leftarrow 0$, and a draw as $r_{t,j} \leftarrow 0.5$. For a two-player game, propagating the values back in the tree can be performed similar to negamax (Knuth and Moore 1975).

MCTS is able to handle games with multiple players and even multiple winners (Sturtevant 2008). The result is then backpropagated as a tuple of size N, where N is the number of players. For instance, if Player 1 and Player 3 both reach a winning condition in a three-player game, then result is returned as the tuple $\left(\frac{1}{2}, 0, \frac{1}{2}\right)$. For multiplayer games, propagating the values back in the tree is performed similar to max^n. As it also covers the two-player case, this approach is used for the remainder of the chapter.

In some games such as Chinese Checkers and Go, it is beneficial to add a small bonus to a win/draw/loss result. Such a bonus can favor shorter playouts or terminal positions with a larger score (Enzenberger et al. 2010). This plays more human-like moves by preferring wins with fewer moves and by maximizing the final score even if moves are all wins or all losses. It also can improve the playing strength (Pepels et al. 2014). Playouts with long sequences and wins with a small score have a larger error. Preserving a larger advantage increases the margin of safety against errors later in the game.

Finally, a *backpropagation strategy* is applied to the *value* v_i of a node i. Usually, it is computed by taking the average of the results of all simulated games made through this node (Coulom 2007b), i.e., $v_i \leftarrow R_{i,j}/m_i$, where j is the player to move in its parent node p and $R_{i,j} \leftarrow \sum_t r_{t,j}$ the cumulative score of all the simulations.

Final Move Selection

The four steps are repeated either a fixed number of simulations or until time runs out. After the search is finished, one of the children of the root is selected as the best move to be played in the actual game. There are several *final move selection* strategies that determine the best child (Chaslot et al. 2008a).

1. *Max child*. The max child is the child that has the highest value.
2. *Robust child*. The robust child is the child with the highest visit count.
3. *Robust-max child*. The robust-max child is the child with both the highest visit count and the highest value. If there is no robust-max child at the moment, more simulations are played until a robust-max child is obtained (Coulom 2007b).
4. *Secure child*. The secure child is the child that maximizes a lower confidence bound, i.e., which maximizes the quantity $v + \frac{A}{\sqrt{m}}$, where A is a parameter (e.g., 4), v is the node's value, and m is the node's visit count.

In practice, the difference in performance between these strategies is limited when a sufficient number of simulations for each root move have been played. In case there is a certain amount of time for the whole game and the player manages its own time for each move, the robust-max child is the most promising.

Pseudocode

The pseudocode for MCTS is given in Algorithm 1. In the pseudocode, $\mathcal{ST}$ is the search tree; *Select*(*Node p*) is the method that returns the most promising child of node *p* to explore. *PlayOut*(*Node p*) is the method that plays a simulated game beyond the search tree and returns the result *r* of this simulated game. *Expand* (*Node p*) is the method that adds one node to the tree. Dependent on the implementation, its siblings can be added as well. *Backpropagate*(*Result r*) is the method that updates the value of the node depending on the result *r* of the last simulated game. *A*(*Node p*) is the set of moves for a node *p*.

```
Data: root
Result: bestMove
while (timeLeft()) do
    currentNode ← root
    /* The tree is traversed                                   */
    while (currentNode ∈ ST) do
        lastNode ← currentNode
        currentNode ← Select(currentNode)
    end
    /* A simulated game is played                              */
    r ← PlayOut(currentNode)
    /* A node is added                                         */
    lastNode ← Expand(lastNode, currentNode)
    /* The result is backpropagated                            */
    currentNode ← lastNode
    while (currentNode ∈ ST) do
        Backpropagation(currentNode, r)
        currentNode ← currentNode.parent
    end
end
return bestMove ← argmax_{a∈A}(root)
```

Algorithm 1 Pseudocode for Monte-Carlo tree search

MCTS-Solver

Although MCTS is unable to *prove* the game-theoretic value, in the long run, MCTS equipped with the UCT formula is able to *converge* to the game-theoretic value. For example, in endgame positions of fixed-termination games like Go or Amazons, MCTS is often able to find the optimal move relatively fast. But in a tactical game, where the main line towards the winning position is typically narrow with many non-progressing alternatives, MCTS may often lead to an erroneous outcome because the nodes' values in the tree do not converge fast enough to their game-theoretic value.

In this section, the MCTS variant called MCTS-Solver is explained (Winands et al. 2010), which is able to prove the game-theoretic value of a position with a binary outcome (i.e., win or loss). MCTS-Solver has been successfully applied in Breakthrough, Hex, Havannah, Lines of Action, and Shogi. The backpropagation and selection steps are modified for this variant, as well as the procedure for choosing the final move to play. The changes are described in subsections "Backpropagation," "Selection," and "Final Move Selection," respectively.

Backpropagation

Proving a win works as follows if at one of the children a win is found for the player who has to move in the current node, then this node is a win for this player. If all children lead to a win for the same opponent, then the current node is also labeled as a win for this opponent. However, if the children lead to wins for different opponents, then only the simulation result r is used ($\{1, 0.5, 0\}$) for simulations ending in a win, draw, and loss, respectively.

An example for a three-player game is given in Fig. 2 (Nijssen and Winands 2013). Assume node E, which is a terminal node where Player 1 has won, is added to the tree. It means that node B is a mate-in-one for Player 1, regardless of the value of node F. This node receives a game-theoretic value of $(\infty, -\infty, -\infty)$. Nodes G, H, and I all result in wins for Player 2. Then parent node D receives a game-theoretic value of $(-\infty, \infty, -\infty)$, since this node always leads to a win for the same opponent of Player 1. The game-theoretic value of node A cannot be determined in this case, because both Player 1 and Player 2 can win and there is no win for Player 3. As the added node E was a (proven) win for Player 1, the simulation result (1,0,0) is further propagated.

Selection

As seen in the previous subsection, a node can have a proven game-theoretic value of ∞ or $-\infty$. The question arises how these game-theoretic values affect the selection strategy. When entering a node with such a proven value, that value can simply be returned without any selection taking place. A more interesting case is when the node itself has a nonproven value, but some of its children have.

Assume that one or more moves of node p are proven to lead to the loss for the player to move in p. It is tempting to discard them in the selection step based on the argument that one would never pick them. However, this can lead to overestimating the value of node p, especially when moves are semi-randomly selected by the simulation strategy. For example, Fig. 3 has three one-ply subtrees. Leaf nodes B and C are proven to be a loss (for player to move in A), indicated by $-\infty$; the numbers below the other leaves are the *expected* pay-off values (also from the perspective of the player to move in A). Assume moves are selected with the same likelihood (as could happen when a simulation strategy is applied). If the loss nodes

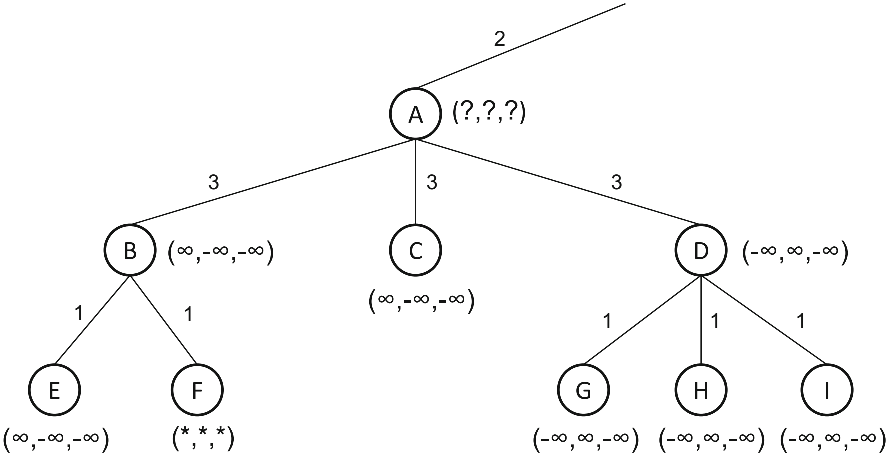


Fig. 2 Example of backpropagating game-theoretic values in an MCTS tree

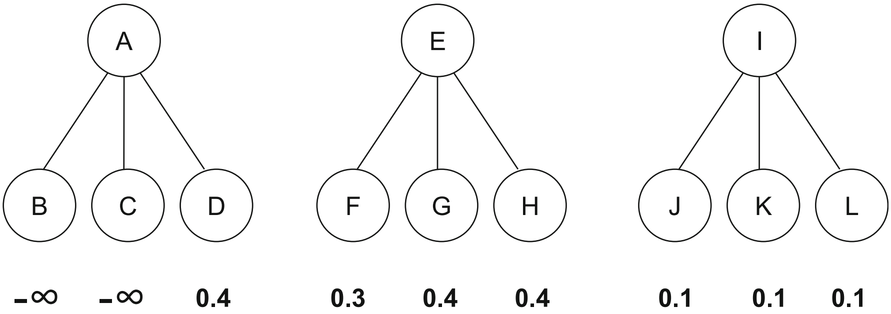


Fig. 3 Monte-Carlo subtrees

would be pruned, node *A* would be preferred above *E*. The average of *A* would be 0.4 and 0.37 for *E*. It is easy to see that *A* is overestimated because *E* has more good moves.

Conversely, if proven loss nodes are not pruned, there is the risk of underestimation. Especially, when there is a strong preference for certain moves (because of a bias) or when the selection strategy is still in its explorative phase (because of the UCT formula), positions could be underestimated. Assume that there is a strong preference for the first move in the subtrees of Fig. 3. Node *I* would be preferred than *A*. It is easy to see that *A* is underestimated because *I* has no good moves at all.

The most effective selection is performed in the following way (Winands et al. 2010). In case Formula 1 is applied, moves leading to a loss for the player are never selected. For nodes that instead select moves according to a simulation strategy, that is, nodes having the visit count below the preset threshold, moves leading to a loss *can* be selected.

One additional improvement is to perform a one-ply lookahead at leaf nodes (i.e., where the visit count equals one) (Winands et al. 2010). It is checked whether they lead to a direct win for the player to move. If there is such a move, the playout can be skipped, the node is labeled as a win, and the backpropagation step is started. If it were not for such a lookahead, it could take many simulations before a child leading to a mate-in-one is selected and the node proven.

Final Move Selection

For MCTS-Solver, final move selection does somewhat matter. Because of the backpropagation of game-theoretic values, the score of a move can suddenly drop or rise. Therefore, the move with the highest value has to be chosen. Finally, when a win can be proven for the root node, the search is immediately terminated, and the winning move is played. In case of a proven loss, the move with the highest number of visits has to be chosen. This move will likely delay the forced loss the most, therefore increasing the probability that a fallible opponent makes a mistake.

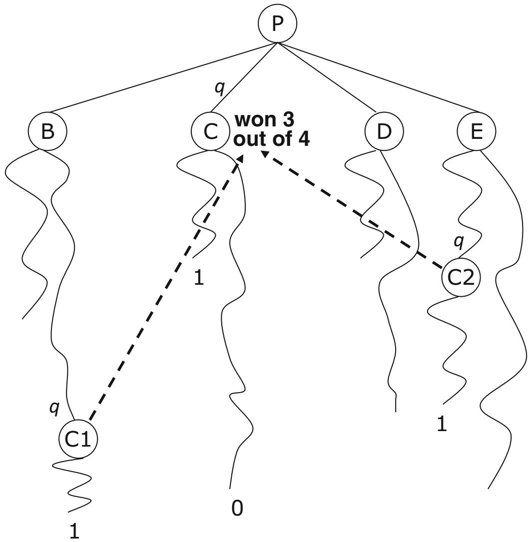
Selection Strategies

When nodes are visited for the first time or if the number of visits is still small, the UCT mechanism gives unreliable results. Several strategies have been developed to solve this problem, which are discussed in this section. These selection strategies can be categorized either as *domain independent* such as "RAVE" and "Progressive History" or *domain dependent* such as "Progressive Bias," "Prior Knowledge," "Implicit Minimax," and "Progressive Widening." For the former, information from the simulations is extracted. For the latter, (handcrafted) heuristic knowledge such as move patterns and static board evaluators is used. When a few of these enhancements are successfully added to MCTS, the UCT constant C becomes usually very small or even 0.

RAVE

Results from simulations can be acquired quicker by the "all-move-as-first heuristic" (AMAF) (Brügmann 1993). For a given position p, AMAF assigns each move a with a value. This value is computed in the following way, which considers each move of the simulated game as important as the first move. For every simulated game S_t played starting from p, in which the move a has been played, $S_t(p, a) = 1$ if the player who played a won the simulated game, and $S_t(p, a) = 0$ if the player who played a lost the simulated game. The AMAF value is then the average over t of the $S_t(p, a)$ values. The AMAF values can be computed incrementally.

Fig. 4 AMAF update



An example is provided in Fig. 4. Node P has four children B, C, D, and E, which have been sampled each twice. For node C, the first simulation led to a win and the second one to a loss, resulting to $v_C = 0.5$. Its corresponding move q has actually been played four times, twice in the simulation going through node C, but also twice in simulations going through node B and E. These latter two simulations scored a win. The AMAF score for node C is three wins out of four simulations, thus 0.75.

Rapid Action-Value Estimator (RAVE) combines UCT with AMAF values (Gelly and Silver 2007). The goal of RAVE is to increase the amount of usable information if the number of visits at a node is small. As RAVE biases a move based on its historical performance, it is closely related to the history heuristic in $\alpha\beta$ search (Schaeffer 1983). RAVE has many implementations for different games (Browne et al. 2012); for this chapter, the following is used.

In the MCTS tree, the AMAF values of a move a at node p are stored at its corresponding child node i (i.e., $AMAF_i$). The selected node b has now to satisfy the following modified UCT formula:

$$b \in \text{argmax}_{i \in I}\left(\left(1 - \beta\left(m_p\right)\right) \times \left(v_i + C \times \sqrt{\frac{\ln m_p}{m_i}}\right) + \beta\left(m_p\right) \times \text{AMAF}_i\right) \quad (2)$$

In Formula 2, p is the position associated with the current node and β is a coefficient, which goes to 0 as the number of simulations grows. For instance, in the Go program MoGo, $\beta(m) = \sqrt{\frac{K}{3m + K}}$ was used (Gelly and Silver 2007). Good results were achieved for all values of K from 3 to 10,000, with the best result obtained with $K = 1{,}000$. The RAVE enhancement has been successfully applied in Go, Havannah, and Hex.

Progressive Bias

UCT can be enhanced with progressive bias (PB) (Chaslot et al. 2008a). PB is a technique to embed domain-knowledge bias into the UCT formula. It has been successfully applied in Go, Hex, and Lines of Action.

UCT with PB works as follows. Let I be the set of nodes immediately reachable from the current node p. The selection strategy selects the child b of node p that satisfies Formula 3:

$$b \in \text{argmax}_{i \in I}\left(v_i + C \times \sqrt{\frac{\ln m_p}{m_i}} + \frac{H_i}{l_i + 1}\right) \tag{3}$$

where $\frac{H_i}{l_i+1}$ is the PB part of the formula. H_i represents the heuristic knowledge. Instead of dividing the PB part by the visit count m_i as done originally (Chaslot et al. 2008a), it is here divided by the number of losses l_i (draws are counted as a half loss) (Nijssen and Winands 2011). In this approach, nodes that do not perform well are not biased too long, whereas nodes that continue to have a high score continue to be biased. To ensure that there is no division by 0, a 1 is added in the denominator. For nodes that have not been visited yet, the default score is set to $1 + H_i$.

Progressive History

A disadvantage of progressive bias is that heuristic knowledge is required. In progressive history (Nijssen and Winands 2011), the heuristic knowledge H_i of progressive bias is replaced with (relative) history score (Schaeffer 1983). The idea is to exploit the fact that moves that are good in a certain position are also good in other positions. For each move that has been performed for each player during all the simulations, the number of games and the total score are stored. This information is then used to compute the history score. It represents the average score of all simulated games for that particular move. This domain-independent variant has been successfully applied in Chinese Checkers, Focus, Havannah, Rollit, and Scotland Yard.

Progressive history has similarities with RAVE as both are based on the history heuristic. For progressive history, the move data is stored, for each player separately, in a global table, while RAVE keeps track of the AMAF values in every node (or edge). An advantage of progressive history is that keeping track of the values globally instead of locally at every node saves memory space. Because the size of the table does not grow as the MCTS tree becomes larger, its memory complexity is $O(1)$. The memory complexity of RAVE is $O(V)$, because the amount of memory necessary to store the AMAF values increases linearly with the number of nodes V as the tree grows larger. Another advantage of using a global table is that, contrary to RAVE, information may be immediately available when a node is visited for the first time. A possible disadvantage of using a single history table is that the available information may be less accurate for that particular part of the tree. In order to prevent outdated information from influencing the performance of progressive history, the history table is emptied after a move is played in the actual game.

Prior Knowledge

An alternative for biasing the selection strategy with heuristic knowledge is so-called prior knowledge (Gelly and Silver 2007). It uses the concept of "virtual" wins and losses. Instead of initializing R and m to 0 for a node i, they are initialized as follows:

$$R_i \leftarrow Q_i \times m_{\text{prior}} \tag{4}$$

$$m_i \leftarrow m_{\text{prior}} \tag{5}$$

where Q_i, which is in [0, 1], is the prior estimation of the position and m_{prior} is the weight that controls the influence of the prior knowledge. After this initialization, R and m are updated in the default way. For example, if the prior estimation of as position is 0.75 and the weight is 100, 75 virtual wins and 25 virtual losses are assumed. Because the heuristic knowledge estimates are typically more reliable than the MCTS value estimates resulting from only a few samples, this prior helps to guide the search into a promising direction. If the node is visited frequently though, the influence of the prior progressively decreases over time, as the virtual playout results represent an increasingly smaller ratio of the total number of playout results stored in the node. Thus, MCTS playouts progressively override the heuristic knowledge. This weight has to be tuned experimentally. The prior knowledge enhancement has been successfully applied in Breakthrough, Go, Hex, and Settlers of Catan.

Implicit Minimax

Implicit minimax (Lanctot et al. 2014) is based on the following principle: if an evaluation function is available, then it should be possible to augment MCTS to make use of it for a potential gain in performance. Suppose an evaluation function $eval(i)$ is available whose range is the same as that of the simulation result r. Implicit minimax uses the following simple and elegant solution: add another value to maintain at each node the *implicit minimax evaluation with respect to player j*. This new value at node *i only* maintains a heuristic minimax value built from the evaluations of its subtrees. During backpropagation, R and m are updated in the usual way, and additionally IM_i is updated using minimax backup rule based on children values. Then, similarly to RAVE, rather than using the average score for selection in Formula 1, the following formula is used instead to compute v_i:

$$v_i \leftarrow (1-\alpha)\frac{R_{i,j}}{m_i} + \alpha \times IM_i, \tag{6}$$

where α weights the influence of the heuristic minimax value. During backpropagation, the implicit minimax evaluations IM_i are updated based on the children's values in a $\max^n$ fashion.

Implicit minimax assumes that all children are added at once and subsequently scored by the static evaluation function. A more memory-efficient implementation could add just a single child without fundamentally changing the algorithm. When not all children of a node p have been evaluated, the following precaution has to be taken. In case their implicit minimax values are lower than the evaluation score of p, the latter should be used.

In essence, implicit minimax is a new *information scheme* where each node is augmented with heuristic estimates which are backed up differently than the Monte-Carlo statistics. When MCTS-Solver is enabled, proven values take precedence in the selection policy, and the resulting scheme is informative and consistent, so the algorithm converges to the optimal choice eventually. However, before a node becomes a proven win or loss, the implicit minimax values act like a heuristic approximation of MCTS-Solver for the portion of the search tree that has not reached terminal positions. The implicit minimax enhancement has been successfully applied in Breakthrough and Lines of Action.

Progressive Widening

The previous strategies biased the selection in the direction of the seemingly promising moves while neglecting the unpromising moves. These unpromising moves are still selected a few times. However, for games with a large branching factor (>100), this is still insufficient for achieving a large lookahead. A soft pruning mechanism can be applied that postpones unpromising moves being selected and searched. Such a mechanism is progressive widening (Chaslot et al. 2008a; Coulom 2007a), which has been applied in large games such as 19 × 19 Go and Amazons.

Progressive widening consists of (1) reducing the move set artificially when the selection strategy is applied and (2) increasing its size progressively (given enough time). When the number of visits n_p of a node p equals the threshold T, progressive widening prunes most of the children. The children, which are not pruned from the beginning, are the k_{init} children with the highest heuristic scores H_i. Next, the children of the node i are progressively "unpruned" ranked to their heuristic score H_i. Every time k nodes are unpruned when the number of simulations in the parent surpasses $A \times B^{k-k_{init}}$ (e.g., $A = 50$ and $B = 1.3$). An outline of progressive widening is given in Fig. 5.

Progressive widening alleviates the risk that the best move may have been prematurely pruned and removed from consideration. However, it requires reliable domain knowledge, which is not always available when using MCTS. Even so, this

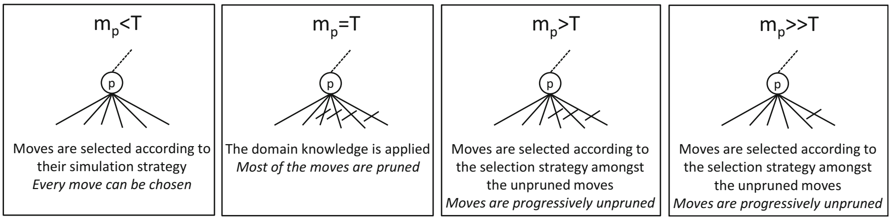


Fig. 5 Progressive widening

kind of domain knowledge can be slow to compute, and therefore, progressive widening waits until a node had been visited a fixed number of times T before calculating H_i. So storing the result and limiting the number of nodes that use expensive domain knowledge reduce the overhead.

Simulation Strategies

MCTS can already work quite well with only playing random moves in the playout. To get more out of the simulations, they have to mimic intelligent play. Simulation strategies aim to have more realistic and therefore informative playouts. This section discusses first domain-independent strategies, "MAST" and "N-Grams," and subsequently domain-dependent ones, applying move ordering "Static Move Categories" and evaluation functions "Integrating Evaluation Functions."

Move-Average Sampling Technique

The Move-Average Sampling Technique (MAST) (Björnsson and Finnsson 2009) is based on the principle that moves good in one position are likely to be good in other positions as well. The history heuristic (Schaeffer 1983), which is used to order moves in $\alpha\beta$ search (Knuth and Moore 1975) and to bias moves in the selection step of MCTS (see subsection "Progressive History"), is based on the same principle. For each move a, a global average $Q_h(a)$ is kept in memory. It is the average of the results of the playouts in which move a has occurred. These values are used to bias the selection of moves, primarily in the playout step but also for tie-breaking unexplored positions in the selection step. In the original version of MAST, a softmax-based Gibbs measure is used:

$$P(s,a) = \frac{e^{Q_h(a)/\tau}}{\sum_{b \in \mathcal{A}(s)} e^{Q_h(b)/\tau}} \tag{7}$$

where $P(s, a)$ is the probability that move a will be selected in position s. Moves with a higher $Q_h(a)$ value are more likely to be selected. How greedy the selection is can be tuned with the τ parameter. In order to bias the selection of unexplored moves, the initial $Q_h(a)$ value is set to the maximum possible score.

A disadvantage of the Gibbs measure is that the probability of selecting the move with the highest $Q_h(a)$ score is not fixed and unknown. Consequently, it is not assured that moves with the highest $Q_h(a)$ scores will be even chosen at all. Therefore, a different exploration technique has been proposed also based on the $Q_h(a)$ values, namely, ϵ-greedy (Sturtevant 2008; Sutton and Barto 1998). With a probability of $1 - \epsilon$, the move with the highest $Q_h(a)$ value is selected, and with a probability of ϵ, a legal move is chosen uniformly at random.

MAST generalizes the merits of moves from one position to the next without considering the context in which the moves are made. Nonetheless, despite its simplicity, this technique has been quite successful in General Game Playing (Björnsson and Finnsson 2009) and Lord of the Rings: The Confrontation (Powley et al. 2013).

N-Grams

The *N-Gram Selection Technique* (*NST*) keeps track of move sequences instead of single moves (Tak et al. 2012). It offers context in a way that is more favorable in terms of added benefits vs. computational overhead. NST variants have been applied successfully in Havannah (Stankiewicz et al. 2012), Lord of the Rings: The Confrontation (Powley et al. 2013), and General Game Playing (Tak et al. 2012).

NST is based on N-Gram models, often employed in computer games for predicting the next move of the opponent (Millington and Funge 2009). The N-Grams in NST consist of consecutive move sequences z of length 1, 2, and 3. Similar to MAST, the average of the results of the playouts is accumulated. However, the average, here called $Q(z)$, is kept also for longer move sequences as opposed to only single moves.

The N-Grams are formed as follows. After each simulation, starting at the root of the tree, for each player all move sequences of length 1, 2, and 3 that appeared in the simulated game are extracted. The averages of these sequences are updated with the obtained result from the playout.

The move sequences consist of moves from the current player and moves from the opponent(s). Role numbers $0, 1, 2, \cdots, n - 1$, which are assigned to the players at the beginning of a game with n players, are employed in order to determine the move of which opponent to include in the sequences. Suppose that the current player has role number j and there are n players, then the sequences are constructed as follows. A sequence of length 1 consists of just one move of the current player. A sequence of length 2 starts with a move of player with role $(j + n - 1) \bmod n$ and ends with a move of the current player. A sequence of length 3 starts with a move of player with role $(j + n - 2) \bmod n$, followed by a move of the player with role $(j + n - 1) \bmod n$ and ends with a move made by the current player. The moves in these sequences are consecutive moves. Figure 6 gives an example of a simulation.

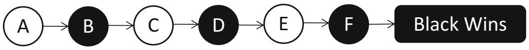


Fig. 6 An MCTS simulation

The sequences of length 3 that have to be updated are A-B-C, B-C-D, C-D-E, and D-E-F. For the ones of length 2, they are A-B, B-C, C-D, D-E, and E-F.

In the playout, and at the nodes of the MCTS tree containing unvisited legal moves, the N-Grams are used to determine which move to select. For each legal move, the player determines which sequence of length 1, which sequence of length 2, and which sequence of length 3 would occur when that move is played. The sequence of length 1 is just the move itself. The sequence of length 2 is the move itself appended to the last move played by player with role $(j + n - 1) \bmod n$. The sequence of length 3 is the move itself appended to the last move played by player with role $(i + n - 1) \bmod n$ and the last move played by the player with role $(\mathrm{j} + n - 2) \bmod n$. Thus, in total three sequences could occur. The player then calculates the score $G(a)$ for the move a by taking the unweighted average of the $\mathrm{Q}(z)$ values stored for these sequences.

In this calculation, the $\mathrm{Q}(z)$ values for the move sequences of length 2 and length 3 are only taken into account if they are visited at least g times. If a move has been played at least once, but the sequences of length 2 and length 3 occurred less than g times, then the $\mathrm{Q}(z)$ value of the move sequence of length 1 (which is the move itself) is returned. The g parameter thus prevents move sequences with only a few visits from being considered.

If a move has never been played before, then there exists no move sequence, and the calculation outlined above is not possible. In that case the score is set to the maximum possible value in order to bias the selection towards unexplored moves.

In this manner, a score $G(a)$ is assigned to each legal move in a given position. These scores are then used with an ϵ-greedy strategy to determine which move to select. With a probability of $1 - \epsilon$, the move with the highest $G(a)$ value is selected; otherwise, a legal move is chosen uniformly at random.

Furthermore, the entries of the N-Gram tables are multiplied by a decay factor γ after every move played in the actual game, where $0 \leq \gamma \leq 1$ (Tak et al. 2014). This ensures that, as the game progresses, new move sequences are tried as well, instead of only playing the move sequence over and over again.

Static Move Categories

Instead of playing completely randomly or using domain-independent techniques such as MAST or N-Grams, light heuristic knowledge can be used to select a move in the playout. Moves can be categorized, for instance, on their board location, capturing or blocking potential, and proximity to the last move. These move categories have weights indicating how promising they are. These weights can be manually determined by a human expert or fine-tuned by a machine-learning approach.

Another approach is to determine for each move category the probability that a move belonging to that category will be played in a real game. This probability is called the *transition probability* (Tsuruoka et al. 2002). This statistic is obtained from game records of matches played by expert players (either human or machine). The transition probability P for a move category mc is calculated as follows:

$$P_{mc} \leftarrow \frac{f_{\text{played}(mc)}}{f_{\text{available}(mc)}} \tag{8}$$

where $f_{\text{played}(mc)}$ is the number of game positions in which a move belonging to category mc was played and $f_{\text{available }(mc)}$ is the number of positions in which moves belonging to category mc were available.

The main idea is to play now interesting moves based on the move categories. A straightforward approach is to exclude moves who belong to a category that is almost always bad (e.g., a suicide move), or to enforce moves who belong to a category that is almost always the best (e.g., capturing a large number of stones).

In case such an approach is not possible, an ϵ-greedy strategy can be applied, in which the most greedy move is chosen (the move with the highest weight according to the move categories) with a probability $1 - \epsilon$. A random move is selected with a probability of ϵ. An alternative is roulette-wheel selection where the moves are sampled based on their weights. The pseudocode for roulette-wheel selection is given in Algorithm 2.

```
Data: moveList
Result: bestMove
U ← 0
foreach Move a ∈ moveList do
  | U ← U + weight[a]
end
U ← U × random(0, 1)
foreach Move a ∈ moveList do
  | U ← U − weight[a]
  | if U ≤ 0 then
  |   | return a
  | end
end
```

Algorithm 2 Roulette-wheel selection

Integrating Evaluation Functions

Adding sophisticated heuristic knowledge to the simulation strategy increases the accuracy and reliability of each playout. However, if the heuristic knowledge is too computationally expensive, the number of playouts per second decreases, offsetting the benefits. In this subsection, four strategies to integrate evaluation functions that assess complete board positions are discussed in detail.

Evaluation Cutoff

Instead of simulating the game until the end, an evaluation function could be used to assess the game prematurely. Such an evaluation cutoff strategy has improved the performance of MCTS engines in Amazons, Breakthrough, Kriegspiel, and Lines of Action. There are two basic approaches to apply an evaluation function as a cutoff mechanism, *dynamic* and *fixed length*.

The *dynamic cutoff* strategy (Winands et al. 2010) stops a simulated game before a terminal position is reached if, according to the heuristic knowledge, the game is judged to be effectively over. In general, once a position gets very lopsided, an evaluation function can return a quite trustworthy score, more so than even elaborate simulation strategies. The game can thus be (relatively) safely terminated both earlier and with a more accurate score than if continuing the simulation (which might, e.g., fail to deliver the win). This is somewhat analogous to the "mercy rule" in computer Go. For instance, in Fuego a simulation can also end when the difference in number of captured stones exceeds 30 % of the board (Enzenberger et al. 2010). When the evaluation function gives a value that exceeds a certain threshold, the game is scored as a win. Conversely, if the evaluation function gives a value that is below the negated threshold, the game is scored as a loss. If on average, the evaluation function is computationally cheaper than playing out the rest of the simulation, this strategy can also increase the number of simulations per second. It might be even a better choice to use a more aggressive cutoff threshold even though being occasionally wrong. The added number of simulations might more than offset the errors introduced by the occasional erroneous termination decisions. This trade-off has to be determined empirically.

An alternative is the *fixed-length cutoff* strategy (Lorentz 2008). Here the difference is that the simulation stops after a fixed number of d moves have been executed in the playout, whereas in the previous approach the simulation may terminate at any time. The simulation is either scored as a win for the player ahead according to the evaluation function or using directly the *value* of the evaluation function (scaled to [0,1]). An improved version of this strategy also accounts for quiescence and allows a simulation to run past the limit of d (e.g., in the case of a capture sequence for the game of Kriegspiel (Ciancarini and Favini 2010)). When the evaluation function is quite accurate in comparison to the default simulation strategy, d will be small (<5).

Corrective

One known disadvantage of simulation strategies is that they may draw and play a move which immediately ruins a perfectly healthy position. In the *corrective* strategy (Winands et al. 2010), the evaluation function is used to further bias the move selection towards minimizing the risk of choosing an obviously bad move.

This is done in the following way. First, the position for which a move is chosen is evaluated. Next, the moves are generated and scanned to get their weights. If the move leads to a successor that has a lower evaluation score than its parent, the weight of that move is set to a preset minimum value (close to 0). If a move leads to a win, it is immediately played. The effectiveness of the algorithm is partially determined by how efficiently game positions and moves are evaluated.

Greedy

In the *greedy* strategy (Winands et al. 2010), the evaluation function is more directly applied for selecting moves: the move leading to the position with the highest evaluation score is selected. However, because evaluating every move is time consuming, only moves that have a good potential for being the best are evaluated. For this strategy, it means that only the k-best moves according to their move ordering (e.g., the move categories) are fully evaluated. As in the dynamic cutoff strategy, when a move leads to a position with an evaluation over a preset threshold, the playout is stopped and scored as a win. Finally, the remaining moves, which are not heuristically evaluated, are checked for a mate.

A potential weakness of the greedy strategy is that despite a small random factor in the evaluation function, it is too deterministic. A solution is to use an ϵ-greedy strategy as seen in subsections "Move-Average Sampling Technique," "N-Grams," and "Static Move Categories." The risk is again that the players select a move that spoils a healthy position. As the opponent will likely play greedy, a mistake is immediately punished. This may happen a lot in a sudden-death game. An alternative is the *mixed* strategy (Winands et al. 2010) that combines the corrective strategy and the greedy strategy. The corrective strategy is used in the selection step, i.e., at tree nodes where a simulation strategy is needed (i.e., $m < T$), as well as in the first positions entered in the playout step. For the remainder of the playout, the greedy strategy is applied.

Choosing a move based on a selective one-ply search equipped with a static evaluation function (i.e., the greedy strategy) has been shown to perform better than drawing the moves based only on light knowledge items (i.e., the move categories) in several board games such as Chinese Checkers, Focus, and Lines of Action.

Minimax Search in the Playout Step

In the previous subsection, selective one-ply search was introduced for the playout. This direction can be further taken, by applying relatively small d-ply selective minimax-based searches to choose the moves in the playout step (Nijssen and Winands 2012b; Winands et al. 2011). A schematic sketch is depicted in Fig. 7. However, to reduce the computational overhead, this can be done selectively, by fully exploring only the first k moves at every node. The $\alpha\beta$ mechanism (Knuth and Moore 1975) is applied to prune the branches irrelevant for the search tree. The success of $\alpha\beta$ search is strongly dependent on move ordering (Marsland 1986). A simple but effective scheme is first to try two killer moves. These are the last two

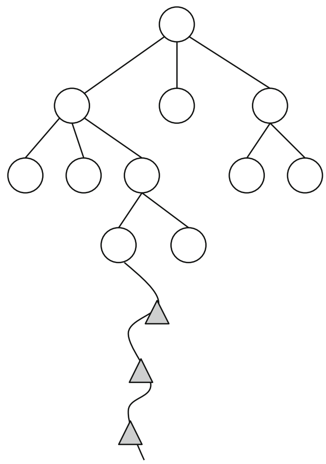


Fig. 7 Search in the playout step

moves that were best, or at least caused a cutoff, at the given depth. The remaining moves can be sorted by the static move ordering using the history tables as the tiebreaker. Finally, an aspiration window when invoking the search can be used to prune even more branches (Marsland 1986). This window is based on the threshold configuration that is used to stop the playout dynamically. When the value of this $\alpha\beta$ search lies outside the window, the simulation is stopped prematurely. If the value exceeds the preset threshold, the simulation is scored as a win. Conversely, if the search gives a value that is below the negated threshold, the simulation is scored as a loss.

For determining moves based on a small lookahead in multiplayer games and stochastic games, paranoid search and expectimax search can be applied, respectively. When a sufficient amount of thinking time per move is available (5–30 s), this strategy is successful in the games of Chinese Checkers, Focus, and Lines of Action. With current hardware means only a two- or three-ply search is feasible in games with a modest branching factor (<30). However, it is not unlikely that deeper searches can be successfully utilized in the near future as machines become faster. Finally, a performance increase due to search lookahead can also be obtained without a static board evaluation function. Checking for a one-ply win or avoiding a two-ply loss in the playout has increased performance in the game of Havannah (Teytaud et al. 2010).

Parallelization

The development of computer hardware has gone into the direction that nowadays even personal computers contain several cores. To get the most out of the available computing resources, MCTS has to be parallelized. There are three main approaches for parallelization (Chaslot et al. 2008b), depending on which part of

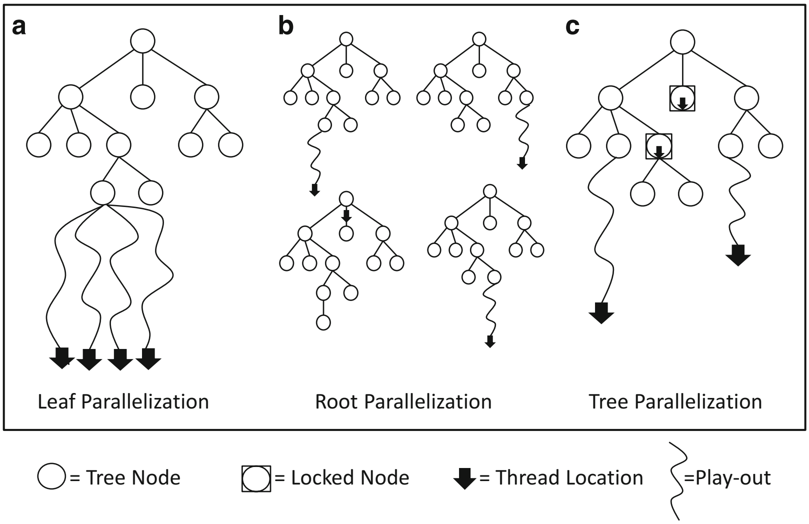


Fig. 8 (**a**) Leaf parallelization (**b**) Root parallelization (**c**) Tree parallelization

MCTS is parallelized: "Leaf Parallelization," "Root Parallelization," and "Tree Parallelization,"

Leaf Parallelization

Leaf parallelization (Cazenave and Jouandeau 2007) is a quite easy approach to parallelize MCTS. Only one thread traverses the tree and adds a leaf node to the tree. Next, starting from the leaf node, independent playouts are executed for each available thread. When all simulated games are finished, the results are propagated backwards through the tree by one single thread. Leaf parallelization is depicted in Fig. 8a.

Leaf parallelization is interesting because its implementation is easy and does not require any locks. However, two problems emerge. First, simulating using t different threads takes more time on average than simulating one single game using one thread, since the program has to wait for the longest simulated game. Second, information is not shared. For instance, if 16 threads are available, and 8 (faster) finished games are all losses, it is highly probable that most games lead to a loss. Therefore, playing eight more games is a waste of computational power. To decrease the waiting time, the program might stop the simulations that are still running when the results of the finished simulations become available. This strategy would enable the program to traverse the tree more often, but some threads would still be idle. Leaf parallelization can be performed on a multiprocessor/core machine with shared memory, or even on a cluster using MPI (Message Passing Interface) communication.

Root Parallelization

A second parallelization (Cazenave and Jouandeau 2007) approach is *root parallelization*. It consists of building multiple MCTS trees in parallel, with one thread per tree. Similar to leaf parallelization, the threads do not share information with each other. When the available time is up, all the root children of the separate MCTS trees are merged with their corresponding clones. For each group of clones, the scores of all games played are added. The best move is selected based on this grand total. This parallelization method only requires a minimal amount of communication between threads, so the parallelization is easy, even on a cluster. Root parallelization is depicted in Fig. 8b. For a small number of threads, root parallelization performs remarkably well (Cazenave and Jouandeau 2007; Chaslot et al. 2008b). However, root parallelization does not scale well for a larger number of threads (Soejima et al. 2010).

Tree Parallelization

The third parallelization method is called *tree parallelization*. On shared memory systems, tree parallelization is the natural approach that takes full advantage of the available bandwidth to communicate simulation results. It uses one shared tree from which several simultaneous games are played. Each thread could modify the information contained at any node, which may lead to data corruption if certain parts of the tree are not locked from time to time. This locking should be done *locally*, not *globally* (i.e., the whole tree). When a thread visits a node, it should be locked. At the moment it leaves the node, it should be unlocked. Tree parallelization with locking is depicted in Fig. 8c. For a medium number of threads (<8), there is no strong need for locks in practice as MCTS is able to handle noise rather well (Enzenberger and Müller 2010).

If several threads start from the root at the same time, it is possible that they traverse the tree for a large part in the same way. Playouts might start from nodes, which are in the neighborhood of each other. It can even happen that playouts start from the same node. To prevent that only a small part of the search space is explored, a solution is to assign a virtual loss when a node is visited by a thread (i.e., in the selection step). Hence, the value of this node is decreased. The next thread will only select the same node if its value remains better than its siblings' values. The virtual loss is removed when the thread that gave the virtual loss starts propagating the result the finished simulated game (i.e., in the backpropagation step). Due to this mechanism, nodes that are clearly better than others are still explored by all threads, while nodes for which the value is uncertain are not explored by more than one thread. Therefore, this enhancement keeps a certain balance between exploration and exploitation in a parallelized MCTS program.

Uncertainty

Uncertainty can be introduced in a board game in two ways: (1) by hiding certain parts of the position for the player (e.g., hiding the rank of piece in Stratego) and (2) by introducing randomness (e.g., rolling a die in Backgammon). The first is called imperfect information (or a partial observable environment); the second is called stochasticity. This section discusses determinization for handling imperfect information in MCTS (subsection "Determinization") and chance nodes for stochasticity (subsection "Chance Nodes").

Determinization

To deal with board games having imperfect information, determinization can be applied in the MCTS engine. The principle behind determinization is that, at the start of each simulation at the root, the hidden information is filled in, while being consistent with the history of the game.

Determinization has been called "averaging over clairvoyance" (Russell and Norvig 2010), where players never try to hide or gain information, because in each determinization, all information is already available. Despite these shortcomings, it has produced strong results in the past, for instance, in Monte-Carlo engines for the trick-based card game Bridge (Ginsberg 1999), the card game Skat (Buro et al. 2009), Scrabble (Sheppard 2002), and Phantom Go (Cazenave 2006).

Determinization in the MCTS framework has been applied in games such as Scotland Yard and Lord of the Rings: The Confrontation. It works as follows: for each MCTS simulation starting at the root, the missing information is filled in a random manner. The determinization is used throughout the whole simulation. There are two approaches to build and traverse the search tree.

The first approach is by generating a separate tree for each determinization (also called determinized UCT (Cowling et al. 2012)). After selecting a determinization at the root node, the corresponding tree is traversed. Based on majority voting (Soejima et al. 2010), the final move is selected. Each candidate move receives one vote from each tree where it is the move that was played most often. The candidate move with the highest number of votes is selected as the best move. If more moves are tied, the move with the highest number of visits over all trees is selected. The concept of separate-tree determinization is similar to root parallelization.

The second approach is using single-tree determinization (Ciancarini and Favini 2010; Cowling et al. 2012; Nijssen and Winands 2012a). When generating the tree, all possible moves from all possible determinizations are generated. When traversing the tree, only the moves consistent with the current determinization are considered. An example is given in Fig. 9. The advantage of this technique is that information is shared between different determinizations, increasing the amount of usable information. This type of determinization is also called Single-Observer Information Set Monte-Carlo Tree Search (Cowling et al. 2012).

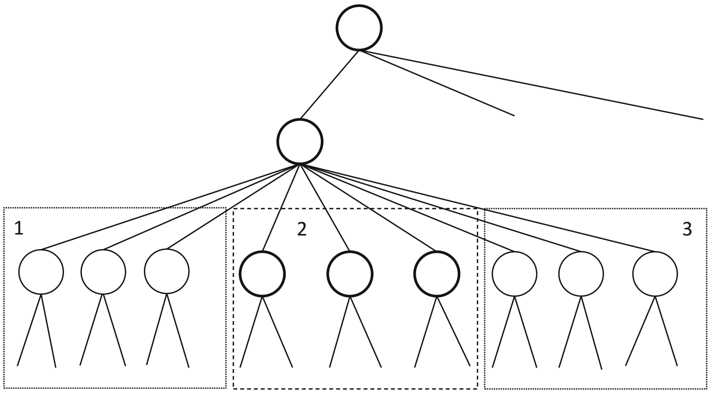


Fig. 9 Example of a determinization with a single tree. In this example, $D = \{1,2,3\}$ and the selected determinization is 2

A potential weakness is that quite a number of the determinizations generated is unrealistic (Ciancarini and Favini 2010). Strong opponents would not risk to put themselves in a dangerous position. Examples of such are not protecting pieces in Kriegspiel, putting the flag in the front row in Stratego, having Mr. X very close by the Detectives in Scotland Yard. Additionally, opponents could take advantage of the fact they possess information that the player does not have and may direct play in another direction.

Instead of sampling determinizations randomly, a model of belief distributions can be used. A straightforward approach is to define a certain number of features (e.g., King in danger, location of the flag, distance between Mr. X and the Detectives) and gather statistics about their frequency in real tournament games. These statistics can be used to bias determinizations and have shown to improve performance against different opponents in Scotland Yard (Nijssen and Winands 2012a).

Chance Nodes

In classic adversarial search algorithms such as expectimax, chance nodes can be added to the tree to model stochasticity in games. A chance node represents a chance event in a game such as a die role. The children of a chance node indicate the possible outcomes. A similar approach can be used in MCTS as well. The value of a chance node is computed by taking the average of the results of all simulated games made through this node. Selecting a child is typically performed by sampling a single outcome based on a given probability distribution.

The question remains whether it is always useful to explicitly implement a chance node in MCTS. In the case the outcomes of a chance event lead to different board positions, a chance node has to be used. For example, a battle's outcome in the game of Risk is decided by rolling a die, leading to outcomes where board positions have a different number of pieces. However, there are also games where

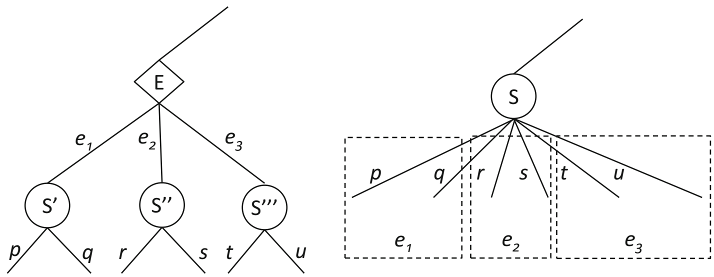


Fig. 10 (**a**) Chance node *E* leading to the same board configuration *S* but with different moves (**b**) Determinization of the *left* figure

the chance event leads to the same board configuration but with a different move set. An example is the game of Backgammon where players can only move their checkers according to the number shown on each die. A chance node is not necessary; a similar approach to the determinization mechanism seen previously can be used instead. Every time a regular node is entered, an outcome is sampled, and subsequently only moves associated to this outcome can be selected by the UCT mechanism (Lorentz 2012). This approach is depicted in Fig. 10.

To improve the performance of MCTS in densely stochastic games, progressive widening can be applied at chance nodes as well. This enhancement is called double progressive widening (Couetoux et al. 2011). It applies progressive widening at regular (decision) and chance nodes. When a decision or chance node is entered, double progressive widening computes a maximum number of moves or outcomes to consider $k \leftarrow \lceil B \times m^{\alpha} \rceil$, where B and α are parameter constants and m represents a number of visits to the node. At a decision node, the enhancement acts in the way as described in subsection "Progressive Widening." At a chance node, a set of outcomes is stored and incrementally grown. If k is larger than the size of the current set of outcomes, an outcome is sampled. If an already added outcome is sampled, then the search selects the associated child. Otherwise, a new child is created and added to the current set; the simulation continues in that direction. Otherwise, double progressive widening samples from existing children at chance nodes in the tree, where a child probability is computed with respect to the current children in the restricted set. As a chance node is visited more frequently, more chance outcomes are gradually added to the set. When the branching factor at chance nodes is extremely high, double progressive widening prevents MCTS from degrading into one-ply Monte-Carlo planning. This enhancement has been applied in stochastic games such as Can't Stop and Ra (Lanctot et al. 2013).

Conclusion

Monte-Carlo Tree Search (MCTS) has caused a revolution in AI for board games due to its three main selling points. (1) As MCTS already works with little domain knowledge required, it is good candidate if handcrafting an evaluation function is difficult or when rapid prototyping is required. (2) The MCTS framework can be easily modified to be successful in a whole range of board games which may have two or more players, chance or determinism, or being fully or only partial observable. (3) MCTS is relatively easy to parallelize and scales quite well, making it an attractive technique for multi-core machines.

However, for MCTS to work effectively, search-control knowledge is required to guide the simulations. Techniques such as RAVE, progressive bias, progressive widening, and others have to be applied in order to get an MCTS agent on expert level for many domains. It is crucial that the playouts are informative. If possible, applying a static evaluation function or a small $\alpha\beta$ search in MCTS can boost its performance even further. It is therefore a myth that classic adversarial techniques will become obsolete. Even if it is not possible for a particular domain to use them directly, enhancements such as static move ordering, history tables, and transposition table can be applied in MCTS in some way or another. Properly understanding ideas from other adversarial search techniques is beneficial for developing a state-of-the-art MCTS board game playing agent.

References

B. Abramson, Expected-outcome: a general model of static evaluation. IEEE Trans. Pattern Anal. Mach. Intell. **12**(2), 182–193 (1990)

B. Arneson, R.B. Hayward, P. Henderson, Monte Carlo tree search in Hex. IEEE Trans. Comput. Intell. AI Games **2**(4), 251–258 (2010)

P. Auer, N. Cesa-Bianchi, P. Fischer, Finite-time analysis of the multi-armed bandit problem. Mach. Learn. **47**(2–3), 235–256 (2002)

Y. Björnsson, H. Finnsson, CadiaPlayer: a simulation-based general game player. IEEE Trans. Comput. Intell. AI Games **1**(1), 4–15 (2009)

B. Bouzy, B. Helmstetter, Monte-Carlo Go developments, in *Advances in Computer Games 10: Many Games, Many Challenges*, ed. by H.J. van den Herik, H. Iida, E.A. Heinz (Kluwer, Boston, 2003), pp. 159–174

C.B. Browne, E. Powley, D. Whitehouse, S.M. Lucas, P.I. Cowling, P. Rohlfshagen, S. Tavener, D. Perez, S. Samothrakis, S. Colton, A survey of Monte Carlo tree search methods. IEEE Trans. Comput. Intell. AI Games **4**(1), 1–43 (2012)

B. Brügmann, *Monte Carlo Go*. Technical report, Physics Department, Syracuse University, Syracuse, NY, 1993

M. Buro, J.R. Long, T. Furtak, N.R. Sturtevant, Improving state evaluation, inference, and search in trick-based card games, in *IJCAI 2009, Proceedings of the 21st International Joint Conference on Artificial Intelligence*, ed. by C. Boutilier (AAAI Press, Menlo Park, CA, 2009), pp. 1407–1413

T. Cazenave, A Phantom Go program, in *Advances in Computer Games (ACG 11)*, ed. by H.J. van den Herik, S.-C. Hsu, T.-S. Hsu, H.H.L.M. Donkers. Lecture Notes of Computer Science, vol. 4250 (Springer, Berlin, 2006), pp. 120–125

T. Cazenave, N. Jouandeau, On the parallelization of UCT, in *Proceedings of the Computer Games Workshop 2007 (CGW 2007)*, ed. by H.J. van den Herik, J.W.H.M. Uiterwijk, M.H.M. Winands, M.P.D. Schadd (Universiteit Maastricht, Maastricht, 2007), pp. 93–101

G.M.J.-B. Chaslot, M.H.M. Winands, H.J. van den Herik, J.W.H.M. Uiterwijk, B. Bouzy, Progressive strategies for Monte-Carlo tree search. New Math. Nat. Comput. **4**(3), 343–357 (2008a)

G.M.J.-B. Chaslot, M.H.M. Winands, H.J. van den Herik, Parallel Monte-Carlo tree search, in *Computers and Games (CG 2008)*, ed. by H.J. van den Herik, X. Xu, Z. Ma, M.H.M. Winands. Lecture Notes in Computer Science (LNCS), vol. 5131 (Springer, Berlin, 2008b), pp. 60–71

P. Ciancarini, G.P. Favini, Monte Carlo tree search in Kriegspiel. AI J. **174**(11), 670–684 (2010)

A. Couetoux, J-B. Hoock, N. Sokolovska, O. Teytaud, N. Bonnard. Continuous upper confidence trees, in *Learning and Intelligent Optimization – 5th International Conference (LION 5)*. Lecture Notes in Computer Science, vol 6683 (Springer Berlin Heidelberg, 2011), pp. 433–445

R. Coulom, Computing "Elo ratings" of move patterns in the game of Go. ICGA J. **30**(4), 199–208 (2007a)

R. Coulom, Efficient selectivity and backup operators in Monte-Carlo tree search, in *Computers and Games (CG 2006)*, ed. by H.J. van den Herik, P. Ciancarini, H.H.L.M. Donkers. Lecture Notes in Computer Science (LNCS), vol. 4630 (Springer, Berlin, 2007b), pp. 72–83

P.I. Cowling, E.J. Powley, D. Whitehouse, Information set Monte Carlo tree search. IEEE Trans. Comput. Intell. AI Games **4**(2), 120–143 (2012)

M. Enzenberger, M. Müller, A lock-free multithreaded Monte-Carlo tree search-algorithm, in *Advances in Computer Games (ACG 2009)*, ed. by H.J. van den Herik, P.H.M. Spronck. Lecture Notes in Computer Science (LNCS), vol. 6048 (Springer, Berlin/Heidelberg, 2010), pp. 14–20

M. Enzenberger, M. Müller, B. Arneson, R. Segal, Fuego – an open-source framework for board games and Go engine based on Monte Carlo tree search. IEEE Trans. Comput. Intell. AI Games **2**(4), 259–270 (2010)

S. Gelly, D. Silver, Combining online and offline knowledge in UCT, in *Proceedings of the International Conference on Machine Learning (ICML)*, ed. by Z. Ghahramani (ACM, New York, 2007), pp. 273–280

S. Gelly, L. Kocsis, M. Schoenauer, M. Sebag, D. Silver, C. Szepesvári, O. Teytaud, The grand challenge of computer Go: Monte Carlo tree search and extensions. Commun. ACM **55**(3), 106–113 (2012)

M.L. Ginsberg, Gib: steps toward an expert-level bridge-playing program, in *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (IJCAI-99)*, ed. by T. Dean, vol. 1 (Morgan Kaufmann, San Francisco, 1999), pp. 584–589

D.E. Knuth, R.W. Moore, An analysis of alpha-beta pruning. Artif. Intell. **6**(4), 293–326 (1975)

L. Kocsis, C. Szepesvári, Bandit based Monte-Carlo planning, in *Machine Learning: ECML 2006*, ed. by J. Fürnkranz, T. Scheffer, M. Spiliopoulou. Lecture Notes in Artificial Intelligence, vol. 4212 (Springer, Berlin, 2006), pp. 282–293

M. Lanctot, A. Saffidine, J. Veness, C. Archibald, M.H.M. Winands. Monte Carlo *-minimax search, in *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence (IJCAI)* (AAAI Press, Menlo Park, CA, USA, 2013), pp. 580–586

M. Lanctot, M.H.M. Winands, T. Pepels, N.R. Sturtevant. Monte carlo tree search with heuristic evaluations using implicit minimax backups, in *2014 I.E. Conference on Computational Intelligence and Games, CIG 2014* (IEEE, Piscataway, NJ, USA, 2014), pp. 341–348

R.J. Lorentz, Amazons discover Monte-Carlo, in *Computers and Games (CG 2008)*, ed. by H.J. van den Herik, X. Xu, Z. Ma, M.H.M. Winands. Lecture Notes in Computer Science (LNCS), vol. 5131 (Springer, Berlin, 2008), pp. 13–24

R.J. Lorentz, An MCTS program to play EinStein Würfelt Nicht! in *Advances in Computer Games (ACG 2011)*, ed. by H.J. van den Herik, A. Plaat. LNCS, vol. 7168 (Springer, Berlin, 2012), pp. 52–59

T.A. Marsland, A review of game-tree pruning. ICCA J. **9**(1), 3–19 (1986)

I. Millington, J. Funge, *Artificial Intelligence for Games, Chapter 7*, 2nd edn. (Morgan Kaufmann, Burlington, 2009), pp. 579–665

J.A.M. Nijssen, M.H.M. Winands, Enhancements for multi-player Monte-Carlo tree search, in *Computers and Games (CG 2010)*, ed. by H.J. van den Herik, H. Iida, A. Plaat. Lecture Notes in Computer Science (LNCS), vol. 6151 (Springer, Berlin, 2011), pp. 238–249

J.A.M. Nijssen, M.H.M. Winands, Monte-Carlo tree search for the hide-and-seek game Scotland Yard. IEEE Trans. Comput. Intell. AI Games **4**(4), 282–294 (2012a)

J.A.M. Nijssen, M.H.M. Winands, Playout search for Monte-Carlo tree search in multi-player games, in *13th International Conference on Advances in Computer Games (ACG 2011)*, ed. by H.J. van den Herik, A. Plaat. Lecture Notes in Computer Science, vol. 7168 (Springer, Berlin, 2012b), pp. 72–83

J.A.M. Nijssen, M.H.M. Winands, Search policies in multi-player games. ICGA J. **36**(1), 3–21 (2013)

T. Pepels, M.J.W. Tak, M. Lanctot, M.H.M. Winands. Quality-based rewards for Monte-Carlo tree search simulations, in *Proceedings of the 21st European Conference on Artificial Intelligence (ECAI 2014)*, ed. by T. Schaub, G. Friedrich, B. O'Sullivan. Frontiers in Artificial Intelligence and Applications, vol 263 (IOS Press, Amsterdam, The Netherlands 2014), pp. 705–710

E.J. Powley, D. Whitehouse, P.I. Cowling. Bandits all the way down: UCB1 as a simulation policy in Monte Carlo tree search, in *Computational Intelligence in Games (CIG), 2013 I.E. Conference on* (IEEE, Piscataway, NJ, USA, 2013), pp. 81–88

H. Robbins, Some aspects of the sequential design of experiments. Bull. Am. Math. Soc. **58**(5), 527–535 (1952)

S.J. Russell, P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd edn. (Prentice-Hall, Upper Saddle River, 2010)

J. Schaeffer, The history heuristic. ICCA J. **6**(3), 16–19 (1983)

B. Sheppard, World-championship-caliber Scrabble. Artif. Intell. **134**(1–2), 241–275 (2002)

Y. Soejima, A. Kishimoto, O. Watanabe, Evaluating root parallelization in Go. IEEE Trans. Comput. Intell. AI Games **2**(4), 278–287 (2010)

J.A. Stankiewicz, M.H.M. Winands, J.W.H.M. Uiterwijk, Monte-Carlo tree search enhancements for Havannah, in *Advances in Computer Games (ACG 13)*, ed. by H.J. van den Herik, A. Plaat. LNCS, vol. 7168 (Springer, Berlin, 2012), pp. 60–71

N.R. Sturtevant, An analysis of UCT in multi-player games. ICGA J. **31**(4), 195–208 (2008)

R.S. Sutton, A.G. Barto, *Reinforcement learning: an introduction* (MIT Press, Cambridge, MA, 1998)

M.J.W. Tak, M.H.M. Winands, Y. Björnsson, N-grams and the last-good-reply policy applied in general game playing. IEEE Trans. Comput. Intell. AI Games **4**(2), 73–83 (2012)

M.J.W. Tak, M.H.M. Winands, Y. Björnsson, Decaying simulation strategies. IEEE Trans. Comput. Intell. AI Games **6**(4), 395–406 (2014)

F. Teytaud, O. Teytaud. On the huge benefit of decisive moves in Monte-Carlo tree search algorithms, in *2010 I.E. Conference on Computational Intelligence and Games (CIG 2010)*, ed. by G.N. Yannakakis, J. Togelius (IEEE, Piscataway, NJ, USA 2010), pp. 359–364

Y. Tsuruoka, D. Yokoyama, T. Chikayama, Game-tree search algorithm based on realization probability. ICGA J. **25**(3), 132–144 (2002)

M.H.M. Winands, Y. Björnsson, J.-T. Saito, Monte Carlo tree search in Lines of Action. IEEE Trans. Comput. Intell. AI Games **2**(4), 239–250 (2010)

M.H.M. Winands, Y. Björnsson. $\alpha\beta$-based play-outs in Monte-Carlo tree search, in *2011 I.E. Conference on Computational Intelligence and Games (CIG 2011)*, ed. by S-B. Cho, S.M. Lucas, P. Hingston (IEEE, Piscataway, NJ, USA, 2011), pp. 110–117

Physics Simulation Games

4

Jochen Renz and Xiaoyu Ge

Contents

Abstract

Building Artificial Intelligence (AI) that can successfully interact with the physical world in a comprehensive and human-like way is a big challenge. Physics simulation games, i.e., video games where the game world simulates real-world physics, offer a simplified and controlled environment for developing

J. Renz (✉) • X. Ge
Artificial Intelligence Group, Research School of Computer Science, The Australian National University, ANU College of Engineering and Computer Science, Canberra, Australia
e-mail: jochen.renz@anu.edu.au

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_29

and testing Artificial Intelligence. It allows AI researchers to integrate different areas of AI, such as computer vision, machine learning, knowledge representation and reasoning, or automated planning in a realistic setting and to solve various problems that occur in the real world without having to consider all of its complexity at once. This chapter first outlines the main categories of physics simulation games, some of which have become increasingly popular in recent years with the widespread availability of handheld touchscreen devices. It then discusses the motivation and rationale for conducting Artificial Intelligence research on these games and highlights the main research goals. Some of the underlying AI problems and recent advances are discussed and exemplified using a popular physics simulation game. Finally, an overview of current research in related areas is given.

Keywords
Angry Birds game • General game playing (GGP) • Physics mixed reality games • Puzzle games • Research problems • Simulation games • Physics simulation games • Artificial intelligence • Procedure content generation (PCG) • Serious game

Introduction

Physics simulation games have been around since the beginning of video games. Even some of the very first games (e.g., breakout) (Weiss 2009) on commercial game consoles fall under this category. Such games consist of objects, liquids, or other entities that behave according to the laws of physics and they often use an underlying physics simulator that computes the correct physical behavior. These games look and feel very realistic as all actions a player performs have outcomes that are more or less consistent with what one would expect to happen in the real world. What this requires is that all physical properties of all game entities and the game world, such as mass, density, friction, gravity, metric, angles, or locations, are exactly known to the game. Then each action and each movement can be exactly and deterministically computed by the physics simulator.

Implementing the physics of these games is quite a standard task, and the biggest advance over the years has been the more and more realistic and sophisticated graphics. These games form a very popular game category, particularly through the rise of touchscreen devices that allow easy manipulation of the game world and easy execution of actions by the players. Interaction with the game via touching is particularly suitable for physics games as it feels like interacting with real objects.

Physics simulation games and Artificial Intelligence have always had a close and fruitful relationship. The goal of Artificial Intelligence is to develop systems or agents that act, think, and behave like humans, like intelligent beings. This is particularly useful for physics simulation games since other entities in the game world should behave intelligently; they should behave like they are controlled by other human players. This is part of the field of Game AI (McShaffry 2009;

Rabin 2013; Millington and Funge 2009) which tries to achieve a very realistic and smart behavior of other game characters.

What this chapter is about is a much more recent research trend in Artificial Intelligence. Its goal is to build systems or agents that can play physics simulation games as good as or better than human players. This is a very different problem from traditional Game AI and most probably a much harder one. The main difference is that for Game AI, all physical parameters and the complete information of the game world are known to the AI. What is unknown is the behavior of the human player who could be an opponent or a partner, or who could be ignored, depending on the game. In this chapter, the AI knows only as much about the game world as it can see, the physical parameters are unknown, the exact angles and locations of objects are unknown, it is even unknown what the objects are. That is, computer vision should be employed to detect objects and tell the AI where they are and what they are. While this gives us uncertainty about what and where the objects are, another major problem is that the outcome of actions is unknown. Simulating the effects of an action is easy when all physical parameters are known, but if they are not exactly known then a simulation does not produce accurate results and one has to find other ways of predicting the outcome of actions. Humans are very good at predicting physics thanks to a lot of practice and experience in interacting with the real world. For AI, this is still a very difficult problem that needs to be solved in order to build AI that can successfully interact with the real world.

Now why should this be an interesting and an important problem to solve and a research area worth considering? There are certainly enough human players to play these games, why to develop AI to play them too? The motivation for this is quite unexpected and has significant implications for the whole of AI. Much of AI research over the past decades has been divided into specific research areas devoted to specific sub-problems of AI, such as Machine Learning, Computer Vision, Knowledge Representation, etc. Most research in these AI areas is so focused on their particular sub-problems that the big vision of AI has been ignored. It is not possible to simply plug and play modules from these areas and to obtain a complete AI system. Also, much of the research is focused on solving toy problems that may have nothing to do with the real world or with real problems. This was necessary as the real world is simply too complex. There are too many distractions, such as reflections, other people, unexpected activities, etc., that make it very difficult to focus on the particular problem one tries to solve.

This is where physics simulation games become very interesting: They allow AI researchers to develop methods and solve problems in a simplified and controlled environment where all these distractions can be removed, but that is still realistic enough and similar enough to the real world and to problems that need to be solved in the real world. In these controlled environments, it is then possible to integrate methods from the different AI areas to solve realistic problems. It is possible to see which existing methods work and which do not and what still needs to be developed in order to solve problems in the real world. Due to the difficulty of these problems, it is necessary to start simple, with games that only have a few features of the real world, and once these can be solved, to move to more realistic game worlds. Currently, the AI community aims to develop agents that can successfully play the game Angry Birds (Ge et al. 2014a; Renz et al. 2013), which is clearly at the lower end of this journey, but already requires us to solve some very realistic and very hard problems.

A further benefit of using games for developing AI is that it allows us to easily evaluate how well AI can already solve problems compared to human players. It also allows us to easily set up competitions, which is a very good way of achieving fast scientific progress in the area of the competition.

The remainder of this chapter is structured as follows. The next section introduces different kinds of physics simulation games. Section "Research Problems in PSG" then summarizes some of the important problems that need to be solved in order to build AI that can successfully play these games and also mentions some successful approaches to these problems. Section "Angry Birds: An Example" then explains some of the problems using the Angry Birds game, where a current AI competition has created some attraction. Section "Other Related Research" briefly summarizes some related research areas.

Physics Simulation Games

A physics simulation game (PSG) is a video game where the game world simulates real-world physics (Newtonian physics). In the following, different types of such games are discussed.

Puzzle Games

This genre has gained in popularity in recent years, especially on mobile devices. In the game, the player needs to solve a physical puzzle by obeying various laws of physics. These games have the following common features:

- Simple Physics. The game physics is simple in comparison to large commercial games. The laws of physics in these games are based on some basic principles of classical mechanics. Physical properties of the game objects are also simplified so that the entire game environment can be completely parameterized using a relatively small number of parameters. Besides, the game world in most of these games is two-dimensional, which further simplifies the underlying physics simulation in which the z-axis is ignored.
- Simple Actions. The player manipulates game objects via very simple actions such as finger gestures (tap, drag, swipe, etc.). Any chosen action can be simulated using an underlying physics simulator which makes the execution and the consequences of actions look very real. Despite a huge, and often infinite, action space (the number of ways to execute the action), it is common to have a small number of different actions in these games. For example, there are two actions in Angry Birds, namely, to drag and release a bird as well as to

tap to trigger a special ability of the bird. In Cut the Rope, the only action is to swipe to cut a rope. The actions are simple also in the sense that the outcome of these actions is deterministic.

- Simple (Commonsense) Knowledge to solve the puzzle. To be able to solve the puzzle, the player needs to predict the outcome of actions and plan a sequence of actions to achieve the desired result. Accurately predicting the outcome of each action is not hard for human players. In fact, most of these puzzles could be solved using knowledge of Naive Physics (Hayes et al. 1978).

These games can be further classified according to the types of puzzles.

- Build a Stable Structure. These games often require the player to remove or add objects to form a stable structure. Typical games: Super Stacker (http://www.thegamehomepage.com/play/super-stacker-2/), Moonlights (http://www.bonuslevel.org/moonlights/), World of Goo (http://worldofgoo.com/).
- Dismantle a Stable Structure. The player needs to remove or destroy objects to dismantle a structure while achieving specific goals. For example, the goal of Angry Birds is to kill all the pigs which are usually protected by a sheltering structure. In Jungle Bloxx, the goal is to collect all the diamonds within a structure. Typical Games: Angry Birds (http://www.angrybirds.com/), Jungle Bloxx (http://www.gamespot.com/jungle-bloxx/).
- Build a Rube Goldberg Machine. The player needs to create or use existing game objects to build a Rube Goldberg device that performs a certain task by triggering a chain of interactions. Typical games: The Incredible Machine (http://www.mobygames.com/game-group/incredible-machine-series), Bad Piggies (http://www.badpiggies.com/), Amazing Alex (http://www.amazingalex.com/), Crayon Physics Deluxe (http://www.crayonphysics.com/).
- Modifying a Physical System. In the game, the player needs to make some changes to the existing physical system to achieve a particular goal. Typical games: Cut the Rope (http://www.cuttherope.net/), Feed me Oil (http://holywatergames.com/), God of light (http://www.playmous.com/).

Simulation Games

This category contains all the video games wherein a physics engine is used to simulate the game world. Compared with the puzzle games, these games typically use extensive physics engines (e.g., Bullet (http://www.bulletphysics.org/)) to create more realistic game play. Unlike a puzzle game of which the actions and goals are strictly defined, a simulation game usually allows the player to freely control a game character to perform various “real-world” tasks in a simulated environment. The tasks include playing sports (e.g., playing tennis, snooker, bowling), racing, and combating. The simulated physical system sometimes can be extremely realistic and extremely complicated (e.g., Microsoft’s Flight Simulator (http://www.microsoft.com/games/fsinsider/)).

Physics Mixed Reality Games

A mixed reality game is a game played in both reality and simulated environment simultaneously. It is typically played on a mobile device equipped with a camera that captures the real-world images. The game uses techniques from computer vision and augmented reality to allow real and visual objects interact physically on the device's screen. For example, in the mixed reality driving game (Oda et al. 2008), the player needs to drive a simulated car around on the player's physical desk using real and virtual obstacles. One challenge in implementing a mixed reality game lies in modeling the physical interactions between virtual and physical objects. Recent years also have seen a rapid advancement in haptic technology (e.g., Kinect (Zhang 2012)) that enables players to physically interact (moving their bodies) with the game world.

Research Problems in PSG

PSG is a new benchmark for Artificial Intelligence. The following features make PSGs an excellent test bed for research in physical reasoning.

- The research problems identified in PSGs domain are the same problems that need to be solved by AI systems that can successfully interact with the physical world. Just like the real-world physics, the physics simulator of a PSG works as a black box that hides all the equations and parameters from the player. To be able to play these games well, the player does not need to understand how the black box functions and neither does the player need to perform numerical calculations. In fact, what the player does in the game is essentially the same as what the player does in performing daily tasks – solve problems by intuition and qualitative reasoning (aka naive physics).
- Video games have the advantage that all images are generated and rendered using computer graphics where objects typically do not have the complexity and diversity that can be found in real world images. Therefore, working in the game domain allows researchers to focus on the problems independent of the computer vision problems.
- Intelligent agents can be tested by comparing their performance against the human performance in these games.
- Human computation in computer games has received considerable attention for the past decade. Many research efforts (Kuo et al. 2009; Von Ahn et al. 2006; Lieberman et al. 2007; Speer et al. 2009) have been devoted to extracting commonsense knowledge from computer games. In PSGs games, one can easily record player-generated behavioral data and use it as an excellent source for learning and reasoning algorithms. PSGs have another big advantage over the existing research-directed or serious games (Zyda 2005). Most of the existing research-directed games were developed with a particular research task in mind, which usually sacrifices fun in the game; therefore the game can hardly attract

the public. In fact, most players are from the research area, and the number of players is on average far less than a normal popular video game. In contrast, PSGs are games that are truly popular among the general public of various educational background, which may generate meaningful and more general behavioral data.

The ultimate goal is to develop an intelligent agent that can play the physics simulation games as well as or better than human players. Unlike the traditional in-game AI that has complete knowledge about the game world, the agent is only allowed to access the same information the human player can obtain from the game. The following is a selection of research problems that need to be solved in order to achieve the goal of AI playing PSG better than human players and ultimately for AI to be successful in interacting with the physical world.

Visual Detection of Objects

Object detection (Papageorgiou et al. 1998) and recognition (Grauman and Leibe 2011) are two major research problems in the field of computer vision (see Forsyth and Ponce (2002) for a survey) which has been extensively studied in the last two decades. A lot of efforts have been devoted to solving these two problems on real-world images. Given an image, object detection concerns the question: where a particular object is in the image? The detection focuses on a certain class of objects, such as human faces (Zhang and Zhang 2010) and pedestrians (Dollar et al. 2012), and the algorithm is usually trained on the annotated images containing objects of the same class.

Instead of detecting objects from a previously known class, object recognition targets on identifying and classifying both known and unknown objects in an image, which finds applications in various areas such as robotics and scene understanding. The recognition techniques can be appearance (Belongie et al. 2002) and/or features (Lowe 1999), supervised (Branson et al. 2010) and/or unsupervised (Niebles et al. 2008). One challenging problem in object recognition is to reliably detect and classify unknown objects while assuming minimal prior information, which is also known as category learning (Lee and Grauman 2012).

Most of the vision research has been done within the real-world context. In contrast, vision problems in the game domain have been largely ignored. The nature of the illumination, texture, and perceptual noise in a real scene is completely different from the artificial setting of the game world which is rendered by a graphics engine. It is not surprising that vision researchers are rarely interested in solving problems in such artificial domain because of its "simplicity." What is surprising is that one can hardly find a suitable solution within the state-of-the-art that can reliably detect and categorize unknown objects in different games without human intervention (annotation and feature selection). One reason is a lack of generality of those techniques which are specialized for handling real-world images rather than game images. Another reason is that most of the state-of-the-art computer vision methods are purely statistical methods that operate at pixel level. What they miss are mechanisms to model rich spatial information and reasoning mechanisms to figure out what and where an object should be. It turns out that video games is a suitable test bed for developing and testing computer vision algorithms that focus on a high level understanding of the scene (which is more close to the way humans "see" the world) by modeling and reasoning about spatial information (Christensen and Nagel 2006; Cohn et al. 2003).

Physical Reasoning

Physical reasoning is an essential capability for AI agents seeking to interact with physical systems. In PSG games, there is a diverse set of tasks that require physical reasoning and planning for solving them while physical reasoning in the PSG domain remains an open problem. There are two main research paradigms in physical reasoning, namely, qualitative physics and simulation-based reasoning.

Qualitative physics has been extensively studied in the area of knowledge representation and reasoning (see Davis (2008a) for a survey). Much of the research in qualitative physics focuses on the development of qualitative representations and reasoning mechanisms that enable AI to make commonsense inferences about physical systems. For example, Forbus (1984) developed the qualitative process theory that allows one to specify dynamic processes in formal languages. This theory has been successfully applied to various applications, such as Cogsketch (a sketch understanding system) (Forbus et al. 2008), PEARS (Eckstein et al. 2015) (a physics engine that simulates complex physical phenomena), and analogical reasoning (Friedman and Forbus 2009). Davis (2008b, 2011) proposed a framework based on first order logic for reasoning about the physical properties of solid objects and liquids. Davis et al. (2013) invented a mechanism for qualitative reasoning about containers. The major advantage of the qualitative methods is that they are able to draw some useful conclusions or inferences rapidly from incomplete and noisy data. However, most methods are based on domain specific theories and cannot be easily generalized.

On the other hand, simulation-based reasoning methods can offer precise predictions when the environment is fully observable. This family of methods make physical inferences based on probabilistic simulations of Newtonian mechanics. Simulation-based reasoning methods have been mainly used for solving prediction problems. For example Battaglia et al. (2013) applied probabilistic simulations based on "intuitive physics engine" to verify the stability of a tower of blocks. Johnston and Williams (2007) applied quantitative simulation to solve the egg-cracking problem (a benchmark problem in commonsense reasoning). Similarly, in robotics, Nyga and Beetz (2012) and Kunze et al. (2011) used simulation physical reasoning to perform daily tasks such as housework. Davis and Marcus (2013) offers a detailed discussion on the role of simulation-based reasoning in solving physical reasoning problems.

The physical reasoning problems in the PSG domain are challenging: When playing a PSG game, the agent can only receive visual inputs from the game environment. While the physical properties of the game environment and the game objects are initially unknown, the tasks in these games are complicated and usually require multiple actions to complete, and it is difficult to formalize the knowledge about the PSG domain. Therefore, one possible direction to tackle the physical reasoning problems is to integrate the qualitative methods and simulation-based methods, which is known as hybrid reasoning. For instance, Johnston and Williams (2008) combined simulation with tableaux reasoning to solve some commonsense reasoning problems. Abdo et al. (2015) developed a hybrid reasoning approach that guides robots to clean up shelves according to users' preferences.

Predict the Outcome of Actions

The ability of reasoning about action and change (RAC) (Prendinger and Schurz 1996) is essential for an intelligent agent to adapt to a dynamic environment. RAC has been addressed as a knowledge representation and reasoning problem, which is a central topic in Artificial Intelligence. There have been various formalisms proposed for the representation and reasoning (Brachman and Levesque 2004) about a dynamic environment, of which the important ones are situation calculus (McCarthy 1963; Levy and Quantz 1998), fluent calculus (Thielscher 1998), event calculus (Kowalski and Sergot 1989), and action languages (Giunchiglia and Lifschitz 1998). Another related research stream is qualitative spatial and temporal reasoning (QSTR) (Cohn and Renz 2008) that aims to mimic the human commonsense knowledge about space and time. There have been quite a few calculi and reasoning mechanisms developed and applied in various areas ranging from GIS to computer-aided design (see Wolter and Wallgrün (2012) for a recent survey on QSTR applications). One way to predict the outcome of an action is to know which object before the action corresponds to which object after the action. To be able to establish a correct correspondence, an agent can track the objects on its continuous observations, which can be treated as a tracking problem (see Yilmaz et al. (2006) for a survey). When observations are not continuous, a proper reasoning mechanism becomes necessary (Ge and Renz 2014).

Automated Planning

Equally important is automated planning (see Ghallab et al. (2004) for a survey) which is about the derivation of sequence of actions that lead to an optimal result. Planning has been widely applied to video games since the last decade for various tasks such as controlling intelligent nonplayer characters (NPCs) and generating stories. The commonly used techniques include STRIPS planning (Fikes and Nilsson 1972), hierarchical planning (Kelly et al. 2008; Li and Riedl 2010), and behavior trees (Lim et al. 2010). There are two planning paradigms, namely, online planning and offline planning. Online planning assumes minimal prior knowledge of the game environment and computes an optimal plan in real time. In contrast, offline planning has complete knowledge of the game environment and generates plans offline ahead of time. Therefore, online planning algorithms surpasses their offline counterparts in handling unforeseen situations and works better in dynamic environments while offline planning algorithms use much less computational resources during the game play.

Planning in PSG remains a big challenge due to the fact that the action and state space in these games are huge, possibly infinite while the outcome of actions is unknown, which renders a brute-force search or simulation implausible. Over the past 10 years, a wide range of techniques have been developed for planning in huge, uncertain, and partially observable environments (see Vaccaro (2010) for a survey). Learning strategies from human experts is also a well-established topic (Khardon 1996). PSGs are great learning platforms where human players' strategies can be digitized and made available for learning algorithms.

Learning Properties of the Game World

Learning properties of game objects and the game environment is another fundamental problem in implementing AI in PSG. The properties of game objects consist of two types: (1) physical properties such as density, friction, elasticity, and strength; (2) object affordances (Gibson et al. 1990) and functional features. While the visual detection of objects answers what the object is, object properties learning answers how this object can be used. There has been a significant amount of work within the robotics community in learning object affordances. Some approaches identify object affordances by observing visual clues (Sun et al. 2010), a combination of visual and physical attributes (Hermans et al. 2011), or geometrical properties (Aldoma et al. 2012) of objects, which are usually applied to static images. Some methods learn object functions through robots' exploratory actions (Moldovan et al. 2012; Montesano et al. 2008) in real-world scenarios or simulated environments. There are also some learning techniques that focus on learning object functionalities and learning to perform actions from human demonstrations Koppula et al. (2013); Saxena et al. (2008); Li et al. (2010).

Angry Birds: An Example

This section explains the aforementioned problems using the Angry Birds game and reviews the techniques that have been proposed to tackle these problems. The AI Birds competition was founded in 2012 as an initiative to encompass a variety of AI areas to achieve its long-term goal – "build an intelligent Angry Birds playing agent that can play new levels better than the human players" (Renz 2015). Angry Birds is a typical physics simulation game wherein the game world is completely parameterized, which is simulated by the Box2D (http://www.box2d.org) physics engine.

Visual Detection of Objects

To solve new levels, the agent has to be able to detect and classify both known and unknown objects as well as localize foreground objects amid an intricate background. There are no existing off-the-shelf computer vision solutions for solving this problem; Ge et al. (2014b) took this challenge and developed a novel method based on qualitative stability analysis. The method infers the existence of yet undetected objects by observing that other objects that have already been detected are physically unsupported and therefore must be supported by some object still to be detected. The method was tested on 444 available Angry Birds levels (http://chrome.angrybirds.com); Initially, only the green pig is known to the algorithm. After a few iterations for all the levels, all the game objects are detected.

Physical Reasoning

In the past three competitions, almost all the angry birds playing agents have been endowed with a certain degree of capability for physical reasoning. For instance, Brovicka et al. (2014) performs spatial reasoning to find a connected block structure (distinguished between Pyramid, Rectangle, and Skyscraper) near pigs and selects the most suitable block (often the weak point of the structure) by considering the supporters, reachability, and the shape of the block. Calimeri et al. (2013) uses a declarative, logic-programming based module to model the domain knowledge and compute optimal shots based on spatial configurations of the current game scene. Narayan-Chen et al. (2013; Tziortziotis et al. (2014) are machine learning agents that preserve essential structural and spatial information in the feature space, and "learn" to solve the puzzle by analyzing the structures. Walega et al. (2014) proposed a qualitative physics approach to evaluate a shot regarding its reachability and the scale of the damage to the target. Zhang and Renz (2014) developed a spatial calculus to represent the game objects and used it to identify the weak points to hit. Similarly, Ge and Renz (2013) devised a qualitative spatial representation for general solid rectangles (GSR), i.e., rectangles that can have an arbitrary angle, and cannot be penetrated, which can be used to analyze the stability of structures.

Polceanu and Buche (2014) tackle the problem by advanced simulation. The agent first detects all the objects in the game by the provided software and then uses these objects to construct an "Imaginary World" in which mental simulations can be performed. The objects' motion in the world is governed by Newtonian physics.

Predict the Outcome of Actions

The core action in Angry Birds is firing a bird. Some agents use simulation to estimate the consequence of a shot. For instance, Polceanu and Buche (2014) proposed an agent that performs multiple parallel simulations to test different courses of actions and chooses the best action. A good simulation relies on an accurate knowledge of the underlying physical system. However, the parameters of the physical system are invisible to the agent, which adversely affects the simulation result.

Another approach is to identify how objects are affected by a shot, specifically, is to determine which objects before a shot correspond to which objects after a shot. To be able to identify the correct correspondences, the agent can track objects through the before-and-after observations. The problem becomes challenging when the observations are not continuous (the time gap between the observations is greater than 50 ms), and when those objects are perceptually indistinguishable (i.e., have the same appearance). Ge and Renz (2014) developed a spatial reasoning based tracking method that can accurately track perceptually indistinguishable objects in discrete observations with large time gaps.

Other Related Research

Physics Simulation in Serious Game

A serious game, as defined in Zyda (2005), is "a mental contest, played with a computer in accordance with specific rules, that uses entertainment to further government or corporate training, education, health, public policy, and strategic communication objectives." Serious games have been used for physics education since the early 1980s (White 1984; Lee et al. 1993). Compared to commercial physics puzzle games, these serious games feature a more accurate simulation of Newtonian physics. The game Newton's Playground (Ventura et al. 2013) has been used to help secondary school students understand qualitative physics (naive physics). In the game, the instructor can create a simple physical machine to illustrate certain physical concepts to students in a qualitative way. Similar games are Coller and Scott (2009; Squire et al. (2004). There are also 3D platform games that emphasize realistic simulations for training purpose (Davis 2004). Some researchers (Erignac 2001; Cavazza et al. 2004) proposed the use of qualitative simulation to simulate the way objects behave according to naive physics. Limited research has been done in those games to qualitatively simulate (Zhou and Ting 2006; Lugrin and Cavazza 2007) the physical behavior of game objects.

General Game Playing

General game playing (GGP) (Genesereth et al. 2005) aims at developing intelligent agents that can play a class of previously unknown games effectively. Given an arbitrary game, the agent can access a formal description (written in logic) file of the game and it needs to figure out the legal actions, winning strategy, and winning goals by itself. Considerable research (Finnsson and Björnsson 2008; Banerjee and Stone 2007; Cerexhe et al. 2014) has been done in this area, and a number of successful playing agents (Kuhlmann and Stone 2006; Geier et al. 2014; Kirci et al. 2011) have been developed and evaluated in the GGP competition at the annual AAAI conference (Genesereth and Bjrnsson 2013). The competition tests the performance of GGP agents in abstract strategy games such as chess-like games and card games.

Atari-GGP (Bellemare et al. 2013) shifts the focus from abstract strategy games to video games. An Atari game is a video game that has simple graphics and game settings. There is a finite set of discrete actions (e.g., move the game character in different directions) available to the player. Recently, Mnih et al. (2013) developed an agent that can play a range of Atari games with minimal domain specific knowledge. The technique is based on the integration of reinforcement learning and deep learning. During playing, the agent only receives the screenshot from the current game screen and the agent is able to figure out the game dynamics over time. There are few Atari games that involve some physics while they do not require sophisticated physical reasoning. For example, the task in Breakout is to destroy all bricks by bouncing a ball on them. In this scenario, bouncing is one of the few physical phenomenon an agent should "understand" in order to complete the task. However, PSG games usually have more physical elements and the tasks are much more complicated. There is no existing technique for developing general game playing agents in the PSG domain.

Procedure Content Generation

Procedure content generation (PCG) (see Hendrikx et al. (2013) for a survey) refers to the development of an automated or semiautomated procedure for game content generation. The game content refers to various aspects of a game ranging from game levels to game stories. Nowadays, it is not uncommon that a physics puzzle game such as Angry Birds has hundreds of different levels. Therefore, PCG becomes one of the major efforts in the game industry in part due to the need for reducing time consumption as well as budget for content generation and in part due to the need for increasing game content variations. In the meantime, PCG has received increasing attention from the AI community, and a variety of AI techniques have been developed for solving PCG problems such as modeling behavior of game objects (Hastings et al. 2009; Hidalgo et al. 2008), generating levels (Shaker et al. 2010; Dormans 2010), and creating puzzles (Iosup 2011; Ashlock 2010). Generating levels for physics puzzle games have attracted a growing interest in recent years. This genre provides an interesting test bed for level generation techniques as it imposes certain physics constraints that are essential for evaluating the quality of the generated levels. The state-of-the-art is dominated by evolutionary algorithms (Cardamone et al. 2011; Cook and Colton 2011; Mourato et al. 2011). For example, Shaker et al. (2013) combined an evolution algorithm with a grammatical representation to generate levels in Cut the Rope; Ferreira and Toledo (2014) viewed level generation in Angry Birds as an optimization problem and developed a search approach (known as Search Based Procedural Content

Generation, for a survey see Togelius et al. (2011)) built on an evolution algorithm. All these techniques are entirely application specific and therefore not reusable. Another issue to overcome when developing level generators in these games is playability evaluation (Shaker et al. 2013) which has to be done through physics simulation. As there is no source code available for most commercial games, game researchers have to develop their own simulators that "clone" the physical behaviors of the simulator of the original game.

Recommended Reading

N. Abdo, C. Stachniss, L. Spinello, W. Burgard, Robot, organize my shelves! Tidying up objects by predicting user preferences, in *Proceedings of International Conference on Robotics and Automation (ICRA) (2015)*, Seattle, 2015

A. Aldoma, F. Tombari, M. Vincze, Supervised learning of hidden and non-hidden 0-order affordances and detection in real scenes. in *Robotics and Automation (ICRA), 2012 IEEE International Conference on, IEEE*, (St. Paul, Minnesota, USA, 2012), pp. 1732–1739

D. Ashlock, Automatic generation of game elements via evolution. in *Computational Intelligence and Games (CIG), IEEE Symposium on,* IEEE, (Copenhagen, Denmark, 2010), pp. 289–296

B. Banerjee, P. Stone, General game learning using knowledge transfer, in *IJCAI 2007, Proceedings of the 20th International Joint Conference on Artificial Intelligence*, Hyderabad, 6–12 January 2007, pp. 672–677 (2007)

P.W. Battaglia, J.B. Hamrick, J.B. Tenenbaum, Simulation as an engine of physical scene understanding. Proc. Natl. Acad. Sci. **110**(45), 18,327–18,332 (2013)

M.G. Bellemare, Y. Naddaf, J. Veness, M. Bowling, The arcade learning environment: an evaluation platform for general agents. J. Artif. Intell. Res. **47**(1), 253–279 (2013)

S. Belongie, J. Malik, J. Puzicha, Shape matching and object recognition using shape contexts. IEEE Trans. Pattern Anal. Mach. Intell. **24**(4), 509–522 (2002)

R. Brachman, H. Levesque, *Knowledge Representation and Reasoning* (Elsevier, Amsterdam, 2004)

S. Branson, C. Wah, F. Schroff, B. Babenko, P. Welinder, P. Perona, S. Belongie, Visual recognition with humans in the loop. in *Computer Vision–ECCV 2010*, Springer, (Heraklion, Crete, Greece, 2010), pp. 438–451

T. Brovicka, R. Spetlik, K. Rymes, Datalab Birds Angry Birds AI (2014), http://aibirds.org/2014-papers/datalab-birds.pdf

F. Calimeri, M. Fink, S. Germano, G. Ianni, C. Redl, A. Wimmer, Angryhex: an artificial player for angry birds based on declarative knowledge bases. in *Proceedings of the Workshop Popularize Artificial Intelligence co-located with the 13th Conference of the Italian Association for Artificial Intelligence (AI*IA 2013)*, Turin, 5 Dec 2013, pp. 29–35

L. Cardamone, D. Loiacono, P.L. Lanzi, Interactive evolution for the procedural generation of tracks in a high-end racing game. in *Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation*, ACM, (Dublin, Ireland, 2011), pp. 395–402

M. Cavazza, S. Hartley, J.L. Lugrin, M. Le Bras, Qualitative physics in virtual environments. in *Proceedings of the 9th International Conference on Intelligent User Interfaces*, ACM, (Funchal, Madeira, Portugal, 2004), pp. 54–61

T. Cerexhe, D. Rajaratnam, A. Saffidine, M. Thielscher, A systematic solution to the (de-) composition problem in general game playing, in *Proceedings of the European Conference on Artificial Intelligence (ECAI)* (IOS Press, Prague, 2014)

H.I. Christensen, H.H. Nagel, *Cognitive Vision Systems* (Springer, Berlin/Heidelberg, 2006)

A.G. Cohn, J. Renz, Qualitative spatial representation and reasoning, in *Handbook of Knowledge Representation*, vol. 3 (Elsevier, Amsterdam, 2008), pp. 551–596

A.G. Cohn, D.R. Magee, A. Galata, D.C. Hogg, S.M. Hazarika, Towards an architecture for cognitive vision using qualitative spatio-temporal representations and abduction. in *Spatial Cognition III*, Springer, (Berlin Heidelberg, 2003), pp. 232–248

B.D. Coller, M.J. Scott, Effectiveness of using a video game to teach a course in mechanical engineering. Comput. Educ. **53**(3), 900–912 (2009)

M. Cook, S. Colton, Multi-faceted evolution of simple arcade games. in *2011 I.E. Conference on Computational Intelligence and Games, CIG 2011*, Seoul, 31 Aug – 3 Sep, 2011, pp. 289–296

K.J.W. Craik, *The Nature of Explanation*. CUP Archive (Cambridge University Press, Cambridge, 1967)

M. Davis, *America's Army pc Game Vision and Realization* (US Army and the Moves Institute, San Francisco, 2004)

E. Davis, Physical reasoning. Found. Artif. Intell. **3**, 597–620 (2008a)

E. Davis, Pouring liquids: a study in commonsense physical reasoning. Artif. Intell. **172**(12), 1540–1578 (2008b)

E. Davis, How does a box work? A study in the qualitative dynamics of solid objects. Artif. Intell. **175**(1), 299–345 (2011)

E. Davis, G. Marcus, The scope and limits of simulation in automated reasoning. Artif. Intell. (to appear) (2013)

E. Davis, G. Marcus, A. Chen, Reasoning from radically incomplete information: the case of containers, in *Proceedings of the Second Annual Conference on Advances in Cognitive Systems ACS*, vol. 273 (2013), p. 288

J. De Kleer, J.S. Brown, A qualitative physics based on confluences. Artif. Intell. **24**(1), 7–83 (1984)

P. Dollar, C. Wojek, B. Schiele, P. Perona, Pedestrian detection: an evaluation of the state of the art. IEEE Trans. Pattern Anal. Mach. Intell. **34**(4), 743–761 (2012)

J. Dormans, Adventures in level design: generating missions and spaces for action adventure games. in *Proceedings of the 2010 Workshop on Procedural Content Generation in Games*, ACM, (Monterey, California, 2010), p. 1

E.Giunchiglia, V. Lifschitz, An action language based on causal explanation: preliminary report. in *Proceedings of the Fifteenth National Conference on Artificial Intelligence and Tenth Innovative Applications of Artificial Intelligence Conference, AAAI 98*, IAAI 98, Madison, 26–30 July 1998, pp. 623–630

B. Eckstein, J.L. Lugrin, D. Wiebusch, M. Latoschik, Pears – physics extension and representation through semantics. IEEE Trans. Comput. Intell. AI Games **PP**(99), 1–1 (2015)

C.A. Erignac, Interactive semi-qualitative simulation for virtual environments. PhD thesis, University of Pennsylvania, 2001

B. Falkenhainer, K.D. Forbus, D. Gentner, *The Structure-Mapping Engine*. Department of Computer Science, University of Illinois at Urbana-Champaign, Philadelphia, Pennsylvani, USA, 1986

L. Ferreira, C. Toledo, A search-based approach for generating angry birds levels. in *2014 IEEE Conference on Computational Intelligence and Games (CIG)*, IEEE. Dortmund, Germany, 2014, pp. 1–8

R.E. Fikes, N.J. Nilsson, Strips: a new approach to the application of theorem proving to problem solving. Artif. Intell. **2**(3), 189–208 (1972)

H. Finnsson, Y. Björnsson, Simulation-based approach to general game playing, in *Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence, AAAI 2008*, Chicago, 13–17 July 2008 (AAAI Press, 2008), pp. 259–264

K.D. Forbus, Qualitative process theory. Artif. Intell. **24**(1), 85–168 (1984)

K.D. Forbus, P. Nielsen, B. Faltings, Qualitative spatial reasoning: the clock project. Artif. Intell. Annecy, France, **51**(1), 417–471 (1991)

K.D. Forbus, J.M. Usher, A. Lovett, K. Lockwood, J. Wetzel, CogSketch: open-domain sketch understanding for cognitive science research and for education, in *SBM* (2008), Annecy, France, pp. 159–166

D.A. Forsyth, J. Ponce, *Computer Vision: A Modern Approach*. Prentice Hall Professional Technical Reference, (Englewood Cliffs, Prentice-Hall, 2002)

S. Friedman, K.D. Forbus, Learning naive physics models and misconceptions, in *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (2009)

S. Friedman, K.D. Forbus, Learning naive physics models and misconceptions. in *Proceedings 31st Annual Conference of the Cognitive Science Society*, 2009

X. Ge, J. Renz, Representation and reasoning about general solid rectangles. in *IJCAI 2013, Proceedings of the 23rd International Joint Conference on Artificial Intelligence*, Beijing, 3–9 Aug 2013

X. Ge, J. Renz, Tracking perceptually indistinguishable objects using spatial reasoning. in *PRICAI 2014: Trends in Artificial Intelligence – Proceedings of the 13th Pacific Rim International Conference on Artificial Intelligence*, Gold Coast, 1–5 Dec 2014, pp. 600–613

X. Ge, S. Gould, J. Renz, S. Abeyasinghe, J. Keys, A. Wang, P. Zhang, Angry Birds Game Playing Software Version 1.3: Basic Game Playing Software. URL (2014a), http://www.aibirds.org

X. Ge, J. Renz, P. Zhang, Visual detection of unknown objects in video games using qualitative stability analysis," in *Computational Intelligence and AI in Games, IEEE Transactions on*, 99, 1 (2014b)

F. Geier, T. Keller, R. Mattmller, Past, present, and future: an optimal online algorithm for single-player gdl-ii games, in, *Proceedings of the 21st European Conference on Artificial Intelligence (ECAI)*, Prague, Czech Republic, (IOS Press, 2014), pp. 357–362

M. Genesereth, Y. Bjrnsson, The international general game playing competition. AI Mag. **34**(2), 107–111 (2013)

M. Genesereth, N. Love, B. Pell, General game playing: overview of the AAAI competition. AI Mag. **26**(2), 62 (2005)

M. Ghallab, D. Nau, P. Traverso, *Automated Planning: Theory & Practice*. San Francisco, California, USA, (Elsevier, 2004)

J.J. Gibson, W. Bricken, S. Uselton, S.S. Fisher, A. Kay, T. Oren, G. Salomon, K. Kreitmann, A. Don, B. Laurel, et al., The ecological approach to visual perception. SIGGRAPH'90 Show Daily. **24**, 165–170 (1990)

K. Grauman, B. Leibe, *Visual Object Recognition*, vol. 11 (Morgan & Clay-pool Publishers, San Rafael, 2011)

E.J. Hastings, R.K. Guha, K.O. Stanley, Automatic content generation in the galactic arms race video game. IEEE Trans. Comput. Intell. AI Games **1**(4), 245–263 (2009)

P.J. Hayes et al., *The Naive Physics Manifesto* (Université de Genève, Institut pour les études sémantiques et cognitives, Genève, 1978)

M. Hendrikx, S. Meijer, J. Van Der Velden, A. Iosup, Procedural content generation for games: a survey. ACM Trans. Multimed. Comput. Commun. Appl. (TOMCCAP) **9**(1), 1 (2013)

T. Hermans, J.M. Rehg, A. Bobick, Affordance prediction via learned object attributes. in *International Conference on Robotics and Automation: Workshop on Semantic Perception, Mapping, and Exploration,* (IEEE. Shanghai, China, 2011)

J.L. Hidalgo, E. Camahort F., Abad, M.J. Vicent, Procedural graphics model and behavior generation. in *Computational Science–ICCS 2008*, Springer, (Kraków, Poland, 2008), pp. 106–115

A. Iosup, Poggi: generating puzzle instances for online games on grid infrastructures. Concurr. Comput. Pract. Exp. **23**(2), 158–171 (2011)

B. Johnston, M.A. Williams, A generic framework for approximate simulation in commonsense reasoning systems, in *Logical Formalizations of Commonsense Reasoning, Papers from the 2007 AAAI Spring Symposium*, Technical Report SS-07-05, Stanford, 26–28 March 2007. (AAAI, 2007), pp. 71–76

B. Johnston, M.A. Williams, Comirit: commonsense reasoning by integrating simulation and logic, in *Proceedings of the 2008 conference on Artificial General Intelligence 2008: Proceedings of the First AGI Conference* (IOS Press, 2008), pp. 200–211

J.P. Kelly, A. Botea, S. Koenig, Offline planning with hierarchical task networks in video games. in *AIIDE,* (AAAI Press. Stanford, California, USA, 2008)
R. Khardon, Learning to take actions. in *Machine Learning*, (AAAI Press, Portland, Oregon, 1996), pp. 787–792
M. Kirci, N.R. Sturtevant, J. Schaeffer, A GGP feature learning algorithm. KI **25**(1), 35–42 (2011)
H.S. Koppula, R. Gupta, A. Saxena, Learning human activities and object affordances from rgb-d videos. Int. J. Rob. Res. **32**(8), 951–970 (2013)
R. Kowalski, M. Sergot, A logic-based calculus of events. in *Foundations of Knowledge Base Management*, (Springer, Tokyo, Japan, 1989), pp. 23–55
G. Kuhlmann, P. Stone, Automatic heuristic construction in a complete general game player, in *Proceedings, The Twenty-First National Conference on Artificial Intelligence and the Eighteenth Innovative Applications of Artificial Intelligence Conference*, Boston, 16–20 July 2006 (AAAI Press, Boston, Massachusetts, USA 2006), pp. 1457–1462
B. Kuipers, Qualitative reasoning: modeling and simulation with incomplete knowledge. Automatica **25**(4), 571–585 (1989)
L. Kunze, M.E. Dolha, M. Beetz, Logic programming with simulation-based temporal projection for everyday robot object manipulation, in *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on, IEEE*, San Francisco, California, USA (2011), pp. 3172–3178
L.Von Ahn, M. Kedia, M. Blum, Verbosity: a game for collecting common-sense facts. in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, (Montréal, Québec, Canada, 2006), pp. 75–78
Y.J. Lee, K. Grauman, Object-graphs for context-aware visual category discovery. IEEE Trans. Pattern Anal. Mach. Intell. **34**(2), 346–358 (2012)
J. Lee, M. Trigueros, J. Tagüeña, R.A. Barrio, Spectrum: an educational computer game. Phys. Educ. **28**(4), 215 (1993)
F. Levy, J. Quantz, Representing beliefs in a situated event calculus. in *ECAI*, (John Wiley. Brighton, UK, 1998), pp. 547–551
B. Li, M.O. Riedl, An offline planning approach to game plotline adaptation. in *AIIDE*, (Curran Associates. Vancouver, British Columbia, Canada, 2010)
C. Li, A. Kowdle, A. Saxena, T. Chen, Towards holistic scene understanding: feedback enabled cascaded classification models. in *Advances in Neural Information Processing Systems*, 2010, pp. 1351–1359
H. Lieberman, D. Smith, A. Teeters, Common consensus: a web-based game for collecting commonsense goals. in: *ACM Workshop on Common Sense for Intelligent Interfaces*, (ACM. New York, USA, 2007)
C.U. Lim, R. Baumgarten, S. Colton, Evolving behaviour trees for the commercial game defcon, in *Applications of Evolutionary Computation* (Springer, Berlin, 2010), pp. 100–110
D.G. Lowe, Object recognition from local scale-invariant features. in *The Proceedings of the Seventh IEEE International Conference on Computer Vision*, vol 2 IEEE, (Kerkyra, Corfu, Greece, 1999), pp. 1150–1157
J.L. Lugrin, M. Cavazza, Making sense of virtual environments: action representation, grounding and common sense. in *Proceedings of the 12th International Conference on Intelligent User Interfaces*, ACM, (Honolulu, Hawaii, USA, 2007), pp. 225–234
J. McCarthy, Situations, actions, and causal laws. Technical report, DTIC Document, (1963)
M.L. McShaffry, *Behavioral Mathematics for Game AI* (Cengage Learning, Boston, 2009)
I. Millington, J. Funge, *Artificial Intelligence for Games* (CRC Press, Boca Raton, 2009)
V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, M.A. Riedmiller, Playing atari with deep reinforcement learning. CoRR abs/1312.5602 (2013)
B. Moldovan, M. van Otterlo, P. Moreno, J. Santos-Victor, L. De Raedt, Statistical relational learning of object aff ordances for robotic manipulation. in *Latest Advances in Inductive Logic Programming*, (Springer-Verlag Berlin Heidelberg, 2012), p. 6
L. Montesano, M. Lopes, A. Bernardino, J. Santos-Victor, Learning object affordances: from sensory–motor coordination to imitation. IEEE Trans. Robot. **24**(1), 15–26 (2008)

F. Mourato, M.P. dos Santos, F. Birra, Automatic level generation for platform videogames using genetic algorithms. in *Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology*, ACM, (Lisbon, Portugal, 2011), p. 8

A. Narayan-Chen, L. Xu, J. Shavlik, An empirical evaluation of machine learning approaches for angry birds. in *IJCAI 2013 Symposium on AI in Angry Birds,* (IJCAI. BeiJing, China, 2013)

J.C. Niebles, H. Wang, L. Fei-Fei, Unsupervised learning of human action categories using spatial-temporal words. Int. J. Comput. Vis. **79**(3), 299–318 (2008)

D. Nyga, M. Beetz, Everything robots always wanted to know about housework (but were afraid to ask), in *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on, IEEE* Vilamoura, Algarve, Portugal (2012), pp. 243–250

O. Oda, L.J. Lister, S. White, S. Feiner, Developing an augmented reality racing game. in *Proceedings of the 2nd International Conference on Intelligent Technologies for Interactive Entertainment*, (ACM. Cancun, Mexico, 2008), p. 2

P.Walega, T. Lechowski, M. Zawidzki, Qualitative physics in angry birds: first results. in *ECAI 2014 Symposium on AI in Angry Birds*, (IOS Press. Prague, Czech Republic, 2014)

C.P. Papageorgiou, M. Oren, T. Poggio, A general framework for object detection. in *Computer Vision, 1998. Proceeding fo the Sixth International Conference on*, IEEE, (IEEE Computer Society. Bombay, India, 1998), pp. 555–562

M. Polceanu, C. Buche, Towards A theory-of-mind-inspired generic decision-making framework. CoRR abs/1405.5048, (2014)

H. Prendinger, G. Schurz, Reasoning about action and change. J. Log. Lang. Inf. **5**(2), 209–245 (1996)

S. Rabin, *Game AI Pro: Collected Wisdom of Game AI Professionals* (CRC Press, Natick, Massachusetts, USA, 2013)

J. Renz, The angry birds artificial intelligence competition (to appear). in *Proceedings of the 29th AAAI Conference,* (AAAI. Austin Texas, USA, 2015)

J. Renz, S. Gould, X. Ge, Angry Birds: AI Competition (2013), http://www.aibirds.org, URL http://www.aibirds.org

E. Sacks, L. Joskowicz, Automated modeling and kinematic simulation of mechanisms. Comput.-Aided Des. **25**(2), 106–118 (1993)

A.N. Sanborn, V.K. Mansinghka, T.L. Griffiths, Reconciling intuitive physics and newtonian mechanics for colliding objects. Psychol. Rev. **120**(2), 411 (2013)

A. Saxena, J. Driemeyer, A.Y. Ng, Robotic grasping of novel objects using vision. Int. J. Rob. Res. **27**(2), 157–173 (2008)

D.L. Schwartz, Physical imagery: kinematic versus dynamic models. Cogn. Psychol. **38**(3), 433–464 (1999)

N. Shaker, G.N. Yannakakis, J. Togelius, Towards automatic personalized content generation for platform games. in *Proceedings of the Sixth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, AIIDE, Stanford, 11–13 Oct 2010

M. Shaker, M. Sarhan, O. Al Naameh, N. Shaker, J. Togelius, Automatic generation and analysis of physics-based puzzle games. in *2013 I.E. Conference on Computational Intelligence in Games (CIG)*, (IEEE. Niagara Falls, Ontario, Canada, 2013), pp. 1–8

K.A. Smith, E. Vul, Sources of uncertainty in intuitive physics. Top. Cogn. Sci. **5**(1), 185–199 (2013)

R. Speer, J. Krishnamurthy, C. Havasi, D. Smith, H. Lieberman, K. Arnold, An interface for targeted collection of common sense knowledge using a mixture model. in *Proceedings of the 14th International Conference on Intelligent User Interfaces*, ACM, (Sanibel Island, Florida, USA, 2009), pp. 137–146

K. Squire, M. Barnett, J.M. Grant, T. Higginbotham, Electromagnetism supercharged!: learning physics with digital simulation games. in *Proceedings of the 6th International Conference on Learning Sciences, International Society of the Learning Sciences*, ICLS ‘04, (Santa Monica, California, USA, 2004), pp. 513–520

J. Sun, J.L. Moore, A. Bobick, J.M. Rehg, Learning visual object categories for robot affordance prediction. Int. J. Rob. Res. **29**(2–3), 174–197 (2010)

M. Thielscher, Introduction to the fluent calculus. Comput. Inf. Sci. **3**(14) (1998)

J. Togelius, G.N. Yannakakis, K.O. Stanley, C. Browne, Search-based procedural content generation: a taxonomy and survey. IEEE Trans. Comput. Intell. AI Games **3**(3), 172–186 (2011)

N. Tziortziotis, G. Papagiannis, K. Blekas, A bayesian ensemble regression framework on the angry birds game. CoRR abs/1408.5265, (2014)

J.M. Vaccaro, Automated planning in very large, uncertain, partially observable environments. PhD thesis, University of California, San Diego, 2010

M. Ventura, V. Shute, Y.J. Kim, Assessment and learning of qualitative physics in newtons playground. in *Artificial Intelligence in Education*, Springer, (Springer Berlin Heidelberg, 2013), pp. 579–582

B. Weiss, *Classic Home Video Games, 1985–1988: A Complete Reference Guide* (McFarland and Company, McFarland, 2009)

B.Y. White, Designing computer games to help physics students understand newton's laws of motion. Cogn. Instr. **1**(1), 69–108 (1984)

D. Wolter, J.O. Wallgrün, Qualitative spatial reasoning for applications: new challenges and the sparq toolbox. in *Qualitative Spatio-temporal Representation and Reasoning: Trends and Future Directions,* IGI Global, Hershey, (Pennsylvani, USA, 2012) doi:10978–1

Y.l. Kuo, J.C. Lee, K.Y. Chiang, R. Wang, E. Shen, C.W. Chan, J.Y.J. Hsu, Community-based game design: experiments on social games for commonsense data collection. in *Proceedings of the ACM SIGKDD Workshop on Human Computation*, ACM, (Paris, France, 2009), pp. 15–22

A. Yilmaz, O. Javed, M. Shah, Object tracking: a survey. Acm Comput Surv (CSUR) **38**(4), 13 (2006)

Z. Zhang, Microsoft kinect sensor and its effect. IEEE Multimedia **19**(2), 4–10 (2012)

C. Zhang, Z. Zhang, A survey of recent advances in face detection. Technical report, Microsoft Research (2010)

P. Zhang, J. Renz, Qualitative spatial representation and reasoning in angry birds: the extended rectangle algebra. in *Principles of Knowledge Representation and Reasoning: Proceedings of the Fourteenth International Conference*, KR, Vienna, 20–24 July 2014

S. Zhou, S.P. Ting, Qualitative physics for movable objects in mout. in *Proceedings of the 39th Annual Symposium on Simulation*, IEEE Computer Society, (Huntsville, Alabama, 2006), pp. 320–325

M. Zyda, From visual simulation to virtual reality to games. Computer **38**(9), 25–32 (2005)

Part II

BCI and Games

Action Games, Motor Imagery, and Control Strategies: Toward a Multi-button Controller

5

Damien Coyle, Jacqueline Stow, Karl. A. McCreadie, Chen Li, Jhonatan Garcia, Jacinta McElligott, and Aine Carroll

Contents

Abstract

Motor imagery is a skill that can be learned to maximize the accuracy of sensorimotor rhythm (SMR)-based brain-computer game interaction (BCGI). Strategies for learning to intentionally modulate sensorimotor cortex activity

D. Coyle (✉) • K.A. McCreadie • C. Li • J. Garcia
Intelligent Systems Research Centre, Ulster University, Derry, Northern Ireland, UK
e-mail: dh.coyle@ulster.ac.uk; mccreadie-k1@email.ulster.ac.uk; li-c@email.ulster.ac.uk; jgarcia.sp@gmail.com

J. Stow • J. McElligott • A. Carroll
National Rehabilitation Hospital, Dun Laoghaire, Republic of Ireland
e-mail: jacqui.stow@nrh.ie; jacinta.mcelligott@nrh.ie; ainem.carroll@hse.ie

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_1

have been developed, using computer games as a training paradigm and gameplay characteristics to motivate and challenge players. These range from one-dimensional movement of a game object to single-button or multi-button BCGI controllers. This chapter overviews SMR-based BCGI focusing on a number of studies to illustrate the key concepts, principles, and methodologies. Examples drawn from the action genre, the most popular BCI game genre, with progressive difficulty and challenges, are presented, including a classic ball-basket game, a spaceship game involving asteroid avoidance, and a platform-based combat-fighter game. A focus is on elucidating the prospects and challenges for BCGI. Preliminary results from a proof-of-concept study of a BCGI multi-button controller referred to as the "CircleTime" controller are presented. The CircleTime controller offers the user the option of selecting between six separate buttons using just two motor imagery tasks. Results involving five able-bodied and seven physically impaired users are presented to provide evidence that the games are accessible even without motor control and the typical levels of control accuracy given the length of time played. The CircleTime controller is tested within combat-fighter game which requires higher cognitive processes to determine commands and select actions as well as completion of short-term and longer-term time-critical actions. The chapter covers basic SMR-based BCI signal processing and performance assessment, progressive learning across games, and camouflaging prolonged training.

Keywords
Brain-computer interface • Electroencephalography • EEG • BCI • Sensorimotor • Games • Controller • Motor imagery • Action games • Game genre

Introduction

A brain-computer interface (BCI) translates the thoughts or mental activities of a person, reflected in neurophysiological signals, into control signals that enable a movement-free communication and interaction with technology (Wolpaw et al. 2002; Mason et al. 2007). Although BCI performance results from able-bodied users are most frequently reported in the literature, the most envisioned benefactors of the technology are those for whom the natural communication pathway from the brain has been damaged (Birbaumer et al. 1999; Silvoni et al. 2011; Pfurtscheller et al. 2000; Coyle et al. 2011b; Stow et al. 2012; Enzinger et al. 2008) as BCI bypass the normal muscular pathways. This breakdown in the communication pathway between the brain and muscles can limit the means of communication prompting the need for a BCI; however, there is also an increasing interest in using BCI for direct brain-computer game interaction (BCGI) and for entertainment. Incidentally, computer games (or video games) have been and are being used to engage BCI users, while they train to control BCI (Marshall et al. 2013; Lalor et al. 2005; Kreilinger et al. 2011; Pfurtscheller et al. 2010; Coyle et al. 2011a; Royer et al. 2010; McFarland et al. 2010), and therefore, using BCI as another interaction method for computer gaming has a number of potential rewards including (1) improved learning and BCI control; (2) enhanced user experience, motivation, and engagement; (3) entertainment; and (4) enablement of the many physically impaired users who still wish to play computer games and for whom BCI can offer an alternative approach. In particular BCI has the potential to put physically impaired users on an "even playing field" with able-bodied users when a physically impaired user competes against an able-bodied user in a competitive gaming scenario. This is particularly pertinent to someone who, for example, has suffered a high-level spinal cord injury (SCI) but who still desires to play computer games and wishes to do so, competitively, even with their disability. There is therefore a demand for BCGI for entertainment and a critical need for BCGI in terms of enablement of the physically impaired. However, there still remain a number of challenges to address (Krusienski et al. 2011).

Signals collected from the brain are inherently highly nonstationary in nature, and considerable effort is involved in decoding these noisy signals into a usable output which is stable and consistent over time. User acceptance is key to the uptake of any new technology and hinges on a number of factors when considering a BCI. The setup time involved when using large numbers of electrodes can diminish the appeal for many users. Additionally, control limitations or deficits in comparison to standard game controllers, in terms of the accuracy and speed of interaction and the number of selectable options, need to be addressed. The games and methods presented in this chapter have been developed with these limitations and challenges in mind with specific focus on keeping up training.

Spontaneous modulation of the sensorimotor rhythms (SMRs) of the brain is possible by performing imagined movement (Chatrian et al. 1959; Pfurtscheller et al. 1994, 1997, 1998), and this forms the basis of the BCGI control strategy focused on in this chapter. Several BCI surveys have analyzed and reviewed BCI games in terms of HCI (Bos et al. 2009), multiplayer gaming (Bos et al. 2010), and virtual reality (VR) applications (Lécuyer et al. 2008), and there have been a number of interesting 2D, 3D, and hybrid-based SMR-controlled BCI games presented recently (Royer et al. 2010; Bordoloi et al. 2012; Leeb et al. 2013; Bonnet et al. 2013).

According to a review by Marshall et al. (2013), the action genre is the most popular genre employed in BCI computer games. As action games typically include fast-moving gameplay, it is surprising that the action genre is popular in BCGI. Gameplay is the specific way in which players interact with a game and is associated with the features of a computer game, such as its plot and the way it is played, as distinct from the graphics and sound effects. Motor imagery is the most popular BCI control method to use within the action genre; this is because motor imagery often provides a method for continuous control of onscreen object or character. A continuous control paradigm enables a player to have control of an element of the game (e.g., player movement) continuously over an extended period, not simply discrete response control.

Motor imagery-based BCI however requires user training in order to be effective. Sessions typically last an hour or perhaps longer, and users may be required to partake in multiple sessions before a reasonable performance level is attained. Feedback, either visual or audible, is critical to the learning process. Attempting to maintain a user's interest over many sessions can also prove difficult as the repetitive nature of the tasks involved can become boring for many participants. The introduction of more varied methods of feedback in the form of different game scenarios can alleviate the monotony experienced by some users during the training periods. It is also critical that user training is camouflaged or disguised as much as possible, in particular for commercial gaming applications, and therefore, accurate feedback in the form of motivating and exciting games should be delivered almost immediately for novice users, with game scenario, difficulty, and challenge changing as the user gains in control proficiency.

This chapter uses examples of multiple games that can take the naïve BCI user from the standard simple ball-basket game (Vidaurre et al. 2006) to a continuously controlled spaceship game (Coyle et al. 2011a) involving asteroid avoidance as BCI proficiency improves. These games are aimed at enhancing SMR modulation proficiency as well as entertaining and challenging the user but are somewhat limited, relative to modern game designs which have multiple control options and involve much greater information transfer between the player and the user.

The basic control strategies in modern games involve interacting with games through multi-button game controllers. Normally to access multiple buttons using an SMR-BCI, the user is required to perform multiple communications to select the required button or spend a substantial time training to produce multiple distinguishable brain responses or SMRs. This is often not possible from a user perspective and is certainly not practical from a game perspective; therefore, there is a need to innovate in terms of the way multiple options and controller buttons can be selected using a BCI that requires a limited amount of training but is accurate, i.e., two classes as opposed to multiple classes.

A proof-of-concept study of a multi-button controller to address these issues is presented. The CircleTime controller allows access to any one of six buttons using one of two motor imagery classes therefore offering a possible solution. The CircleTime controller can follow a natural progression of usage from the simpler 1D continuous control strategies of the ball-basket and spaceship game.

This approach of progressively increasing the number of selectable BCGI controller options without changing the actual BCI control strategy, along with increasing the amount of distractions and items to observe in the game environment in an attempt to draw the player into the gameplay and hold the player's interest for longer, is likely to challenge their capability and lead to increased BCGI control proficiency. The preliminary results provide evidence that the more engaging spaceship game may improve performance over a basic game and user learning. The concept for applying the novel multi-button BCGI controller for more advanced action games such as the proposed combat-fighter game is investigated.

Importantly, as BCIs open the possibility that movement-free gaming improves accessibility of games for the physically impaired, the games presented in this work are tested with able-bodied and physically impaired users.

A major aim of the chapter is to provide the reader with an overview of concepts, principles, and methodologies associated with this SMR (or motor imagery)-based BCI control strategies using illustrative examples and evidence from proof-of-concept studies, including evidence from experiment from able-bodied and physically impaired users. Results from both of these groups are presented to illustrate the typical performance achievable at varying stages of BCGI experience. Section "Methods" describes the methodology employed in each of these examples and their interrelationship along with a brief overview of a typical BCI signal processing framework and performance assessment. Section "Results and Discussion" outlines the results and discusses them in the context of user training, BCI performance, and game performance metrics. Limitations and conclusions are discussed in section "Conclusion" in the context of the illustrative examples presented and the prospects for SMR-based BCGI in the future.

Methods

SMR-BCI Control Strategies for Simple Game Control

Before describing the games that have been used to illustrate motor imagery-based BCGI, it is important to illustrate the types of control that are possible using a motor imagery-based BCI. Voluntary modulation of sensorimotor rhythms (SMRs) forms the basis of noninvasive (EEG-based) motor imagery (MI) BCIs. Planning and execution of hand movement are known to block or desynchronize neuronal activity which is reflected in an EEG spectral power decrease in mu band (8–12 Hz). Inhibition of motor behavior synchronizes neuronal activity (Chatrian et al. 1959). During unilateral hand imagination, the preparatory phase is associated with a contralateral mu (8–12 Hz) and beta (13–30 Hz) event-related desynchronization (ERD) that is preponderant during the whole imagery process (Pfurtscheller et al. 1997, 1998). More recently studies have shown activation of other bands during movement related such as theta (4–8 Hz) and lower gamma (30–100 Hz), particularly in intracortical EEG (or electrocorticography (ECoG)) (Miller et al. 2007; Blakely et al. 2014; Bradberry et al. 2010). BCIs utilize a number of self-directed neurophysiological processes during MI. The dynamical and nonstationary patterns in the EEG signals (or time series) must be dealt with to ensure information can be discriminated and classified consistently. Maximizing the capacity for computer algorithms to separate noise from source, distinguished between two or more different mental states or one mental task (intentional control (IC) state) from all other possible mental states (no control (NC) state), has been the goal of many BCI-focused researchers for the past 20 years. Linear and nonlinear approaches to classification have been applied to classifying EEG signals (Müller et al. 2003; Krusienski et al. 2011;

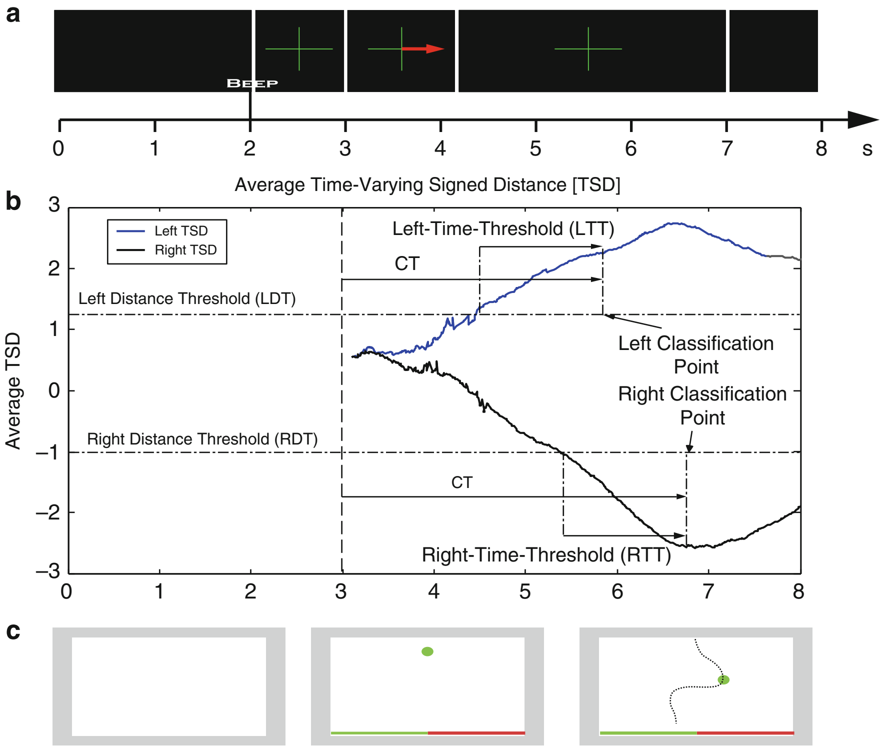


Fig. 1 (**a**) Timing of arrow cue training paradigm with no feedback (*NoFB*), (**b**) an exemplary set of time-varying signed distance (*TSD*)/control signal examples for *left* and *right* motor imagery in accordance with the timing of the paradigm (hypothetical time and distance thresholds for binary decision making are also shown), and (**c**) illustration of the ball-basket game training paradigm referenced to timing of the cue training paradigm and control signal

Pfurtscheller et al. 1998; Lotte et al. 2007; Bashashati et al. 2007; Coyle 2009; Coyle et al. 2005a, 2009; Satti et al. 2009a) (section "Signal Processing" overviews the signal processing strategy used in this work).

Normally the SMR response to motor imagery is analyzed within a window and classified, e.g., a left motor imagery is 1 and right motor imagery is −1. Alternatively, data within the window is continuously classified to produce a continuous control signal or distance value where $D > 0$ for class 1 and $D < 0$ for class 2. The control signal can be considered a magnitude value reflecting the user's ability to modulate a brain response, e.g., a larger consistent deviation of the control signal from a resting state point during motor imagery reflects the users' ability to modulate brain activity. This signal which is continuous in magnitude and in time is often referred to as the time-varying signed distance (TSD) (Pfurtscheller et al. 1998; Schlögl et al. 1997, n.d.). An exemplary (TSD) is shown in Fig. 1b. A typical training and basic game feedback paradigm is also shown reflecting a standard timing of a trial in BCI. During training the user is presented with an arrow pointing either left or right indicating which of the two motor imageries to perform (left/right hand/arm). This arrow appears at $t = 3$ s and the user normally maintains the motor imagery for 4–5 s during which time the user normally has a point at which they can effect a maximum deviation from the resting state.

This normally occurs more than two 2 s after the motor imagery initiation. In the case of this example, the signal increases with imagined left-hand movement and decreases for right-hand movement imagery. If the signal exceeds a distance threshold, referred to here as a left distance threshold (LDT) or a right distance threshold (RDT), for a certain duration or beyond a time threshold, referred to here as a left or right time threshold, a binary class decision can be made. Additional thresholds can be deployed for rising and falling edge control signals to improve performance (Satti et al. 2009a). An alternative to binary classification is to exploit the continuous control signal to effect the movement of an onscreen object as shown in Fig. 1c. This shows the typical ball-basket game where the user is presented with two colored baskets at the bottom of the screen and has to place a ball, which appears at the top of the screen and falls continuously for 3 s, into the green basket. The green basket can be on the left or right of the screen. The user manipulates the one-dimensional horizontal position of the ball as it falls from top to bottom, using the continuously derived control signal which is modulated using motor imagery. This procedure of training with no feedback followed by the traditional ball-basket feedback paradigm has been used extensively in user training. In the following section a more challenging game is introduced.

A BCI-Controlled Spaceship Game

The ball-basket paradigm described in the previous section has clear rest periods between each basket scoring attempt where no target basket or ball feedback is displayed. This section presents a game which allows the user to continuously control a spaceship where the challenge is to manipulate the position of a spaceship using left/right motor imagery to avoid asteroids. The participant is given control of a spaceship sprite located at the bottom of the screen (as shown in Fig. 2). Similar to the ball in the ball and basket paradigm, the lateral movement of the spaceship is directly translated from the TSD (control) output from the BCI classifier. However, whereas the ball would only appear on screen for 4 s and fall from the top to the bottom of the screen (i.e., has a varying x- and y-coordinate), the spaceship is displayed continuously during gameplay and has a constant y-coordinate. The user can affect the x position in a continuous or self-paced control; however, the movement objective is to avoid asteroids which appear on the left or right of the screen and fall toward the bottom, i.e., toward the y-coordinate on which the spaceship moves. The spaceship game timings are consistent with those of both the ball-basket game and the arrow cue training (NoFB) timings. As the player moves through the levels, there are more asteroids which move over and back horizontally as they fall toward the bottom of the

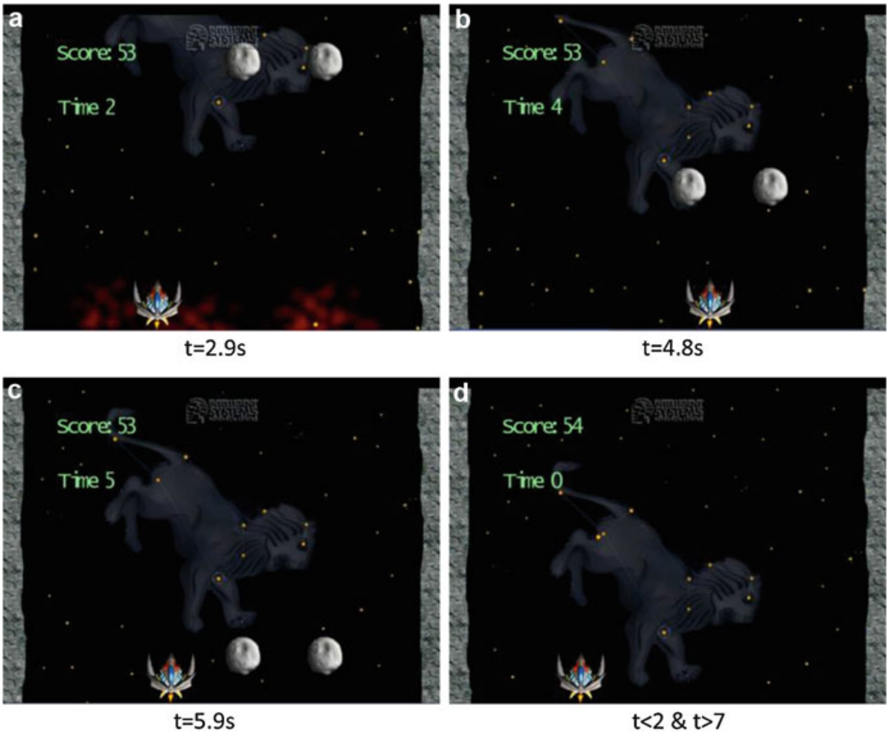


Fig. 2 Example of spaceship game and graphics in difficulty level 2. Approx. timings also shown. Spaceship moves from *left* to *right* at the *bottom* of the screen. (**a**) Asteroids appear at the *top* of the screen, *right-hand side*; therefore cue the subjects to imagine left-hand/arm movement to move the spaceship to the left to avoid asteroids, (**b**) asteroids fall toward the spaceship, (**c**) spaceship moves to left to avoid asteroids, and (**d**) rest period for short time after asteroids disappear and new asteroids appear

screen, increasing the difficulty. This consistency of the timings across the three paradigms is more easily visualized in Fig. 3. Videos showing real-time control and a number of difficulty levels can be viewed here (Coyle et al. 2010, 2011). The space game also contains a number of background images which scroll as if to give the impression of moving through the cosmos. These background images, which include depictions of the constellations, serve not only to give the impression of movement but also to increase the mental load on the participant and are aimed at forcing the user to pay more attention to the game while playing it. The player is also provided with audio. There is a continuous looping of an etheric soundscape, and an explosion is triggered if contact with an asteroid is made. This again adds another dimension to the gameplay and helps to add interest. The game is cognitively more challenging than the basic ball-basket game for a number of reasons: (1) additional moving objects such as shooting stars are visible; (2) objects that look like the asteroids to be avoided have to be detected and ignored; (3) sounds of,

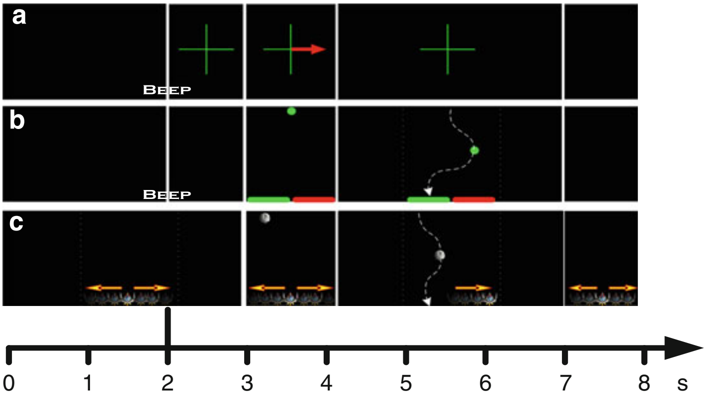


Fig. 3 Illustration of consistency of timing across (**a**) the training with NoFB, (**b**) the ball-basket, and (**c**) spaceship game. Each of the paradigms has 3 s prior to the cue to move *left* or *right* where the participant can relax after which motor imagery is performed for 4 s following the cue to perform *left* or *right* motor imagery to achieve the goal of the training or games

e.g., jet stream and explosions cause distraction and further cognitive processing; (4) other interesting graphics such as animals overlaid on star constellations and associated with star signs, space nebulae, galaxies, and planets which are deliberately there to cause distraction to the users but for which the user must ignore; and (5) there are at least two moving objects on the screen which the user has to pay attention to (i.e., the spaceship and the asteroid) and prevent these from colliding. An advantage of the game is that the user has continuous influence on the position of the spacecraft, and asteroids appear to them randomly to enhance the game-playing experience. A deliberate design decision for the spaceship game was to ensure consistency with earlier training games in terms of timing. As outlined and as can be seen in Fig. 3, the timing is consistent which not only makes transitioning from one game to the next less difficult for the user but also allows acquisition of cued data that can be used to update the BCI translation algorithms offline or online adaptively. Even though the spaceship game provides the user with an illusion of continuous self-paced control, the timing is dictated and the actual targets (class labels) are dictated by the position of the asteroids (left or right of the screen) in the earlier difficulty levels within the game.

CircleTime Controller Technology for Brain-Computer Game Interaction (BCGI)

The games described previously are very limited compared to modern games (in fact they are somewhat similar to some of the very earliest known computer/video games such as Pong and Space Invaders). However, as BCGI presents the player with a new challenge and gameplay experience, these games are appealing to gamers.

Nevertheless, it is important that communication bandwidth in BCGI is improved so that as many games as possible can be controlled with BCI and made accessible to physically impaired users. This would require control options ranging from single-button/switch control to multi-button controllers. Providing users with an array of buttons than can be accessed with minimal latency using motor imagery but at the same time not increasing the interaction difficulty for the user is a challenge. In addition, control strategies and BCIs that require learning additional motor imagery tasks which increase the probability of error and require training over extended periods are not desirable. Approaches using visual evoked potentials (VEPs) have been proposed (Beveridge et al. 2015; Marshall et al. 2015; Wang et al. 2006); however, the control strategies differ from motor imagery and are therefore not reviewed here (for a review see (Marshall et al. 2013)). A number of attempts to achieve greater information transfer rates between the BCI user and application being controlled without increasing the learning effort or increasing the error by introducing more classes/mental tasks have been investigated for a typewriter application (Blankertz et al. 2006), navigation through virtual and real environments (Velasco-Álvarez et al. 2013), and mobile robot navigation (Geng and Gan 2008).

Here we propose a motor imagery-based controller which allows the user to access any one of six buttons in less than 2.5 s at any time during gameplay using one of two motor imageries. The controller comprises up to six buttons located circumferentially on two rotating circles. Each circle has three buttons (small colored circles) located equidistant from each other, and each circle is associated with one of the two motor imageries. In Fig. 4 the left-hand circle is associated with left-hand motor imagery and rotates continuously so that each cursor does one full rotation every 2.4 s. If the user wishes to select the red cursor in the left-hand circle and thus the function associated with that cursor, the user must stop the circle rotating when the red button is on the circumference of the green-shaded segment inside the inner circle (which in this case is the top one third of the circumference of the inner circle). To stop the button rotating, the user must perform left motor imagery and increase the feedback bar (Fig. 4) above a preset threshold (red line). If the green feedback bar exceeds the threshold, the circle is stopped. If the red circle button is located on the circumference of the green area in the inner circle, the function or command associated with that circle is chosen.

Fig. 4 Circle time controller including feedback bar, threshold (*red line*), rotating *circle* (*white*), buttons (*colored circles*), and *green* segment (*stop area*). *Left-hand-side* feedback bar is activated using left-hand motor imagery and *right-hand-side* feedback bar is activated using right-hand motor imagery

The main advantage of the proposed controller is that the user can transition from the earlier games involving a two-class paradigm to a more advanced game requiring up to six commands without the need to learn additional motor imagery commands or significantly change motor imagery strategy. For example, with two-class motor imagery BCI, the maximum probability of misclassification (or error) is 0.5 (probably of selecting one of the classes randomly), whereas if a BCI involves six motor imageries, the maximum probability of error is significantly increased to 5/6 or 0.833, much higher than two-class error probability. Four motor imagery classes are the maximum used in SMR-BCIs to date although some studies have used additional tasks such as simple calculation tasks (Obermaier et al. 2001) and/or imagination of visual auditory or tactile sensation (Dornhege et al. 2004) to make up to five or six classes. The optimal number of classes to use is still an open question. Assuming the user can master the timing to extend the feedback above a threshold with two-class BCI accuracy at the appropriate time, the user can select up to six options with two-class error probability using the proposed controller. This can have a significant impact on the information transfer (IT) rate between the user and the game. IT rate is the bit rate in terms of the number of bits that can be communicated by the user in a given time period, e.g., bits per trial/decision (Shannon 2001; Wolpaw et al. 1998), can be expressed as

$$B = \log_2 N + P\log_2 P + (1-P)\log_2\left[\frac{(1-P)}{(N-1)}\right] \tag{1}$$

where N is the number of possible commands (choices) and P is the probability of selecting the correct command. If we assume a worst-case scenario where the user must select up to six choices using six motor imageries, then $N = 6$ and $P = 0.166$ (chance level accuracy) and then the minimum bit rate B could be as low as $2.3*10^{-6}$ bits. Assuming worst-case scenario with the CircleTime controller, $N = 6$, yet $P = 0.5$, then the minimum bit rate will be 0.432 bits which is a significant improvement, i.e., with the six-button controller, the lower bound on the bit rate is much higher when using a direct six-class BCI, yet the upper bound is not affected. Given that the maximum accuracy reported for a four-class motor imagery-based BCI is normally less than 70 % (Obermaier et al. 2001; Dornhege et al. 2004; Pfurtscheller et al. 2006; Coyle et al. 2007; Schlögl et al. 2005), it can be assumed that accuracy with a six-class SMR-BCI would be significantly lower, whereas accuracy levels with two-class BCIs can reach 100 % accuracy. With the CircleTime controller, a bit rate of approximately 2.6 bits/decision can be achieved assuming 100 % accuracy levels for two-class motor imagery, whereas if we assume a single command can be selected every 4.8 s (every two full rotations of the circles), then a maximum bit rate of 32 bits/min is possible. Assuming a motor imagery classification every 4.8 s is reasonable as in many studies, it has been shown that accurate motor imagery classification is greatest between 1 and 4 s after movement imagery initiation and then additional time is needed for the subject to relax, observe the results of the intended action, and prepare for the next imagery.

Although 100 % accuracy is unlikely, it is much more probable to attain close to this accuracy with a two-class BCI than a six-class BCI, while the CircleTime controller offers six options (classes) with the benefits of two-class BCI performance. Performance is dependent on user and BCI translation algorithms; however, it does illustrate the potential improvement in IT rate given by the CircleTime controller, i.e., a normal two-class controller operating 1 bit/decision every 4.8 s would only achieve 12.5 bits/min. Additionally, the CircleTime controller incorporates gameplay elements, e.g., timing and planning, which in itself improves interaction qualities and challenges and entertains the user. Subjects use the information on timing provided by the rotation of the circle to predict when the target marker/cursor will be located in selection segment of the circle. The user may use more than one full rotation or begin motor initiation in advance of a rotation, targeting a selection in the next rotation during a preceding rotation. Subjects normally learn over time the delay or latency motor imagery response, but there is no doubt there is an added challenge with the timing required for the CircleTime controller.

The following section highlights how the CircleTime controller has been interfaced with a game.

A Combat-Fighter Game with CircleTime Controller

A platform-based combat-fighter style game which gives the user control over a humanlike character has been interfaced with the CircleTime Controller presented in the previous section. The CircleTime controller is presented at the left and right bottom corners of the screen (the control area), one circle corresponding to each of the hand/arm imageries. Again, the control is made up of two parts: a rotating dial and a feedback bar. The outer ring of each dial is continually rotating clockwise at a rate of one revolution every 2.4 s and contains three-colored spheres (buttons): red, blue, and yellow spaced equally. Each color corresponds to a desired command located to the left or right of each dial. The inner circle contains a section highlighted in green located at the top encompassing one third of the circle. The imagination of movement causes the corresponding green feedback bar to increase toward a threshold marked as a red horizontal line. When this threshold is met, the outer ring ceases to rotate, and whichever button appears within the green selection sector is the selected command. Left imagery is used to select one of the three commands, either *forward*, *jump*, or *punch*, while right imagery is used to select between *hard kick*, *super move 1*, and *super move 2*. The game with integrated controller is shown in Fig. 5.

The game is divided into two distinct stages. The goal of the player is to survive while encountering various obstacles which are placed in the character’s way. Stage 2 is a one-on-one fighter game scenario, the aim of which is to defeat the enemy by performing a succession of commands within a limited time frame. A video of the first stage of the game is available to view here (Coyle 2012).

Fig. 5 A screenshot of the combat-fighter game with the CircleTime controller. To defeat the enemy the player has to select punch as instructed in the text above the enemy and highlighted as the shining *yellow button* on the *left circle* by extending the *green feedback bar* on the *left* above the *red threshold line* when the *yellow button* rotates over the circumference of the *green-shaded area* in the inner circle. This must be achieved before the timer (in *red* again above the enemy's position) counts down from 15 to 0 s

Stage One

The player's character begins with a 100 % energy bar which decreases over the period of gameplay, i.e., if the tasks dictated by the game are not completed within a certain time. For example, initially the player must select the forward command in order to progress to the next obstacle in the game, but if they are unable to select this within 30 s, the game will perform an "auto-move" which ensures that they do not become trapped in one position for longer than 30 s. However, this auto-move function will deduct 10 points every time it is invoked. Stage one contains four different types of obstacles: enemies, crates, bombs, and holes/gaps.

Enemies: The player's character encounters five enemies over the course of stage one, and the goal is to defeat each one before they can progress to the next obstacle. The aim is to attack the enemy before being attacked themselves by performing a particular command selected at random from each of the available attack commands, i.e., the game dictates which action will successfully defeat the enemy. All other commanded actions at this point are ineffectual. The enemy will attack every 15 s and so it is important to perform the requested command quickly or points will be lost with each of the enemy's counterattacks.

Crates: The player will encounter a stack of crates which serve to block further progress. The objective is to knock them over in order to clear the path.

It is possible to do this by performing any move except the forward or jump which are ineffective in this instance. If a less powerful move (such as those located on the left-hand/arm imagery control) is selected, only one of the crates will fall leaving another crate to remove. Alternatively, right-hand/arm imagery will select one of the three more powerful moves and cause both crates to fall away from the path. The player is not informed of this and it is left up to them to work out this tactic over multiple plays and is intended to add to the gameplay and to the interest of the game.

Gap: A gap in the path forces the player to come to a stop, and the only way it is possible to overcome this obstacle is by selecting the jump command. It is possible to fall into the gap by incorrectly selecting the forward command, or if jump is not selected within 30 s, the auto-move function will cause the player to also fall into this gap. This will cause the player to lose, not only 10 points for invoking auto-move but an additional 10 points for failure to jump the gap making a total loss of 20 points in all.

Bombs: Stage one contains two bombs which the player must avoid contact with. Each bomb moves between two points vertically at regular intervals and moves to block the player's path from a position just above this point. There is a 5 s gap in which the player has to move underneath the bomb and on to the next obstacle. If the player touches the bomb, it will explode and deduct points and hence timing is critical to avoid this obstacle.

Stage Two

Stage 2 is reached after the successful completion of stage 1 which means that the player was able to pass all obstacles and has some points remaining. Stage 2 takes the form of a one-on-one or faceoff style combat scenario. The aim here is similar to the enemy confrontations in stage 1 in that the player must perform the correct attack command before being attacked themselves or they will lose points. The loser is the character whose energy is depleted first. This can be varied, but in this study if the player gets attacked five times by the enemy before the player has correctly attacked the enemy five times, the player loses and the game ends.

The obstacles and positions at which BCI commands are required in stage 1 are shown in Fig. 6, while stage 2 is depicted in Fig. 7. In essence, the user has to complete the stage in minimum time without running out of energy. Stage 2 requires the user to deplete the enemy's energy five times before the player's energy is depleted five times, each time the enemy gives the player 8 s to attack.

The game has been developed specifically for motor imagery-based BCI and has a number of unique characteristics which make it suitable for BCI control. It is an action game, but it could be considered a "turn-based" action game in that at each point the player is given a period in which to perform the action and then the game-controlled character gets "a turn." Integrating turn-based gameplay slows the game sufficiently for BCI control but also gives the player a feeling of performing time-critical actions. Therefore, there is timing and control on two levels, short-term timing to estimate when the correct command button is located

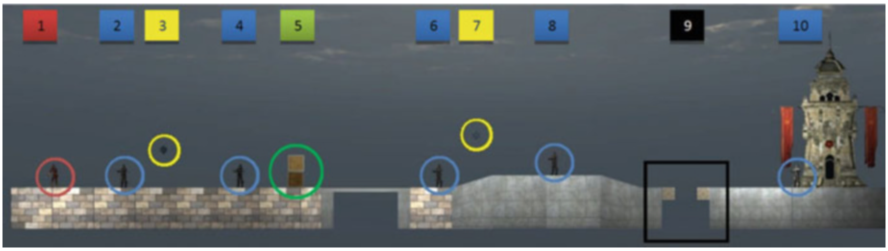


Fig. 6 Combat game map for stage 1 of combat-fighter game. *Circles* indicate points where player has to perform one or more BCI commands. *Numbers* correspond to command action/point details in Table 1

Fig. 7 Stage 2 – one-on-one fighting

at the top segment of the correct circle in the CircleTime controller to select the desired command and the long-term countdown before the enemy attacks. Also, as the user has access to six commands at any time, the player is given the impression he/she has full control of a controller that offers a similar number of control options as standard handheld game controller technology. Free play, where boxes can be hit with any action alongside dictated actions near enemies, is aimed at giving the user a feeling of control of the game and is expected to improve the game immersivity. The following section outlines briefly how the EEG is translated into the control signal to achieve the movement of the ball and

Table 1 Details of actions to be selected and respective number of correct BCI commands to be performed to complete stage 1 flawlessly. Numbers in column 1 correspond to numbered points in the map shown in Fig. 6

Command/ action point	Command 1	Command 2	Command 3	Command 4	Total BCI commands	Cumulative total BCI commands
1	Forward	–	–	–	1	1
2	Attack enemy	Forward (time to prevent hit bomb at 3)	–	–	2	3
3	No action – avoid bomb (see 2)	–	–	–	–	3
4	Attack enemy	Forward	–	–	2	5
5 (option 1)	Hit boxes light	Forward	Hit boxes	Forward	4	See 5 option 2
5 (option 2)	Hit boxes SuperMove	Forward	–	–	2	7
6	Attack enemy	Forward (time to prevent hit bomb at 7)	–	–	2	9
7	No action – avoid bomb (see 6)		–	–	–	9
8	Attack enemy	Forward	–	–	2	11
9	Jump gap	Forward	–	–	2	13
10	Attack enemy	Forward	–	–	2	15

spaceship in the games presented in sections “SMR-BCI Control Strategies for Simple Game Control” and “A BCI-Controlled Spaceship Game” and the movement of the feedback bar in the platform-fighter game.

Signal Processing

The BCI signal processing and setup strategy is briefly described below. The reader is referred to Coyle (2009), Coyle et al. (2005b, 2009), and Satti et al. (2008, 2009a) for a more detailed coverage of the methods applied; however, many of the techniques used in this framework are common to many other motor imagery BCI strategies.

Preprocessing, Feature Extraction, Classification, and Parameter Optimization

A preprocessing framework, referred to as neural time-series prediction preprocessing (NTSPP) (Coyle 2009; Coyle et al. 2005b, 2009), incorporating a collection of specialized prediction neural networks, is employed to enhance EEG signal separability. NTSPP is deployed along with common spatial patterns (CSP), a technique used to maximize the ratio of class-conditional variances (Ramoser et al. 2000; Tomioka and Lemm n.d.; Blankertz et al. 2008) and spectral filtering in subject-specific frequency bands to produce a set of surrogate signals which are much more separable than the raw EEG signals, from which to extract features. Particle swarm optimization (PSO) (Kennedy and Eberhart 1995; Satti et al. 2009b) is implemented to autonomously select subject-specific frequency bands offline. Sensorimotor modulation alters rhythms in the mu (8–12 Hz) and central beta (18–26 Hz) (Pfurtscheller et al. 1998; Coyle et al. 2005a; Satti et al. 2009b; Herman et al. 2008) bands, and hence, all possible bands within the range (8–30 Hz) are included within the search space, e.g., 8–9, 8–10 . . .. 8–30; 9–10, 9–11, . . .9–30; 10–11,. . .; and so on. In order to determine the ideal frequencies for each subject, an inner/outer cross-validation (CV) is performed. A sliding window (width 1–2 s) is then employed on the preprocessed signals in order to extract features using the log variance of each signal within the window. The sampling interval dictates the rate at which features are classified using linear discriminant analysis (LDA). A comparative study of the benefits of using spectral filtering (SF) only, SF with CSP, NTSPP with SF, and a combination of SF with CSP and NTSPP can be found in Coyle (2009), and an overview of the BCI setup can be found in Coyle et al. (2011a).

Online Postprocessing

The control signals shown for each type or motor imagery in Fig. 1 are ideal (an average TSD across multiple motor imagery trials); however, the online, single-trial, continuous control signal suffers a number of problems including biasing toward one of the classes, irregular transients, spikes, and oscillations or jitter which increase the challenge for many users (Satti et al. 2009a). To account

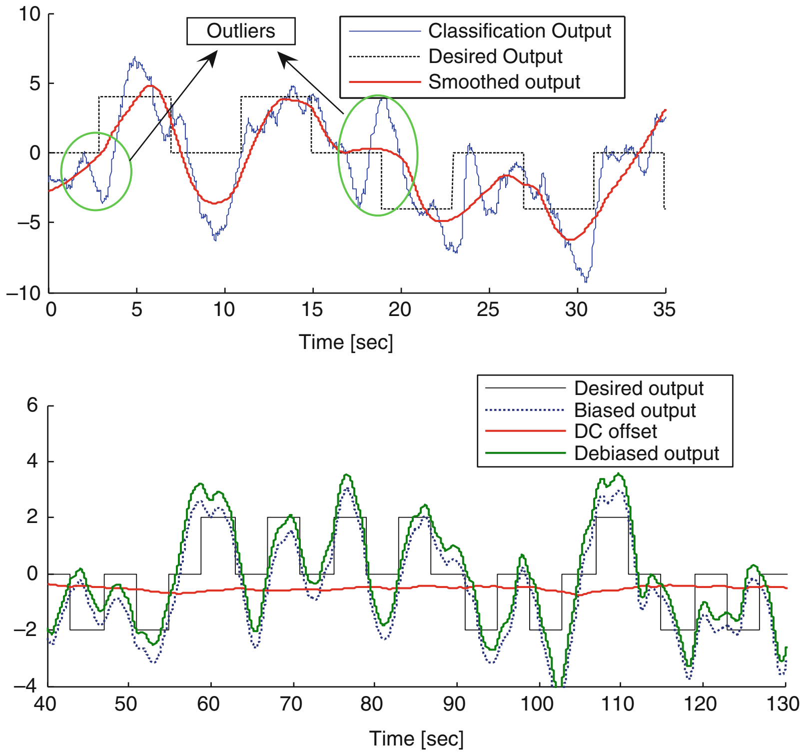


Fig. 8 An illustration of Savitzky-Golay smoothing for removing the outliers from the control signal (classification output) and minimizing jitter and debiasing to remove the class bias (Satti et al. 2009a)

for this and minimize the adverse effects on the user-perceived control, an online postprocessing stage is implemented during online trials in order to account for biasing behavior between classes and is performed by ensuring the mean is removed from the continuous classifier output, i.e., the control signal should have zero mean over a period of event-related and resting state data. A Savitzky-Golay filter is used to remove outliers and smooth this output, chosen as an alternative to the commonly used moving average filter, due to its capacity to retain features such as peak height and width. Figure 8 shows the control signal and postprocessed signal over the duration of control where a particular imagery is required during the period in which the square wave shown is negative or positive.

Experiments

All games were tested using just three bipolar channels of EEG over the sensorimotor cortex (CP3-FC3, CPz-FCz, CP4-FC4, and a reference at FPz) (Fig. 9).

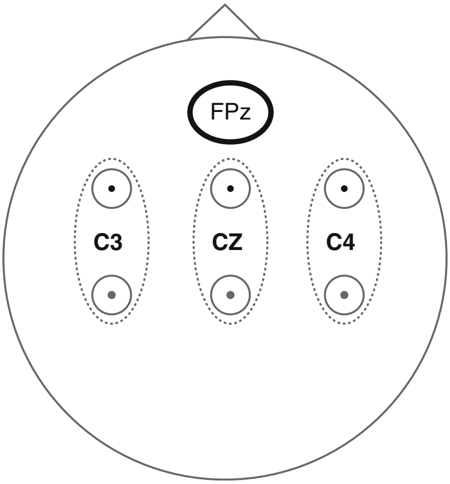


Fig. 9 Three channel montages at positions over the sensorimotor areas

Subjects were instructed to perform hand or arm motor imagery corresponding to the target direction indicated visually in either training or game sessions involving one of the three types of games described previously. The games were assessed in a number of different studies involving different groups of participants as described below.

Assessment 1: Spaceship Assessment with Able-Bodied Users

Assessment 1 involved three able-bodied (AB) right-handed males (30–40 years old) using passive electrodes. EEG was recorded using the EasyCap (www.easycap.de) system through a g.tec gBSamp (www.gtec.at). Participant AB1 was BCI proficient having achieved >90 % in a number of BCI sessions prior to partaking in sessions to assess the spaceship game. Subjects AB1 and AB2 had been involved in ten and three BCI sessions, respectively, 2 months prior to the experiments presented in this work. Participant AB3 was a novice BCI experiment participant. Each participant underwent five or more training/feedback sessions (each of 160 trials) prior to taking part in between three and four spaceship game runs consisting of 80 trials, comprising two levels of difficulty. Some participants attempted the more challenging levels involving laterally and vertically moving asteroids, for fun. The training session involved arrow cue with no feedback and ball-basket paradigm as shown in Fig. 1.

Assessment 2: Spaceship Assessment with Physically Impaired Users

Assessment 2 (A2) included three physically impaired (PI) individuals M, G, and N (all with high-level spinal cord injuries) – participant M has fractures at C3/C4/C5/C6 and an American Spinal Injury Association Impairment Scale (ASIA) classification A, i.e., complete SCI; participant G has an ant wedge fracture of C5 and a compression fracture of C4; and participant N has factures at C5/C6. Data was collected with g.tec gMOBIlab. All PI participants were quadriplegic having sustained spinal injuries within 4 months to 35 years prior to participation. Experiments were conducted in the hospital or patients' home. Three of the PI users initially took part in 10 sessions spread over 4 months involving a training run with no feedback and ball-basket feedback runs to familiarize them with the technology and 4 months later took part in an intensive three to four sessions over 1 week using an active electrode system, same-day classifier recalibration, and mixture of ball and basket paradigm and spaceship games. After a further 1 month, two of the PI participants participated in a further three to four intensive sessions over 1 week, again involving the basket and ball game and spaceship game. Some participants also attempted the combat-fighter game as highlighted in assessment 3.

Assessment 3: Spaceship and Combat-Fighter Games Assessment with Able-Bodied and Physically Impaired Users

Assessment 3 involved seven participants, five physically impaired and two able bodied, and focuses on combat-fighter game performance. Participant M from assessment 2 attempted three combat-fighter gameplays, while G and N attempted one combat-fighter gameplay each. These gameplays all followed the later sessions outlined in assessment 2 (phases 2 and 3). Two additional naïve PI participants also took part in tests. Participant AL (aged 41) sustained a C4 traumatic SCI approximately 4 months prior to the trials and was in the hospital undergoing recovery, assessment, and physical therapy. Participant AN (aged 41) sustained a C4/C5 traumatic spinal cord injury 17 years prior to the trials and was readmitted to the hospital to treat a pressure sore. Both participants AL and AN were medicated and had a number of health issues in addition to their physical impairment. Participant AL participated in four combat-fighter gameplays following four sessions of arrow, ball-basket, and spaceship game. Participant AN participated in only one session of arrow, ball-basket, and spaceship game followed by the combat-fighter game. Two naïve able-bodied participants also participated (ST, SH), each partaking in three runs per session consisting of 60 training trials, 60 ball-basket trials, and 60 spaceship game trials before attempting the combat-fighter game, across five sessions. The objective of these combat-fighter game tests was to assess user sentiment toward the games, to determine controllability, to determine the most appropriate performance assessment methods, and to determine what improvements could be implemented. The combat-fighter game is at the proof-of-concept stage and provides some useful insights into the effectiveness of the proposed platform for multi-button interaction with the CircleTime controller.

Results and Discussion

Using the framework outlined above, results are presented from analysis of data collected from both able-bodied (AB) and physically impaired (PI) individuals from each of the three assessments described previously.

Table 2 A comparison of classification accuracy (*CA*), classification time (*CT*), and area under the accuracy curve for all three subjects with different ball feedback (no parentheses) and space game (with parentheses) paradigms. Best results obtained for each subject on each paradigm are shown in bold

Ball feedback (space game)			
Subject	Mean CA [%]	Mean CT [s]	Mean AUAC
AB1	82.8(**92.6**)	2.3(**2.0**)	351(**382**)
AB2	79.7(75.3)	2.6(2.7)	336(302)
AB3	75.2(**79.8**)	2.4(2.6)	313(307)
Mean	79.3(**82.6**)	2.4(2.4)	**333**(330)

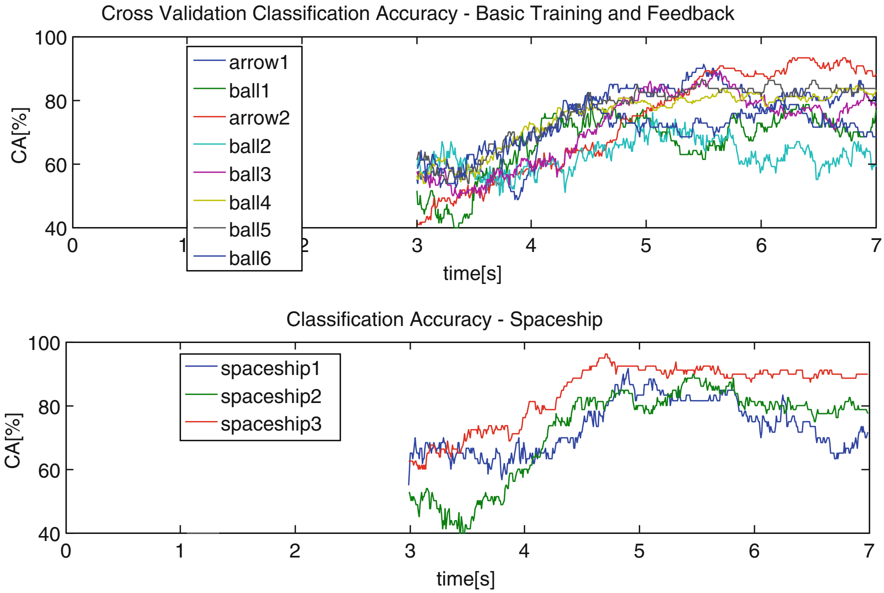
Assessment 1: Space Game Versus Ball and Basket Results (Able-Bodied (AB) Participants)

Three measures of performance are presented here: peak average classification accuracy (CA) (i.e., the percentage of trials classified correctly) obtained from cross-validation, average classification time measured between when the target is known and the point at which accuracy peaks (CT), and the area under the accuracy curve (AUAC), which is proportional to the amount of time spent on the correct side of the screen, i.e., opposite to the asteroids.

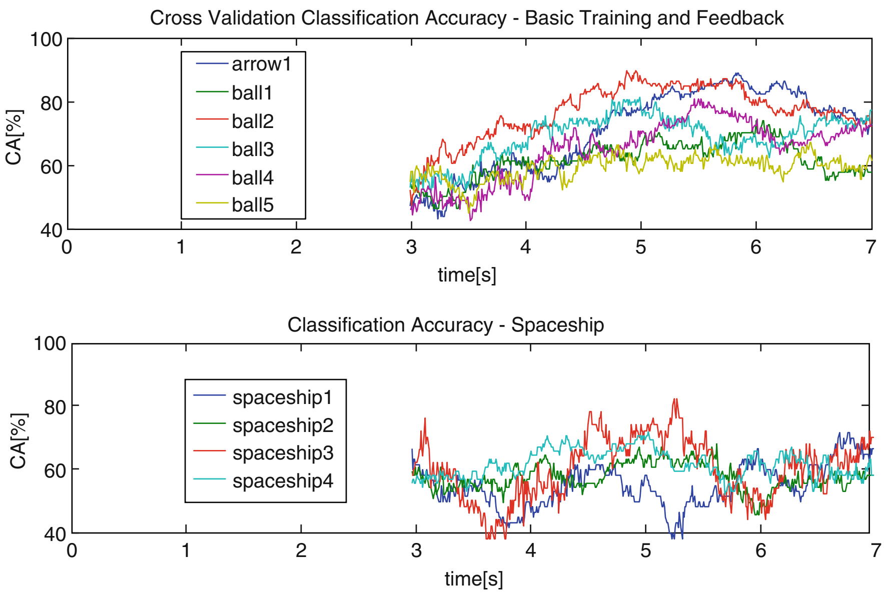
As shown in Table 2, all subjects performed well over the course of the sessions when presented with either the training paradigm or with the classic ball and basket paradigm. Interestingly, when examining the space game results, although subject AB2's CA is lower compared to results obtained in the game, both subjects AB1 and AB3 improved in their CA when presented with this potentially more challenging form of feedback. Subject AB1 who had more BCI experience and facilitated in testing the game during development did not find the space game either appreciably beneficial or detrimental, but both participants AB2 and AB3 reported that they found the space game more challenging but also more engaging. The average accuracy across all space game runs for the three participants is greater than the average across all ball-basket runs, suggesting there is little difference between classification time and AUAC for the three participants. The latter can be measured by estimating the area under the accuracy curve (AUAC) when the accuracy is calculated at every time point during the event-related period as shown in Fig. 10 which shows the time course of accuracy for different sessions for each of the three subjects.

Assessment 2: Spaceship Game Versus Ball-Basket Results (Physically Impaired (PI) Participants)

Figure 11 shows the results over time (session/runs) for the physically impaired participants.



AB1



AB2

Fig. 10 (continued)

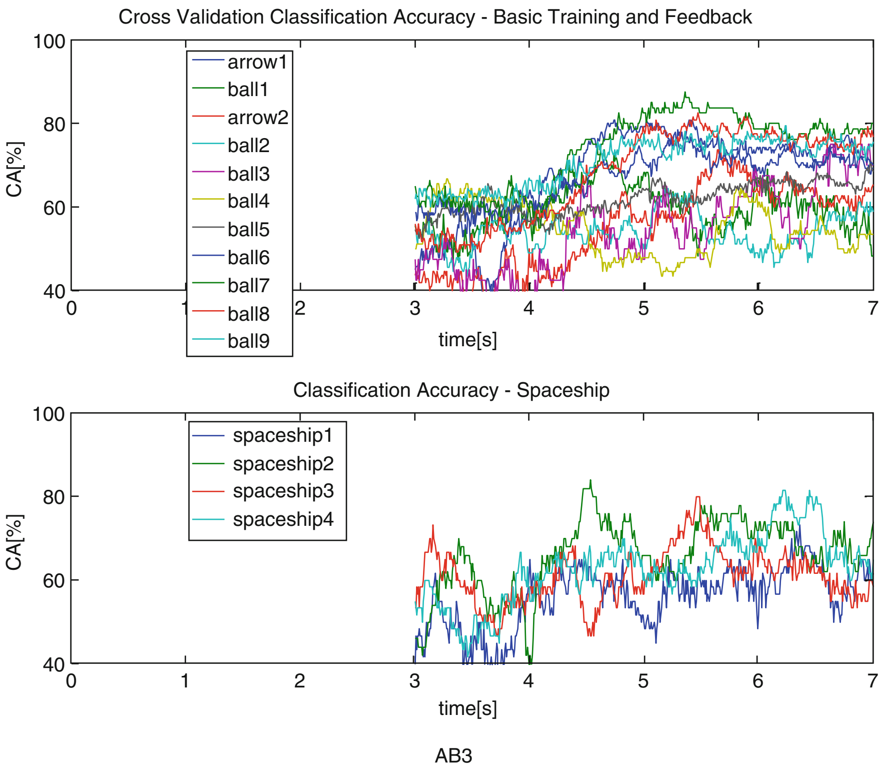


Fig. 10 Time course of classification accuracy cross-validations result for three AB participants (AB1, AB2, and AB3) during sessions without feedback (*arrow* cue for training), with basic feedback (synchronous periodic ball control in the ball-basket paradigm), and the spaceship games (continuous control of the spaceship to avoid moving asteroids). Targets appear (classes are known) at 3 s, at which point the user initiates the correct motor imagery, and disappear at 7 s in every trial according to the timing shown in Fig. 2

In phase 1 (arrow and ball-basket session with session to session transfer of classifier), all participants show a slight improvement, but 70 % criterion level is normally used to determine whether a subject is capable of using a two-class motor imagery BCI (Kübler et al. 2005). In phases 2 and 3 (when the spaceship game is introduced, training is more intensive with within-session classifier setup), participants M and G show significant improvements, whereas participant N's improvement was less pronounced (participant N did not take part in phase 3 due to health issues, and during phases 1 and 2, there were ongoing health issues). Assessing the performance attained by each participant across the phases 2 and 3 and comparing the average performance acquired in the ball-basket runs versus the space game runs, there is an observed average improvement of approximately ∼6 % across the three participants (however this is not significant $p > 0.5$) even though

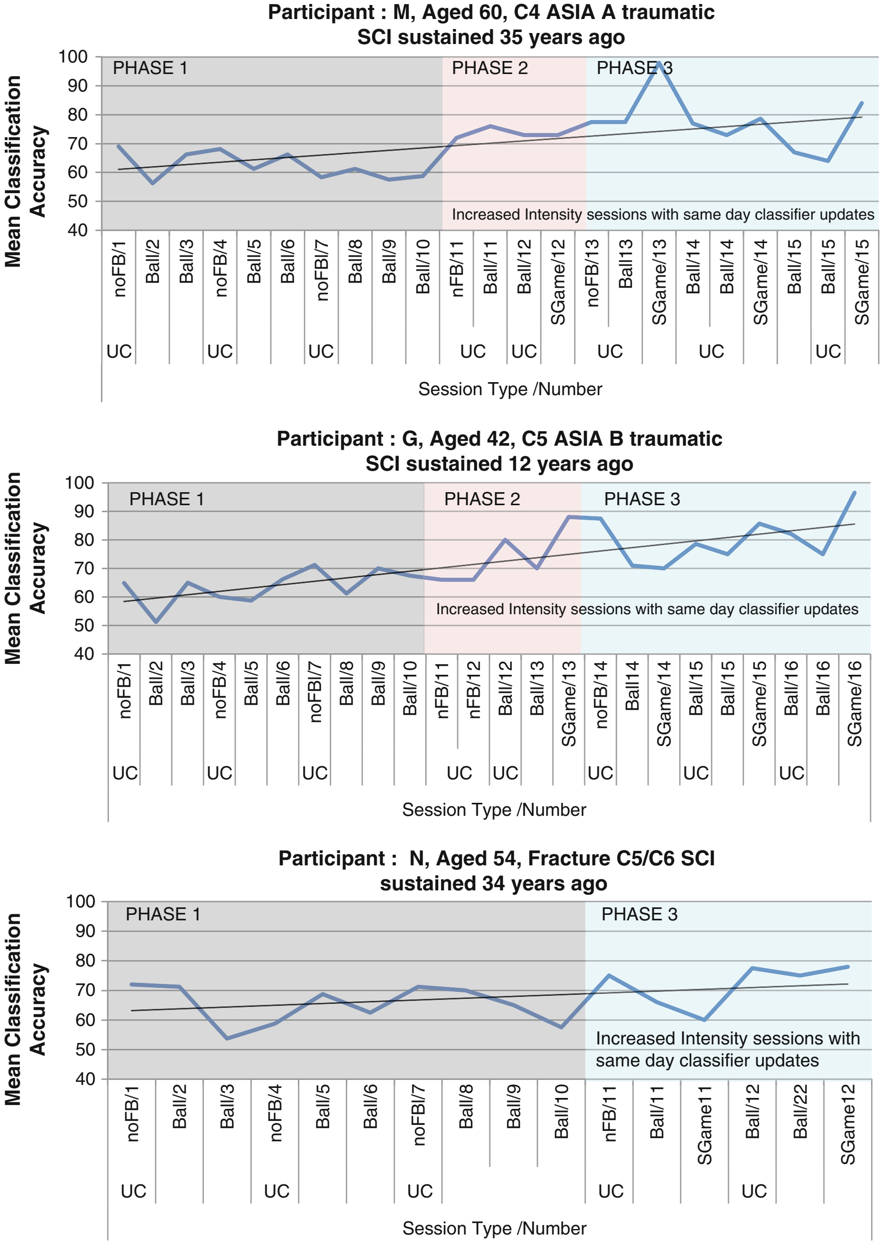


Fig. 11 Assessment 2 results. Three physically impaired participants M, G, and N. UC refers to sessions/run data on which the BCI translation algorithm was updated. The first 10 sessions (phase 1) involved either arrow cue training (*noFB*) or ball-basket feedback, while phases 2 and 3 involved more intensive training with a combination of NoFB, ball, or spaceship game (*SGame*) in each session

Table 3 A comparison of classification accuracy (CA) for three PI participants with different feedback paradigms, ball game (no parentheses) and space game (with parentheses)

	Ball game (spaceship game)
Subject	Mean CA [%]
M	76.0(**85.1**)
G	72.5(**83.4**)
N	**72.8**(69)
Mean	79.3(**82.5**)

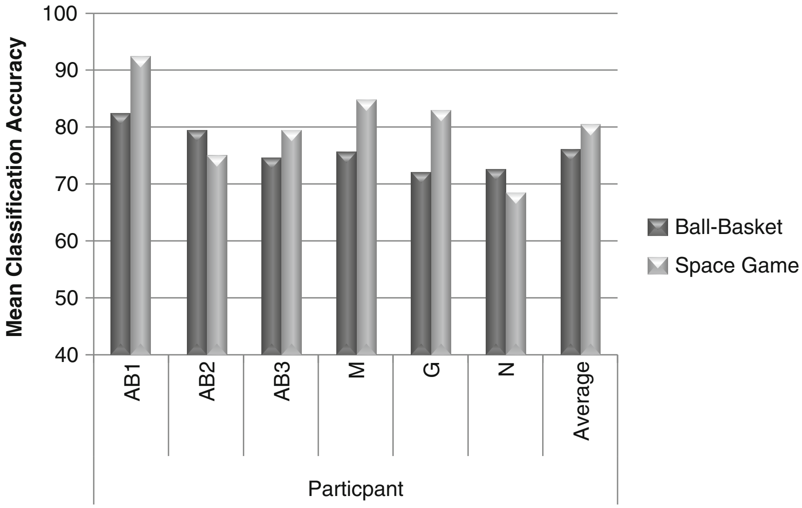


Fig. 12 Performance for ball-basket versus the spaceship games results across the able-bodied and physically impaired participants in assessments 1 and 2

performance of participant N is less with the spaceship game. Participant N only partook in two space game runs. The results are shown in Table 3.

The results from assessment 1 (able bodied) and assessment 2 (physically impaired) are presented in Fig. 12. Overall these results on a small sample of able-bodied and physically impaired users indicate that the space game produces better performance than the ball-basket paradigm, even though all participants were involved substantially in more ball-basket sessions/runs than space game runs (i.e., greater than three times more sessions/runs equating to >1000 trials). In many circumstances the space game test followed ball-basket tests, and therefore, the apparent better performance on the space game may simply be due to the subjects being more experienced at the point at which the space game test was conducted. The results also indicate that there is little difference in performance between able-bodied and physically impaired players of the spaceship game despite long-term spinal cord injury. It must be noted that all results here are indicative but no differences between games and subjects are not shown to be statistically significant. Further investigation with larger sample sizes and stringent experimental setup is necessary to thoroughly investigate the relationship between performance and games or user type.

Assessment 3: A Proof-of-Concept Study of the Combat-Fighter Game (AB and PI Users)

While the ball-basket and the spaceship games follow the same timing paradigm, thus allowing collection of data for easy comparison and performance assessment, the combat-fighter game deviates from the standard setup, and hence, it is not as easy to assess performance or compare results. Performance assessment methods are still being assessed to determine the most appropriate technique. Combat-fighter game-specific performance metrics that can be assessed include correct and incorrect CircleTime stops, average angle of error for stops on the circle but not in the top-shaded area (120°), average time to approach enemies, average time to defeat enemies, time on gaps and boxes, total level 1 and 2 time, and health remaining after each level is completed. During the sessions described in assessment 1, the game performance quantification metrics and some minor details of the combat-fighter gameplay were adapted based on user feedback and observed issues within the games; therefore, only performance metrics which are consistent across all tests are presented here (some actions within combat-fighter game involved auto-move and power loss which could not be used as an appropriate metric for BCI performance but did enhance the game and render it more playable). Also, as there are many variables within the combat-fighter game and some which need to be considered in greater detail and combined into subset of overall metrics, it was decided to focus only on metrics which could be compared consistently across all participants and gameplays and provide comparison of performance in the context of gameplay as well as BCI control. Another reason for this is that some participants only participated in one gameplay session, while others were involved in five gameplays. Collecting data from physically impaired participants does not always run smoothly particularly in a hospital environment where the participant's schedules are not always predicable nor their capacity to partake in trials.

The time taken to defeat each of the five enemies in stage 1 (combat) was consistent across the various gameplays; therefore, assessing these main action points within the game allows assessment of the CircleTime controller as well as providing an indication of the BCI performance. For example, at each enemy action point (points 2, 4, 6, 8, 10 in Fig. 6), the participant has to choose 1 of 6 actions using motor imagery. A correct selection no matter how long it takes equates to approximately 2.6 bits of information transferred between the user and the game according to Eq. 1 where $N = 6$ and $P = 1$ (cf. section "CircleTime Controller Technology for Brain-Computer Game Interaction (BCGI)"). The time it takes to do this is recorded at each enemy, and the average time across the fives enemies is used to calculate the IT rate (bits/min):

$$\text{IT rate} = \frac{60}{\text{mean time per enemy}} \times 2.6 \text{ bits} \tag{2}$$

The results for three PI participants who partook in only one combat-fighter gameplay are presented in Table 4. As can be seen there are variations on the time spent across the different enemies indicating that in some case the participants found it difficult to defeat the enemy; however, as erroneous attempts are ignored, one correct selection indicates a transfer of 2.6 bits of information, and therefore, the overall IT rate for five separate communications is 5 bits/min for participant N, whereas participant G was able to defeat the enemies relatively fast, achieving an IT rates of approximately 22 bits/min across the five enemies.

These results indicate the benefit of employing the CircleTime controller to significantly boost the information transfer between the player and the game even at relatively low accuracies in terms of two-class performance. For two-class motor imagery, the maximum IT rates are normally in the range 5–15 bits/min depending on accuracy and training, e.g., assuming a correct binary classification every 4.8 s, an IT rate of 12.5 bits/min is achieved (see Dornhege et al. 2004; Wolpaw et al. 1998, 2000; Schalk 2008), and bit rates in BCI participant AN achieved 7.53 bits/min in this gameplay even though the combat-fighter attempt was conducted in this participant's first BCI session.

The results for the remaining participants in assessment 3 are presented in Fig. 13 which show the mean time to defeat an enemy and corresponding IT rates across gameplays.

Participant M participated in five gameplays and results show mean time is trending downward, while IT is trending upward. These trends are followed by participants AL and SH, each of whom reaches IT rates of circa 10 bits/min. Participant ST's performance seems to degrade over time from IT rate of approximately 7 bits/min in the first gameplay to circa 3.5 bits/min in the final gameplay. Participants ST's sessions were conducted in disruptive environment due to ongoing construction work in our BCI lab, and this may be the cause of poorer performance.

Although the assessment and results in the chapter are limited by small sample sizes, they do provide a proof of concept for BCGI with the combat-fighter game using the CircleTime controller and it is possible for novice able-bodied and physically impaired users to interact with the game. Although there are erroneous attempts made to defeat some enemies which cause delays, the results show that average communication time and information transfer are sufficient to maintain perceived game playability.

There is a need to improve aspects of the game and fine-tune the CircleTime controller, including subject-specific adjustments to circle rotation speed and feedback gains, and improve performance metrics to ensure everything with the game can be assessed and compared in the context of gameplay and BCI control, e.g., point scoring and energy depleted could be converted to time to render the time to completion of the level as the ultimate performance metric. Other aspects of the

Table 4 Mean time per enemy and corresponding IT rate for three PI participants partaking in one combat-fighter gameplay

Participant	Time[s] enemy 1	Time[s] enemy 2	Time[s] enemy 3	Time[s] enemy 4	Time[s] enemy 5	Mean time[s]	IT rate [bits/min]
N	62.88	4.48	42.40	25.14	18.57	30.69	5.08
G	15.50	0.31	5.20	10.80	4.27	7.22	21.61
AN	26.96	9.61	43.62	4.41	21.56	21.23	7.35

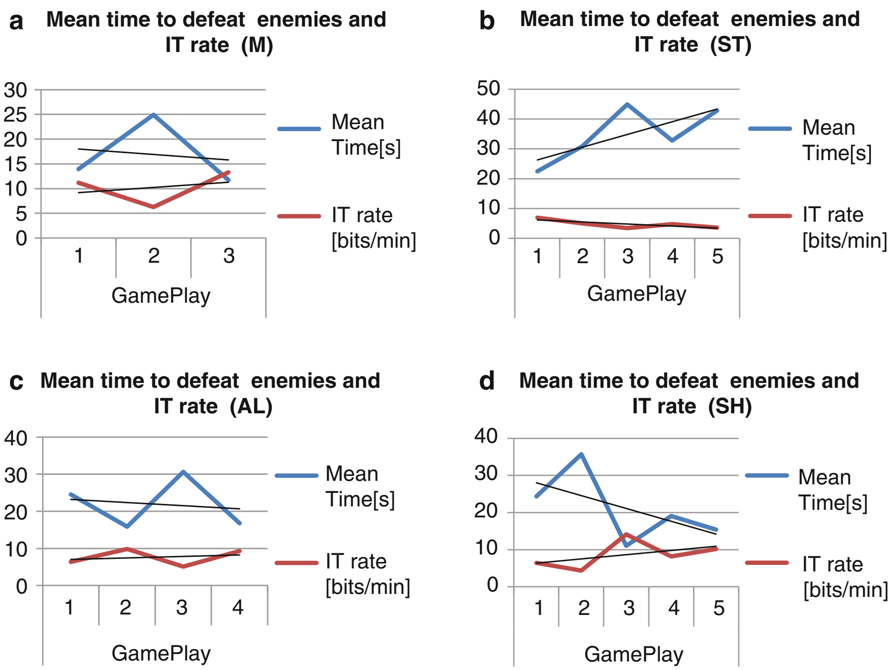


Fig. 13 (**a**) Total stage 1 game time for three subjects. (**b**) Mean time to kill enemies for three subjects. Right and wrong stops using circle time interface (**c**) AL, (**d**) ST, and (**e**) SH

game such as auto-moves and adapting the number of options available to the user on the circle may help to improve playability and increase IT rate. These results serve to prove the concept of both the combat-fighter game and CircleTime controller. Verbal feedback from participants was very positive; even though participants found the game challenging, they found it the most interesting and motivating of all games played. Observers all watched with interest as player progressed through the game, even providing words of encouragement at time-critical moments. The combat-fighter game integrated with the CircleTime controller maintains observers' and players' interest because it offers many of the characteristics of a good game with features such as time-critical actions, rhythmicity and timing, prediction, and multiple different tasks.

Conclusion

This chapter introduces the key concepts, principles, and methodologies associated with motor imagery-based BCGI using a range of action games. A two-class motor imagery BCI that can be used for discrete and continuous 1D object manipulation has been demonstrated. A multi-button BCGI controller (CircleTime) which enables access to six different options at any time (self-paced) using two classes of motor imagery has also been introduced. Results from trials with able-bodied and physically impaired users are presented to demonstrate the type of control performance that is possible. The results suggest that multiple different action games (ball-basket, spaceship, and combat-fighter) requiring different interfaces can be played with one sensorimotor learning strategy. Results from a proof-of-concept assessment CircleTime controller can produce higher IT rates while maintaining the simplicity of two-class motor imagery BCI. The CircleTime controller itself has inherent gameplay characteristics such as timing, prediction, and approximation rendering it well suited for BCGI.

The games presented are progressively more challenging but are also expected to be more appealing to gamers and immersive. An emphasis has been placed on producing progressive game difficulty and disguising SMR-BCI training within the gameplay. It is recommended that the games are played in succession only progressing to the more advanced combat-fighter game when BCI proficiency reaches a criterion level of 70 % (Kübler et al. 2005). The fact that all games use the same motor imagery strategy ensures that participants engage in progressive learning and do not have to change strategy drastically moving from basic training to more advanced multi-button control.

Recommended Reading

A. Bashashati, M. Fatourechi, R.K. Ward, G.E. Birch, A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals. J. Neural Eng. [Online] **4**(2), R32–R57 (2007). doi:10.1088/1741-2560/4/2/R03. Accessed 2 Nov 2012

R. Beveridge, D. Marshall, S. Wilson, and D. Coyle, "Classification Effects on Motion-Onset Visual Evoked Potentials using Commercially Available Video Games," in 20th International Computer Games Conference, 2015, pp. 28–37

N. Birbaumer, N. Ghanayim, T. Hinterberger, I. Iversen et al., A spelling device for the paralysed. Nature. [Online] **398**, 297–298 (1999). doi:10.1038/18581

T.M. Blakely, J.D. Olson, K.J. Miller, R.P.N. Rao et al., Neural correlates of learning in an electrocorticographic motor-imagery brain-computer interface. Brain Comput. Interfaces. [Online] **1**(3–4), 147–157 (2014). doi:10.1080/2326263X.2014.954183. Accessed 5 Jan 2015

B. Blankertz, G. Dornhege, M. Krauledat, M. Schroeder et al., The Berlin Brain-Computer Interface presents the novel mental typewriter Hex-o-Spell. [Online] 2–3 (2006). doi:10.1.1.66.7603

B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K. Muller, "Optimizing Spatial filters for Robust EEG Single-Trial Analysis," IEEE Signal Process. Mag., **25**(1), pp. 41–56, (2008)

L. Bonnet, F. Lotte, A. Lécuyer, Two brains, one game: design and evaluation of a multiuser BCI Video Game based on motor imagery, IEEE transactions on computational intelligence and ai in games. **5**(2), 185–198 (2013)

Bordoloi, S., Sharmah, U. & Hazarika, S.M. (2012) Motor imagery based BCI for a maze game. *4th International Conference on Intelligent Human Computer Interaction (IHCI)*. [Online] 1–6. Available from: doi:10.1109/IHCI.2012.6481848

D.P. Bos, B. Reuderink, B. van de Laar, H. Gürkök, C. Mühl, M. Poel, A. Nijholt, D. Heylen, Brain-computer interfacing and games, in *Brain-Computer Interfaces*, ed. by A. Nijholt, D.S. Tan [Online]. (Springer, London, 2009), p. 149–178. doi:10.1007/978-1-84996-272-8

D.P. Bos, M. Obbink, A. Nijholt, G. Hakvoort, M.C. (2010) Towards multiplayer BCI games, in *BioS-Play* (2010), pp. 1–4

T.J. Bradberry, R.J. Gentili, J.L. Contreras-Vidal, Reconstructing three-dimensional hand movements from noninvasive electroencephalographic signals. J. Neurosci. [Online] **30**(9), 3432–3437 (2010). doi:10.1523/JNEUROSCI.6107-09.2010. Accessed 7 Nov 2012

G.E. Chatrian, M.C. Petersen, J.A. Lazarte, The blocking of the rolandic wicket rhythm and some central changes related to movement. Electroencephalogr. Clin. Neurophysiol. [Online] 11497–11510 (1959). doi:10.1016/0013-4694(59)90048-3

D. Coyle, Neural network based auto association and time-series prediction for biosignal processing in brain-computer interfaces. IEEE Comput. Intell. Mag. (November), **4**(4), 47–59 (2009)

D. Coyle, *Real-time spaceship game control*. [Online]. (2010). Available from: www.youtube.com/watch?v=CSZG_oXf0lg

D. Coyle, *Spaceship game control showing all difficulty levels (speed x 3)*. [Online]. (2011). Available from: http://www.youtube.com/watch?v=j7uOinkVQUY&feature=plcp

D. Coyle, *Brainwave Controlled Combat-Fighter Game*. [Online] (2012). Available from: https://www.youtube.com/watch?v=IiV_Gn3-oo0

D. Coyle, G. Prasad, T.M. McGinnity, A time-frequency approach to feature extraction for a brain-computer interface with a comparative analysis of performance measures. EURASIP J. Adv. Signal Process. [Online] **2005**(19), 3141–3151 (2005a). doi:10.1155/ASP.2005.3141

D. Coyle, G. Prasad, T.M. McGinnity, A time-series prediction approach for feature extraction in a brain-computer interface. IEEE Trans. Neural Syst. Rehabil. Eng. [Online] **13**(4), 461–467 (2005b). doi:10.1109/TNSRE.2005.857690

D. Coyle, T.M. Mcginnity, G. Prasad, A multi-class brain-computer interface with SOFNN-based prediction preprocessing. **44**(0) (2007)

D. Coyle, G. Prasad, T.M. McGinnity, Faster self-organizing fuzzy neural network training and a hyperparameter analysis for a brain-computer interface. IEEE Trans. Syst. Man Cybern. Part B Cybern. Publ. IEEE Syst. Man Cybern. Soc. [Online] **39**(6), 1458–1471 (2009). doi:10.1109/TSMCB.2009.2018469

D. Coyle, J. Garcia, A.R. Satti, T.M. Mcginnity, EEG-based continuous control of a game using a 3 channel motor imagery BCI, in *IEEE Symposium Series on Computational Intelligence,* (2011a), pp. 88–94

D. Coyle, A. Satti, J. Stow, K. Mccreadie et al., Operating a brain computer interface: able bodied vs. physically impaired performance, in *Proceedings of the Recent Advances in Assistive Technology & Engineering Conference*, 2011

G. Dornhege, B. Blankertz, G. Curio, K.R. Müller, Boosting bit rates in noninvasive EEG single-trial classifications by feature combination and multiclass paradigms. IEEE Trans. Biomed. Eng. [Online] 51993–51002 (2004). doi:10.1109/TBME.2004.827088

C. Enzinger, S. Ropele, F. Fazekas, M. Loitfelder et al., Brain motor system function in a patient with complete spinal cord injury following extensive brain-computer interface training. Exp. Brain Res. [Online] 190215–190223 (2008). doi:10.1007/s00221-008-1465-y

T. Geng, J.Q. Gan, Motor prediction in brain-computer interfaces for controlling mobile robots, in *Conference Proceedings: . . . Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*. [Online] 634–637 (2008). doi:10.1109/IEMBS.2008.4649232

I.I. Goncharova, D.J. McFarland, T.M. Vaughan, J.R. Wolpaw, EMG contamination of EEG: spectral and topographical characteristics. Clin. Neurophysiol. [Online] **114**(9), 1580–1593 (2003). doi:10.1016/S1388-2457(03)00093-2

P. Herman, G. Prasad, T.M. McGinnity, D. Coyle, Comparative analysis of spectral approaches to feature extraction for EEG-based motor imagery classification. IEEE Trans. Neural Syst. Rehabil. Eng. Publ. IEEE Eng. Med. Biol. Soc. [Online] **16**(4), 317–326 (2008). doi:10.1109/TNSRE.2008.926694

J. Kennedy, R. Eberhart, Particle swarm optimization, in *Proceedings of ICNN'95 – International Conference on Neural Networks*. [Online] 4 (1995). doi:10.1109/ICNN.1995.488968

A. Kreilinger, V. Kaiser, C. Breitwieser, J. Williamson et al., Switching between manual control and brain-computer interface using long term and short term quality measures. Front. Neurosci. [Online] **5** (January), 147 (2011). doi:10.3389/fnins.2011.00147. Accessed 07 Dec 2012

D.J. Krusienski, M. Grosse-Wentrup, F. Galán, D. Coyle et al., Critical issues in state-of-the-art brain-computer interface signal processing. J. Neural Eng. [Online] **8**(2), 025002 (2011). doi:10.1088/1741-2560/8/2/025002. Accessed 02 Nov 2012

A. Kübler, F. Nijboer, J. Mellinger, T.M. Vaughan et al., Patients with ALS can use sensorimotor rhythms to operate a brain-computer interface. Neurology. [Online] **64**(10), 1775–1777 (2005). doi:10.1212/01.WNL.0000158616.43002.6D. Accessed 01 Aug 2013

E.C. Lalor, S.P. Kelly, C. Finucane, R. Burke et al., Steady-state VEP-based brain-computer interface control in an immersive 3D gaming environment. Eurasip J. Appl.Signal Process. [Online] 3156–3164 (2005). doi:10.1155/ASP.2005.3156

A. Lecuyer, F. Lotte, R. B. Reilly, R. Leeb, M. Hirose, and M. Slater, "Brain-Computer Interfaces, Virtual Reality, and Videogames," Computer (Long. Beach. Calif)., **41**(10), pp. 66–72, (2008)

R. Leeb, M. Lancelle, V. Kaiser, D.W. Fellner et al., Thinking Penguin: multimodal brain–computer interface control of a VR game. IEEE Trans. Comput. Intell. AI Games. [Online] **5** 2), 117–128 (2013). doi:10.1109/TCIAIG.2013.2242072

F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche et al., A review of classification algorithms for EEG-based brain-computer interfaces. J. Neural Eng. [Online] **4**(2), R1–R13 (2007). doi:10.1088/1741-2560/4/2/R01. Accessed 26 Oct 2012

D. Marshall, D. Coyle, S. Wilson, M. Callaghan, Games, gameplay, and BCI: the state of the art. IEEE Trans. Comput. Intell. AI Games. [Online] **5**(2), 82–99 (2013). doi:10.1109/TCIAIG.2013.2263555. Accessed 04 July 2013

D. Marshall, R. Beveridge, S. Wilson, and D. Coyle, "Interacting with Multiple Game Genres using Motion Onset Visual Evoked Potentials," in 20th International Computer Games Conference, 2015, pp. 18–27

S.G. Mason, A. Bashashati, M. Fatourechi, K.F. Navarro et al., A comprehensive survey of brain interface technology designs. Ann. Biomed. Eng. [Online] **35**(2), 137–169 (2007). doi:10.1007/s10439-006-9170-0. Accessed 02 Nov 2012

D.J. McFarland, W.a Sarnacki, J.R. Wolpaw, Electroencephalographic (EEG) control of three-dimensional movement. J. Neural Eng. [Online] **7**(3), 036007 (2010). doi:10.1088/1741-2560/7/3/036007. Accessed 29 Oct 2012

K.J. Miller, E.C. Leuthardt, G. Schalk, R.P.N. Rao et al., Spectral changes in cortical surface potentials during motor movement. J. Neurosci. [Online] **27**(9), 2424–2432 (2007). doi:10.1523/JNEUROSCI.3886-06.2007. Accessed 25 Jan 2013

K.-R. Müller, C.W. Anderson, G.E. Birch, Linear and nonlinear methods for brain-computer interfaces. IEEE Trans. Neural Syst. Rehabil. Eng. Publ. IEEE Eng. Med. Biol. Soc. [Online]. **11**, 165–169 (2003). doi:10.1109/TNSRE.2003.814484

B. Obermaier, C. Neuper, C. Guger, G. Pfurtscheller, Information transfer rate in a five-classes brain-computer interface. IEEE Trans. Neural Syst. Rehabil. Eng. Publ. IEEE Eng. Med. Biol. Soc. [Online] 9283–9288 (2001). doi:10.1109/7333.948456

G. Pfurtscheller, D. Flotzinger, C. Neuper, Differentiation between finger, toe and tongue movement in man based on 40 Hz EEG. Electroencephalogr. Clin. Neurophysiol. [Online] **90**(6), 456–460 (1994). doi:10.1016/0013-4694(94)90137-6. Accessed 07 Apr 2014

G. Pfurtscheller, C. Neuper, D. Flotzinger, M. Pregenzer, EEG-based discrimination between imagination of right and left hand movement. Electroencephalogr. Clin. Neurophysiol. [Online] 103642–103651 (1997). doi:10.1016/S0013-4694(97)00080-1

G. Pfurtscheller, C. Neuper, A. Schlögl, K. Lugger, Separability of EEG signals recorded during right and left motor imagery using adaptive autoregressive parameters. IEEE Trans. Rehabil.

Eng. [Online] **6**(3), 316–325 (1998). Available from: http://www.ncbi.nlm.nih.gov/pubmed/9749909

G. Pfurtscheller, C. Guger, G. Müller, G. Krausz et al., Brain oscillations control hand orthosis in a tetraplegic. Neurosci. Lett. [Online] **292**(3), 211–214 (2000). Available from: http://www.ncbi.nlm.nih.gov/pubmed/11018314

G. Pfurtscheller, C. Brunner, A. Schlögl, F.H. Lopes da Silva, Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks. NeuroImage. [Online] **31**(1), 153–159 (2006). doi:10.1016/j.neuroimage.2005.12.003. Accessed 26 Oct 2012

G. Pfurtscheller, B.Z. Allison, C. Brunner, G. Bauernfeind et al., The hybrid BCI. Front. Neurosci. [Online] **4**(April), 30 (2010). doi:10.3389/fnpro.2010.00003. Accessed 08 Nov 2013

H. Ramoser, J. Muller-Gerking, G. Pfurtscheller, Optimal spatial filtering of single trial EEG during imagined hand movement. IEEE Trans. Rehabil. Eng. [Online] **8**(4), 441–446 (2000). doi:10.1109/86.895946. Accessed 06 Feb 2013

R.G. Robinson, L.B. Starr, J.R. Lipsey, K. Rao et al., A two-year longitudinal study of post-stroke mood disorders: dynamic changes in associated variables over the first six months of follow-up. Stroke J. Cerebral Circ. [Online] **15**(3), 510–517 (1984). doi:10.1161/01.STR.15.3.510

A.S. Royer, A.J. Doud, M.L. Rose, B. He, EEG control of a virtual helicopter in 3-dimensional space using intelligent control strategies. IEEE Trans. Neural Syst. Rehabil. Eng. Publ. IEEE Eng. Med. Biol. Soc. [Online] **18**(6), 581–589 (2010). doi:10.1109/TNSRE.2010.2077654

A. Satti, D. Coyle, G. Prasad, Optimal frequency band selection with particle swarm optimization for a brain computer interface, Workshop/Summer School on Evolutionary Computing Lecture Series by Pioneers. **44**(0), 72–75 (2008)

A. Satti, D. Coyle, G. Prasad, Continuous EEG classification for a self-paced BCI, in *2009 4th International IEEE/EMBS Conference on Neural Engineering*. [Online] 315–318 (2009a). doi:10.1109/NER.2009.5109296

A.R. Satti, D. Coyle, G. Prasad, Spatio-spectral & temporal parameter searching using class correlation analysis and particle swarm optimization for a brain computer interface, in *2009 I.E. International Conference on Systems, Man and Cybernetics*. [Online] 1731–1735 (2009b). doi:10.1109/ICSMC.2009.5346679

G. Schalk, Brain-computer symbiosis. J. Neural Eng. [Online] **5**(1), P1–P15 (2008). doi:10.1088/1741-2560/5/1/P01. Accessed 11 Nov 2012

A. Schlögl, C. Neuper, G. Pfurtscheller, Estimating the mutual information of an EEG-based brain-computer interface. Biomed. Tech. Biomed. Eng. [Online] Biomed. Tech. Eng., **47**(1–2) pp. 3–8, 2002. doi:10.1515/bmte.2002.47.1-2.3

A. Schlögl, D. Flotzinger, G. Pfurtscheller, Adaptive autoregressive modeling used for single-trial EEG classification – Verwendung eines Adaptiven Autoregressiven Modells für die Klassifikation von Einzeltrial-EEG-Daten. Biomed. Tech. Biomed. Eng. [Online] **42**(6), 162–167 (1997). doi:10.1515/bmte.1997.42.6.162. Accessed 11 Feb 2013

A. Schlögl, F. Lee, H. Bischof, G. Pfurtscheller, Characterization of four-class motor imagery EEG data for the BCI-competition 2005. J. Neural Eng. [Online] **2**(4), L14–L22 (2005). doi:10.1088/1741-2560/2/4/L02. Accessed 02 Nov 2012

C.E. Shannon, A mathematical theory of communication. ACM SIGMOBILE Mobile Comput. Commun. Rev. [Online]. **5**, 3 (2001). doi:10.1145/584091.584093

S. Silvoni, A. Ramos-Murguialday, M. Cavinato, C. Volpato et al., Brain-computer interface in stroke: a review of progress. Clin. EEG Neurosci. [Online] **42**(4), 245–252 (2011). doi:10.1177/155005941104200410. Accessed 03 Feb 2013

J. Stow, D. Coyle, A. Carroll, A. Satti et al., Achievable brain computer communication through short intensive motor imagery training despite long term spinal cord injury. Irish Institute of Clinical Neuroscience workshop (abstract) (2012)

R. Tomioka, S. Lemm, Filters for robust EEG. (January 2008), 41–56 (n.d.)

B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K. Muller, “Optimizing Spatial filters for Robust EEG Single-Trial Analysis,” IEEE Signal Process. Mag., **25**(1), pp. 41–56, (2008)

F. Velasco-Álvarez, R. Ron-Angevin, L. da Silva-Sauer, S. Sancha-Ros, Audio-cued motor imagery-based brain-computer interface: navigation through virtual and real environments. Neurocomputing. [Online] 12189–12198 (2013). doi:10.1016/j.neucom.2012.11.038

C. Vidaurre, A. Schlöogl, R. Cabeza, R. Scherer et al., A fully on-line adaptive BCI. IEEE Trans. Biomed. Eng. [Online] 531214–531219 (2006). doi:10.1109/TBME.2006.873542

Y. Wang, R. Wang, X. Gao, B. Hong et al., A practical VEP-based brain-computer interface. IEEE Trans. Neural Syst. Rehabil. Eng. [Online]. 234–239 (2006). doi:10.1109/TNSRE.2006.875576

J.R. Wolpaw, H. Ramoser, D.J. McFarland, G. Pfurtscheller, EEG-based communication: improved accuracy by response verification. IEEE Trans. Rehabil. Eng. Publ. IEEE Eng. Med. Biol. Soc. [Online] **6**(3), 326–333 (1998). http://www.ncbi.nlm.nih.gov/pubmed/9749910

J.R. Wolpaw, N. Birbaumer, W.J. Heetderks, D.J. McFarland et al., Brain-computer interface technology: a review of the first international meeting. IEEE Trans. Rehabil. Eng. [Online] **8**(2), 164–173 (2000). Available from: http://www.ncbi.nlm.nih.gov/pubmed/10896178. Accessed 18 Feb 2013

J.R. Wolpaw, N. Birbaumer, D.J. McFarland, G. Pfurtscheller et al., Brain-computer interfaces for communication and control. Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol. **113**(6), 767–791 (2002)

Brain-Computer Interface Games: Towards a Framework

6

Hayrettin Gurkok, Anton Nijholt, and Mannes Poel

Contents

Abstract

The brain-computer interface (BCI) community has started to consider games as potential applications, while the game community has started to consider BCI as a game controller. However, there is a discrepancy between the BCI games developed by the two communities. This not only adds to the workload of developers but also damages the reputation of BCI games. As a response to this issue, in this chapter, a BCI game framework is presented that was constructed with respect to the research conducted in both the BCI and the game communities. Developers can situate their BCI games within this framework, benefit from the provided guidelines, and extend the framework further.

H. Gurkok • A. Nijholt (✉) • M. Poel
Department EWI Research Group, Human Media Interaction (HMI), Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, Enschede, The Netherlands
e-mail: hgurkok@utwente.nl; anijholt@cs.utwente.nl; mpoel@cs.utwente.nl

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_5

Keywords
Brain-computer interface games • Flow • Challenge • Fantasy • Sociality

Introduction

A brain-computer interface (BCI) is an input modality that can infer certain actions, intentions, and psychological (e.g., cognitive, emotional) states by analyzing the brain activity it captures. Besides its classical purpose of redressing the communication and mobility of disabled people (Wolpaw et al. 2002), BCI has been proposed as a candidate modality for a range of recreational HCI applications to be used by the general population (Tan and Nijholt 2010). Among these, BCI games (Plass-Oude Bos et al. 2010) lead the way. They attract the interest of researchers and developers from both BCI and game communities. However, a discrepancy between the BCI games developed by the two communities is observed.

Many of the BCI games developed by the BCI community aim at testing some psychological hypotheses or evaluating the performance of signal analysis and classification techniques. Thus, less attention is paid to game characteristics – such as challenges presented, which actions are available to overcome the challenges and what are the interaction mechanisms (Marshall et al. 2013) – than to technical aspects. Moreover, these games do not usually have any narrative or rich feedback or visuals and user experience or game experience, and payability evaluations are almost never carried out. This leads to BCI games that are reliable but often not enjoyable or entertaining from a gaming perspective. On the contrary, BCI games from the game community are developed with respect to the game design principles. However, the neurophysiology and signal analysis techniques they rely on are largely unknown, because these games mostly make use of the commercial BCI headsets that have their private technical details. This leads to BCI games that are potentially entertaining but unsatisfactory in terms of feeling of control. In either case, not only are the researchers and developers wasting their efforts in building incomplete, inconsistent BCI games, but also the community keeps on perceiving BCI games as futile applications.

The goal of this chapter is to transfer some knowledge from the games and the BCI communities into a shared framework to make both communities aware of each other's research. From the game community, some game-playing motivations which can be satisfied by the features of BCIs are presented. From the BCI community, the current general interaction paradigms are presented, and the ways they can be used in games are discussed. This will contribute to bridging the gap between the two communities and promoting the development of entertaining, playful, and reliable BCI games.

Background: Brain-Computer Interface Application

A BCI application (or system) can be represented as three procedural blocks (Gürkök and Nijholt 2012) (see Fig. 1). The interaction block manages the high-level interaction between the user and the BCI. It is responsible for evoking or

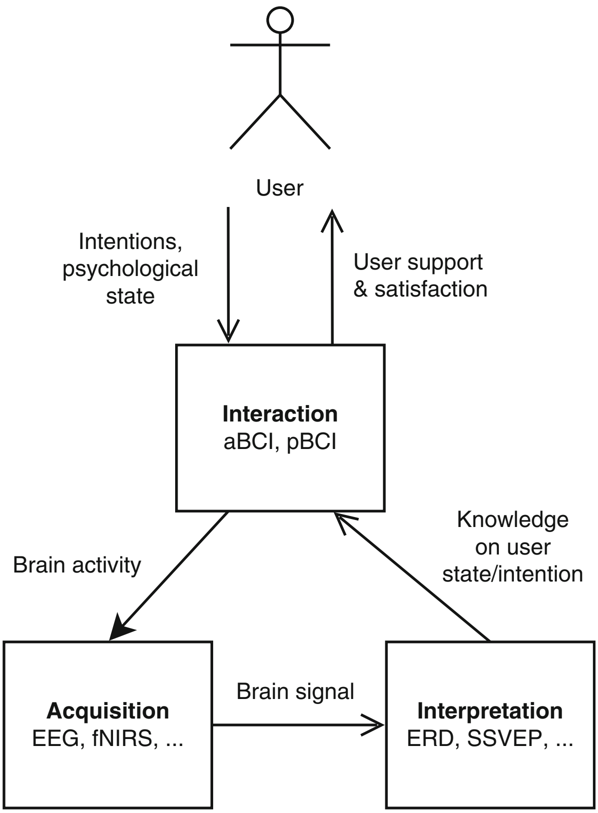


Fig. 1 A brain-computer interface application model

instructing the user to generate the brain activity required for the BCI application to operate. In return, it supports and satisfies the users with respect to their intentions and psychological states. If the application is an active BCI (aBCI), the brain activity that the user has deliberately generated is converted to a control command, such as providing a direction or making a choice. If it is a passive BCI (pBCI), then unintentionally generated brain activity, such as mental workload, is used to optimize the user's well-being by, for example, changing the visuals of a game or adjusting the difficulty.

The acquisition block acquires the user's brain activity. In human-computer interaction (HCI) applications, this is usually the electrical activity captured by an electroencephalograph (EEG), but there are examples with blood movement detected by functional near-infrared spectroscopy (fNIRS) (e.g., Girouard et al. 2010). An EEG is a device that measures voltage changes on the brain surface via electrodes in contact with the scalp and outputs digital signals. Being portable, plug and play, inexpensive, and capable of conveying fast-occurring brain activity, it is preferred in HCI applications.

The interpretation block interprets the digital signals generated by the acquisition block and outputs a prediction on user action, intention, or state. The extent and quality of this prediction are bounded by our knowledge on human neurophysiology. It is only known what happens in the brain with respect to specific, well-defined user actions and external events. For example, it is known what happens when one moves a limb or imagines moving it (called ERD, event-related desynchronization; Pfurtscheller and Lopes da Silva 1999). ERD or motor imaginary is a popular BCI game control strategy (Marshall et al. 2013) and one of the most studied BCI control signals (Coyle et al. 2011). The brain's response when one looks at flickering light (called SSVEP, steady-state visually evoked potential; Herrmann 2001) is known; the frequency of the flickering light or stimulus is reflected in the visual region of the brain. If one pays attention to an infrequent-occurring target image in a sequence of images, the brain response is also known (called a P300 of P3; Polich 2007). So, if one designs a BCI application around these cues and assigns a unique meaning (e.g., a command) to each of them, the application can infer one's actions or intentions within a predefined set. The situation is similar for psychological states. It is not known what happens in the brain when one feels sorry for someone in trouble, but it is known when particular pictures or sounds negatively affect people (Chanel et al. 2009). Thus, when using a BCI, one cannot know everything, but within certain specific contexts, one can give good predictions about someone's actions, intentions, or psychological states.

Besides the unique opportunities BCI provides in sensing the human, which are explained above, it has weaknesses compared to other HCI input modalities. Some of these are not plug and play, use of gel or water-based electrodes, unaesthetic EEG caps, and its low accuracy. Concerning the latter, this is partly due to our limited knowledge about human neurophysiology and to the contamination in the measured brain signal, for instance, due to bodily movement, eye gaze, blinking, or other artifacts. The many ongoing brain studies, in particular source imaging, will keep enhancing the knowledge of the neurophysiology. Moreover, a line of BCI research is trying to remove or correct for contamination due to bodily movement (Fatourechi et al. 2007). Another reason for contamination is the spatial mixing of brain signals sourcing from different cortices (brain areas), masking the signal of interest. Another line of research is working on separating these mixed signals into source components using machine-learning techniques (Bashashati et al. 2007).

A Framework of BCI Games

In the framework a BCI game is represented using two descriptors. The first descriptor specifies the player motivation(s) the BCI game satisfies. In other words, it answers the question: *Why is the game played?* The second descriptor specifies the interaction paradigm(s) the BCI game is built upon. So, it answers the question: *How is the game played?* Next, these descriptors are discussed and some guidelines are provided for each. For the first descriptor, use will be made of research in the game community, while for the latter of research in the BCI community. Finally, the relation between the two descriptors will be discussed.

Motivations Satisfied by the BCI Game

People play games to experience positive affect (Johnson and Wiles 2003), and positive affect is significantly correlated with psychological needs, such as autonomy, competence, relatedness, pleasure, and self-esteem (Hassenzahl et al. 2010; Sheldon et al. 2001). Therefore, people play games which tend to fulfill their psychological needs. Indeed, there is a correspondence between some psychological needs (Sheldon et al. 2001) and some game-playing motivations (Rouse 2005), such as competence and challenge or relatedness and socialization. In this section, some example game-playing motivations (or needs) are provided in which BCI can make a difference, and ways to make the difference are discussed.

Challenge. When someone achieves a goal or when someone feels that he or she is doing something well, positive affect is experienced (Kubovy 1999). That is why people, and particular game players, enjoy doing things that challenge them. Challenge is one of the elements of flow, which is the optimal experience for any activity and described as “so gratifying that people are willing to do it for its own sake, with little concern for what they will get out of it, even when it is difficult, or dangerous” (Csíkszentmihályi 1990). Many researchers have shown the link between flow and games (Cowley et al. 2008). Sweetser and Wyeth (2005) proposed a model describing which elements a game should have in order to provide flow. Their model suggests that a game should offer challenges matching player skills and both must exceed a certain threshold. Similarly, Carroll and Thomas (1988) suggest that “examples of fun indeed must have sufficient complexity or they fall flat (jokes that are too obvious, games that are not challenging).” Moreover, “things are fun when we expect them to be of moderate complexity (interesting but tractable) and then in fact find them to be so (i.e., not too difficult or too easy).”

As pointed out in section “Background: Brain-Computer Interface Application”, BCI is an imperfect technology. Nijholt et al. (2009) suggest that BCI, even when it is still an imperfect technology, can make a perfect game. Players of a BCI game need to show continuous effort to prevent errors and even repeat their actions until they are understood by the BCI. Not only the purposeful repetition brings fun (Blythe and Hassenzahl 2003), but also the successful end result leads to a positive affect. Based on this point of view, one can think that challenge is inherit in any BCI game, because any game requires voluntary control (Sweetser and Wyeth 2005) and any BCI provides imperfect control. However, the technological shortcomings of BCI cannot provide a positively affective challenge if the game cannot conceal them through the tasks the players need to perform. The game should provide the players with the illusion that when they encounter an error, this is not simply a technological fault of BCI but rather an expected situation or challenge given the player tasks and game environment. Players should not be frustrated by the incapability of BCI but should see it as challenge which can be met. This way, they should be motivated to repeat the tasks until they learn how to generate the desired brain activity. Meanwhile, the BCI should also learn how to interpret the generated brain activity.

A similar idea is exploited by the Nintendo DS game Nintendogs in which the players teach dogs some commands, such as sitting, using the built-in microphone. In reality, it is not unusual that it takes multiple trials until a dog learns a command. Moreover, the commander also needs to show effort to provide standard, consistent, and clear commands. So, unsuccessful commands in the game would not frustrate the player but rather motivate and also guide them to provide better commands.

While posing challenges, BCI performance should not be completely neglected. Players should experience uncertainty and curiosity (Malone 1982), but they should be able to predict, to some extent, the game response. Continuous recognition feedback plays an important role in this. Also, the challenge posed by the game should be dynamic. That is, "the level of challenge should increase as the player progresses through the game and increases their skill level" (Sweetser and Wyeth 2005). It has been shown that people can manipulate and learn to improve their voluntary mental actions as well as involuntary reactions as they keep interacting with a BCI that provides accurate feedback (Wolpaw et al. 2002). So, BCI control can be regarded as a skill.

Fantasy. Games let players do things that they cannot do – at least safely or without being criticized – in real life, such as flying or smashing cars. However, in a virtual world, it is not trivial to provide the very same sensation resulting from doing something in the real world. Such a sensation is known as presence and defined as "the perception in which even though part or all of an individual's current experience is generated by and/or filtered through human-made technology, part or all of the individual's perception fails to accurately acknowledge the role of the technology in the experience" (International Society for Presence Research 2000). Riva (2009) claims that rather than our perception, it is our chain of actions that create the presence. He explains that a user "is more present in a perceptually poor virtual environment . . . where he/she can act in many different ways than in a real-like virtual environment where he/she cannot do anything." Actually, 'to act' is not the ultimate goal, the aim is 'to be' in the virtual world and to act is one way of satisfying that aim. So, a player is more present in a virtual world in which he can represent himself more. At this point, the means or actions with which the player can drive the game become crucial.

One of the immersive game activities is role-playing, that is, feeling like oneself and game character are one (Lazzaro 2004; Yee 2006). In role-playing games, the amount and, more importantly, the quality of the self-representation the player can convey to the game are of utmost importance.

Traditional game controllers, such as a game pad or joystick, restrict the information flow from the player to the game. Firstly, the number of buttons or degrees of freedom provided by these controllers is insufficient to satisfy the infinitely large amount of information that could be transferred from the player. And secondly, the idea of representing oneself using buttons or a joystick is not an intuitive one since the player has to spend an effort to learn and memorize the mapping of his or her intentions to controller actions. Tremendous research and development has been going on to alleviate this HCI bottleneck

(Sharma et al. 1998). One example is the work on motion-capturing techniques and devices (such as a Kinect), which enable one-to-one correspondence between player actions (as well as reactions) in the real world and those in the game world.

There are times when the players may need a deeper representation of themselves, rather than their overt actions; consider a life-simulation game, The Sims. In this game, the player controls the life of a character (or several characters) that can be customized to look like the player in terms of outfit or bodily and facial features. The character is also autonomous, and its behavior is influenced by the personality assigned to it by the player at the beginning of the game. It is inevitable that, at some time during play, this virtual look-alike of the player will not act or interact with other characters in congruence with the player's feelings or thoughts, because of either the inaccuracy of the player's personality assignment or the imperfection of the game to produce a desirable action. Consequently, this incongruence will hamper the player's sense of presence. In cases such as these, BCI can provide a translation between the psychological state of the player in the real world and the dynamics of the game world, just as a Kinect provides correspondence between real-world and game-world actions. So, the additional inner state information can strengthen the feeling of presence.

Sociality. Some people enjoy playing computer games with other people (Rouse 2005). They play not necessarily for the challenge but just to be with others. They enjoy spending time with friends, seeing their reactions and expressions, and gloating or feeling proud upon winning (Sweetser and Wyeth 2005). Any multiplayer version of a BCI game can provide such an interactive environment. Players may cooperate or compete using BCI, or they can share their experiences, such as difficulties or enjoyment with control, while playing the game. These are, of course, not specific to BCI games. But, there are other ways in which BCI can provide sociality and which cannot be replicated easily or at all by other controllers.

Many social actions are related to expressing and perceiving emotions. Previous studies have shown that communication of heartbeat, which is a reflection of emotional activity, can improve the copresence (Chanel et al. 2010) and intimacy (Janssen et al. 2010) of players. Heartbeat is certainly not the only nor the best indicator of emotion. BCI can recognize certain psychological states and let us share them. According to neurobiological emotion theories (e.g., LeDoux 1995), the brain is involved in the production and conscious registration of emotional states. So, BCI can provide quick and direct information about our emotional state. Since involuntary brain activity, such as emotional response, is not easily controllable, BCI can provide objective information about our emotional state. For this reason, BCI can also be used in game situations where players would like to hide their psychological states from each other. For example, in a bluffing game, players can restrict their bodily movements and to some extent even their physiological activity but not their brain activity. So, BCI can be used for emotion awareness or, more generally, psychological awareness in two opposite game logics.

Going one step further than emotional awareness, emotional contagion theory states that people tend to converge emotionally and, for this, they tend "to automatically mimic and synchronise expressions, vocalisations, postures, and movements with those of another person" (Hatfield et al. 1994). Research has confirmed that synchronization contributes to coherence (Wiltermuth and Heath 2009) and can be used as a measure of the intensity of the interaction between people (Hatfield et al. 1994). It has therefore been used in some game experience research (Ekman et al. 2012). This suggests that synchronization games can strengthen the interaction between players. In such a game, BCI can enable synchronization of psychological states, in the form of emotional synchrony (Kühn et al. 2011) or mental synchrony (Sobell and Trivich 1989); this adds another dimension of synchrony and can provide a deeper and personal interaction between two players compared to physical synchrony.

Interaction Paradigms Used by the BCI Game

BCI applications rely on brain signals originating from player actions and reactions to events. These actions and reactions are called interaction paradigms and are divided into three categories: mental state regulation, movement imagery, and evoked response generation. In this section an overview is given of BCI games in order to analyze the interaction paradigms used in the interaction design. An extensive survey of BCI games and applied BCI interaction paradigms can be found in Plass-Oude Bos et al. (2010) and Marshall et al. (2013).

Mental State Regulation. Mental state games are usually played via two activities: relaxing or concentrating. These activities stem from clinical practice, such as relaxing to reduce anxiety or concentrating to reduce attention deficiency, but they are used in BCI games for very different purposes. Most of the mental state games allow players to move physical (Hjelm 2003) or virtual (Oum et al. 2010) objects, but there are other mechanisms such as changing the game avatar (Plass-Oude Bos et al. 2010).

Relaxing is a preferable activity in a game as it leads to a positive affective state that players would like to reach while playing games (Lazzaro 2004). Therefore, even if the game environment is not an affective one, people may play such games for the end effect of being relaxed. Moreover, they might easily refer their acquaintances and even children to play such games.

Concentration is also a preferable game activity due to its absorbing effect. According to the flow (Csíkszentmihályi 1990; Sweetser and Wyeth 2005) and immersion theories (Brown and Cairns 2004), concentration is the key to successful games. Therefore, games requiring concentration or paying attention, which is one of the activities leading to concentration, ought to provide a positive play experience. It should be noted that this does not immediately equate the concentration required by a BCI game to that naturally occurring during flow or immersion. However, it suggests that the former can contribute to the establishment of the latter, provided it is integrated into the game in a natural manner. The concentration-related activities should not isolate the player from the game world, but only from the real world. They should make sense in the game context and the story line. The flow should not be broken when the concentration-related activity ends.

Mental state regulation games should either be slow paced or in these games BCI should be used as an auxiliary controller along with a primary controller which is faster than BCI. The speed with which one can change his or her state of relaxedness or concentration is much slower than the speed with which one can press buttons or use any other modality. Mental state games usually allow only binary control. For example, in a relaxation game, players can either be relaxed or not relaxed so they can communicate a maximum of two discrete commands. It is possible to fit a continuous scale between these two states but validating such a scale is nontrivial. Therefore, mental state regulation is less suitable for games that require large numbers of distinct commands. In mental state regulation games, positively affective activities should be preferred. Otherwise, long durations of play may change brain functioning in an unwanted direction.

Movement Imagery. Movement imagery games require no physical movement but imagery of limb movements, mostly the hands, fingers, or feet. Players imagine movements to navigate, as in driving a virtual car (Krepki et al. 2007), or to make selections, as in playing pinball (Tangermann et al. 2009). While in mental state games the capability of BCI (i.e., detecting relaxation or concentration) is irreplaceable by another modality, movement imagery games can be played by more precise modalities. So, players who are not disabled may not enjoy movement imagery games if the interaction is not carefully designed. In this case, providing an intuitive interaction can immerse the players. To provide intuitive interaction, the mapping of imaginary movements to game commands should be coherent. For example, grasping an object at the left- or right-hand side can be matched to left- or right-hand imagery, while walking can be matched to foot imagery. These intuitive mappings can create the illusion that the game is recognizing player's actions, even before they move due to the ERD which happens before the actual or imagined movement.

Movement imagery can be recognized quite quickly, without needing to average the signal (Ramoser et al. 2000; Tangermann et al. 2009). Therefore, movement imagery games are suitable for fast interaction. On the other hand, the number of commands in these games is limited to the number of distinguishable imaginary actions players can perform. Using other modalities in combination with BCI can increase the number of commands. However, the movements made to control other modalities, such as pushing a button or speaking, might contaminate the movement imagery signal with artifacts related to the other movements.

Evoked Response Generation. This class of games is dominated by SSVEP games, accompanied by rare examples of P300 games (e.g., Finke et al. 2009). SSVEP is a brain response to flickering light or images (Herrmann 2001). For example, when one observes a visual stimulus, say an image, that is constantly reappearing, say at a frequency of f, then the amplitude of the signals measured from the visual cortex are found to be enhanced at frequency f and its harmonics ($2f$, $3f$, and so on). This way, if there is a single stimulus, then one can recognize whether someone is observing it. If there are multiple stimuli, then it is possible to recognize which of them someone is observing. Based on these possibilities, SSVEP games have been developed in two ways.

The first approach is to map the strength of SSVEP which is evoked by single stimulus to game actions. For example, a weak SSVEP can steer a virtual plane to the left, while a strong one to the right (Middendorf et al. 2000). Players can manipulate SSVEP strength in different ways. One way is to close and open the eyes to produce weak and strong SSVEP. But this would probably be too trivial to produce challenge in a game. Another way is to regulate SSVEP strength by the amount of attention paid. Research has shown that sustained attention can enhance SSVEP (di Russo and Spinelli 2002). That is, one can infer whether a person is simply exposed to a stimulus or if the person is actually paying attention to it. Sustained attention is an activity which can lead to a state of concentration (Mateer and Sohlberg 2001). This makes SSVEP suitable for concentration games, the advantages of which were discussed before.

The second approach, which is the more popular one, is to use multiple stimuli each of which is associated with a command. In almost all games built with this approach, BCI is used to select a direction. Players can select a direction to aim their gun in a first-person shooter game (Moore Jackson et al. 2009) or to steer their car in a racing game (Martinez et al. 2007). With this approach, a greater number of commands can be issued.

Evoked response generation is less suitable for fast games due to the signal-averaging process, which requires signals to accumulate for some time. But they are suitable for multimodality thanks to their high signal to noise ratio. The number of distinct commands in evoked potential games depends only on the number of stimuli. So, as long as the stimulation space is large enough to accommodate, (finitely) many stimuli can be presented to the player. On the other hand, a computer screen is a limited space so the number of stimuli that can be placed on the screen is also limited. Moreover, as their number increases, the frequency and position of the different stimuli come closer to each other. This makes paying attention difficult for the user as multiple stimuli could interfere with each other. Furthermore, a screen cluttered with attention-demanding flickering stimuli would clutter the visual channel and prevent the player from enjoying primary game elements, such as its visuals or the story line. Therefore, the stimuli necessary for evoked response games should seamlessly be integrated into the game environment. Stimulus properties, such as size, shape, or intensity, should suit the game visuals. They should make sense to the player within the game elements and the storyline. For example, an SSVEP stimulus can be integrated into the game as the wings of a virtual butterfly, which also flicker in reality (Legény et al. 2012).

Guidelines

The framework proposed in section "A Framework of BCI Games" offers BCI game developers a common basis to situate their games. Next, some guidelines will be provided based on the framework so that the developers can further make use of the framework.

Guidelines Regarding Playing Motivations

Regarding virtuosity, the BCI should not introduce challenge merely due to its technical shortcomings but rather in terms of the activities it demands the player to perform. Players should work toward finding the right activities to succeed in the game. They should feel that when BCI cannot understand them, they also have a role in this. Meanwhile, BCI performance should not be completely neglected. Players should experience uncertainty and curiosity, but they should also be able to predict, to some extent, the game response. Continuous recognition feedback is thus helpful. The challenge posed by the game should be dynamic. That is, "the level of challenge should increase as the player progresses through the game and increases their skill level" (Sweetser and Wyeth 2005). It has been shown that people can manipulate and learn to improve their voluntary mental actions as well as involuntary reactions as they keep interacting with a BCI that provides accurate feedback (Wolpaw et al. 2002). So, one can regard BCI control as a skill.

For fantasy fulfillment, capabilities of BCI that resemble fantastic abilities, such as telepathy or telekinesis, should be exploited where possible. Care should be taken not to make promises to provide such controversial abilities though. BCI should be used in situations where "it" is the best or the only modality, rather than replacing a better modality for satisfying the player's motivations or needs.

Regarding sociality game designers should take into account that BCI changes the social bodily interaction; players used bodily interaction to communicate their thoughts and brain state to other players and spectators. Moreover, in order to achieve high level of control, players sometimes behave like paralyzed in order to prevent the contamination of the EEG signal with movement artifacts (O'Hara et al. 2011). But on the other hand, multiplayer BCI games with co-located BCI gamers can enhance the social experience due to collaboration and synchrony of brain states (Nijholt 2015).

As disclaimed before, the motivations provided to play BCI games do not constitute an absolute list, but an illustrative one. Moreover, games can certainly satisfy more than one motivation simultaneously. For example, it is not difficult to imagine a challenging multiplayer fantasy (BCI) game. Based on game-playing motivations, below are some recommendations for developing BCI games in general:

- BCI should not introduce challenge merely in terms of its imperfect recognition performance but rather in terms of the activities the player performs to use it. Players should work toward finding the right mental activity to play the game. They should feel that when BCI cannot understand them, they also have a role in this.
- BCI performance should not be completely neglected. Players should experience uncertainty and curiosity, but they should also be able to predict, to some extent, game response. Continuous recognition feedback is thus helpful.

– The challenge that the BCI game poses should increase as the player skills increase. Players should be aware of the increasing challenge and their increasing skills through feedback.
– BCI training should be done while the player is also training. Repetition is essential for training and also fun for the player if it is integrated well into the game.
– BCI should be used in situations where "it" is the best or the only modality, rather than replacing a better modality for satisfying the player's motivations or needs.
– Capabilities of BCI that resemble fantastic abilities, such as telepathy or telekinesis, should be exploited where possible. Care should be taken not to make promises to provide such controversial abilities though.
– BCI-based games change the social bodily interaction between players. Bodily interaction will be reduced in order to keep up control. On the other hand, multiplayer BCI games can increase brain-based social interaction due to required brain state synchronization or cooperation.

Guidelines Regarding Interaction Paradigms

Regarding the categorization of BCI games (mental state, movement imagery, and evoked potential games), there is no restriction for any category of game to fulfill a certain motivation (virtuosity, fantasy, or sociality). Thus, a game may satisfy more than one player motivation at the same time. But the claim is that a BCI game can exclusively be either an aBCI or a pBCI application. The latter is harder to achieve; therefore, the former has been the popular approach so far. Since any game requires voluntary control (Sweetser and Wyeth 2005), BCI games are intrinsically aBCI applications. Still, it is possible to design a pBCI game in which the involuntary player state influences the game. But as soon as the player realizes the relation, they will start manipulating their state to gain an advantage in the game. Thus, the game will turn into an aBCI application. So the key point in designing pBCI games is to manipulate the game with respect to player state in such a way that the player does not become aware of it.

– Mental state regulation games are more suitable for slow-paced interaction since a player psychological state does not change instantly. It takes time until a player concentrates to do something or to get rid of frustration. Evoked response games are also more suitable for slow-paced interaction, but this is because of the cumulative duration of stimulation and data collection. Movement imagery games may be used for fast-paced interaction since the neurological correlates of movement imagery are instant (they even precede the imagery; Shibasaki and Hallett 2006) and corresponding data is not averaged during analysis.
– Mental state games are less suitable for games that require large numbers of distinct commands, because the number and intensity of the psychological states that BCI can infer are limited. Similarly, in movement imagery games, the number of limbs that BCI is capable of differentiating is limited. On the other

hand, evoked potential games depend only on the number of stimuli. As long as the stimulation space is large enough to accommodate, (finitely) many stimuli can be presented to the player.
– BCI game activities should be intuitive to keep the player in the flow. For example, steering a car to the left/right would better be matched to left-/right-hand imagery rather than concentration/relaxedness.
– The stimuli necessary for evoked response games should seamlessly be integrated into the game environment. They should make sense to the player within the game elements and the storyline. Stimulus properties, such as size, shape, or intensity, should suit the game visuals.
– Just as in any computer game, the long-term effects of BCI game activities should be carefully considered. Especially in mental state games, positively affective activities should be preferred. Otherwise, long durations of play may change brain functioning in an unwanted direction.

The Complete Picture

In this section, the relation between the two descriptors of the framework will be discussed. Specifically, the emphasis will be on the questions: *which interaction paradigms can satisfy which player motivations and in which ways?* Where possible, existing or hypothetical BCI games supporting our discussion will be mentioned.

As any game can do, a BCI game can satisfy more than one set of player motivations (the first descriptor of the framework) at the same time. It might not satisfy any player motivations, for example, if it is just an experimental game. This is illustrated by the set diagrams in Fig. 2. Similarly, a BCI game can make use of multiple interaction paradigms (the second descriptor of the framework). The font stylings of the games in Fig. 2 indicate the different interaction paradigms and their

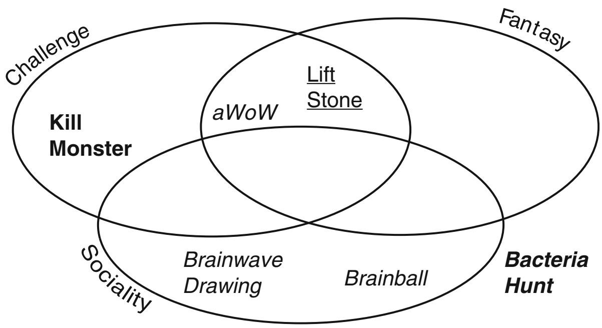


Fig. 2 Diagram showing the set of player motivations and example BCI games. Italic, underlined, and bold fonts represent mental state regulation, movement imagery, and evoked response generation games, respectively

combinations. For example, in the game Bacteria Hunt (Mühl et al. 2010), the players chase fleeing bacteria by controlling an ameba. It is a mental state regulation game because when the players are relaxed, the bacteria flee more slowly. It is also an evoked response game because when the ameba catches a bacterium, an SSVEP stimulus (a circle) appears on the screen, and the players concentrate on this circle to eat the bacterium. Using multiple interaction paradigms is a practice that has been mostly considered in assistive BCI applications to improve the performance, and such applications are called hybrid BCIs (Pfurtscheller et al. 2010). While there is no restriction on using any interaction paradigm to satisfy any player motivation, there might be preferred and non-preferred matchings.

A BCI game can offer a positively affective challenge if the game can hide BCI recognition errors under the player tasks. In the game alpha World of Warcraft (aWoW) (Plass-Oude Bos et al. 2010), the players play a druid that transforms from a bear to a humanoid when the players are relaxed and back again when they are stressed. So, it is a mental state regulation game. The level of relaxedness is determined according to the alpha band power. But the players always think of the high-level tasks of getting relaxed or stressed to transform. There might be times when players cannot transform desirably in the game. Especially for players who are familiar with the game World of Warcraft or similar games, this is not an unexpected situation. These players would consider that it should take some effort and time for a druid to transform. Therefore, they would keep trying to find the right strategy to transform. The same principle may be applied to movement imaginary games. For example, in a game (let us call it Lift Stone for the sake of easy referencing) the players might try to lift a stone by imagining arm movements. Players who are familiar with the phenomenon of telekinesis would find this mapping (imagining arm movement and lifting) an intuitive one and consider that it should take some effort to lift something from far. For evoked potential games, due to the explicit stimuli, offering meaningful challenges is more difficult though not impossible. For example, in a game (let us call it Kill Monster) players might try to kill a monster by concentrating on its heart which flickers like an SSVEP stimulus. People know that, though it does not flicker, the heart beats and it is the organ of vitality. Furthermore, killing a big monster should require some effort. In contrast, for example, in the Bacteria Hunt game, the SSVEP stimulus is not related to game elements (the bacteria), and there is no motivation to keep looking at a flickering circle to eat a bacterium.

To provide fantasy, a BCI game should incorporate additional, inner state information from the players to improve their sense of presence in the virtual world. In the game World of Warcraft, players play the role of a druid. So, the idea is that the players should put themselves in the place of their avatar. The original game translates player actions (e.g., pressing the key W) into game actions (e.g., moving the avatar forward), but it cannot go beyond that. In aWoW, they are represented not only by their actions but also by their psychological states. This way, they can feel more present in the game world. Movement imaginary games cannot convey psychological state information, but they can represent covert player intentions. For example, in the hypothetical Lift Stone game mentioned in the previous paragraph, the players might feel as if the game is understanding their intention without having them move. Going even further, through the readiness potential (Shibasaki and Hallett 2006), the BCI game can actually recognize player intentions "before" the players move or imagine moving their arms.

In a BCI game, sociality might emerge from the explicit social behaviors (e.g., vocalizations, gestures) of co-players while they are competing against or cooperating with each other. In this sense, any interaction paradigm is suitable for use. For example, in the game Brainball (Hjelm 2003), in which co-players regulate their mental states to roll a ball on the table toward each other, players might, for instance, talk to tease each other or generate facial expressions to manifest the difficulty they are experiencing. Sociality might also be formed through communication of implicit player states. For this sort of sociality, mental state regulation is the natural paradigm of choice. For example, in the game Brainwave Drawing (Sobell and Trivich 1989), co-players try to synchronize their mental states in terms of their brain signals in different frequency bands.

Conclusion

In this chapter a framework for BCI games is presented, formed partly by research in the BCI community and partly by research in the game community. The framework lets BCI game developers situate their game ideas among the other BCI games. This way, the developers can take into account the benefits and drawbacks of exploiting particular player motivations and interaction paradigms. As a result, one can have more BCI games which are both enjoyable and reliable.

Within the framework, two descriptors are used to represent a BCI game. One descriptor specifies which player motivation(s) the BCI game can satisfy. Three examples of player motivations are challenge, fantasy, and sociality. The ways BCI games can satisfy each of these motivations are discussed and some guidelines are provided. The other descriptor specifies which interaction paradigm(s) the game is built upon. Three types of interaction paradigms are described: mental state regulation, movement imagery, and evoked response generation. The constructing of these categories is based on the literature on BCI and game research. The interaction paradigm(s) should be chosen to provide intuitive control. For example, steering a car to the left/right would better be matched to left-/right-hand imagery rather than concentration/relaxedness.

But there may be many more player motivations that a BCI game can fulfill or interaction paradigms it can make use of. So, obviously, developers should not feel restricted by the defined categorizations while developing their BCI games. On the contrary, they should investigate the alternative categories. Moreover, the categories for both descriptors are not mutually exclusive. So, a BCI game can satisfy more than one player motivation, or it can accommodate multiple paradigms for control (Mühl et al. 2010).

Last, but not the least, the importance of empirical user experience evaluation should be stressed. While heuristics and recommendations can help to build better BCI games, empirical studies can yield practical information on game characteristics that the players like and dislike.

Recommended Reading

A. Bashashati, M. Fatourechi, R.K. Ward, G.E. Birch, A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals. J. Neural Eng. **4**(2), R32 (2007)

M. Blythe, M. Hassenzahl, The semantics of fun: differentiating enjoyable experiences, in *Funology: From Usability to Enjoyment* (Kluwer, Dordrecht, 2003), pp. 91–100

E. Brown, P. Cairns, A grounded investigation of game immersion, in *CHI' 04 Extended Abstracts* (ACM, New York, 2004), pp. 1297–1300

J.M. Carroll, J.C. Thomas, Fun. ACM SIGCHI Bull. **19**(3), 21–24 (1988)

G. Chanel, J.J. Kierkels, M. Soleymani, T. Pun, Short-term emotion assessment in a recall paradigm. Int. J. Hum. Comput. Stud. **67**(8), 607–627 (2009)

G. Chanel, S. Pelli, N., Ravaja, Kuikkaniemi, K.: Social interaction using mobile devices and biofeedback: effects on presence, attraction and emotions (2010), presented at BioS-Play 2010

B. Cowley, D. Charles, M. Black, R. Hickey, Toward an understanding of flow in video games. Comput. Entertain. **6**(2), 20:1–20:27 (2008)

D. Coyle, J. Garcia, A.R. Satti, T.M. McGinnity, EEG-based continuous control of a game using a 3 channel motor imagery BCI: BCI game, in *Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB), 2011 IEEE Symposium* (IEEE, 2011), Piscataway, USA, pp. 1–7

M. Csíkszentmihályi, *Flow: The Psychology of Optimal Experience* (Harper Perennial, New York, 1990)

I. Ekman, G. Chanel, S. Jàrvelà, J.M. Kivikangas, M. Salminen, N. Ravaja, Social interaction in games measuring physiological linkage and social presence. Simulat. Gam. **43**(3), 321–338 (2012)

M. Fatourechi, A. Bashashati, R.K. Ward, G.E. Birch, EMG and EOG artifacts in brain computer interface systems: a survey. Clin. Neurophysiol. **118**(3), 480–494 (2007)

A. Finke, A. Lenhardt, H. Ritter, The MindGame: a P300-based brain-computer interface game. Neural Netw. **22**(9), 1329–1333 (2009)

A. Girouard, E.T. Solovey, L.M. Hirshfield, E.M. Peck, K. Chauncey, A. Sassaroli, S. Fantini, R.J.K. Jacob, From brain signals to adaptive interfaces: using fNIRS in HCI, in *Brain-Computer Interfaces* (Springer, London, UK, 2010), pp. 221–237

H. Gürkök, A. Nijholt, Brain-computer interfaces for multimodal interaction: a survey and principles. Int. J. Hum. Comput. Interaction **28**(5), 292–307 (2012)

M. Hassenzahl, S. Diefenbach, A. Göritz, Needs, affect, and interactive products – facets of user experience. Interact. Comput. **22**(5), 353–362 (2010)

E. Hatfield, J. Cacioppo, R.L. Rapson, *Emotional Contagion* (Cambridge University Press, New York, 1994)

C.S. Herrmann, Human EEG responses to 1–100 Hz flicker: resonance phenomena in visual cortex and their potential correlation to cognitive phenomena. Exp. Brain Res. **137**(3–4), 346–353 (2001)

S.I. Hjelm, Research + design: the making of Brainball. Interactions **10**(1), 26–34 (2003)

International Society for Presence Research, The concept of presence: explication statement. http://ispr.info/about-presence-2/about-presence/ (2000)

J.H. Janssen, J.N. Bailenson, W.A. IJsselsteijn, J.H.D.M. Westerink, Intimate heartbeats: opportunities for affective communication technology. IEEE Trans. Affect. Comput. **1**(2), 72–80 (2010)

D. Johnson, J. Wiles, Effective affective user interface design in games. Ergonomics **46**(13–14), 1332–1345 (2003)
R. Krepki, B. Blankertz, G. Curio, K.R. Müller, The Berlin brain-computer interface (BBCI) – towards a new communication channel for online control in gaming applications. Multimedia Tools Appl. **33**(1), 73–90 (2007)
M. Kubovy, On the pleasures of the mind, in *Well-being: The Foundations of Hedonic Psychology* (Russell Sage, New York, 1999), pp. 134–154
S. Kühn, B.C.N. Müller, A. van der Leij, A. Dijksterhuis, M. Brass, R.B. van Baaren, Neural correlates of emotional synchrony. Soc. Cogn. Affect. Neurosci. **6**(3), 368–374 (2011)
N. Lazzaro, Why we play games: Four keys to more emotion without story. http://www.xeodesign.com/xeodesign_whyweplaygames.pdf (March 2004)
J.E. LeDoux, Emotion: clues from the brain. Annu. Rev. Psychol. **46**, 209–235 (1995)
J. Legény, R. Viciana Abad, A. Lécuyer, Navigating in virtual worlds using a self-paced SSVEP-based brain-computer interface with integrated stimulation and real-time feedback. Presence **20**(6) (2012)
T.W. Malone, Heuristics for designing enjoyable user interfaces: lessons from computer games, in *Proceedings of the 1982 Conference on Human Factors in Computing Systems* (ACM, New York, 1982), pp. 63–68
D. Marshall, D. Coyle, S. Wilson, M. Callaghan, Games, gameplay, and bci: the state of the art. Comput. Intell. AI Games, IEEE Trans. **5**(2), 82–99 (2013)
P. Martinez, H. Bakardjian, A. Cichocki, Fully online multicommand brain-computer interface with visual neurofeedback using SSVEP paradigm. Comput. Intell. Neurosci. **2007**, 94561 (2007)
C.A. Mateer, M.M. Sohlberg (eds.), *Cognitive Rehabilitation: An Integrative Neuropsychological Approach*, 2nd edn. (Guilford, New York, 2001)
M. Middendorf, G. McMillan, G. Calhoun, K. Jones, Brain-computer interfaces based on the steady-state visual-evoked response. IEEE Trans. Rehabil. Eng. **8**(2), 211–214 (2000)
M. Moore Jackson, R. Mappus, E. Barba, S. Hussein, G. Venkatesh, C. Shastry, A. Israeli, Continuous control paradigms for direct brain interfaces, in *Human-Computer Interaction. Novel Interaction Methods and Techniques* (Springer, Berlin/Heidelberg, 2009), pp. 588–595
C. Mühl, H. Gürkök, D. Plass-Oude Bos, M. Thurlings, L. Scherffig, M. Duvinage, A. Elbakyan, S. Kang, M. Poel, D. Heylen, Bacteria hunt: evaluating multi-paradigm BCI interaction. J. Multimod. User Interfaces **4**(1), 11–25 (2010)
A. Nijholt, Competing and collaborating brains: multi-brain computer interfacing, In *Brain-Computer Interfaces, Intelligent Systems Reference Library*, vol. 74, ed. by A.E. Hassanien, A.T. Azar (Springer International Publishing, 2015), Cham, Switzerland, pp. 313–335, http://dx.doi.org/10.1007/978-3-319-10978-7_12
A. Nijholt, D. Plass-Oude Bos, B. Reuderink, Turning shortcomings into challenges: Brain-computer interfaces for games. Entertain. Comput. **1**(2), 85–94 (2009)
K. O'Hara, A. Sellen, R. Harper, Embodiment in brain-computer interaction, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM (2011), New York, USA, pp. 353–362
K. Oum, H. Ayaz, P. Shewokis, P. Diefenbach, MindTactics: a brain computer interface gaming platform, in *2010 International IEEE Consumer Electronics Society's Games Innovations Conference (ICE-GIC)* (IEEE, Piscataway, 2010)
G. Pfurtscheller, F.H. Lopes da Silva, Event-related EEG/MEG synchronization and desynchronization: basic principles. Clin. Neurophysiol. **110**(11), 1842–1857 (1999)
G. Pfurtscheller, B.Z. Allison, C. Brunner, G. Bauernfeind, T. Solis-Escalante, R. Scherer, T.O. Zander, G. Mueller-Putz, C. Neuper, N. Birbaumer, The hybrid BCI. Front. Neuroprosthet. **4**, 30 (2010)
D. Plass-Oude Bos, B. Reuderink, B. van de Laar, H. Gürkök, C. Mühl, M. Poel, A. Nijholt, D. Heylen, Brain-computer interfacing and games, in *Brain-Computer Interfaces: Applying our Minds to Human-Computer Interaction* (Springer, London, UK, 2010), pp. 149–178

J. Polich, Updating P300: an integrative theory of P3a and P3b. Clin. Neurophysiol. **118**(10), 2128–2148 (2007)

H. Ramoser, J. Muller-Gerking, G. Pfurtscheller, Optimal spatial filtering of single trial EEG during imagined hand movement. IEEE Trans. Rehabil. Eng. **8**(4), 441–446 (2000)

G. Riva, Is presence a technology issue? Some insights from cognitive sciences. Virt. Real. **13**(3), 159–169 (2009)

R. Rouse, *Game Design: Theory & Practice*, 2nd edn. (Wordware, Sudbury, 2005)

F. di Russo, D. Spinelli, Effects of sustained, voluntary attention on amplitude and latency of steady-state visual evoked potential: a costs and benefits analysis. Clin. Neurophysiol. **113**(11), 1771–1777 (2002)

R. Sharma, V. Pavlovic, T. Huang, Toward multimodal human-computer interface. Proc. IEEE **86** (5), 853–869 (1998)

K.M. Sheldon, A.J. Elliot, Y. Kim, T. Kasser, What is satisfying about satisfying events? Testing 10 candidate psychological needs. J. Pers. Soc. Psychol. **80**(2), 325–339 (2001)

H. Shibasaki, M. Hallett, What is the Bereitschaftspotential? Clin. Neuro-physiol. **117**(11), 2341–2356 (2006)

N. Sobell, M. Trivich, Brainwave drawing game, in *Delicate Balance: Technics, Culture and Consequences, 1989*, ed. by N. Sobell (Los Angeles Chapter IEEE SSIT-30, Torrance, 1989), pp. 360–362

P. Sweetser, P. Wyeth, Gameflow: a model for evaluating player enjoyment in games. Comput. Entertain. **3**(3), 1–24 (2005)

D.S. Tan, A. Nijholt (eds.), *Brain-Computer Interfaces: Applying our Minds to Human-Computer Interaction* (Springer, London, UK, 2010)

M. Tangermann, M. Krauledat, K. Grzeska, M. Sagebaum, B. Blankertz, C. Vidaurre, K.R. Müller, Playing pinball with non-invasive BCI, in *Advances in Neural Information Processing Systems*, vol. 21 (The MIT Press, Cambridge, MA, 2009), pp. 1641–1648

S.S. Wiltermuth, C. Heath, Synchrony and cooperation. Psychol. Sci. **20**(1), 1–5 (2009)

J.R. Wolpaw, N. Birbaumer, D.J. McFarland, G. Pfurtscheller, T.M. Vaughan, Brain-computer interfaces for communication and control. Clin. Neurophysiol. **113**(6), 767–791 (2002)

N. Yee, Motivations for play in online games. CyberPsychol. Behav. **9**(6), 772–775 (2006)

Brain-Computer Interfacing and Virtual Reality

7

Doron Friedman

Contents

Abstract

Brain-computer interface (BCI) and virtual reality (VR) are natural companions. BCI provides a new interaction technique for controlling VR, and VR provides a rich feedback environment for BCI while retaining a controlled and safe environment. The combination of VR and BCI allows for providing participants with novel experiences that are impossible otherwise. Both fields still pose many technological challenges to scientists and engineers, but both are making rapid progress.

VR and BCI have been combined in multiple ways: BCI can be used for navigation in VR, for controlling a virtual body, and for controlling the virtual world directly. More recent directions explore the possibilities of using BCI for purposes other than control in VR, such as designing and implementing VR systems that adapt to the participant's cognitive and emotional state.

D. Friedman (✉)
The Advanced Reality Lab, The Interdisciplinary Center, Herzliya, Israel
e-mail: doronf@idc.ac.il

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_2

Keywords
Brain-computer interface • Virtual reality • Embodiment • Avatar • Navigation • EEG • Motor imagery • SSVEP • P300 • Real-time fMRI

Introduction

What is it like to control the world with your mind? Psychokinesis ("mind movement" in Greek) is "an alleged psychic ability allowing a person to influence a physical system without physical interaction" (Wikipedia). While there is no evidence that such parapsychological abilities actually exist, the integration of two technologies – BCI and virtual reality (VR) – now allows a wide range of experiences whereby participants can control various aspects of their environment, using mental effort alone.

This chapter is not intended as a tutorial on BCI nor as a tutorial on immersive virtual reality. Rather, we focus on the outcome of bringing these two disciplines together. For recent reviews on brain-computer interfaces, we recommend other sources (Huggins and Wolpaw 2014; Krusienski et al. 2011; van Gerven et al. 2009), and we only provide a brief introduction. In addition, we focus on the human-computer interface aspects, getting into the BCI engineering aspects only when they are relevant.

Most BCI research in humans is done with electroencephalography (EEG), whereby electrodes are placed on the scalp. Neuroscientific studies overcome the low signal-to-noise ratio of EEG by averaging responses of multiple subjects and multiple events. BCI does not have this luxury, as it requires reasonable accuracy in decoding every single trial, in real time, and thus only a small number of "thought"-based interaction paradigms are possible. In the last two decades, only three EEG-based paradigms have been recruited for BCI. Two of these methods, P300 and SSVEP, are based on evoked potentials and are thus externally driven; i. e., the interaction requires an external stimulus to be provided to the participant, and the participant's commands are inferred from the neural response to this stimulus. The P300 paradigm utilizes the fact that the infrequent events to which the subject is expecting, based on the so-called oddball paradigm, elicit the P300 component of the event-related potential (ERP) (Donchin et al. 2000). The steady-state visually evoked potential (SSVEP) paradigm utilizes the fact that when the retina is excited by a flickering visual stimulus, the brain generates electrical activity at the same (or multiples of) frequency (Cheng et al. 2002). Although these paradigms are based on brain signals, they can be argued to be functionally equivalent to control using eye gaze (Brunner et al. 2010). The third paradigm is based on subjects imagining moving their left hand, right hand, or legs, which is referred to as motor imagery. This paradigm is internally driven and can be used in ways that intuitively map "thoughts" to functionality. However, it is limited in that it requires extensive training, not everyone can use it (Guger et al. 2003), and its information transfer rate is lower than the other two paradigms.

In this chapter, we focus on virtual reality not only as a technology but also as a conceptual framework. The early pioneer Ivan Sutherland envisioned VR as the ultimate display (Sutherland 1965). Brain-computer interface, in theory, has the potential to become the ultimate interaction device – just “think” of something and it happens. Current state of the art in BCI is, of course, very far from that vision; at the moment, BCI should be referred to as “brain reading” rather than “mind reading,” i.e., it is often based on decoding brain waves rather than decoding mental processes (“thoughts”). Eventually, there may be a one-to-one mapping from brain waves to mental processes, but with the current recording techniques, the brain patterns that can be detected are much coarser than specific thoughts.

The relationship between VR and BCI goes further. Recent attempts in explaining the illusions that can be so powerfully induced by highly immersive VR mostly rely on the sensorimotor contingencies between perception and action (Slater 2009). Thus, unlike more traditional interfaces such as keyboard and mouse, VR is based on body-centered interaction and on the immediate feedback that the participants receive when they move their bodies. BCI, however, allows bypassing the muscles and the body, allowing the brain to directly control the environment. The combination of VR and BCI may thus lead to an extreme state of disembodiment – the closest we can get to being a “brain in a vat” (Putnam 1982). Char Davies, with her VR art pieces Osmose and Ephemere, wanted to challenge the “disembodied techno-utopian fantasy,” by controlling VR by breathing – thus bringing the body back into VR (Davies and Harrison 1996; Davies 2004). In this sense, BCI-VR takes us a step backward: while VR attempts to bring back our whole body into the digital realm, BCI attempts to bypass our bodies (Friedman et al. 2009). Until recently, video games have not been played in a highly immersive setup and thus have not utilized the full consequences of VR. However, at the time of writing, the popularity of the low-cost VR devices suggests that this may change.

Why is VR a natural addition for BCI? First, the reasons to use VR for BCI are the same as for using VR in general: it is the best option for exploring and practicing tasks in an environment that is dynamic and realistic yet controlled and safe. For example, VR can be used for evaluating BCI and training paralyzed patients before they attempt to use the BCI in the physical world (Leeb et al. 2007a). In addition, VR can provide motivation for BCI training, which is often lengthy and tedious; motivation has also been shown to play an important role in BCI used by paralyzed patients (Alkadhi et al. 2005). Emotionally relevant stimuli enhance BCI, and this has led some to embed faces in the visual stimuli used for SSVEP and P300 BCIs, rather than just using letters or abstract symbols. Using BCI in VR is expected to lead to higher emotional responses. An interesting finding relates to changes in heart rate in VR BCI. In “typical” BCI, with abstract feedback, heart rate is expected to decrease, but it has been found to increase in VR BCI (Pfurtscheller et al. 2008); this is another evidence that VR feedback has a different physiological effect on subjects than “typical” BCI.

While developers of both VR and BCI still face many technical challenges, both fields may be at the stage of moving out from the research laboratories into the real world. At the time of this writing, low-cost VR devices are becoming available to the mass market. Low-cost EEG devices, such as the Emotiv EPOC or the Interaxon MUSE device, are also available. Most of these EEG devices are limited in signal quality, but they may be at least partially sufficient for BCI (Liu et al. 2012). There are open software platforms for BCI development and customization. The OpenVibe platform may be an easy way to get started, even for nonprogrammers using visual programming, and it is integrated with a VR environment (Renard et al. 2010).

In this chapter we review over 10 years of BCI-VR research. Our focus will be on human-computer interaction paradigms, and our main goal is to highlight both the constraints and the opportunities of BCI and VR combined. Consequently, the chapter will be divided into four themes: (i) navigation, (ii) controlling a virtual body, (iii) controlling the world directly, and (iv) paradigms beyond direct control.

Navigation: Controlling the Viewpoint

Typically, our brain controls our body in an action-perception loop: the brain sends commands to the muscles for generating motor movement, and sensory information provides feedback to the brain regarding the resulting body motion and its effects on the environment. A natural BCI paradigm would therefore aim at substituting the physical body with a virtual body. Such substitution can take place in two ways. The first is by allowing the participant to perform navigation – implicitly controlling the viewpoint; this can be considered a limited form of first-person view. The second is by providing the VR participant with an explicit control over a virtual body – an avatar.

A typical BCI navigation experiment follows three steps: (i) training, (ii) cue-based BCI, and (iii) free choice navigation task. The training stage is used to establish a first model of the user's brain activity: the user is provided with a set of discrete instructions, such as a series of left, right, and forward commands, and no feedback is provided. Cue-based BCI is typically similar, but since a model is already available, feedback is provided about what the system "thinks" that the subject is "thinking," after each trigger. Typically, several sessions of cue-based BCI take place for further training of both the user and the classifier model. Eventually, the goal is to let the users perform a task with free choice, and the subject performs a navigation task. Here, we distinguish between real and fake free choice; in BCI we often prefer fake free choice – we instruct the user to perform specific actions throughout the session – in order to evaluate the BCI performance.

EEG-based BCI suffers from several limitations and constraints as a user input device. Although this varies among the different BCI paradigms, mostly, (i) it is often not 100 % accurate, (ii) it has a long delay, (iii) it has a low information rate, (iv) it requires extensive training, (v) some users cannot perform BCI despite training, (vi) it is difficult to recognize the non-control state, and (vii) it is often synchronous, i.e., the initiation of action and timing are driven by the software.

Most studies to date in BCI-VR used BCI for navigation. The first ever BCI navigation experiment tested whether it can be used in a flight simulator (Nelson et al. 1997). Subjects were trained to control a plane on a single axis in a wide field of view dome display, using a combination of EEG and electrical signals from the muscles – electromyogram (EMG).

In the years 2004–2006, I was fortunate to take part in a set of BCI navigation studies in immersive VR (Friedman et al. 2007a; Leeb et al. 2006; Pfurtscheller et al. 2006). We have integrated the Graz BCI, based on motor imagery, with the VR cave automatic virtual environment (CAVE)-like system (Neira et al. 1992) in UCL, London. We have explored several scenarios. For example, one study included a social scenario whereby the subject sits in a virtual bar room, various virtual characters talk to the subject, and he or she has to rotate left or right to face the character speaking. Rotation was achieved by left- and right-hand imagery, and as a result the virtual bar was rotated. The reason we have eventually focused on a navigation task is that it seemed to provide the best motivation – subjects were competitive and wanted to reach down the virtual street further each time.

Three subjects, already trained with the Graz BCI, performed BCI tasks with three different setups: (i) abstract feedback, (ii) head-mounted display (HMD), and (iii) the CAVE-like system, over a duration of 5 months. In order to assess the impact of the interface on BCI performance, the subjects all went through the order – abstract feedback, HMD, CAVE, HMD, abstract feedback. In order to be able to determine BCI performance, the navigation experiment was trigger based (this is what we referred to as "fake free choice"): the subjects received one of two cues, "walk" or "stop," and had to respond by feet or right-hand imagery, correspondingly. If the cue was "walk" and they correctly activated feet imagery, they moved forward; otherwise, if they activated hand imagery, they stayed in place. If the cue was "stop" and they correctly activated hand imagery, they stayed in place, and if they incorrectly activated feet imagery, they moved backward. Thus, the distance in the virtual street served as a measure of BCI performance (https://www.youtube.com/watch?v=QjAwmSnHC1Q). This study did not find any consistent performance trend related to the type of interface (abstract, HMD, or CAVE), but the event-related synchronization (ERS) was most pronounced in the CAVE (Pfurtscheller et al. 2006).

Self-paced, asynchronous BCI is more difficult, since the system needs to recognize the non-control (NC) state. Leeb et al. first attempted experimenter-cued asynchronous BCI, i.e., the subject was cued when to rest (move into NC state) and when to move (Leeb et al. 2007c). Five participants navigated in a highly immersive setup in a model of the Austrian National Library, using binary classification: one motor imagery class was selected as the most accurate one in training – left hand, right hand, or feet – and this was compared with NC or no activation. The results indicate a very low false-positive rate of 1.8–7.1 %, but the true-positive rate was also low: 14.3–50 %. The authors suggest that the main challenge in this specific study was that keeping imagery for long durations is very difficult for subjects.

Self-paced BCI navigation based on motor imagery was demonstrated for controlling a virtual apartment (Leeb et al. 2007b). Although successful, we also provide details of the limitations of this study, in order to highlight the limitations of BCI, referred to above. After training, subjects performed a free choice binary navigation (left hand vs. right hand). Walking was along predefined trajectories, subjects had to reach specific targets, but the left/right decisions were made freely. Motor imagery recognition was based on offline processing of a training session, taking the duration between 1.5 s and 4.5 s after the trigger. Separating motor imagery from the NC state in real time was done as follows: classification took place at the sample rate, 250 Hz, and only a unanimous classification over a period of 2 s resulted in an action. This study allowed estimating the delay required to classify motor imagery – between 2.06 s and 20.54 s with a mean of 2.88 s and standard deviation (SD) of 0.52 s. The delay was slightly shorter than in cue-based BCI – 3.14 s. Performance in VR was better than cue-based BCI with abstract feedback, and there were no significant differences between a desktop-based virtual environment and an immersive virtual environment (a "power wall" setup) in BCI performance. Despite extensive training, two out of nine subjects were not able to perform the task, and for the rest, mean error was between 7 % and 33 %.

In Leeb et al. (2007a), we showed that a tetraplegic patient could also navigate immersive VR, in the UCL CAVE-like system, in a self-paced study. The subject was trained over 4 months with the Graz BCI until he reached high performance with one class – activating 17 Hz imagining feet movement. Classification was achieved with a simple threshold on the bandpower of a single EEG channel near Cz for determining "go" or NC. Since the subject's control was very good, there was no dwell time (minimum time over threshold to activate motion) or refractory period (minimum time between two activations). The virtual environment included moving along a straight line and meeting virtual female characters on the way (https://www.youtube.com/watch?v=cu7ouYww1RA). The subject performed 10 runs with 15 avatars each and was able to stop in front of 90 % of the avatars. The average duration of motor imagery periods was 1.58 s + − 1.07 s, the maximum 5.24 s, and the minimum 1.44 s.

In a post-experimental interview, the subject indicated that the VR experience was significantly different than his previous BCI training: "It has never happened before, in the sense of success and interaction. I thought that I was on the street and I had the chance to walk up to the people. I just imagined the movement and walked up to them. However, I had the sensation that they were just speaking but not talking to me..." He said that he had the feeling of being in that street and forgot that he was in the lab and people were around him. "Of course the image on the CAVE wall didn't look like you or me, but it still felt as if I was moving in a real street, not realistic, but real. I checked the people (avatars). We had 14 ladies and 1 man" (actually, there were 15 female avatars).

Scherer et al. demonstrated a self-paced four-class motor imagery BCI for navigating a virtual environment (Scherer et al. 2008). They combine two classifiers: one "typical," separating among left-hand, right-hand, and feet imagery, and another to detect motor imagery-related activity in the ongoing EEG. They selected the three top subjects out of eight who performed training, and after three training sessions, they were able to perform cue-based two-class BCI with 71 %, 83 %, and 86 %. The second classifier used two thresholds – one for switching from intentional control (IC) to non-control (NC) and another to switch from NC to IC. The thresholds were applied to the LDA classifier's output vectors. The task was to navigate a virtual environment and reach three targets, including obstacle avoidance. The second classifier, separating NC and IC, resulted in performance of 80 %, 75 %, and 60 %. The mean true-positive (TP) rates for 8 s action period were 25.1 % or 28.4 %. Adapting the thresholds can yield a higher TP rate but at the cost of more false-positives (FPs). Again, we see that keeping motor imagery for long durations is difficult for subjects.

Given the limitation of motor imagery for BCI, Lotte et al. suggested an improvement in the control technique (Lotte et al. 2010): the navigation commands were sorted in a binary tree, which the subjects had to traverse using self-paced motor imagery – left and right to select from the tree and feet for "undo." One branch of the tree allowed selection of points of interest, which were automatically generated based on the subject's location in the VE. Using this interface, users were able to navigate a large environment and were twice faster than when using low-level, "traditional" BCI.

Most BCI-VR navigation studies are aimed at improving the navigation performance. Only a few studies investigate scientific issues around this fascinating setup. In one such example, we compared free choice with trigger-based BCI in the CAVE (Friedman et al. 2010). Ten subjects were split into two conditions: both used left-hand and right-hand imagery to navigate in a VR, but one condition was instructed at each point in time what "to think" and the other condition was not. The subjects in the control condition, which was cue-based, performed significantly better. Post-experimental interviews may have revealed the reason – the subjects were used to being conditioned by the trigger-based training. This highlights the fact that BCI training under strict conditions, while necessary to achieve a good classifier model, might result in mistraining with respect to the target task, which is typically un-triggered.

Larrue et al. compared the effect of VR and BCI on spatial learning (Larrue et al. 2012). Twenty subjects navigated a real city, 20 subjects navigated a VR model of the city using a treadmill with rotation, and eight subjects navigated the same model using BCI. Surprisingly, spatial learning was similar in all conditions. More studies of this type are needed if we want to understand how BCI interacts with cognitive tasks; for example, one limitation of this study is that the BCI required much more time than in the other conditions.

Controlling a Virtual Avatar

VR navigation is equivalent to controlling the virtual camera. This is equivalent to the trajectory of the viewpoint from your eyes when you walk or drive in the physical world. In the physical world, however, you also have a body. In video games, controlling the camera directly is often referred to as “first-person view,” but this is misleading. If you look at yourself now, you will (hopefully) not only see the world around you but also see a body (albeit without a head, unless you are looking at the mirror). The sensation of our own body is so natural that we often forget it, but body ownership has been shown to be highly important for the illusions induced by VR (Maselli and Slater 2013). In this section we focus on studies whereby the visual feedback for the BCI involves a virtual body. Such an experience can be regarded as a radical form of re-embodiment; it is as if the system disconnects your brain from your original body and reconnects your brain to control a virtual body.

Lalor et al. (2005) demonstrated SSVEP control of a virtual character in a simple video game: the subjects had to keep the balance of a tightrope walking character with two checkerboard SSVEP targets. Whenever the tightrope loses balance, a 3 s animation is played, and the subject has to attend to the correct checkerboard to shift the walker to the other side. Thus, the game consists of multiple mini-trials, in controlling two SSVEP targets, with a video game context instead of abstract feedback.

Lalor et al.’s study was a first step, but it did not attempt to provide the subjects with a sense of body ownership, and it was based on arbitrary mapping: gazing at a checkerboard to shift the balance of the character. We have performed a study aimed at checking ownership of a virtual body using motor imagery BCI (Friedman et al. 2007b, 2010). Since this study took place in a CAVE-like system, we opted for third-person embodiment: the subjects sat down on a chair in the middle of the CAVE room and saw a gender-matched avatar standing in front of them, with their back toward the subjects. In one condition the subjects used feet imagery to make the avatar walk forward and right-hand imagery to make the avatar wave its arm, and in the other condition, the control was reversed: hand imagery caused walking and feet imagery caused arm waving. After several training sessions with abstract feedback, three subjects performed the task in eight sessions – four normal and four reversed, in interleaved order. We expected the more intuitive mapping to result in better BCI performance, but the results were not conclusive – one of the subjects did even better in the reverse condition; more studies, with a larger number of subjects, are required to establish the effect of intuitive vs. nonintuitive mapping between imagery and body motion. During the experiment, we have deliberately avoided setting any expectations in the subject regarding body ownership – e.g., in our instructions, we referred to “feet” rather than to “the avatar’s feet” or “your avatar’s feet.” Anecdotally, we have witnessed that one of the subjects, as the experiment progressed, started referring to her avatar as “I” instead of “she.”

A more systematic experiment was carried out by Perez-Marcos et al., intended to induce a virtual hand ownership illusion with BCI (Slater et al. 2009). In the rubber hand illusion (Botvinick and Cohen 1998), tactile stimulation of a person’s hidden real hand in synchrony with touching a substitute rubber hand placed in a plausible position results in an illusion of ownership of the rubber hand. This illusion was reconstructed in virtual reality (Slater et al. 2008), and even a full body illusion was achieved (Ehrsson 2007; Marcos et al. 2009). In the BCI version of this setup, 16 participants went through left-hand vs. right-hand imagery BCI training without receiving any feedback. In the VR stage subjects had their real arm out of view in a hollow box while wearing stereo goggles in front of a "power wall." The subjects saw a virtual arm and used left-hand imagery to open its fingers and right-foot imagery to close the fingers into a fist. Eight subjects experienced a condition whereby motor imagery was correlated to the virtual hand movement, and eight subjects went through a control condition, in which the virtual hand motion was uncorrelated with the motor imagery. The strength of the virtual arm ownership illusion was estimated from questionnaires, electromiogram (EMG) activity, and proprioceptive drift, and the conclusion was that BCI motor imagery was sufficient to generate a virtual arm illusion; this is instead of the "classic" method for inducing the illusion, which is based on synchronous stimulation of the real and virtual arm.

Evans et al. showed that reduced BCI accuracy, resulting in a lower sensory feedback, results in a decrease in the reported sense of body ownership of the virtual body (Evans et al. 2015). Their results also suggest that bodily and BCI actions rely on common neural mechanisms of sensorimotor integration for agency judgments, but that visual feedback dominates the sense of agency, even if it is erroneous.

The combination of VR, BCI, and body ownership is a promising avenue toward stroke rehabilitation. While BCI and rehabilitation are an active area of research (Huggins and Wolpaw 2014), we are only aware of one study attempting to combine these necessary ingredients (Bermúdez et al. 2013). The authors describe a non-immersive desktop-based setup, which includes a first-person view with only virtual arms visible. They compared among several conditions: passive observation of virtual hand movement, motor activity, motor imagery, and simultaneous motor activity and imagery. The BCI phase included three conditions: left arm stretching, right arm stretching, and none. Unfortunately, the subjects were asked to imagine the avatar moving its hands, rather than imagine moving their own hand, which rules out virtual body ownership. In addition, BCI performance results are not reported. We support the authors' assumption that the combination of motor imagery and movement is likely to recruit more task-related brain networks than in the rest of the conditions, making such a setup promising for rehabilitation.

More recently, we have performed several studies using a BCI based on functional magnetic resonance imaging (fMRI) to control avatars. FMRI is expensive, is much less accessible than EEG, and suffers from an inherent delay and low temporal resolution, since it is based on blood oxygen levels rather than directly on electrical brain activity. Nevertheless, fMRI, unlike EEG, has a high spatial resolution: in our typical study using a 3 T fMRI scanner, we perform a whole brain scan every 2 s, and each scan includes approximately 30,000 informative voxels. Our studies aim to show that despite its sluggish signal, fMRI can be used for BCI. We suggest that this method would be extremely useful in BCI for paralyzed patients; due to the limitations of noninvasive BCIs (based on EEG or functional near-infrared spectroscopy – fNIRS), there is a growing effort to opt for invasive BCIs (Hochberg et al. 2012). We suggest that prior to surgery, fMRI-BCI can be used for identifying new mental strategies for BCI, localizing brain areas for implants, and training subjects.

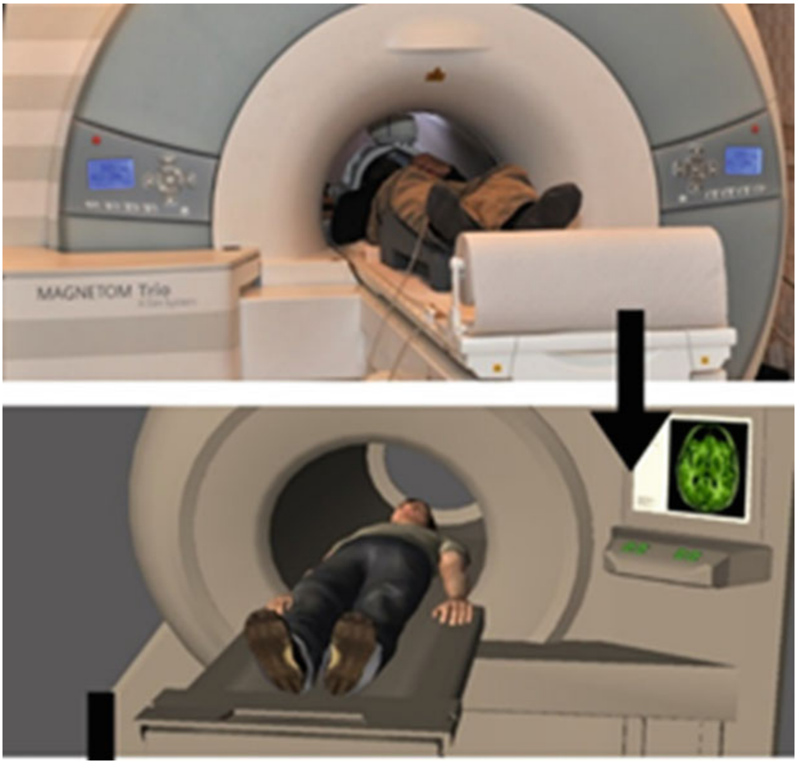


Fig. 1 The subject lying down in the fMRI scanner (*top*) sees an avatar lying down in a virtual fMRI scanner (*bottom*) and controls it using motor movement or imagery

In our studies we have allowed subjects to control a virtual body from a third-person perspective (Cohen et al. 2014b) (https://www.youtube.com/watch?v=rHF7gYD3wI8), as well as a robot from first-person perspective (Cohen et al. 2012) (https://www.youtube.com/watch?v=pFzfHnzjdo4). In our experiments the subject, lying down in the fMRI scanner, sees an image projected on a screen (e.g., Fig. 1). We do not use stereo projection, but since the screen covers most of the field of view, the experience is visually immersive. Our subjects were able to perform various navigation tasks, including walking a very long footpath in the jungle (Video: https://www.youtube.com/watch?v=PeujbA6p3mU). Our first version was based on the experimenter locating regions of interest (ROIs) corresponding to left-hand, right-hand, and feet imagery or movement and a simple threshold-based classification scheme (Cohen et al. 2014b). Recently, we have completed an improved version of fMRI-based BCI, based on machine learning, using information gain (Quinlan 1986) for feature (voxel) selection and a support vector machine (SVM) classifier (Cohen et al. 2014a). This allowed us to test more complex navigation tasks and shorten the delay; we show that subjects can control a

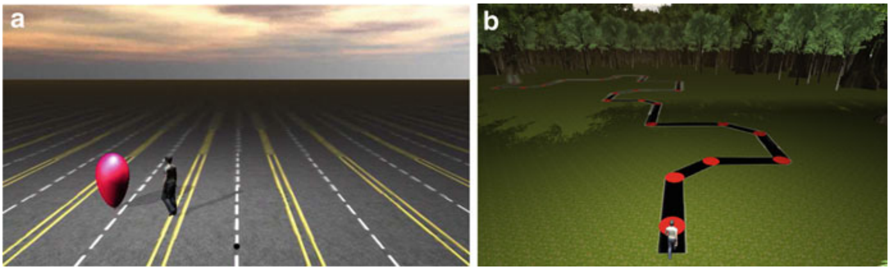


Fig. 2 Snapshots from the fMRI navigation studies: the subjects had to navigate toward a balloon (**a**) or along a trail (**b**) (Journal of Neuroengineering (IOPScience))

four-class (left hand, right hand, feet, or NC state) BCI with a 2 s delay with very high accuracy.

In addition to proving that fMRI-BCI is possible, these studies provided new insights on motor imagery-based BCI. A few anecdotal results came from repeated administration of body ownership questionnaires to the subjects after each experimental session. In one study in which the subjects had to navigate toward a balloon (Fig. 2a) (https://www.youtube.com/watch?v=l1yMd_UFp5s), questionnaires revealed that sense of body ownership over the avatar was significantly higher when using motor imagery as compared to using motor execution for BCI. In another study in which the subjects had to navigate along a footpath (Fig. 2b), subjects seemed to be significantly more confused about their body ownership when the delay was reduced to 2 s; this difference was nearly significant for the question, "I was aware of a contradiction between my virtual and real body," and significant for the question, "It felt like I had more than one body."

Due to fMRI's superior spatial resolution over EEG, it can highlight the differences between motor execution and motor imagery. Figure 3 compares voxels captured by information gain against voxels captured by a general linear model (GLM) analysis, which is typically used in fMRI studies to obtain brain activation patterns. Since each method captures voxels differently, with different thresholds, the figures cannot be directly compared; however, inspection suggests pre-motor cortex activation in motor imagery whereas motor execution was mostly based on the specific body representations in primary motor cortex. In addition, the differential activations were much stronger using motor execution as compared to motor imagery. Figure 4 shows classification results over time comparing motor execution and imagery, showing that using imagery classification accuracy drops faster than it does when using motor execution. The results are based on tenfold cross validation of 150 cues, 50 from each class: left hand, right hand, and feet.

Taken together, these findings suggest that people find it hard to activate motor imagery and especially to keep it active for long durations. Our evidence from fMRI-based BCI thus corresponds to similar evidence obtained in EEG-based BCI. This indicates that these challenges in activating motor imagery are most likely not the result of the limitations of the specific recorded signals but an inherent difficulty

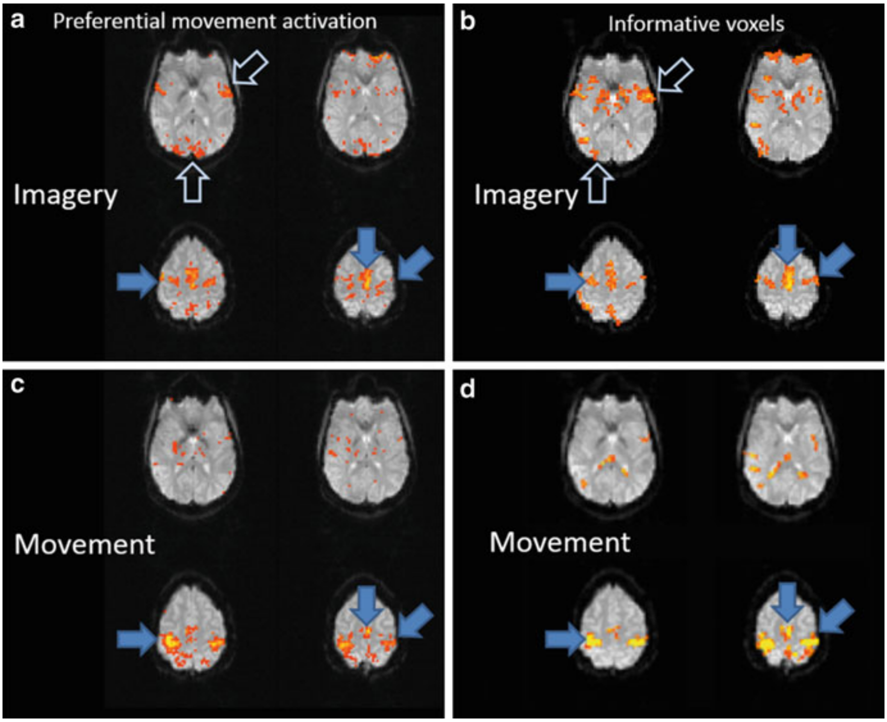


Fig. 3 A subset of corresponding slices from S1. The *left* column shows the GLM contrast (right, left, forward) > baseline (thresholds: t = 4.6 for MM and t = 3.2 for MI), and the *right* column shows the 1024 voxels with highest information gain selected by our algorithm. The *top* row shows imagery and the *bottom* row shows motor movement

in motor imagery. In another study using real-time fMRI, we suggest that there are significant differences in the ways different brain areas lend themselves to internal control (Harmelech et al. 2015); this was demonstrated in the context of neurofeedback, but should equally apply to BCI. Using fMRI, we may be able to extend the repertoire of BCI interaction paradigms and to find the paradigms that are easiest for subjects.

Controlling the World Directly

In the previous sections, we discussed navigation and virtual reembodiment – using BCI to control a virtual body or its position – these interaction paradigms are based on how we interact with the physical world. But in VR we can go beyond – why not control the world directly?

As an example of a practical approach, consider using a P300 BCI matrix to control a room in VR (Edlinger et al. 2009). This is a simulation of the scenario whereby a paralyzed patient can control a smart home. Such as setup can allow

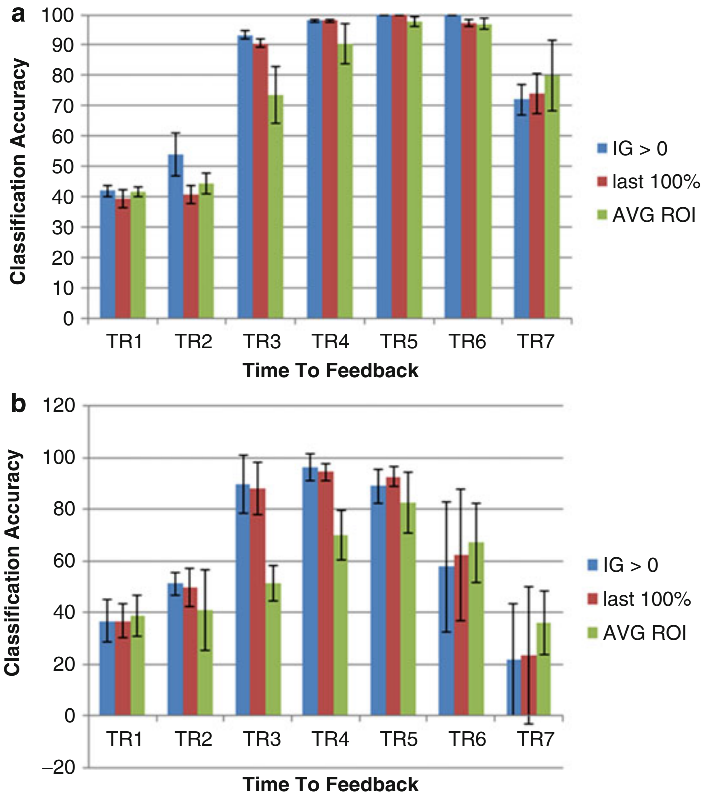


Fig. 4 A comparison of (**a**) motor execution (MM) and (**b**) motor imagery (MI) classification accuracy across six (MM) and three (MI) subjects, between machine learning and ROI-based classification. The TRs have a 2 s duration. Error bars indicate the 95 % confidence interval. The machine learning results were obtained by using either all voxels with information gain above 0 or the smallest number of voxels that permit perfect classification of all training examples. Every repetition time (TR) is 2 s

people to rapidly select the specific command out of many different choices. The study suggests that more than 80 % of the healthy population could use such a BCI within only 5 min of training. In a further study this approach was improved using a hybrid approach: SSVEP was used to toggle the P300 BCI on and off, in order to avoid false-positive classifications (Edlinger et al. 2011).

Using this approach, the P300 matrix serves as a BCI remote control. While this is a practical approach, it goes against VR philosophy. Even the best BCI requires several seconds of attention to the P300 matrix for each selection, which is outside the VR display. This greatly reduces the sense of being present in the VR, as demonstrated in another study by the same authors, after they noted that the subjects reported a very low sense of presence (Heeter 1992; Lombard and Ditton 1997; Sanchez-Vives and Slater 2005; Slater 1993; Witmer and Singer 1998) in post-experiment questionnaires. In this follow-up study (Groenegress et al. 2010), post-experiment questionnaires revealed that subjects reported a significantly higher sense of presence in a gaze-based interface as compared with the P300 interface, for controlling the same virtual apartment in the same VR setup.

In-place Control

Given the limitations arising of having the P300 or SSVEP targets outside the VR, several attempts were made to embed the target visual stimuli more naturally into the VR scene. Imagine what it would be like if you could just focus on an object around you and thereby activate it. In fact, one of the first ever BCI-VR studies used this approach by turning the traffic lights in a driving simulation into P300 targets (Bayliss and Ballard 2000; Bayliss 2003). The setup included a modified go-cart and an HMD. Red stoplight was used as the P300 oddball task: most lights were yellow, and the subject was instructed to ignore green and yellow lights and detect red light, which were less frequent.

Donnerer and Steed (Donnerer and Steed 2010) embedded P300 in a highly immersive CAVE-like system and compared three paradigms: (i) spheres arranged in an array, (ii) different objects cluttered around the virtual room, and (iii) tiles – different areas of the virtual world can be selected, instead of objects. Each sphere, object, or tile flashed separately in order to enable its selection by the subject's P300 response, after eight flashes (16 in the training phase). The setup was successful but results do not show very high accuracy. In addition, the interaction is relatively slow, since sequential flashing of the stimuli is required, as opposed to SSVEP.

Faller et al. have developed such a system using SSVEP, in order to control VR and even augmented reality (Faller and Leeb 2010; Faller et al. 2010). They have achieved high classification results using just two occipital electrodes – O1 and O2. They demonstrate three applications, but in all of them, the BCI is used for navigation rather than for controlling the world. They report an average number of true-positive (TP) events of 8.5, 7.1, and 6.5 per minute.

In a similar study Legeny et al. also demonstrated BCI navigation with embedded SSVEP targets (Legény et al. 2011). They have attempted a more natural embedding, which they call mimesis: rather than controlling buttons or arrows, the SSVEP cues were embedded inside the wings of butterflies. Three butterflies kept hovering around the middle of the screen and were used for navigating forward, left, or right. The wings changed color for SSVEP stimulation and also flapped their wings; the latter did not interfere with SSVEP classification. Feedback about the level of BCI confidence toward one of the classes (distance from separating the hyperplane used by LDA classifier) was also provided in the appearance of the butterflies' antennas. Since the BCI was self-paced, such feedback is useful, especially when none of the classes are activated. The study was carried out in a 2×2 design: overlay/mimesis and feedback/no feedback. Their results indicate that overlay was significantly faster than mimesis, mimesis resulted in higher sense of presence, and feedback had no effect on the sense of presence. The mimesis interaction increased subjective preference and sense of presence, but reduced performance in terms of speed, as compared with a more "standard" SSVEP overlay interface.

The studies by Faller et al. and Legeny et al. used in-place SSVEP, but only for navigation. In my lab we have also developed such in-place SSVEP, but our interaction approach is different – we are interested in using BCI to activate arbitrary objects in the virtual world, as a form of virtual psychokinesis. We have developed a generic system that allows easily turning any object in a 3D scene in the Unity game engine into an SSVEP target. A Unity script is attached to the object, which makes it flicker at a given frequency. Another script connects to the BCI system using user datagram protocol (UDP), assigns different frequencies to different objects, and activates objects in real time based on classifier input. We have shown that this software implementation of SSVEP allows for very high classification rates and robust BCI control.

Given the novel aspect of this interface, we have decided to allow participants to experience a "psychokinesis"-like experience, without telling them that they have such "powers." We have conducted an experiment in which subjects controlled a brain-computer interface (BCI) without being aware that their brain waves were responsible for events in the scenario. Ten subjects went through a stage of model training in steady-state visually evoked potential (SSVEP)-based BCI, followed by three trials of an immersive experience where stars moved as a response to SSVEP classification. Only then the subjects were explained that they were using a BCI, and this was followed by an additional trial of immersive free choice BCI and a final validation stage. Three out of the ten subjects realized that they controlled the interface, and these subjects had better accuracy than the rest of the subjects and reported a higher sense of agency in a post-study questionnaire (Giron and Friedman 2014).

Furthermore, our study shows that subjects can implicitly learn to use a SSVEP-based BCI (Giron et al. 2014). The SSVEP stimuli were presented in a pseudorandom order in an immersive star field virtual environment, and the participants' attention to the stimuli resulted in stars moving within the immersive space (Fig. 5). Participants were asked to view four short clips of the scene and try to explain why the stars were moving, without being told that they are controlling a BCI. Two groups were tested: one that interacted implicitly with the interface and a control group in which the interaction was a sham (i.e., the interface was activated independently of the participants' attention, with the same response frequency). Following the exposure to the immersive scene, the participants' BCI accuracy was tested, and the experiment group showed higher accuracy results. This finding may indicate that implicit SSVEP BCI interactions are sufficient in inducing a learning effect for the skill of operating a BCI.

Hybrid Control

Due to its limitations, a promising direction for BCI is to be used as an additional input channel complementing other interaction devices, rather than replacing them.

Fig. 5 The star field experience, responding to SSVEP-based BCI unbeknown to subjects (From Giron and Friedman 2014)

This is true for able-bodied users – BCI cannot compete with keyboard, mouse, or similar devices in terms of information rate and accuracy. A similar case can be made for paralyzed patients: BCI does not need to compete with other assistive technologies, but can be part of a basket of solutions, such that patients can leverage whatever muscle control works best for them, in parallel to using the brain waves as an input signal.

Leeb et al. demonstrated a hybrid BCI for skiing in a CAVE-like system: steering with a game controller and jumping (to collect virtual fish targets) with a feet motor imagery BCI (Leeb et al. 2013). The joystick controller did not deteriorate BCI performance. The BCI was continuous, based on crossing a threshold for 0.5–1.5 s. The threshold was defined for each subject as the mean plus one standard deviation of the classifier output during the time of the fixation cross, and the dwell time was selected as half of the time over this threshold during the imagery period. The detected events were transferred into control commands for the feedback. After every event, a refractory period of 4 s was applied during which event detection was disabled. The study compared using a push button (94–97 % success) with BCI (45–48 % success).

Another form of hybrid BCI involves the combination of two or more BCI paradigms simultaneously. For example, Su et al. used two-class motor imagery for navigation of a virtual environment and P300 over five targets for controlling a device (Su et al. 2011). The control was toggled between P300 and motor imagery rather than simultaneous, and the toggle was automatically activated based on the subject's location inside the virtual environment: the subject used motor imagery to navigate a virtual apartment and the P300 to control a virtual TV set. Subjects reported that hybrid control was more difficult than standard BCI, but showed no drop in performance.

Beyond Control

So far, we have discussed BCI for direct control of VR, but BCI technologies also allow to be used for other closed-loop interaction paradigms. For example, aspects of the user's cognitive and emotional state can be computed online, and the application can be adapted accordingly. Applications that are based on automatic recognition of emotions have been studied extensively in the field of affective computing (Picard 1997). A more recent term is passive BCIs, referring to applications that respond to online cognitive monitoring (Zander and Kothe 2011). Despite the great promise of this field, there is very little work, and almost none involving VR.

One question is how to extract emotional and cognitive state from brain signals; this is a major challenge that is still open (Berka et al. 2004; Liu et al. 2011). The other challenge is how to adapt the application to the feedback; in the context of VR, this opens up opportunities for new types of experiences. In one such creative example, affective mood extracted from online EEG was coupled to the avatar in the massive multiuser game World of Warcraft (Plass-Oude Bos et al. 2010). The parietal power of the alpha band was mapped to shape shifting between animal forms in the fantasy world: e.g., increase in parietal alpha is related to relaxed readiness and thus was mapped in the game world to transforming to an elf. The authors do not validate or evaluate the brain activity or the accuracy of the BCI but provide some useful lessons regarding interaction – for example, they use hysteresis and some dwell time in order to avoid shape-shifting too frequently.

Finally, Gilroy et al. suggest a new interaction technique incorporating empathy derived from brain signals which drives interactive narrative generation (Gilroy et al. 2013). Subjects used EEG neurofeedback, based on frontal alpha asymmetry (Coan and Allen 2004; Davidson et al. 1990), to modulate empathic support of a virtual character in a medical drama, and their degree of success affected the unfolding of the narrative. FMRI analysis also showed activations in associated regions of the brain during expression of support. This study demonstrates that there are yet many opportunities for integrating real-time information from brain activity into virtual environments and VR. While some progress can be made with peripheral physiological signals, such as heart rate and its derivatives, electrodermal activity (EDA, "sweat response"), or EMG (indicating muscle activity), the information from the central nervous system is expected to contain more information.

Conclusion and Future Directions

BCI still faces many challenges, but it has matured, especially over the last decade. There is now growing interest in getting BCI out of the laboratory and into real-world applications. For paralyzed patients the goal is restoring basic communications and control abilities. For able-bodied participants, it seems that the greatest potential is in hybrid BCI and passive BCI. In all cases VR is a natural partner for BCI.

Due to the limitations of EEG, there is an effort in exploiting other brain signals. For medical applications, methods such as fMRI and electrocorticogram (ECoG) hold much promise for moving BCI forward. For other applications the devices need to be low cost and noninvasive. FNIRS may allow for novel BCI paradigms, instead or in addition to EEG. Furthermore, we see potential in combining brain signals with other signals, such as from the autonomous nervous system – heart rate and its derivatives, electrodermal activity, and respiration – as well as eye tracking. It remains to be seen whether the value of these joint signals would be greater than their sum and if so how this value can be translated into new interaction paradigms and applications.

The combination of VR and BCI offers radically new experiences. Since both of these fields are young, especially BCI, we have only scratched the surface, and we have barely begun to study the resulting psychological impact and user experience. Each breakthrough in BCI would allow us to provide VR participants with novel experiences.

Recommended Reading

H. Alkadhi, P. Brugger, S.H. Boendermaker, G. Crelier, A. Curt, M.C. Hepp-Reymond, S.S. Kollias, What disconnection tells about motor imagery: evidence from paraplegic patients. Cereb. Cortex **15**, 131–140 (2005). doi:10.1093/cercor/bhh116

J.D. Bayliss, Use of the evoked potential P3 component for control in a virtual apartment. IEEE Trans. Rehabil. Eng. **11**(2), 113–116 (2003)

J.D. Bayliss, D.H. Ballard, A virtual reality testbed for brain computer interface research. IEEE Trans. Rehabil. Eng. **8**(2), 188–190 (2000)

C. Berka, D.J. Levendowski, M.M. Cvetinovic, M.M. Petrovic, G. Davis, M.N. Lumicao, ... R. Olmstead, Real-time analysis of EEG indexes of alertness, cognition, and memory acquired with a wireless EEG headset. Int. J. Hum. Comput. Interact. (2004) doi:10.1207/s15327590ijhc1702_3

I. Bermúdez, S. Badia, A. García Morgade, H. Samaha, P.F.M.J. Verschure, Using a hybrid brain computer interface and virtual reality system to monitor and promote cortical reorganization through motor activity and motor imagery training. IEEE Trans. Neural Syst. Rehabil. Eng. **21**, 174–181 (2013). doi:10.1109/TNSRE.2012.2229295

M. Botvinick, J. Cohen, Rubber hands "feel" touch that eyes see. Nature **391**(6669), 756 (1998)

P. Brunner, S. Joshi, S. Briskin, J.R. Wolpaw, H. Bischof, G. Schalk, Does the "P300" speller depend on eye gaze? J. Neural Eng. **7**, 056013 (2010). doi:10.1088/1741-2560/7/5/056013

M. Cheng, X. Gao, S. Gao, D. Xu, Design and implementation of a brain-computer interface with high transfer rates. IEEE Trans. Biomed. Eng. **49**, 1181–1186 (2002). doi:10.1109/TBME.2002.803536

J.A. Coan, J.J.B. Allen, Frontal EEG asymmetry as a moderator and mediator of emotion. Biol. Psychol. (2004). doi:10.1016/j.biopsycho.2004.03.002

O. Cohen, S. Druon, S. Lengagne, A. Mendelsohn, A. Kheddar, R. Malach, D. Friedman, fMRI-based robotic embodiment: a pilot study, in *IEEE International Conference on Biomedical Robotics and Biomechatronics* (Rome, 2012), pp. 314–319

O. Cohen, M. Koppel, R. Malach, D. Friedman, A generic machine learning tool for whole brain classification from fMRI. in *6th International BCI Conference,* Graz, Austria, (2014a)

O. Cohen, M. Koppel, R. Malach, D. Friedman, Controlling an avatar by thought using real-time fMRI. J. Neural Eng. **11**(3), 35006 (2014b)

R.J. Davidson, P. Ekman, C.D. Saron, J.A. Senulis, W.V. Friesen, Approach-withdrawal and cerebral asymmetry: emotional expression and brain physiology. I. J. Pers. Soc. Psychol. **58**, 330–341 (1990). doi:10.1037/0022-3514.58.2.330

C. Davies, Virtual spaces, in *Space: In Science, Art, and Society*, ed. by F. Penz, G. Radick, R. Howell (Cambridge University Press, Cambridge, UK, 2004), pp. 69–104

C. Davies, J. Harrison, Osmose: towards broadening the aesthetics of virtual reality. ACM Comput. Graph. [special Issue on Virtual Reality] **30**(4), 25–28 (1996)

E. Donchin, K.M. Spencer, R. Wijesinghe, The mental prosthesis: assessing the speed of a P300-based brain- computer interface. IEEE Trans. Rehabil. Eng. **8**, 174–179 (2000). doi:10.1109/86.847808

M. Donnerer, A. Steed, Using a P300 brain–computer interface in an immersive virtual environment. Presence Teleop. Virt. Environ. **19**(1), 12–24 (2010). doi:10.1162/pres.19.1.12

G. Edlinger, C. Holzner, C. Groenegress, C. Guger, M. Slater, Goal-oriented control with brain-computer interface, in *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. LNAI, vol 5638 (2009), pp. 732–740. doi:10.1007/978-3-642-02812-0_83

G. Edlinger, C. Holzner, C. Guger, A hybrid brain-computer interface for smart home control, in *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. LNCS, vol 6762 (2011), pp. 417–426. doi:10.1007/978-3-642-21605-3_46

H.H. Ehrsson, The experimental induction of out-of-body experiences. Science **317**(5841), 1048 (2007)

N. Evans, S. Gale, A. Schurger, O. Blanke, Visual feedback dominates the sense of agency for brain-machine actions. PLoS One **10**(6), e0130019 (2015). doi:10.1371/journal.pone.0130019

J. Faller, R. Leeb, Avatar navigation in virtual and augmented reality environments using an SSVEP BCI ICABB-2010. . . . *and Virtual Reality*. (2010) Retrieved from http://brainable.org/Documents/FallerJ-ICABB.pdf

J. Faller, G. Müller-Putz, D. Schmalstieg, G. Pfurtscheller, An application framework for controlling an avatar in a desktop-based virtual environment via a software SSVEP brain–computer interface. Presence Teleop. Virt. Environ. (2010). doi:10.1162/pres.19.1.25

D. Friedman, R. Leeb, A. Antley, M. Garau, C. Guger, C. Keinrath, . . . M. Slater, Navigating virtual reality by thought: what is it like? Presence Teleop. Virt. Environ. **16**(1), 100–110 (2007a)

D. Friedman, R. Leeb, L. Dikovsky, M. Reiner, G. Pfurtscheller, M. Slater, Controlling a virtual body by thought in a highly-immersive virtual environment. Proc. Graph. Appl. Barcelona, Spain, 83–90 (2007b)

D. Friedman, A. Donenfeld, E. Zafran, Neurophysiology-based art in immersive virtual reality. Int. J. Arts Technol. **2**(4), 331 (2009)

D. Friedman, R. Leeb, G. Pfurtscheller, M. Slater, Human–computer interface issues in controlling virtual reality with brain–computer interface. Hum. Comput. Interact. **25**(1), 67–94 (2010)

S.W. Gilroy, J. Porteous, F. Charles, M. Cavazza, E. Soreq, G. Raz, . . . T. Hendler, A brain-computer interface to a plan-based narrative, 1997–2005. (2013). Retrieved from http://dl.acm.org/citation.cfm?id=2540128.2540415

J. Giron, D. Friedman, Eureka: realizing that an application is responding to your brainwaves, in *Universal Access in Human-Computer Interaction. Design and Development Methods for Universal Access* (2014), Springer International Publishing, Crete, Greece, pp. 495–502

J. Giron, M. Segal, D. Friedman, Implicit learning of SSVEP based brain computer interface, in *The 6th International Brain-Computer Interface Conference*, (Graz, 2014)

C. Groenegress, C. Holzner, C. Guger, M. Slater, Effects of P300-based BCI use on reported presence in a virtual environment. Presence Teleop. Virt. Environ. **19**(1), 1–11 (2010). doi:10.1162/pres.19.1.1

C. Guger, G. Edlinger, W. Harkam, I. Niedermayer, G. Pfurtscheller, How many people are able to operate an EEG-based brain-computer interface (BCI)? IEEE Trans. Neural Syst. Rehabil. Eng. **11**, 145–147 (2003)

T. Harmelech, D. Friedman, R. Malach, Differential magnetic resonance neurofeedback modulations across extrinsic (visual) and intrinsic (default-mode) nodes of the human cortex. J. Neurosci. **35**(6), 2588–2595 (2015)

C. Heeter, Being there: the subjective experience of presence. Presence Teleop. Virt. Environ. **1** (2), 262–271 (1992)

L.R. Hochberg, D. Bacher, B. Jarosiewicz, N.Y. Masse, J.D. Simeral, J. Vogel, ... van der P. Smagt, Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. Nature **485**(7398), 372–375 (2012)

J.E. Huggins, J.R. Wolpaw, Papers from the fifth international brain-computer interface meeting. Preface. J. Neural Eng. **11**, 030301 (2014). doi:10.1088/1741-2560/11/3/030301

I.E.E.S. Sutherland, The ultimate display, in *Proceedings of the {IFIPS} Congress 1965* New York: IFIP, **65**(2), 506–508 (1965)

D.J. Krusienski, M. Grosse-Wentrup, F. Galán, D. Coyle, K.J. Miller, E. Forney, C.W. Anderson, Critical issues in state-of-the-art brain-computer interface signal processing. J. Neural Eng. **8**, 025002 (2011). doi:10.1088/1741-2560/8/2/025002

E. Lalor, S.P. Kelly, C. Finucane, R. Burke, R. Smith, R.B. Reilly, G. McDarby, Steady-state VEP-based brain computer interface control in an immersive 3-D gaming environment. EURASIP JASP **19**, 3156–3164 (2005)

F. Larrue, H. Sauzéon, L. Aguilova, F. Lotte, M. Hachet, B.N. Kaoua, Brain computer interface vs walking interface in VR: the impact of motor activity on spatial transfer, in *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST* (2012), pp. 113–120. doi:10.1145/2407336.2407359

R. Leeb, C. Keinrath, D. Friedman, C. Guger, C. Neuper, M. Garau, ... G. Pfurtscheller, Walking from thoughts: not the muscles are crucial, but the brain waves! Presence Teleop. Virt. Environ. (2006). doi:10.1155/2007/79642

R. Leeb, D. Friedman, G.R. Müller-Putz, R. Scherer, M. Slater, G. Pfurtscheller, Self-paced (asynchronous) BCI control of a wheelchair in virtual environments: a case study with a tetraplegic. Comput. Intell. Neurosci. (2007a)

R. Leeb, F. Lee, C. Keinrath, R. Scherer, H. Bischof, G. Pfurtscheller, Brain-computer communication: motivation, aim, and impact of exploring a virtual apartment. IEEE Trans. Neural Syst. Rehabil. Eng. **15**(4), 473–482 (2007b). doi:10.1109/TNSRE.2007.906956

R. Leeb, V. Settgast, D. Fellner, G. Pfurtscheller, Self-paced exploration of the Austrian National Library through thought. Int. J. Bioelectromagn. **9**, 237–244 (2007c)

R. Leeb, M. Lancelle, V. Kaiser, D. Fellner, G. Pfurtscheller, Thinking penguin: multimodal brain-computer interface control of a VR game. IEEE Trans. Comput. Intell. AI Games **5**, 117–128 (2013). doi:10.1109/TCIAIG.2013.2242072

J. Legény, R.V. Abad, A. Lécuyer, Navigating in virtual worlds using a self-paced SSVEP-based brain–computer interface with integrated stimulation and real-time feedback. Presence Teleop. Virt. Environ. **20**(6), 529–544 (2011). doi:10.1162/PRES_a_00075

Y. Liu, O. Sourina, M.K. Nguyen, Real-time EEG-based emotion recognition and its applications, in *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. LNCS, vol 6670 (2011), pp. 256–277. doi:10.1007/978-3-642-22336-5_13

Y. Liu, X. Jiang, T. Cao, F. Wan, P.U. Mak, P.I. Mak, M.I. Vai, Implementation of SSVEP based BCI with Emotiv EPOC, in *Proceedings of IEEE International Conference on Virtual Environments, Human-Computer Interfaces, and Measurement Systems, VECIMS* (2012), pp 34–37. doi:10.1109/VECIMS.2012.6273184

M. Lombard, T.B. Ditton, At the heart of it all: the concept of presence. J. Comput. Mediat. Commun. **3**(2), 20 pp (1997). doi:10.1093/cid/cir583

F. Lotte, A. van Langhenhove, F. Lamarche, T. Ernest, Y. Renard, B. Arnaldi, A. Lécuyer, Exploring large virtual environments by thoughts using a brain–computer interface based on motor imagery and high-level commands. Presence Teleop. Virt. Environ. **19**(1), 54–70 (2010). doi:10.1162/pres.19.1.54

D.P. Marcos, H. Ehrsson, M.V. Sanchez, Inducing illusory ownership of a virtual body. Front. Neurosci. **3**(2), 214–220 (2009)

A. Maselli, M. Slater, The building blocks of the full body ownership illusion. Front. Hum. Neurosci. **7**, 83 (2013). doi:10.3389/fnhum.2013.00083

C.C. Neira, D.J. Sandin, T.A. DeFanti, R.V. Kenyon, J.C. Hart, The CAVE: audio visual experience automatic virtual environment. Comm. ACM **35**(6), 65–72 (1992).

W. Nelson, L. Hettinger, J. Cunningham, M.R. Nelson, Navigating through virtual flight environments using brain-body-actuated control, in *Proceedings of IEEE Virtual Reality Annual International Symposium* (1997), pp. 30–37

G. Pfurtscheller, R. Leeb, C. Keinrath, D. Friedman, C. Neuper, C. Guger, M. Slater, Walking from thought. Brain Res. **1071**(1), 145–152 (2006)

G. Pfurtscheller, R. Leeb, D. Friedman, M. Slater, Centrally controlled heart rate changes during mental practice in immersive virtual environment: a case study with a tetraplegic. Int. J. Psychophysiol. **68**(1), 1–5 (2008). doi:10.1016/j.ijpsycho.2007.11.003

R.W. Picard, Affective computing. Pattern Recogn. **73**. (1997) doi:10.1007/BF01238028

D. Plass-Oude Bos, B. Reuderink, B. Laar, H. Gürkök, C. Mühl, M. Poel, … D. Heylen, Brain-computer interfacing and games, in *Brain-Computer Interfaces* (2010), pp. 149–178. doi:10.1007/978-1-84996-272-8

H. Putnam, Brain in a vat, in *Reason, Truth, and History* (1982), Cambridge University Press, pp. 5–8

J.R. Quinlan, Induction of decision trees. Mach. Learn. **1**(1), 81–106 (1986)

Y. Renard, F. Lotte, G. Gibert, M. Congedo, E. Maby, V. Delannoy, … A. Lécuyer, OpenViBE: an open-source software platform to design, test, and use brain–computer interfaces in real and virtual environments. Presence Teleop. Virt. Environ. (2010). doi:10.1162/pres.19.1.35

M.V. Sanchez-Vives, M. Slater, From presence to consciousness through virtual reality. Nat. Rev. Neurosci. **6**(4), 332–339 (2005)

R. Scherer, F. Lee, A. Schlogl, R. Leeb, H. Bischof, G. Pfurtscheller, Toward self-paced brain–computer communication: navigation through virtual worlds. IEEE Trans. Biomed. Eng. **55**(2), 675–682 (2008)

M. Slater, Presence in immersive virtual environments, in *Proceedings of the IEEE Virtual Reality 1993* (Seattle, 1993)

M. Slater, Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. Philos. Trans. R. Soc. **364**(1535), 3549–3557 (2009)

M. Slater, D. Perez-Marcos, H.H. Ehrsson, M.V. Sanchez-Vives, Towards a digital body: the virtual arm illusion. Front. Hum. Neurosci. **2**, 6 (2008). doi:10.3389/neuro.09.006.2008

M. Slater, D. Perez-Marcos, H.H. Ehrsson, M.V. Sanchez-Vives, Inducing illusory ownership of a virtual body. Front. Neurosci. **3**(2), 214 (2009)

Y. Su, Y. Qi, J. Luo, B. Wu, F. Yang, Y. Li, … W. Chen, A hybrid brain-computer interface control strategy in a virtual environment. J. Zhejiang Univ. Sci. C (2011). doi:10.1631/jzus.C1000208

M. Van Gerven, J. Farquhar, R. Schaefer, R. Vlek, J. Geuze, A. Nijholt, … P. Desain, The brain-computer interface cycle. J. Neural Eng. **6**, 041001 (2009). doi:10.1088/1741-2560/6/4/041001

B.G. Witmer, M.J. Singer, Measuring presence in virtual environments: a presence questionnaire. Telepresence Teleop. Virt. Environ. **7**(3), 225–240 (1998)

T.O. Zander, C. Kothe, Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general. J. Neural Eng. **8**(2), 25005 (2011)

Games for BCI Skill Learning

8

Reinhold Scherer, Gernot Müller-Putz, Elisabeth V. C. Friedrich, Viktoria Pammer-Schindler, Karin Wilding, Stephan Keller, and Johanna Pirker

Contents

R. Scherer (✉) • G. Müller-Putz
Institute for Knowledge Discovery, Graz University of Technology, Graz, Austria
e-mail: reinhold.scherer@tugraz.at

E.V.C. Friedrich
Department of Cognitive Science, University of California San Diego, La Jolla, CA, USA

V. Pammer-Schindler
Know-Center GmbH, Graz, Austria

K. Wilding • S. Keller
Knowledge Technologies Institute, Graz University of Technology, Graz, Austria

J. Pirker
Institute of Information Systems and Computer Media, Graz University of Technology, Graz, Austria

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_6

Abstract

A brain–computer interface (BCI) is a device that translates the users' thoughts directly into action. Brain signal patterns used to encode messages are user specific. However, experimental paradigms used to collect neurophysiological trials from individuals are typically data-centered and not user-centered. This means that experimental paradigms are tuned to collect as many trials as possible – which is indeed important for reliable calibration of pattern recognition – and are generally rather demanding and not very motivating or engaging for individuals. Subject cooperation and their compliance with the task may decrease over time. This leads in turn to a high variability of the collected brain signals and thus results in unreliable pattern recognition. One solution to this issue might be the implementation of engaging games instead of the use of standard paradigms to gain and maintain BCI control. This chapter first reviews basic principles and standard experimental paradigms used in BCI training that detect messages expressed by spontaneous electroencephalogram (EEG) rhythms. Users can independently modulate oscillations by performing appropriate mental tasks. Then, requirements for successful connection of games and these BCI paradigms are outlined in order to provide users with engaging methods to acquire the BCI skill. Last, a novel training concept for BCI in the framework of games is proposed. A recently introduced communication board for users with cerebral palsy is described as example to illustrate game-inspired training paradigms.

Keywords

Brain–computer interface • User-centered design • Man–machine learning • People with special needs • Skill learning

Introduction

A brain–computer interface (BCI) is a device that translates brain activity directly to action without the need of motor activity. Initially, BCI technology was developed with biomedical applications in mind, such as restoring communication capabilities for paralyzed individuals (Neuper et al. 2003) and replacing lost motor function (Pfurtscheller et al. 2008). More recent applications explore the use of BCI as input device for entertainment and gaming for able-bodied individuals and users with disability alike (e.g., Lecuyer et al. 2008; Nijholt and Tan 2007). The information transfer rates of BCIs are low compared to manual control. Typically less than 30 binary selections per minute can be executed. One may argue that these rates are too low for allowing meaningful interaction. However, for users with severe physical disability that cannot use standard human–computer interaction (HCI) devices or assistive technologies, even a slow communication channel is key. A barrier to more widespread adoption of BCI as input technology is that users must first acquire the "BCI skill": Operating a BCI is a skill that a user must learn and train just like any other skill. Again, for users with limited HCI options, the added value of BCI is sufficient to undergo this training. The necessity of BCI skill training shows that BCI technology does not allow mind reading and inferring thoughts in general as often depicted in science fiction movies.

Depending on the class of brain pattern used to encode the users' communication intention into messages for the BCI, training may take up to several months. The user must learn to reliably generate brain signal patterns that the BCI can robustly translate. Games form a good basis for keeping users motivated and engaged in the training process. However, so far games have mostly been used as end effector in BCI. In other words, users already need the skill to control a BCI for playing the game.

The goal of this chapter is to introduce readers to the field of BCI, outline the common training approaches, and illustrate how games can be used as engaging and motivating means for learning the "BCI skill," i.e., the skill to use a BCI effectively. To gain a better understanding of the challenge, some of the basic principles and concepts are reviewed. For a comprehensive review on BCI technology, please refer to relevant literature (e.g., Pfurtscheller and Neuper 2001; Wolpaw et al. 2002; Mason et al. 2007; Millán et al. 2010).

Background

Man–Machine Learning Dilemma

BCI involves two intelligent and highly adaptive systems: the human brain and the machine. Both are strongly interdependent but have to be adapted for optimal performance (Pfurtscheller and Neuper 2001). Figure 1 illustrates their relationship. Note that also the application software or device to be controlled can be adaptive (e.g., due to built-in context-dependent intelligence or by execution of automated workflows). Brain signals can only be translated into messages for applications when users are able to generate a very specific and stable brain signal pattern. Otherwise, translation may be erroneous and unreliable. The brain is a big interconnected network consisting of billions of neurons. Neurons are cells that are specialized in signal processing and signal transmission. These cells, at a local level, do not have a fixed wiring plan. During early development, molecular cues guide nerve fibers to target regions and initiate the formation of synaptic connections. Synapses are the interface between neurons. Synaptic development depends on the coordination of

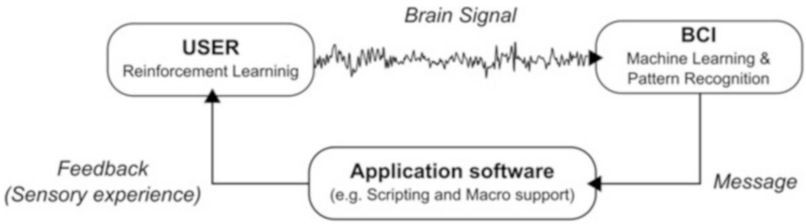


Fig. 1 User-BCI-app closed-loop feedback

neural activity between neurons. Sensory experience, for example, due to learning, allows a neural circuit to refine itself to suit a particular environment (Kandel et al. 2014). In short, learning new and reinforcing learned skills induce some sort of cortical reorganization. Since neural circuits depend on the sensory experience gained in the past, every brain is – strictly speaking – different. And as such, also the resulting brain signal patterns may vary significantly.

Operant conditioning (feedback or reinforcement learning) was the first approach used to train the BCI skill. In this case the translation rule is fixed a priori, and the user has to learn to generate patterns by trial and error. A more common approach is to distribute the BCI skill learning task between the brain and the machine. Machine learning and pattern recognition methods are used to find user-specific brain signal patterns and infer translation rules. This typically leads to BCI control in shorter periods of time. During the initial skill learning stage, however, the brain does not know how to reliably generate a unique pattern, which may lead to a large variability in the data. Consequently, gradual coadaptation of brain signal pattern generation and machine learning is required. In other words periods of feedback training and machine learning alternate or take place simultaneously (online learning). For coadaptation to happen in the first place, however, having a good set of data that allows inferring robust translation and prediction models is fundamental. In this chapter, the focus is on user-related and data collection aspects of coadaptation. For readers interested in machine learning aspects, please refer to the appropriate literature (e.g., Bishop 2006).

Brain Signal Recording Techniques

To establish a closed-loop interaction between the human brain and the machine (Fig. 1), the user's brain activity has to be first monitored. Common techniques for recording brain activity include functional magnetic resonance imaging (fMRI), positron emission tomography (PET), near-infrared spectroscopy (NIRS), electrocorticography (ECoG), and magnetoencephalography (MEG). These techniques measure either metabolic or electrophysiological correlates of cortical activity with different temporal and spatial resolutions and have different levels of invasiveness. As a result, each method has distinct advantages and drawbacks.

The Electroencephalogram (EEG)

The focus in this chapter is on the noninvasive EEG, which has emerged as an important noninvasive signal source for BCI and functional brain mapping in humans. EEG signals are potential differences recorded from electrodes placed on the scalp and reflect the temporal and spatial summation of electrical inputs to large groups of neurons lying beneath the recording electrode. EEG is, compared to MEG or fMRI, widely available, inexpensive, compact, and portable and offers a reasonable trade-off between temporal and spatial resolution (Niedermeyer 1999; Wolpaw et al. 2002). The temporal resolution of EEG is in the millisecond range and the spatial resolution in the centimeter range (Nunez and Srinivasan 2006). The coarse spatial resolution is originated by the different layers between the source of the signal in the cortex and the sensor placed on the scalp. Layers include the meninges, cerebrospinal fluid, skull, and scalp. These layers act as a volume conductor and effectively low-pass filter and smear the signal sources. To get a better picture of the level of activity that can be recorded by EEG, the following analogy can be used: imagine flying in a helicopter over a sport stadium where the finale of the soccer world championship takes place. The helicopter is equipped with a microphone (corresponds to the EEG sensor). The microphone will not be able to record conversations between single individuals (correspond to cortical neurons). The microphone, however, may be able to pick up the noise of the crowd (corresponds to millions of cortical neurons) when a goal is scored. Accordingly, the major drawback of EEG is the low signal-to-noise ratio (SNR) with the resulting susceptibility to noise and artifacts (Fatourechi et al. 2006). Absolute EEG signal amplitudes are typically less than 100 μV (10^{-6} V). Artifacts are signals that do not originate from the brain, but which may influence BCI output. The most frequent artifacts generated by the user are muscle activities and eye movements. Both signal amplitudes can take values up to the millivolt (10^{-3} V) range, i.e., they are about 1000 times larger than the EEG and may mask EEG activity. There can also be technical artifacts in the EEG signal, such as those caused by power line interface at either 50 or 60 Hz.

In summary, EEG data is very noisy, and hence, a lot of data may be necessary to train the machine learning algorithms. This also means that users will have to spend a substantial amount of time learning the BCI skill. Consequently, having engaging and fun experimental paradigms or games will keep users focused and motivated.

Evoked and Induced EEG Activity Patterns

Two different characteristic EEG signatures (patterns) are typically used to encode messages for BCI. These are evoked potentials (EPs) and event-related changes in spontaneous EEG rhythms, also known as event-related synchronization and desynchronization (ERS and ERD) (Pfurtscheller and Lopes da Silva 1999). Event-related potentials (ERPs) are characteristic EEG responses to sensory, cognitive, or motor events. EPs are ERPs that are evoked by external sensory stimuli that the user can modulate by cognitively attending the stimulus. An example is the steady-state visual EP (SSVEP) (Müller-Putz et al. 2005). Whenever a user attends flashing objects, for example, on a computer screen, the EEG recorded over visual cortex shows the same rhythm, given the blinking rate is above 5 Hz. Another example is the P300 (Bell et al. 2008). The P300 is so named because it is characterized by a positive potential shift about 300 ms after the presentation of a perceptually significant event embedded within a series of routine stimuli. EPs are time locked to external events, have a user-specific but defined shape, and can, assuming the stimulation details are known, be detected with very high accuracies (>90 %). Users that operate modern EP-based BCIs do not need much training time. Typically, the system is calibrated after minutes of time. Photosensitivity may be an issue for some user and ongoing stimulation (visual, auditory, or tactile) is not always well accepted by users.

This chapter focuses on spontaneous EEG activity, more precisely on BCIs that detect changes in spontaneous EEG rhythms, which individuals modulate by executing specific mental activities. The most commonly performed mental activity is motor imagery, that is, the mental imagination of body movements (Neuper et al. 2005). The kinesthetic imagination of a hand opening and closing movement modulates sensorimotor rhythms over corresponding cortical hand areas that the BCI can detect. Typically, ERD and ERS changes in the mu (10–13 Hz range) and beta (13–30 Hz) band range are induced. ERD is an amplitude decrease and ERS an amplitude increase, respectively, in specific oscillatory components over defined brain areas. ERS can be interpreted as a synchrony of the activity of cortical neurons. Since large populations of neurons behave in the same way, it is hypothesized that no signal processing takes place and that these neurons are at rest (idling). ERD describes an amplitude decrease, which indicates that populations of neurons are not cooperating (not in synchrony). When so, then it is hypothesized that cortical signal processing takes place. Since a user can perform motor imagery at will at any time and therewith modulate EEG rhythms that induce specific EEG patterns, this class of BCIs provides the users on-demand access to communication. The challenge for such self-paced BCIs – self-paced because the users themselves are in command and decide on the pace of communication – is to correctly identify and translate brain activity patterns in the ongoing EEG (Scherer et al. 2007). The noisy, nonlinear superposition of the electrical activity of large populations of neurons as measured on the scalp can mask the underlying neural patterns and hamper their detection. Consequently, there is a high probability that the temporal and spatial summation of different cortical activations generates very similar EEG activity patterns. The literature rarely provides specific details on the motor imagery tasks individual users perform to gain BCI control. Common motor imagery tasks are the kinesthetic imagination of movements of the left or right hand (e.g., wrist extension and flexion or squeezing movements) or both feet (e.g., dorsiflexion or foot pedal pressing tasks). Users are typically asked to repetitively perform the mental motor task at a comfortable speed for a given period of time with the aim to induce sustained ERD and/or ERS patterns. Note that also other mental tasks such as mental arithmetic, mental word generation, or spatial navigation induce ERD and/or ERS patterns (Friedrich et al. 2012, 2013; Scherer et al. 2015a).

BCI Skill Learning from a Machine Learning Point of View

Signal processing and statistical machine learning methods are commonly used to translate induced EEG patterns into messages (Wolpaw et al. 2002; Lotte et al. 2007; Mason et al. 2007). First-time BCI users achieve average binary accuracies, for example, by using right hand and foot motor imagery, of approximately 75 % (Blankertz et al. 2010). This performance level, however, does not allow useful and fast interaction. Moreover, conventional training approaches fail in about 40 % of the users. Fail means that users achieve accuracies <70 %. This phenomenon is called "BCI inefficiency" (Allison and Neuper 2010; Blankertz et al. 2010). Since the brain can only learn to reliably generate patterns by reinforcement or feedback learning, one can imagine that, given the <30 % mistranslation rates, BCI skill acquisition can become burdensome and demanding for the user. Mistranslation is partly due to the non-stationarity and inherent variability of spontaneous EEG, which, along with limited sample sizes and limited knowledge about the underlying signal, makes BCIs a challenging domain for signal processing. Non-stationarity and inherent variability mean that the statistical properties of EEG signals change over time. This can have several reasons such as changes in electromechanical properties of the electrodes (e.g., changing electrode impedances which results in changing signal amplitudes) or differences in electrode placement between BCI uses. Limited sample size refers to the fact that the collection of one single mental activity trial takes about 8 s in conventional training paradigms. Given that the human brain learns by doing – reinforcement learning – hundreds of repetitions are required. Additionally, since statistical methods are used to translate EEG patterns, a reasonable large number of trials have to be collected to calibrate the model parameters before BCI feedback training can begin. The number of training trials required depends on the model choice. Linear models typically require less training trials than more complex nonlinear models (Müller et al. 2003). Assuming the number of training trials is predefined, the predictive power of a method is reduced as the dimensionality of the feature space increases (Hughes effect and curse of dimensionality). A feature is an individual measurable property of the phenomenon being observed (Bishop 2006). A rule of thumb for linear models is to have at least ten trials per mental task for each feature of the pattern that is considered for translation. Regularization, cross-validation, and feature subset selection are typically applied to find reasonable solutions for the model parameter optimization problem. All these measures together contribute to minimizing the generalization error and lead to models that predict mental activity from EEG patterns with high accuracy. Generalization is the ability of a predictive model to perform well on unseen and new data. Since EEG properties change over time, model parameters may need regular updating. There are several ways how this can be achieved. Please see Vidaurre et al. (2011), Faller et al. (2012), and Samek et al. (2013) for more information on coadaptation, online adaptation, and transfer learning approaches in BCI.

Standard Graz BCI Training Paradigm

EEG trials of mental imagery have to be collected from users to setup predictive models prior to BCI use. In the following the standard training paradigm that is used in many BCI laboratories around the world is described in more detail (Neuper et al. 2005; Müller-Putz et al. 2010). Figure 2a summarizes the temporal structure of an individual trial, which consists of the following blocks: At second 0 a cross appears in the middle of the black computer screen. Users are asked to fixate the cross and thus prevent eye movement artifacts. After 2 s a beep is sounded to catch the user's attention. A visual cue indicating the requested motor imagery task appears at second 3 and stays on the screen until second 4.25. Arrows, starting in the middle of the screen and pointing toward the left side or right side or toward the bottom of the screen, indicate left hand and right hand of foot motor imagery, respectively. Visual cues are presented in pseudorandom order to ensure that the same number of trials is available for each motor imagery class. Users are asked to repetitively perform kinesthetic motor imagery according to the presented cue for 5 s until second 8. At second 8 the cross disappears and the screen is blank again. A variable pause (i.e., intertrial interval, ITI) lasting between 2.5 and 3.5 s is added before the next trial begins. Users are asked to relax and restrict movements during the 8-s trials and to move (e.g., blink or take a deep breath) during the ITI. For two classes, as depicted in Fig. 2a, typically, 160 trials are collected in total, subdivided into four runs of 20 trials per class each. This amounts to a data collection time of about 50 min. A break of several minutes between the runs is inserted. The collected EEG is subsequently used to calibrate BCI model parameters. After the calibration, feedback training is performed. The experimental timing is identical to the above training without feedback timing. Additionally, from second 3 to second 8, visual feedback in the form of a bar graph is presented to the user (Fig. 2a). The length of the bar graph corresponds to the quality of the prediction. Whenever the predicted motor imagery task conforms with the predefined mental task, the bar graph moves in the direction of the cue-arrow. Otherwise, the bar extended toward the other direction. Users are hereby enabled to train to generate EEG patterns that lead to high correct detections, i.e., to originate patterns by performing motor imagery and move the bar into the direction of the arrow. The collected data is then again used to update model parameters toward an increase in correct translations. This training paradigm was designed with the aim to systematically collect as much data as possible. However, visual and auditory feedbacks are not appealing and handling the recurrent tasks may be perceived as boring, frustrating, and unfruitful from users. As results they may not comply with the rules and do not follow the instructions. Here, game-based training comes into play. Games may be a better way to motivate users and keep them engaged in performing the very same task over and over again and still keeping motivation and concentration high.

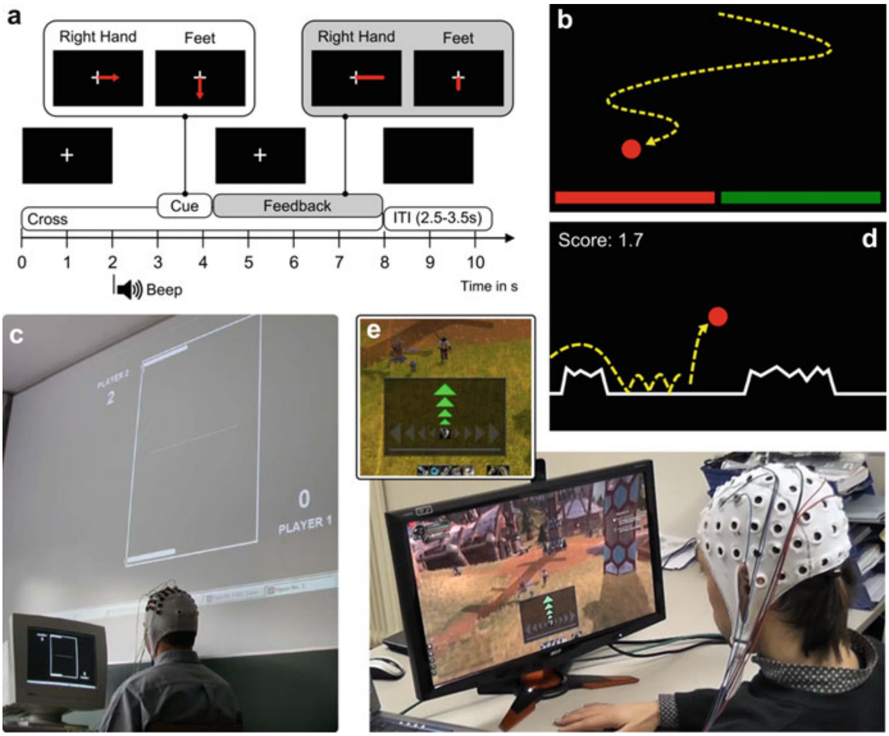


Fig. 2 Graz BCI training paradigms. (**a**) Cue-guided 2-class Graz BCI standard training paradigm. (**b**) Basket paradigm. (**c**) Pong game presented at the 1st International BCI Conference 2002 in Graz, Austria. (**d**) Bouncing ball game. (**e**) BCI user playing World of Warcraft. User can control the game avatar (available commands are "move forward," "turn left," and "turn right") via in-game add-on. Interaction with non-player character and monsters is triggered by waiting close to the characters for a predefined dwell time

Design Flaws from a Game Point of View

There are several differences between the standard training paradigm described above and games in general. Firstly, the graphics design used in BCI is minimalistic and cannot be seen as an eye-catcher that attracts the users' attention and makes them want to train. Secondly, the gameplay – if one can call it that – is straightforward and does not allow alternative or different solutions. Users have to keep up with the predefined timing and perform mental activity based on the presented cue information. After a certain period, this may result in boredom and disinterest and the users might start mind wandering. Thirdly, during the initial training, no whatsoever feedback is presented to the users. This again potentially leads to disinterest and indifference. Cooperation is crucial in this stage as future performance strongly depends on the users' behavior and compliance with the task. Fourthly, the goal of the training is clear: gaining BCI control. However, users have a hard time in estimating the training progress as well as the resulting application of their control.

Quantification of the performance during runtime would help users to generate patterns that can later be used to control the desired real-world application. As the training paradigm and the application the users want to control are often not directly linked, integrating components of the application into the training would allow the user to better understand training progress and performance. Training should be tailored to the application and be based on sequential modules that clarify the situation concerning the progress made. Fifthly, training without and with feedback is mostly visually different. The visualization of the feedback bar, for example, can induce additional brain activity that results in alternated EEG patterns (e.g., stress or a feedback-induced co-activation of the visual cortex). Thus, new BCI paradigms provide sham feedback in the screening and calibration runs in which usually no real-time feedback is shown, to minimize differences in brain activity between sessions with and without feedback (Friedrich et al. 2013).

Gaming Concepts for BCI Training

To overcome some of the above-listed design flaws and with the aim to create more motivating and entertaining paradigms, self-designed gaming concepts and game-like interactions to BCI training that fulfilled the most important requirements were applied. The following examples illustrate developments made in the Graz BCI laboratory. Krausz et al. (2003) proposed the "basket" paradigm. Instead of a bar graph that was moving left or right, a ball was falling with a predefined speed from the top of the screen toward the bottom. The user could manipulate the horizontal position of the ball on the screen by performing the appropriate mental tasks, for example, left hand and right hand, respectively, for moving the ball toward the right and left border of the screen. At the bottom of the screen, two baskets were positioned. The task of the user was to drop the ball into the target basket. The target basket had the same color as the ball (Fig. 2b). In this way, the user could experiment with ball control in the upper part of the screen but had a goal to achieve, i.e., collect points for hit target baskets, in the lower part of the screen. The score was shown to the user at the end of each experimental run. The next step was the implementation of a Pong game-based training paradigm. Pong is a simulated table tennis arcade video game released by Atari, Inc., Sunnyvale, CA, USA, in 1972. Figure 2c shows a screenshot taken at the 1st International BCI Workshop and Training Course that took place in Graz, Austria, in 2002. The user could either play against the computer in the single-player mode or against another BCI player in the two-player mode. This brought social interaction and competition into the training. Users were engaged and motivated to score points and to win the game. To train users to generate EEG patterns with different durations, a training paradigm was implemented that was inspired by the bouncing pumpkin character from "The Cauldron II: The Pumpkin Strikes Back" game that was released from Palace Software, London, UK, in 1985. The user had to manipulate the height of the ball's bounce and jump over objects of different lengths that were entering the screen on the right side and moving with a constant speed to the left. The ball was always bouncing in the middle of the screen. Figure 2d shows a diagram of the screen. Cue information used to categorize and analyze the data was implicitly the onset of the obstacle objects. By using this paradigm, users successfully learned to generate EEG patterns of two different durations by using one single mental task only. With this temporal coding, users were able to control two independent degrees of freedom of an artificial robotic arm (Müller-Putz et al. 2010). The graphics of the game were rather simple; however, the user was motivated and enjoyed playing games by thought alone.

More recently, the massively multiplayer online role-playing game "World of Warcraft" (Blizzard Entertainment, Irvine, CA, USA) (Scherer et al. 2013a) (Fig. 2e) and the first-person puzzle-platform game "Portal 2" (Valve Corporation, Bellevue, WA, USA) were interfaced. The original idea was to use these graphically interesting and motivating games for training. However, for them to be really useful for training, certain aspects of the gameplay should have been changed. For example, one first simple training exercise could have been to map left and right hand motor imagery, respectively, to left and right hand movements of the game avatar. Controlling individual limbs of the avatar, however, was not supported. The standard command repertoire included navigation commands such as move forward or turns left/right and cast spells. A certain degree of BCI control is required to use these commands in a meaningful way. In other words, users already need the BCI skill.

BCI Skill Learning from a Game Point of View

Requirements for Games To Be Useful for BCI Training

If games are a great way to learn the BCI skill, why is not everybody already using games for BCI training? Certainly, there is a huge variety of engaging games out there. The problem with off-the-shelf games is, as already indicated above, that it is typically not possible to modulate them for specific research or application purposes (Ninaus et al. 2014). Games in general have to fulfill certain requirements in order to be useful for BCI training. First and crucial for BCI implementation, the game metrics must provide the possibility for real-time operation and closed feedback-loop systems. Additionally, it should be possible to change and adjust the game metrics as needed. The game should be compatible with different recording soft- and hardware, and the game control should be possible with different parameters and signals (e.g., EEG rhythms in alpha (8–12 Hz), theta (4–7 Hz), and beta (13–30 Hz) band, electromyogram, electrooculogram). Marker or triggers have to be stored in the game and physiological recordings to ensure temporal synchronization between game events and neuro-recordings. Thresholds should be individually adjustable for control channels as the ability to control brain rhythms is different between users and also changes between as well as during sessions depending, for example, on mood or impedances. Second, the game content should be adjustable for the target group concerning age, sex, interests, or neurophysiological disorders. The difficulty and story plot must also be suitable for age and IQ. Third, the game design should make specific feedback possible such as that the feedback is specific to the significance of the signals being trained.

Classification of BCI Skill Training Games

The motivation to embed BCI skill training in games is to make such training more engaging for end users. In terms of HCI research, this means that BCI skill training games are games for a purpose (Carter et al. 2014). More precisely, the specific purpose is to support users in learning to interact with computers via brain signals. In terms of a game's classification that was recently proposed in the HCI community (Carter et al. 2014), BCI skill training games fall into the category of "games for a purpose," but would span two subcategories, namely, "games for data collection" if the focus is on collecting users' brain signals to train the BCI system and "games for learning and education" if the focus is on the users learning to interact with the BCI system.

Another option to make BCI training more engaging is the use of game design elements. Instead of designing an entire game, which can be an expensive and challenging task, the original training process can be enhanced with gamification strategies such as the use of game-based feedback types, rewards, and playful task design (Deterding et al. 2011).

What Do Good Games Offer to the User?

Good games offer two important aspects to players in the context of BCI learning. First is **an engaging, immersive user experience.** The game's visual design, its rules and goals, and the rewards are engaging and invite users to immerse themselves in the game. Naturally, different users and different user groups have diverging preferences. Related to the points of critique for standard BCI skill training software, this means that a good BCI skill training game would be visually engaging and give users (players) freedom over their actions in the confines of each game's rules (Oblinger (2006): "require the player to make frequent, important decisions"). Moreover, the standard paradigms could be replaced by control strategies that are user specific, i.e., individually adapted to the user's skills and preferences (Friedrich et al. 2012, 2013). Some people might find the imagination of a tune more engaging than motor imagery, for example, and thus, this mental activity could be taken as control strategy. Additionally, their immersive character can lead to a more focused and concentrated training sequence. A well-balanced game design, which takes into account both the user's skills and the current level of difficulty, can introduce a very constant focus of the user on a task. This feeling is also referred to as "flow" state and is explained as a mental state in which a person is fully immersed in a task. To design such an experience, a well-designed balance between the difficulty of the game/level and the player's skill is important. This flow experience is not only important for the fun but also facilitates the transfer of the learned content from the virtual game into real-world behavior (Slater et al. 1996). The level of immersion can also be measured by event-related potentials (ERPs) (Kober and Neuper 2012). In the case of such sensible training operations such as BCI skill training, an adaptive difficulty level to meet the players' skills is essential. As the level of immersion can be measured by ERPs – which do not interfere with the BCI control by oscillatory activity – it might be possible to adapt the game in real time accordingly in order to keep the person "in the game." Designing games and activities to achieve flow experiences requires the paradigm (game) designer to design an experience with clear rules and simple goals aligned with the player's skill level, to diminish distractions to facilitate focus, and to give in-time feedback (Csikszentmihalyi 1991; Schell 2014).

Second, games constitute **an environment in which what must be learned is directly relevant**. Games typically have a clear set of goals (Oblinger 2006). Consequently, games provide direct feedback about the skill(s) to be learned. At the same time, they invite users to set themselves challenges (e.g., increase difficulty, experiment with various problem-solving strategies). Both aspects (feedback, inviting to self-directed exploration) are crucial for learning, as they "give players a sense of self-efficacy" (Breuer and Brente 2010), which ultimately is a prerequisite for self-motivated and self-directed learning. Games are valuable tools to learn patterns to solve in-game problems. Well-designed games are engaging the player to learn those patterns and to recognize the required pattern in various situations to complete and master the game. This property can also be used to master BCI skills in games (Koster 2013).

The mental activities used for control could be selected so that they are meaningful for the specific application. For example, if someone wants to jump in the game, the control strategy could be "motor imagery with the feet," or if someone wants to start a dialogue in the game, the control strategy might be "word association." This might lead to a more intuitive feeling of control and thus might increase again the flow, the feeling of "being in the game." However, if users are not able to produce an analogous mental image, they could also produce a different mental image, for instance, they think of moving the right hand for "jumping" in the game. Finding the appropriate task for the individual is the key for successful BCI control. We currently assume that over time the user will then learn to directly associate the mental image that was originally "move the right hand" with "jumping in the game." Preliminary tests with patients suggest this. Other mental tasks that could be used for BCI control are, for example, spatial navigation, auditory imagery, mental rotation, word association, and mental arithmetic (Friedrich et al. 2012, 2013). However, there are substantial interindividual differences in the performance of the mental tasks. That means that the tasks that achieve the highest accuracy vary significantly across users (Friedrich et al. 2012, 2013; Scherer et al. 2015a).

The Impact of Feedback

Feedback should be direct and objective and thus easy to interpret and clear to understand for the user. Feedback in BCI, however, can be difficult to interpret for users. Typically, the output of the selected machine learning and pattern recognition model (classification or regression) is shown to users. More precisely, one single value (discrete value or numbers) is mostly presented to the user. Since commonly a number of features that characterize the EEG pattern to be translated enter the model, users face the challenge that they have to learn the mapping of the BCI model and thus identify the feature(s) that mostly contributes to the current feedback behavior. For high-dimensional feature, spaces with nonlinear mapping feedback interpretation may become very tricky, which results in lengthy trial-and-error training. One improvement could be that the feedback could be specific to the significance of the signals being trained (Friedrich et al. 2014). For example, if the user desynchronizes alpha or beta over the motor areas, the feedback shows fast activity in the game, movement, or imitation behavior, and if the user synchronizes, the game gets slower, and the feedback indicates relaxation. The connection of the signals being trained with functional-significant feedback could further improve the learning.

Use Case: Row–Column Scanning Communication Board for Individuals with Cerebral Palsy

This section illustrates the above-described more abstract discussion of issues in using games for BCI training in a concrete use case. More precisely, it is about training the BCI skill in people with special needs.

Cerebral palsy (CP) is a nonprogressive condition caused by damage to the brain during early developmental stages (Holm 1982). Individuals with CP may have a range of problems related to motor control, speech, comprehension, or mental retardation. Due to lack of movement and posture control, a lot of individuals with CP cannot use standard HCI devices and rely on the help of others. Providing them a though-based communication device will give them more independence and provide outer social inclusion (e.g., social networks). Very recently, Scherer et al. (2015b) introduced a BCI-based communication board for users with CP. The board implements the one-switch row–column scanning principle to select items that are arranged in a grid. A binary value can easily be assigned to a switching function. Whenever users want to make a selection, executing the appropriate mental activity activates the brain switch. Row–column scanning means that each row within the grid is sequentially highlighted until the user selects the row containing the desired item by activating the switch. The columns within the selected row are then scanned until the target item is highlighted and can be selected by activating the switch a second time (Fig. 3a). Figure 3b shows CP users during the testing phase. One user is located in front of the communication board during the EEG sensor montage. Sensors are integrated in a cap. Conductive gel is applied to enhance signal quality. Combining a reliable thought-activated switch with augmentative and alternative

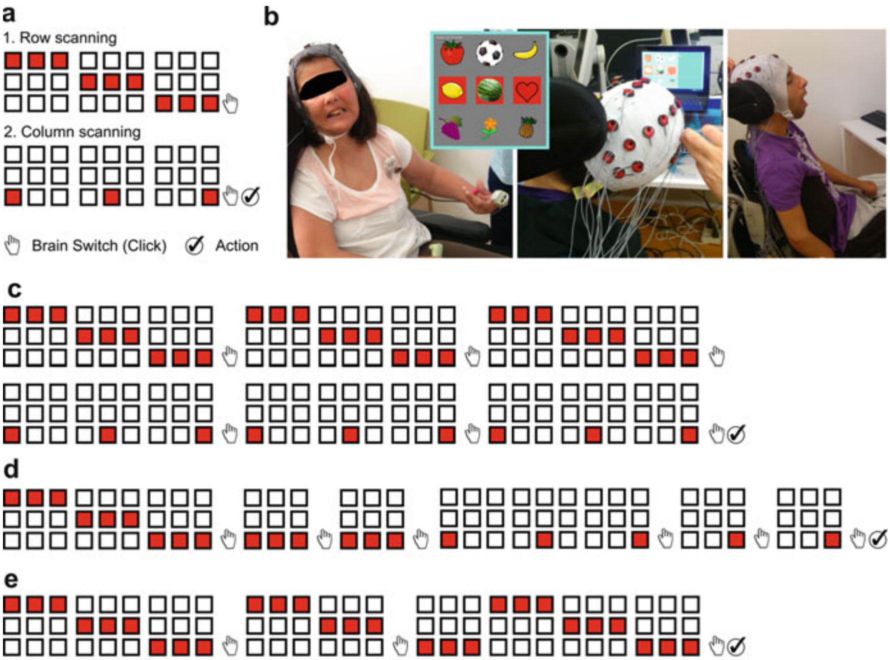


Fig. 3 Row–column communication board. (**a**) Basic principle. Selection of the item in the third row and third column. The "Finger" symbol represents the activation of the brain switch. The "Action" symbol shows that the selected item is executed. (**b**) CP users. The middle image shows one user in front of the communication board. (**c**) Evidence accumulation. Click on a row (item) three times before the selection is accepted. (**d**) Optimized evidence accumulation with row (item) scanning repeat. The previously selected row (item) is highlighted again and can be reselected. (**e**) Evidence accumulation m out of N selection. In the example $m = 2$ out of the past $N = 3$ scan cycles, the row has to be clicked before it is accepted by the system

communication software such as the Grid (Sensory Software International, Malvern, UK) enables users to autonomously communicate and operate computers.

In order to design appropriate game-inspired training paradigms, it is important to know and understand the system dynamics of the row–column scanner and the operation principle. The following timing was used during the pilot experiments in Scherer et al. (2015b). Each row (item) was highlighted for 4 s (trial). To select the marked row (item), users had to perform the predefined mental task for the whole duration of the highlighting period. During this period, users did not receive any feedback. To notify the user that a 4-s segment was classified as "target mental task," an auditory beep was presented, and the appropriate action was executed. When the current 4-s EEG segment was translated as "nontarget mental task," a 2-s break (intertrial interval) was presented and scanning continued with the next row (item). When users did not want or need to make a selection, users were asked not to perform the predefined mental task. Spasms or involuntary movement may prevent users from looking at the graphical user interface – preventing interaction – or even induce EEG patterns that may mistakenly be interpreted as imagery. To minimize erroneous selections, evidence accumulation was implemented. This means that users had to confirm a selection several times before it was selected by the system. This measure also meant that selection time is increased. Figure 3c, d shows two different evidence accumulation selection mechanisms. The example in Fig. 3e shows that erroneous translations do not lead to wrong selections. First, results suggest that seven out of eleven CP users performed better than chance and would consequently be able to communicate to some degree by using the developed system (Scherer et al. 2015b).

Another important concept when designing HCI systems is user-centered design. CP users have different levels of motor and cognitive impairment. Due to these issues, the user may not be able to go to school and hence use a different vocabulary – assuming they can speak – and have different knowledge levels such as writing and math skills. Some individuals can write or do basic math, while others cannot. However, it is important to explain to users the basic concepts of BCI technologies, and user cooperation is required for machine learning calibration. The vocabulary and the way this technology needs to be explained to users are highly variable. Therefore, personalized instructions are required that support the user during the learning and training process. Making BCI technology available to children and adolescents may have a significant impact on their quality of life. The ability to communicate and express their emotions may allow them to participate in special school programs and as such become better integrated in social life.

Two Games on Trial: A Jigsaw Puzzle and Tic-Tac-Toe

Jigsaw Puzzle

One popular game in which scanning would seem natural is the genre of jigsaw puzzles. Scherer et al. 2015c introduced a prototype jigsaw puzzle game for android devices that was designed with CP users for CP users (user-centered design).

Overall Game Design

A puzzle is subdivided into pieces that are arranged as a grid on the screen. Different images and grid sizes can be selected which makes the training more interesting and challenging. The puzzle game with the default grid size of 3×3 is shown in Fig. 4a. The game randomly selects one piece and the user has to place it at the correct position by first selecting the row and then the column. The piece to be placed is shown in the bottom right corner of the screen. Additionally, to help users, a preview of the whole puzzle image is displayed in the top right corner. The aim of the game is to piece together the puzzle to unmask the selected puzzle image. Users can see their progress by the number of pieces that was already placed. Once a piece is set correctly, the selection process starts again from the beginning with a new piece to place. If a row (item) is already set, it will be skipped in the scanning process. This helps to speed up the game and keeps the user concentrated and focused on the task.

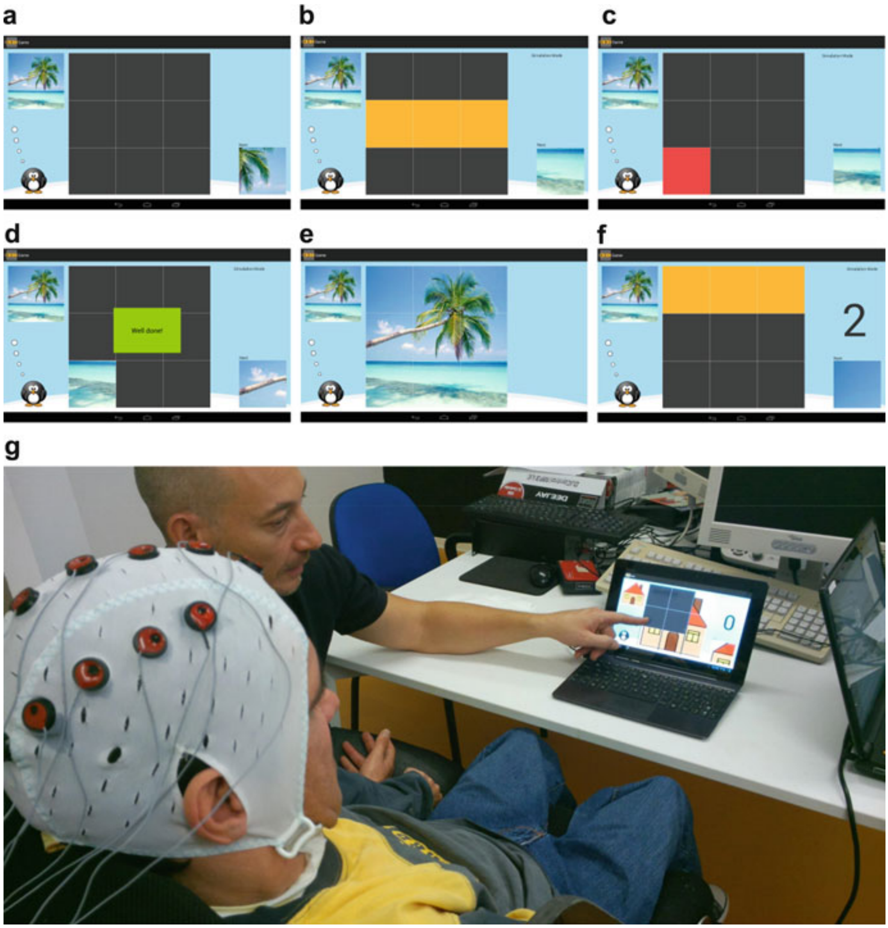


Fig. 4 Jigsaw puzzle game. (**a**) Empty 3 × 3 grid. The puzzle image is shown in the *left upper corner* of the screen. The piece to be placed is shown in the *lower right corner* of the screen. (**b**) Row scanning mode. (**c**) Column scanning mode. (**d**) Piece placed correctly. (**e**) Completed puzzle. The correct position of the jigsaw puzzle piece (*bottom right*) in the overall picture (*top left*) has been selected. (**f**) Evidence accumulation counter on the *right side*. (**g**) Picture of pilot testing with a CP user. The caregiver shows the user where to place the puzzle piece

The attention span of users may vary significantly. Overall, better quality of EEG training data is collected if users play for a shorter time but are more concentrated than vice versa. Figure 4b–d shows screenshots of the row and column scanning process and the feedback presented to the user for placing a piece into the correct grid position. Figure 4e shows the selected palm beach image fully assembled.

Since evidence accumulation was implemented, the counter of how many times the brain switch for a row (item) was mentally pressed is shown on the right side of the screen (Fig. 4f). A $m = 2$ out of $N = 3$ past detections was used (see Fig. 3e).

Modes of Play

Several operation modes were implemented (simulation, beginner, expert) to give players the possibility to smoothly move from getting familiar and learning the game to learning BCI interaction and to playing the game via BCI control.

Simulation mode: In the simulation mode, the game automatically selects the puzzle pieces correctly. This mode is comparable to a movie which users "watch" to get familiar to the system dynamics and to train performing imagery without the need of the BCI ("dry training"). Additionally, since users are different and need different instructions, personalized audio cues can be presented to support and instruct the user on what he or she needs to succeed in the game. The audio cues are comparable to a narrator in a movie, which comments the actions of the player. Since users may have vision and/or hearing problems, providing multimodal cues (visual and auditory) is important as this allows providing clear instructions that user should be able to follow easily. This mode was implemented to give users time to get familiar with the application and give them the chance to practice imagery before the BCI training starts. This should make them feel more confident with performing the task and perceive the training less stressful.

This mode is also used to setup the machine learning model parameters. Assuming that users were motivated and compliant with the task, EEG trials (segments) recorded can be labeled either as target or nontarget trials. Predictive models can then be trained by using this labeled data.

Beginner (error-ignoring) mode: After BCI calibration with the simulation mode, users can operate the thought-based switch. The beginner mode, however, accepts only correct (true positive, TP) selections. Remember that at this stage, the system is already pre-trained (vs. simulation). In addition, the game is easy enough, and users are motivated enough, that we assume that users always try to make a correct selection. False positive (FP) activations therefore stem from the BCI system's inaccuracy and are ignored. This keeps motivation high as users can only make correct selections; however, the user can practice and get positive reward. Data collected can again be used to adapt and update BCI model parameters. As above, labels are generated by assuming that users only make correct selections.

Expert mode: The expert mode is the regular gameplay without any support from the system.

Tic-Tac-Toe

Tic-tac-toe is a paper-and-pencil game for two players, "*X*" and "*O*," who take turns marking the spaces in a grid composed of three rows and three columns. The player who succeeds in placing three respective marks in a horizontal, vertical, or diagonal row wins the game. In addition to the abovementioned features, the tic-tac-toe game introduces competitive elements (player vs. player or computer vs. player). The user interface of the game is shown in Fig. 5a, b. In order to be more attractive for younger players, tic-tac-toe has been designed with a lion and an elephant instead of "X" and "O" game figures.

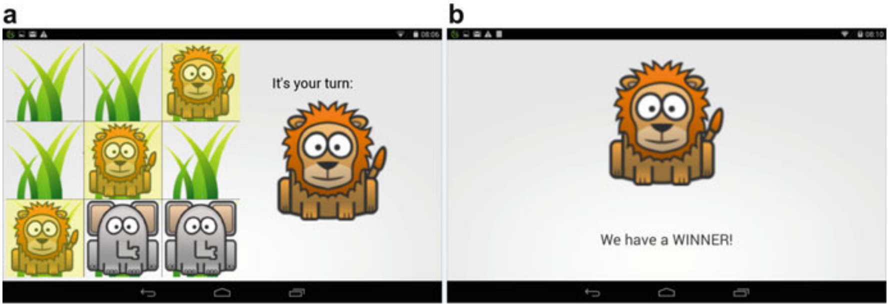


Fig. 5 Tic-tac-toe with an elephant and a lion. (**a**) Screenshot of the game board. (**b**) Bouncing lion wins the game

Insights on BCI Games for Training for CP Users

Feedback from first pilot tests with CP users (Fig. 4g) are very encouraging, and users like interacting with the games (Scherer et al. 2015c). Please note that testing involved the simulation mode and the beginner (error-ignoring) mode. The implemented puzzle and tic-tac-toe games are graphically more appealing than the cross, arrow, and bar graph graphics used in the standard training paradigm (see Fig. 1a). The implemented puzzle and tic-tac-toe games are moreover motivating and fun to interact with when compared to standard paradigms presented.

End users and caregivers were involved in the design process, which is fundamental for identifying the needs of users and implementing the required functionalities. Most important was the simulation with audio cue operation mode concept. This allows users to familiarize with the application and with the idea of BCI control. Visiting patients at home or in clinical environments puts pressure on them and on their relatives. "Dry training" is one way to reduce this inner pressure and allows users to mentally prepare themselves for the upcoming experiment (where scientists show up and bring a lot of unfamiliar devices and run experiments and ask questions).

Thinking Aloud: Other Options of Game-Based Training

So far, games mostly have been used as end effector and for user motivation. BCI control – to some minimal degree – was already necessary and related training paradigms helped to fine-tune the BCI skill.

The jigsaw puzzle and tic-tac-toe introduced a novel concept for BCI training in general, namely, that the training paradigm is an agent that teaches both the user and the machine. The idea of an evolving training paradigm, which introduces a story to the users and collects their reactions and brain responses, which are used to calibrate the pattern recognition methods, is compelling.

Other games that might be very suitable for this training paradigm are, for instance, simple story-driven games focusing on the narrative aspects instead of the interactive and action-packed experience as they can be played in a slow and controlled pace. This enables the players to focus on the experience and makes the training process an immersive and self-paced adventure. Narrative or adventure games can be traditional text-based adventures or more visually appealing modern games. An excellent example for this concept is the interactive story modification game “The Stanley Parable,” which was developed by Davey Wreden and released in July 2011. The player guides the game’s protagonist through a surreal environment and has the opportunity to make numerous decisions on which paths to take. Every choice made by the player is commented on by the narrator, and depending on the choices the player makes, they will encounter different endings to the game before it restarts. Basically, the player can play for hours and listen to an interesting and surreal story, by mostly binary decisions. Based on this principle, BCI training storyboard games can be designed that encourage user to perform imagery at defined periods of time, allowing for the collection of the data required to calibrate the BCI.

A wider variety of games that need only slow and low-dimensional input are available and limited only by the imagination of game creators: The requirement to create games with a very simple interface supporting people with special needs was spotlighted in the framework of the Global Game Jam 2015, a worldwide developer event to create games on a specific theme with various optional constraints. One of these optional constraints was the design aspect to create a game, which can be played and navigated with only one button and without the need for timely and quick reactions. A total of 359 games have been developed, which are in line with this diversifier. Many of them can be used as inspiration to support BCI training scenarios.

In particular in BCI training, the learners’ skill levels are very different. This is the reason why meeting the skill level of the learner, which is necessary to maximize the learner’s engagement and concentration (too easy is boring; too difficult is frustrating), is challenging in BCI skill training games. Adaptive game/level design with a parameterized version of the game allows analyzing and predicting the player’s skill level. Such techniques can be used to create adept levels based on runtime. In Flappy Bird (developed by Vietnam-based developer Dong Nguyen and released in May 2013), for example, a simple one-button side-scroller game, such variable parameters could be the position, width, or height of objects or the physics behavior of the player. All those setups will affect the gameplay and the game difficulty. Introducing such an adaptive game behavior in the context of BCI skill training can maximize the training experience to create experiences designed for the very specific learner requirements (Isaksen et al. 2015).

Another future vision for BCI skill training can be a self-representation of the learner in an avatar-based environment, such as a virtual world. Remember, this was the initial intention with World of Warcraft. Such representations can enhance the user’s self-awareness, the feeling of presence and immersion. Learners can train the BCI skills with interactions with objects or other inhabitants of the virtual world. In a multiuser virtual world, users can train together to make the training experience more fun and engaging. Such environments support distant training scenarios with other learners, but also with the trainer (Pirker et al. 2012).

The example of World of Warcraft illustrates that in the future, the BCI community should foster closer cooperation with the game industry. Commercial games might consider a research version of their game in which it is easier to interface the game with BCIs. The BCI community would greatly benefit from having access to high-end games, and the game industry could expand their potential customers and benefit from increased publicity. Cooperation not only between research and industry but also within the scientific community should be increased in order to provide game developers with the necessary requirements which are needed for a game suitable for BCI training as well as to encourage the BCI community to make more effort in replacing the standard paradigms with more motivating games.

A future challenge in the context of CP users – but also in general – is the identification of the most meaningful mental task for encoding messages. A user that never moved is likely not able to perform motor imagery and imagine moving and perceiving their limbs. Also, other mental tasks such as mental mathematics and mental word generation might be difficult, since users may not be able to do math or write. Therefore, it is crucial to select appropriate tasks individually from a range of reliable control strategies to modulate EEG rhythms (e.g., Friedrich et al. 2012, 2013). Combining games and storytelling in a creative way may be used to screen users and to identify appropriate activity patterns. Users themselves can decide which kind of imagery is appropriate and easy for them.

Recently, the gaming concept was applied also to the functional brain-mapping domain (Scherer et al. 2013b). Results from an experimental study with able-bodied subjects playing a virtual ball game suggest that the Kinect motion sensing input device (Microsoft, Redmond, WA, USA) is useful for isolating specific movements during the interaction with the game and that computed EEG pattern for hand and foot movement is in agreement with results described in the literature. The use of modern game hardware and technology such as the Kinect or the Oculus Rift virtual reality head-mounted display (Oculus VR, Irvine, CA, USA) may significantly contribute to advancements in neuroscience and help gain more comprehensive understanding of brain functioning. Future experimental paradigms will be designed to better reflect real-world settings and hence allow collecting brain signals from behaving brains in more natural but well-controlled conditions.

Conclusion

This chapter illustrated how games can be used to increase user motivation and involvement during the potentially tedious and time-consuming BCI skill training period. Moreover, the application of presented concepts and ideas to BCI training should allow collecting enough high-quality data from motivated and engaged users as needed for reliable pattern recognition. Educational psychology and instructional design (Lotte et al. 2013) will provide further knowledge on how to expand and refine game-based training. Adding online coadaptation (e.g., Faller et al. 2012) is another important step toward user-friendly plug and play BCIs. Finally, it is still unclear how BCIs can be used most efficiently for communication and control. It can be expected that the interchange of ideas between the BCI and the gaming community may help BCIs to evolve to a more intuitive and natural form of interaction, such as the use of a pen or a computer keyboard.

Cross-References

- ► Action Games, Motor Imagery, and Control Strategies: Toward a Multi-button Controller
- ► Brain-Computer Interface Games: Towards a Framework
- ► Brain-Computer Interfacing and Virtual Reality
- ► Towards Serious Games for Improved BCI
- ► User-Centered BCI Videogame Design

Recommended Reading

B.Z. Allison, C. Neuper, Could anyone use a BCI? in *Brain-Computer Interfaces* (Springer, London, 2010), pp. 35–54

B. Blankertz, C. Sannelli, S. Halder, E.M. Hammer, A. Kübler et al., Neurophysiological predictor of SMR-based BCI performance. Neuroimage **51**, 1303–1309 (2010)

C.J. Bell, P. Shenoy, R. Chalodhorn, R.P.N. Rao, Control of a humanoid robot by a noninvasive brain-computer interface in humans. J. Neural Eng. **5**(2), 214–220 (2008). doi:10.1088/1741-2560/5/2/012

C.M. Bishop, *Pattern Recognition and Machine Learning* (Springer, New York, 2006)

J. Breuer, G. Brente, J. Comput. Game Culture **4**(1), 7–24 (2010)

M. Carter, J. Downs, B. Nansen, M. Harrop, M. Gibbs, Paradigms of games research in HCI: a review of 10 years of research at CHI, in *Proceedings of the First ACM SIGCHI Annual Symposium on Computer-Human Interaction in Play* (CHI PLAY'14) (ACM, 2014)

M. Csikszentmihalyi, *Flow: The Psychology of Optimal Experience*, vol. 41 (Harper Perennial, New York, 1991)

S. Deterding, D. Dixon, R. Khaled, L. Nacke, From game design elements to gamefulness: defining gamification, in *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments* (ACM, 2011), pp. 9–15

J. Faller, C. Vidaurre, T. Solis Escalante, C. Neuper, R. Scherer, Autocalibration and recurrent adaptation: towards a plug and play online ERD-BCI. IEEE Trans. Neural Syst. Rehabil. Eng. (2012). doi:10.1109/TNSRE.2012.2189584

M. Fatourechi, A. Bashashati, R.K. Ward, G.E. Birch, EMG and EOG artifacts in brain computer interface systems: a survey. Clin. Neurophysiol. **118**(3), 480–494 (2006)

E.V.C. Friedrich, R. Scherer, C. Neuper, The effect of distinct mental strategies on classification performance for brain-computer interfaces. Int. J. Psychophysiol. **84**, 86–94 (2012)

E.V.C. Friedrich, C. Neuper, R. Scherer, Whatever works: a systematic user-centered training protocol to optimize brain-computer interfacing individually. PLoS One **8**(9), e76214 (2013)

E.V.C. Friedrich, N. Suttie, A. Sivanathan, T. Lim, S. Louchart, A. Pineda, Brain-computer interface game applications for combined neurofeedback and biofeedback treatment for children on the autism spectrum. Front. Neuroeng. **7**, 21 (2014). doi:10.3389/fneng.2014.00021

V.A. Holm, The causes of cerebral palsy: a contemporary perspective. JAMA **247**(10), 1473–1477 (1982)

A. Isaksen, D. Gopstein, A. Nealen, Exploring game space using survival analysis. Foundations of Digital Games. Best Paper in Artificial Intelligence and Game Technology (2015)

E.R. Kandel, J.H. Schwartz, T.M. Jessel, S.A. Siegelbaum, J. Hudspeth, *Principles of Neural Science*, 5th edn. (McGraw-Hill Medical, New York, 2014)

S.E. Kober, C. Neuper, Using auditory event-related EEG potentials to assess presence in virtual reality. Int. J. Hum. Comput. Stud. **70**, 577–587 (2012)

R. Koster, *Theory of Fun for Game Design* ("O"Reilly Media, Sebastopol, 2013)

G. Krausz, R. Scherer, G. Korisek, G. Pfurtscheller, Critical decision-speed and information transfer in the Graz brain–computer interface. Appl. Psychophysiol. Biofeedback **28**(3), 233–240 (2003)

A. Lecuyer, F. Lotte, R.B. Reilly, R. Leeb, M. Hirose, M. Slater, Brain-computer interfaces, virtual reality, and videogames. Computer **41**(10), 66–72 (2008). doi:10.1109/MC.2008.410

F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, B. Arnaldi et al., A review of classification algorithms for EEG-based brain–computer interfaces. J. Neural Eng. **4**, R1–R13 (2007)

F. Lotte, F. Larrue, C. Mühl, Flaws in current human training protocols for spontaneous brain-computer interfaces: lessons learned from instructional design. Front. Hum. Neurosci. **7**, 568 (2013)

S. Mason, A. Bashashati, M. Fatourechi, K. Navarro, G. Birch, A comprehensive survey of brain interface technology designs. Ann. Biomed. Eng. **35**, 137–169 (2007)

J.D. Millán, R. Rupp, G.R. Müller-Putz, R. Murray-Smith, C. Giugliemma, M. Tangermann, C. Vidaurre, F. Cincotti, A. Kübler, R. Leeb, et al., Combining brain–computer interfaces and assistive technologies: state-of-the-art and challenges. Front. Neurosci. **4**, 161 (2010). doi:10.3389/fnins.2010.00161

K. Müller, C.W. Anderson, G.E. Birch, Linear and nonlinear methods for brain-computer interfaces. IEEE Trans. Neural Syst. Rehabil. Eng. **11**(2), 165–169 (2003). doi:10.1109/TNSRE.2003.814484

G.R. Müller-Putz, R. Scherer, C. Brauneis, G. Pfurtscheller, Steady-state visual evoked potential (SSVEP)-based communication: impact of harmonic frequency components. J. Neural Eng. **2**(4), 123–130 (2005). doi:10.1088/1741-2560/2/4/008

G.R. Müller-Putz, R. Scherer, G. Pfurtscheller, C. Neuper, Temporal coding of brain patterns for direct limp control in humans. Front. Neurosci. **4**, 34 (2010). doi:10.3389/fnins.2010.00034

C. Neuper, G.R. Müller, A. Kübler, N. Birbaumer, G. Pfurtscheller, Clinical application of an EEG-based brain-computer interface: a case study in a patient with severe motor impairment. Clin. Neurophysiol. **114**(3), 399–409 (2003)

C. Neuper, R. Scherer, M. Reiner, G. Pfurtscheller, Imagery of motor actions: differential effects of kinesthetic and visual-motor mode of imagery in single-trial EEG. Brain Res. Cogn. Brain Res. **25**(3), 668–677 (2005). doi:10.1016/j.cogbrainres.2005.08.014

M. Ninaus, S.E. Kober, E.V.C. Friedrich, I. Dunwell, S. Freitas et al., Neurophysiological methods for monitoring brain activity in serious games and virtual environments: a review. Int. J. Technol. Enhanc. Learn. **6**(1), 78 (2014). doi:10.1504/IJTEL.2014.060022

E. Niedermeyer, The normal EEG of the waking adult, in *Electroencephalography: Basic Principles, Clinical Applications and Related Fields* (1999), Williams & Wilkins, Baltimore, pp. 149–173

A. Nijholt, D. Tan, Playing with your brain: brain-computer interfaces and games. In Proceedings of the international conference on Advances in computer entertainment technology (ACE '07). ACM, New York, pp 305–306 (2007)

P.L. Nunez, R. Srinivasan, *Electric Fields of the Brain: The Neurophysics of EEG*, 2nd edn. (Oxford University Press, New York, 2006)

D.G. Oblinger, Games and learning. Educ. Q. **29**(3), 5–7 (2006)

G. Pfurtscheller, F.H. Lopes da Silva, Event-related EEG/MRG synchronization and desynchronization: basic principles. Clin. Neurophysiol. **110**(11), 1842–1857 (1999). doi:10.1016/ S1388-2457(99)00141-8

G. Pfurtscheller, C. Neuper, Motor imagery and direct brain-computer communication. Proc. IEEE **89**(7), 1123–1134 (2001)

G. Pfurtscheller, G.R. Müller-Putz, R. Scherer, C. Neuper, Rehabilitation with brain-computer interface systems. Computer **41**(10), 58–65 (2008). doi:10.1109/MC.2008.432

J. Pirker, S. Berger, C. Gütl, J. Belcher, P.H. Bailey, Understanding physical concepts using an immersive virtual learning environment, in *Proceedings of the 2nd European Immersive Education Summit* (2012)

W. Samek, F.C. Meinecke, K.R. Müller, Transferring subspaces between subjects in brain–computer interfacing. IEEE Trans. Biomed. Eng. **60**(8), 2289–2298 (2013). doi:10.1109/TBME.2013.2253608

J. Schell, *The Art of Game Design: A Book of Lenses* (CRC Press, Boca Raton, 2014)

R. Scherer, A. Schlögl, F.Y. Lee, H. Bischof, D. Grassi, G. Pfurtscheller, The self-paced Graz brain-computer interface: methods and applications. Comput. Intell. Neurosci. **2007**, 79826 (2007)

R. Scherer, J. Faller, D. Balderas, E.V. Friedrich, M. Pröll, B. Allison, G. Müller-Putz, Brain–computer interfacing: more than the sum of its parts. Soft Comput. **17**(2), 317–331 (2013a)

R. Scherer, G. Moitzi, I. Daly, G.R. Müller-Putz, On the use of games for noninvasive EEG-based functional brain mapping. IEEE Trans. Comput. Intell. AI Games **5**(2), 155–163 (2013b). doi:10.1109/TCIAIG.2013.2250287

R. Scherer, J. Faller, E.V.C. Friedrich, E. Opisso, U. Costa, A. Kübler, G.R. Müller-Putz, Individually adapted imagery improves brain-computer interface performance in end-users with disability. PLoS One **10**(5), e0123727 (2015a). doi:10.1371/journal.pone.0123727

R. Scherer, M. Billinger, J. Wagner, A. Schwarz, D.T. Hettich, E. Bolinger, M. Lloria Garcia, J. Navarro, G.R. Müller-Putz, Thought-based row-column scanning communication board for individuals with cerebral palsy. Ann. Phys. Rehabil. Med. (2015b). doi:10.1016/j.rehab.2014.11.005

R. Scherer, A. Schwarz, G.R. Müller-Putz, V. Pammer-Schindler, M. Lloria Garcia, Game-based BCI training: interactive design for individuals with cerebral palsy, in *Proceedings of the IEEE SMC* (2015c), in press

M. Slater, V. Linakis, M. Usoh, R. Kooper, in *Immersion, Presence and Performance in Virtual Environments: An Experiment with Tri-Dimensional Chess*, ed. G. Mark, in Proceedings of the ACM symposium on Virtual reality software and technology (VRST), 1996, pp. 163–172

W.O. Tatum, B.A. Dworetzky, D.L. Schomer, Artifact and recording concepts in EEG. J. Clin. Neurophysiol. **28**(3), 252–263 (2011)

C. Vidaurre, C. Sannelli, K.R. Müller, B. Blankertz, Co-adaptive calibration to improve BCI efficiency. J. Neural Eng. **8**(2), 025009 (2011)

J. Wolpaw, N. Birbaumer, D.J. McFarland, G. Pfurtscheller, T.M. Vaughan, Brain-computer interfaces for communication and control. Clin. Neurophysiol. **113**, 767–791 (2002). doi:10.1016/S1388-2457(02)00057-3

Towards Serious Games for Improved BCI

9

Brent J. Lance, Jon Touryan, Yu-Kai Wang, Shao-Wei Lu, Chun-Hsiang Chuang, Peter Khooshabeh, Paul Sajda, Amar Marathe, Tzyy-Ping Jung, Chin-Teng Lin, and Kaleb McDowell

Contents

B.J. Lance (✉) • J. Touryan • A. Marathe • K. McDowell
Translational Neuroscience Branch, U.S. Army Research Laboratory, Aberdeen Proving Ground, Aberdeen, MD, USA
e-mail: brent.j.lance.civ@mail.mil; jonathan.o.touryan.civ@mail.mil; amar.marathe.civ@mail.mil; kaleb.g.mcdowell.civ@mail.mil

Y.-K. Wang • S.-W. Lu • C.-H. Chuang • C.-T. Lin
Brain Research Center, National Chiao Tung University, Hsinchu, Taiwan
e-mail: yukaiwang@cs.nctu.edu.tw; swlucifer@mail.nctu.edu.tw; chchuang@mail.nctu.edu.tw; ctlin@mail.nctu.edu.tw

P. Khooshabeh
Cognitive Sciences Branch, U.S. Army Research Laboratory, Aberdeen Proving Ground, Aberdeen, MD, USA
e-mail: peter.khooshabehadeh2.civ@mail.mil

P. Sajda
Department of Biomedical Engineering, Columbia University, New York, NY, USA
e-mail: psajda@columbia.edu

T.-P. Jung
Swartz Center for Computational Neuroscience, University of California San Diego, San Diego, CA, USA
e-mail: jung@sccn.ucsd.edu

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_4

Abstract

Brain-computer interface (BCI) technologies, or technologies that use online brain signal processing, have a great promise to improve human interactions with computers, their environment, and even other humans. Despite this promise, there are no current serious BCI technologies in widespread use, due to the lack of robustness in BCI technologies. The key neural aspect of this lack of robustness is human variability, which has two main components: (1) individual differences in neural signals and (2) intraindividual variability over time. In order to develop widespread BCI technologies, it will be necessary to address this lack of robustness. However, it is currently unknown how neural variability affects BCI performance. To accomplish these goals, it is essential to obtain data from large numbers of individuals using BCI technologies over considerable lengths of time. One promising method for this is through the use of BCI technologies embedded into games with a purpose (GWAP). GWAP are a game-based form of crowdsourcing which players choose to play for enjoyment and during which the player performs key tasks which cannot be automated but that are required to solve research questions. By embedding BCI paradigms in GWAP and recording neural and behavioral data, it should be possible to much more clearly understand the differences in neural signals between individuals and across different time scales, enabling the development of novel and increasingly robust adaptive BCI algorithms.

Keywords

Games with a purpose (GWAP) • Human variability • Brain-computer interface (BCI) • Electroencephalography (EEG)

Introduction

Brain-computer interface (BCI) technologies, or technologies that use online brain signal processing, have a great promise to improve human interactions with computers, their environment, and even other humans (Lance et al. 2012). In particular, BCI technologies have a strong likelihood for enhancing or revolutionizing several fields, including training and education, medicine, and communication technologies. Broadly speaking, there are four classes of BCI systems that utilize neural data to estimate or predict behavior: active, reactive, passive, and hybrid (Lance et al. 2012; Wolpaw et al. 2002; Zander and Kothe 2011). Active BCIs rely on neural signals intentionally induced by the operator, typically after some degree of training or associative mapping. These volitionally generated patterns of brain activity are then used as an input modality to directly control a computer, prosthetic limb, or some specific function of a given technological system (e.g., sensorimotor rhythms for control of virtual movement (McFarland et al. 2010)). Reactive BCIs rely on neural responses elicited by an external stimulus or event, which are then used as an input modality to a technological system (e.g., P300 spellers (Krusienski et al. 2006) and for multimedia image triage and search systems (Sajda et al. 2010; Pohlmeyer et al. 2011)). Passive BCIs, including cognitive monitoring systems, utilize implicit or ongoing neural responses for the purposes of detecting an operator's current cognitive or affective state. Indicators of fatigue (Balasubramanian et al. 2011), mental workload (Mühl et al. 2014), and emotional state (Bos et al. 2010) can be used as supplemental information to adaptively enhance or augment performance of human-system interactions. Finally, hybrid BCIs may utilize any combination of active, reactive, and passive approaches. In addition, hybrid BCIs may integrate other signals reflective of operator state, including eye movement dynamics or behavioral performance, as well as incorporate context awareness information obtained from the environment or technological system itself (Zander and Jatzev 2012; Pfurtscheller et al. 2010; Jangraw et al. 2014).

This chapter is written from the viewpoint of having a strong interest in the development of serious (i.e., primarily nonentertainment in purpose), nonmedical BCI technologies that improve the performance of healthy individuals. As a result, this chapter focuses on BCI technologies that are noninvasive, i.e., based on sensors which are not within the user's body. While some medical BCI technologies are invasive, i.e., based on sensors which are actually within the user's body, we consider noninvasive methods to be a more practical and user-acceptable approach for nonmedical BCIs. In addition, this chapter focuses on mobile neuroimaging technologies, primarily EEG. While the ability and precision of measuring brain activity have been dramatically improved in recent years due to the advances in neuroimaging technology, including functional magnetic resonance imaging (fMRI), magnetoencephalography (MEG), and functional near-infrared spectroscopy (fNIRS); electroencephalography (EEG) remains the most promising near-term method for realizing broadly available BCI. For all classes of BCI, the majority of approaches, algorithms, and systems continue to utilize EEG as the neuroimaging modality of choice.

BCI technologies have seen expanding usage primarily within the medical and entertainment domains (van Erp et al. 2012). For example, BCI technologies have been used in the development of medical technologies that provide a clinical population with a capability that they have lost, such as the BCI-controlled wheelchair (Vanacker et al. 2007) or P300 speller (Krusienski et al. 2006), which enables the disabled to slowly type via neural signals alone. They have also been applied to the development of entertainment technologies such as the Emotiv EPOC, a low-cost EEG headset designed to be used as a game controller, or the Force Trainer toy by Uncle Milton, which uses a single EEG electrode to control a fan and blow a light plastic ball through an obstacle course. Despite these successes, there are essentially no current BCI technologies in widespread use. This is because developing serious, nonmedical BCI technologies is a fundamentally different problem space than developing either medical or entertainment BCIs. For example, medical BCI technologies do not have to perform to the level of a healthy individual's motor control system to be useful. When comparing normal mouse movement to BCI-controlled mouse movement, Hochberg et al. (2006) found that users could reach a target on a screen 100 % of the time in under 1.3 s with a mouse but could only reach the target 75 % of the time in under 7 s using a BCI to control the mouse cursor. While this may be acceptable for some clinical populations, entertainment BCIs that perform poorly provide for frustrating gameplay (Dakan 2010), but have limited real-world effect. Serious, nonmedical BCI technologies will need to perform to a higher standard than either medical or entertainment BCI technologies and may also need to emphasize different approaches (see Box 1).

Box 1: A Brief Overview of Relevant BCI Paradigms

There are many different BCI paradigms, several of which are described elsewhere in this volume. The current most commonly used BCI paradigm is the motor imagery paradigm, where a subject attempts to directly control the movement of an effector by imagining the movement of an extremity such as a hand or a foot. We take the view that this paradigm is poorly suited to serious, nonmedical BCI technologies due to its poor performance when compared to existing control technologies (Hochberg et al. 2006). As a result, in this chapter, we focus on three different BCI paradigms: the steady-state visually evoked potential (SSVEP) paradigm, the rapid serial visual presentation (RSVP) paradigm, and a specific driving performance estimation paradigm. This section provides a brief overview of each of these paradigms, including analytical methods and potential applications.

The SSVEP paradigm is a reactive BCI paradigm and involves detecting which one of a number of objects the subject is attending to on a screen by rapidly flickering the objects at distinct frequencies and phases. While the subject's attention is focused on a specific object flickering at a specific frequency, a phase-locked increase in the spectral power of that frequency can be detected in the visual cortex. Even through using simple methods such as a fast Fourier transform (FFT) followed by classification, SSVEP can be detected fairly reliably, rapidly, and with minimal time spent training either the player or the classifier, due to its high signal-to-noise ratio (SNR). More advanced methods, such as canonical correlation analysis, allow for faster processing, a broader range of frequencies, and multiple stimuli at the same frequency but differing phases (Bin et al. 2009). One area in which SSVEP-based BCI technologies have been successfully used is in spelling applications that allow paralyzed patients to communicate.

The RSVP paradigm is also a reactive BCI paradigm, but it consists of subjects looking for targets in a rapidly presented stream of images. In RSVP, images are presented to a subject at a rate of about 5–10 Hz. These images

(continued)

either belong to a sparsely appearing target class or are distracter images which do not belong to that target class. When the subject sees an image in the target class, their brain will generate an evoked neural response which can be detected using a machine learning classifier. Successful approaches include ensembles of logistic regression classifiers or spatial filtering followed by binary classification. Generally, the target images can make up no more than 15–20 % of the total number of images. Otherwise, two things occur: First, the strength of the neural signal will begin to attenuate, and as a result, classifier performance will begin to degrade. Second, the subject will begin to increasingly make mistakes as a result of an effect known as the "attentional blink." What this means is that, if two target images occur too close together in time, the subject has a drastically increased probability of missing the second target. RSVP has two key strengths: first, the neural signal evoked when the subject sees a target can be recognized by a classifier with a high degree of accuracy. Second, the neural signal can be evoked by a variety of targets, leading to a potentially broad space of applications. One area in which RSVP has been successfully applied is in the domain of satellite image analysis. By using RSVP to pre-triage satellite images, targets such as airfields can be detected more rapidly than with manual search (Sajda et al. 2010). Figure 1 shows the neural response from an RSVP task. Within this neural response can be seen an SSVEP-like neural signal induced by the flashing of the images in the RSVP task.

While there are many passive BCI paradigms for neurally estimating performance at simulated driving tasks, we will be focusing on the paradigm described by Lin et al. (2005). In this task, the subject drives a car down a simulated road and attempts to stay in the lane through periodic but random "lane-deviation events," which consist of a lateral force that attempts to push the driver off of the road. The intended goal of this paradigm is to estimate when driving performance decrements are more likely to occur, potentially enabling them to be mitigated and thus decreasing the likelihood of accidents. These lane-deviation events provide a probe to evaluate the performance of the driver as their fatigue and attention fluctuate and as time-on-task increases.

In order for BCI technologies to achieve this higher standard of performance, they must be able to handle or adapt to human variability, both in terms of individual differences and intraindividual variability. Every individual is different, as are their neural signals, even those generated by individuals in similar mental states or being exposed to the same stimuli. In practice, what this means is that the feature extraction and classification methods underlying a BCI technology will perform very poorly unless they are calibrated or customized for each individual user. In fact, a percentage of the population may even be "BCI illiterate" in that BCI technologies cannot be effectively trained on the neural signals of those individuals

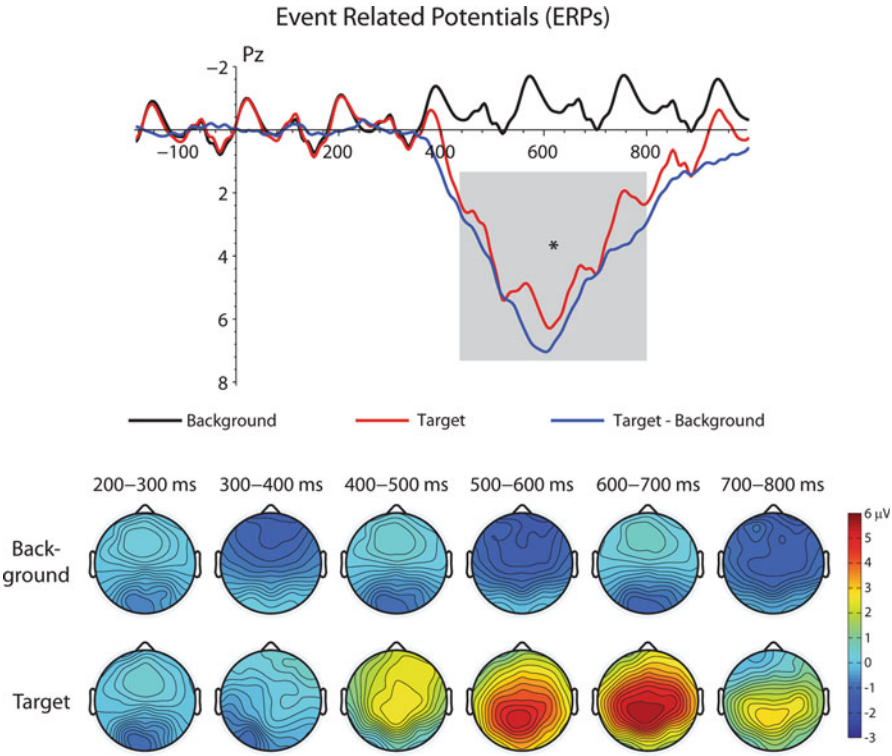


Fig. 1 Example neural signals from a 5-Hz RSVP task. In this task, a person is attempting to discriminate target images from background images in a continuous image stream. Target images are sparsely appearing images that the subject is tasked with identifying. Background images are commonly appearing images that the subject is tasked with ignoring. The three curves show 1 s of averaged neural response time-locked to image presentation. The *black line* shows neural response to a background image, the *red line* shows neural response to a target image, and the *blue line* shows the difference between the two. Slight but regular positive and negative deflections induced by image presentation which are similar to an SSVEP response can be seen in both the target and background neural responses. In addition, a large positive target-evoked inflection that peaks around 600 ms can be seen in the target neural response. The difference wave clarifies this target-evoked response by subtracting out the SSVEP-like response. The scalp maps show the location of the averaged neural responses as recorded on the scalp. Further description of this specific task and data can be found in Touryan et al. (2014)

(Pfurtscheller et al. 2010). People also have significant intraindividual variability, which can lead to differences in neural signals within the same individual during similar mental states or in reaction to the same stimuli at different points in time (Wu et al. 2010). Therefore, static feature extraction and classification methods that are customized to the individual user will still have large fluctuations in performance over time. However, there are currently few BCI technologies that attempt to adapt to the neural states of the individual users.

For serious, nonmedical BCI technologies to be viable, it will be crucial to develop algorithms that can address both aspects of human variability by adapting to the individual user and to the changes in their neural signals at multiple time scales. To develop these algorithms, it will be necessary to study how individual differences and variability in neural response affect BCI performance, which will require data from large numbers of individuals using BCI technologies over considerable lengths of time. One promising method for obtaining this needed data is through the use of BCI technologies embedded into *games with a purpose* (GWAP). As defined by von Ahn (2006), GWAP are a game-based form of human computation which players choose to play for enjoyment. As a side effect of the gameplay, in GWAP the player performs key tasks which cannot be automated but that are required to solve research questions (von Ahn 2006). In this case, the player will interact with BCI paradigms embedded into the gameplay while recording neural and behavioral data. Through the analysis of this data, it should be possible to much more clearly understand the differences in neural signals between individuals and across different time scales.

However, the development of BCI-based GWAP to study the effects of individual differences and intraindividual variability on BCI performance faces significant challenges. First, among these is the fact that the individual differences and intraindividual variability being studied will have significant negative effects on the performance of the BCI paradigms embedded in the GWAP. The challenge arises from the fact that forcing the player to continually use BCI paradigms that do not function due to low performance will make the game unenjoyable and non-immersive. As a result, the GWAP will not attract a large player base that spends significant time with the game, and there will not be sufficient data to make the GWAP worthwhile. In this chapter, we argue that this challenge can be addressed by making the BCI paradigms peripheral to the main gameplay experience and sufficiently rewarding to the player.

The remainder of this chapter will review the research on individual differences and human variability and their effects on the robustness of BCI technologies, the potential value of GWAP for long-term and large-scale data collection, and the key challenges for developing GWAP for collecting large-scale and long-term BCI data.

Human Variability and BCI Performance

As previously mentioned, humans are highly variable. People display significant individual differences, with each individual generating distinct neural signals, even under highly similar conditions. People also show strong intraindividual variation, with neural signals varying across similar conditions at multiple time scales. A hypothetical example of this can be seen in Fig. 2.

In this section, we will provide a review of human variability, including both individual differences and intraindividual variability, and how this variability affects BCI performance. We will focus on three distinct examples, showing how both individual differences and intraindividual variability affect the underlying neural signals that BCI technologies rely upon. First, we will discuss the effects

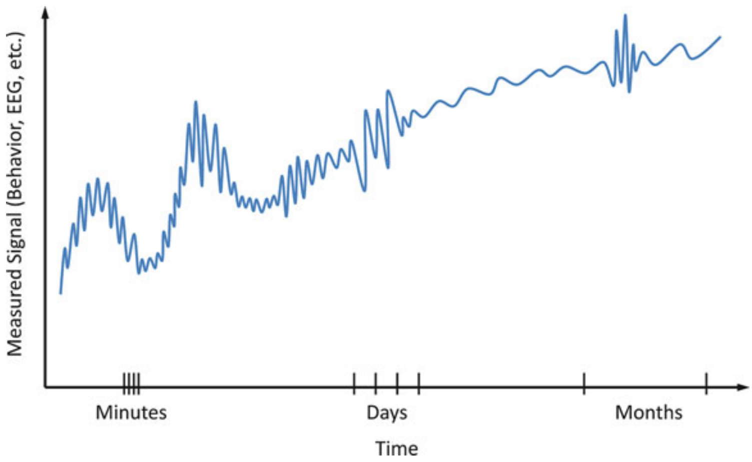


Fig. 2 Hypothetical example of intraindividual variability over time. Variability occurs in the course of minutes and hours, days and weeks, and even over months. Some of this fluctuation is internally driven, such as short-term fluctuations in attentional focus, while some of it is in response to changes in the external environment. Some of this variability is cyclic and predictable, such as circadian rhythm and sleep cycles; most of it is not. The cumulative result of this variability is neural signals that are nonstationary (i.e., the statistical properties of the signal change over time) and non-ergodic (i.e., the signal may not return to previously seen states), negatively impacting BCI performance

of individual differences and intraindividual variation on a three-image-class RSVP task (see Box 1). We will then discuss how the prior state of an individual affects their neural signals in a binary RSVP task and how this effect is modulated by individual differences. Finally, we will give an example of how individual differences and intraindividual variability affect the estimation of task performance from neural data.

Human Variability Example #1: Evoked Potentials

A specific example of both individual differences and intraindividual variability from three different people performing an RSVP task (see Box 1) with three categories of image stimuli, target, nontarget, and background images, can be seen in Fig. 3. In this task, target images are sparsely appearing images that the subject is tasked with identifying. Nontarget images are sparsely appearing images that are visually similar to the targets but which the subject is tasked with ignoring. Background images are commonly appearing images that the subject is also tasked with ignoring. In Fig. 3, the top rasters show the single-trial evoked neural responses time-locked to target images. Each horizontal row of the raster contains the single-trial neural response to one target appearance, and each vertical column contains one sample of EEG data. The color represents the amplitude of the neural response, with blue pixels representing lower-amplitude EEG recordings, red pixels

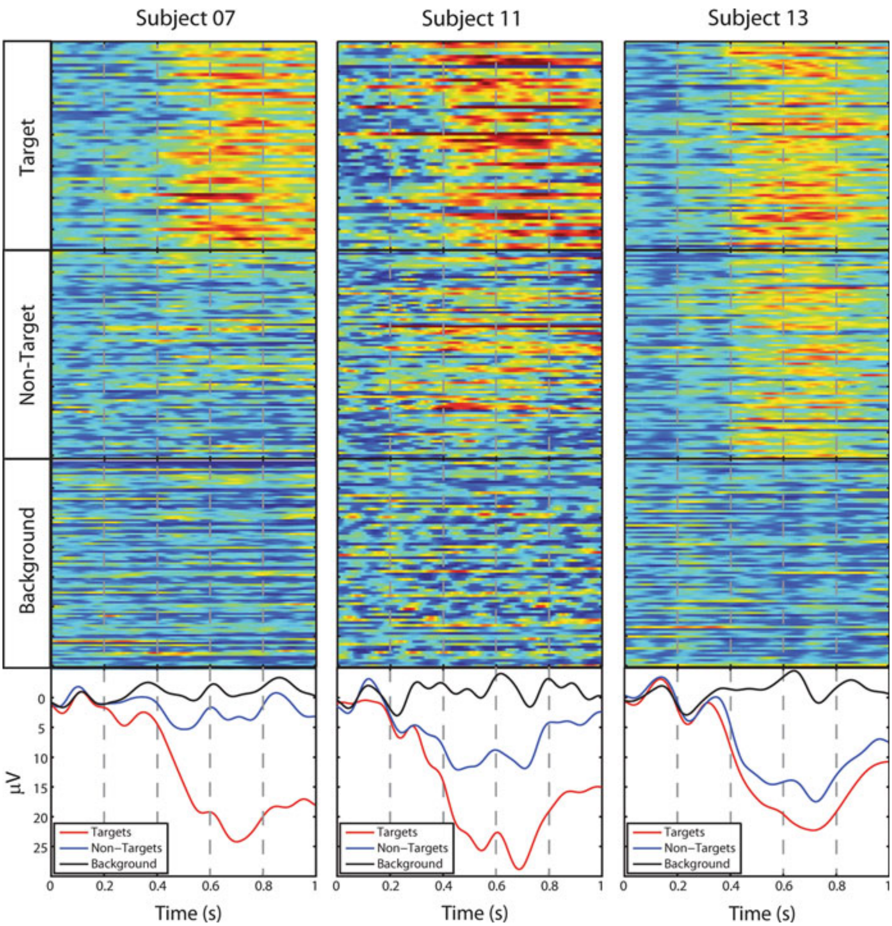


Fig. 3 An example of human variability in both single-trial evoked neural responses and averaged ERPs from an RSVP task. Each column contains the data from a different subject. Top three rows show rasters of the single-trial evoked neural responses time-locked to target, nontarget, and background images, respectively. Each horizontal row of the raster contains the single-trial neural response to one target appearance, and each vertical column contains one sample of EEG data. *Blue* pixels represent lower-amplitude EEG recordings, *red* pixels represent higher amplitude, and *green* and *yellow* represent intermediate amplitude. Bottom row shows the averaged event-related potentials (*ERPs*) from the three preceding rasters

representing higher amplitude, and green and yellow representing intermediate amplitude. Similarly, the next two rasters show the neural responses to nontarget and background images, respectively.

Finally, the bottom images show the averaged event-related potentials (ERPs) from the three preceding rasters. ERPs are a traditional method for handling signal variability in EEG by averaging the EEG signal that occurs immediately after a stimulus (Geisler et al. 1958). Given enough trials where the data is time-locked to stimulus onset, the averaging will smooth out the variation that is not phase-locked to the stimulus and leave only the common phase-locked neural response. The non-phase-locked variation removed by this approach is a combination of intraindividual variation and the low signal-to-noise ratio (SNR) of EEG.

In Fig. 3, intraindividual variability in the recorded neural activity can be seen in the row-to-row differences in both the peak amplitude of the neural signal and the poststimulus latency of that peak within each raster. This row-to-row variability causes the noisy appearance of the rasters. Clear individual differences between the amplitude, latency, and even the amount of variability in the neural signals can also be seen in the different subjects' rasters. For example, Subject 11 has considerably more variability in their rasters than does Subject 13. The averaged ERPs at the bottom of Fig. 3 also reveal a very interesting pattern of individual differences in the neural response to the nontarget images. Specifically, it shows that for some subjects (such as Subject 07), the neural response to the sparse nontarget images was indistinguishable from the response to the sparse target images. For other subjects (Subject 13), the neural response to the nontarget stimuli was indistinguishable from the background images, while other subjects (Subject 11) were between these end points.

Human Variability Example #2: Effect of Prior State on Evoked Neural Signals

As the state of an individual changes due to intraindividual variability, this state affects their neural signals in a way that is modulated by individual differences. The concept of a "state" in this instance is intentionally ambiguous. Consider fatigue as an example of state change. Over time, individuals will have levels of decreased and increased fatigue. Different levels of fatigue have clear effects on an individual's neural signals (Gevins et al. 1977). However, from the perspective of BCI development, coping with these effects is extremely difficult for several reasons. First, fatigue is a very loaded term, potentially being used to mean many different but interrelated states, including sleep deprivation, physical exhaustion, time on task, visual fatigue, mental fatigue, boredom, and inattention (Matthews and Desmond 2002; Jung et al. 1997). While each of these different states will affect the neural response to stimuli, the effects are often concurrent and not clearly distinguishable. In addition, they will have differing effects on different individuals and will even affect the same person differently based on additional factors such as the stressfulness or consequences of their current situation.

Further, BCI research tends to use a limited number of approaches to address fatigue. The most common approach is avoidance. Most BCI research is done with experimental sessions that are sufficiently short and far apart in order to minimize the effects of task-related fatigue. When this does not occur, another approach often used is to discard the data recorded while the subject was fatigued and focus exclusively on the period of time when the subject was alert. Approaches like these are used to minimize the effects of fatigue because the alternative is BCI

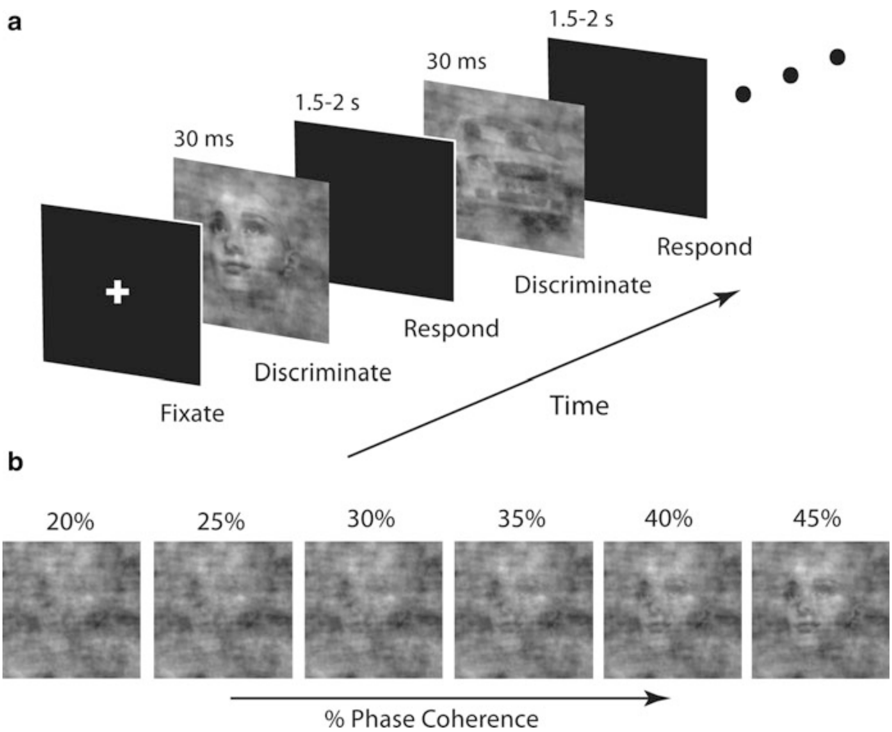


Fig. 4 Pre-stimulus cognitive state modulates sensory encoding of rapidly presented visual stimuli. Rapid serial visual presentation (*RSVP*) task in which subjects must classify the category of an object in an image corrupted by visual noise. The difficulty of the task is manipulated by controlling the noise level, through manipulating the images' phase coherence. Shown are three-phase coherence levels, with a lower coherence resulting in more noise in the image

systems that fail to perform reliably. As a result, the actual quantitative effect of fatigue on BCI algorithms and technologies is poorly understood.

Fatigue is not the only state that a person can be in which affects their neural signals. Many other factors of intraindividual variability, including emotion, stress, and prior experience, can modify neural signals and thereby affect BCI performance. It is not at all clear how these factors all affect neural signals independently and even less understood how they are interrelated. The interrelation and complexity of these factors mean that it can often be easier to consider abstract states, defined as changes in neural signal during a task.

Consider the example shown in Fig. 4, a modified RSVP task where the objective is to classify rapidly presented images masked with visual noise. In this task, the image is "flashed" for 30 ms, and the subject must respond with a button press to determine if the image was from category 1 (face) or from category 2 (car). The goal of the BCI experiment is to classify whether or not the user correctly identified the category of the image, using only the neural data. This is done by

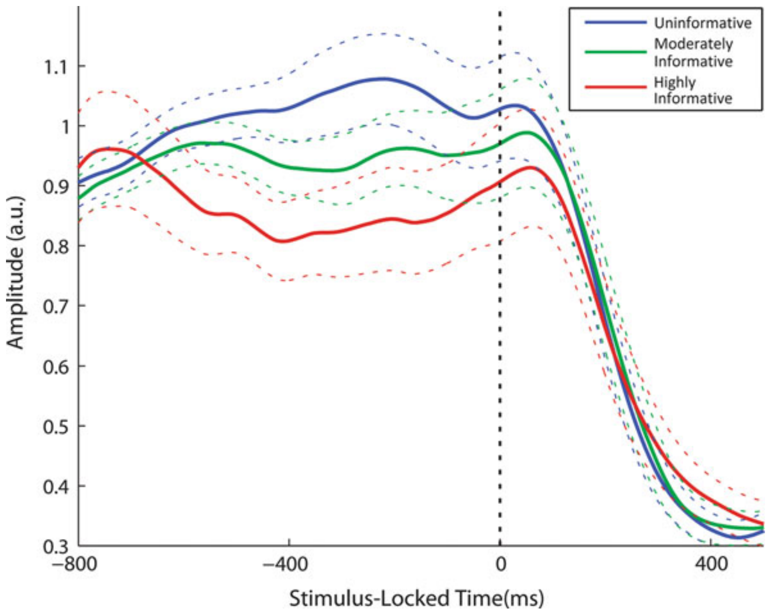


Fig. 5 Peak-to-peak amplitude of alpha power, grouped by the strength of the early poststimulus component. Greater strength of the early poststimulus component is associated with lower pre-stimulus alpha amplitude. Note that shortly after image presentation, alpha power drops sharply and is no longer useful for predicting targets. Instead, it becomes necessary to rely on the discriminating components identified using logistic regression (Reprinted from NeuroImage, 87, Lou et al., Prestimulus Alpha power predicts fidelity of sensory encoding in perceptual decision making, 242–251, Copyright (2014), with permission from Elsevier)

using a windowed logistic regression to identify "discriminating components" in the post-image presentation neural data that can be used to identify the image class at a statistically significant level (Lou et al. 2014). This process identified two primary discriminating components: one that is early (appearing approximately 170 ms after the image was displayed) and one that is late (appearing approximately 350 ms after image presentation).

Interestingly, the strength of these post-image presentation discriminating components is directly affected by the pre-image presentation state of the subject, as defined by the oscillations that occur in the alpha frequency band (7–13 Hz) prior to the display of the stimulus image. These oscillations may reflect idling processes in the brain, with occipital (visual) alpha potentially serving as an index of attention. High alpha power is associated with idling and a resulting reduction in visual attention, while low alpha power represents engagement with the visual stimulus. However, the spectral power of EEG frequencies is highly variable over time. Thus, a subject might have different behavioral and neural responses evoked by the exact same visual stimulus, depending in part on their pre-stimulus alpha state. Figure 5

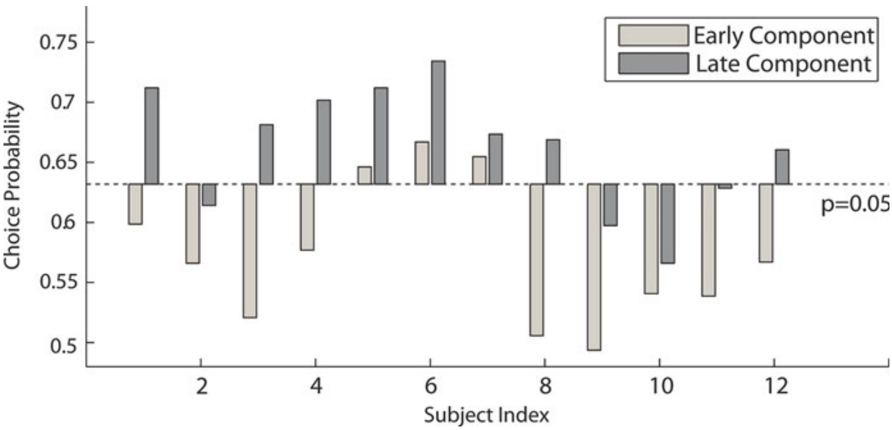


Fig. 6 Decoding of EEG results in two poststimulus evoked components that predict the subjects' choice: an early component and a late component. Shown is the decoding accuracy (chance is choice probability of 0.5) across 12 subjects for the two components. The early component, occurring around 170 ms poststimulus, is less of a predictor of the decision than the late component, which occurs around 350 ms poststimulus (Reprinted from NeuroImage, 87, Lou et al., Prestimulus Alpha power predicts fidelity of sensory encoding in perceptual decision making, 242–251, Copyright (2014), with permission from Elsevier)

illustrates the variability of pre-stimulus alpha oscillations as a function of the strength of the early poststimulus discriminating component. Trials where the pre-stimulus alpha amplitude is low will tend to result in highly informative poststimulus evoked responses (red curve in Fig. 5). Conversely, higher pre-stimulus alpha results in less informative evoked responses (blue curve in Fig. 5). This relationship between the pre-stimulus alpha and the evoked components is true for the early component but not the late component.

Figure 6 shows that, in addition to being affected by the state of the subject, the information content of the two discriminating components also varies substantially across the subjects. While the late component is more informative than the early component for every subject, for some subjects (e.g., S06), both components are significantly predictive. For other subjects (e.g., S09) neither is, and for some subjects (e.g., S01) only the late component significantly predicts performance.

Human Variability Example #3: Estimating Behavioral Performance

For the third example, consider a passive BCI that attempts to estimate performance at a task, based solely on neural signals (see Box 1). Intuitively, it seems like developing a system that can estimate behavioral performance at a task from neural signals should be straightforward. After all, the neural signals are the underlying driver that causes the behavioral task performance. However, due to the effects of human variability; the behavioral task performance, the neural data being used to estimate the behavior, and the relationship between the two are all widely variable, both within and across individuals. As a result of this variability, robustly predicting task performance using neural signals remains a promising but elusive goal.

Figure 7a demonstrates the variability of behavioral performance by showing the change in average relative reaction times to a lateral force across 80 90-min simulated driving sessions. This behavioral reaction time was smoothed using a causal 90-s bell-shaped moving average filter to eliminate variance at cycle lengths shorter than 1–2 min. The results follow a well-known trend of vigilance data, with three distinct states clearly visible in the data: initial near-perfect (relative RT = 1) performance begins to decay after about 1 min. Thereafter, reaction time to lane-deviation event rate rises steadily until ∼10 min into the task, after which it remains more or less stable near three to four times the optimal RT. While this group mean shows a strong time-on-task effect in this sustained attention task, the fluctuations in task performance during a single session could largely differ from the group trend and from those of other sessions.

Figure 7b shows that the behavior was highly variable both by individual and over time. For example, some subjects had poor performance in the first half of the session (e.g., S07), while others had exactly the opposite (e.g., S44). Similarly, the relationship between neural activity and behavior was also highly variable across subjects and over time. Figure 7c shows the correlations between EEG log power spectrum and behavioral performance at the simulated driving task at each EEG frequency. Some subjects had higher correlations between neural signals and behavior in the lower frequency bands (e.g., S49), while, for others, the higher frequency bands were more informative (e.g., S75). In addition, as shown in Fig. 7d, the relationship between EEG spectra and the behavioral reaction time varies as the subject performs the task. While increases in both the theta (4–7 Hz) and alpha (8–15 Hz) bands are generally predictive of periods of slower reaction time, they do not seem to predict just how slow the reaction time will be and, thus, how problematic the period of slow reaction times is for the driving task.

Across all three of these examples, both individual differences and intraindividual variability significantly affect the neural signals of BCI users. In terms of implications for robust and reliable BCI, this means that systems must be able to be calibrated to the individual user and must be able to adapt to the variability that occurs in both the individual's neural signals and task performance. These are known problems in many areas of science and engineering, and several techniques have been developed in order to address them, often by attempting to capitalize on the similarities between and across individuals.

These include machine learning approaches such as transfer learning, active learning, online learning, deep learning, and data mining approaches such as collaborative filtering. Transfer learning is a class of machine learning methods where data, features, or models from one domain are used to make predictions in another domain and which can be used to transfer information learned from one subject to help understand data collected from another subject or the same subject at a later time (Pan and Yang 2010). Active learning methods select the most

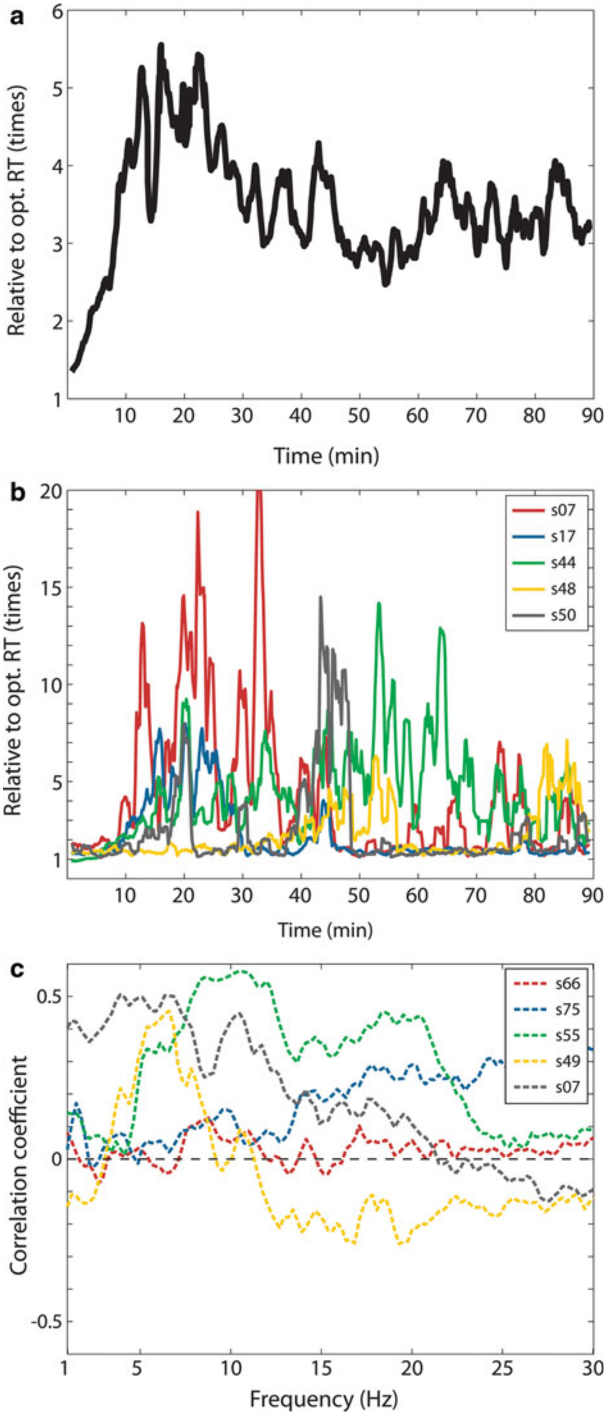


Fig. 7 (continued)

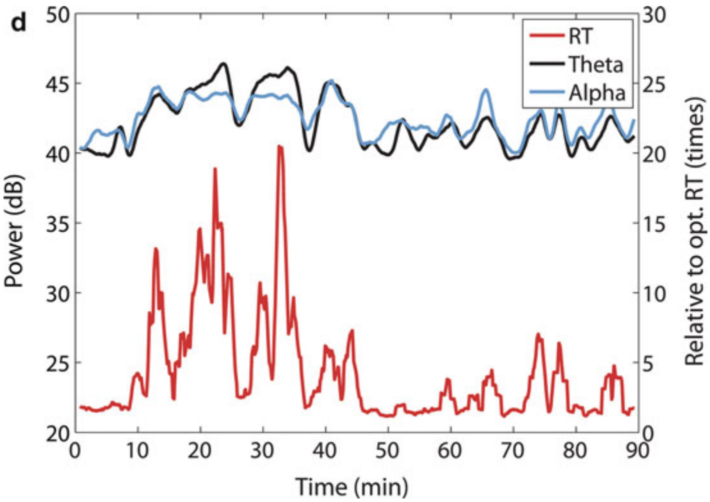


Fig. 7 (**a**) Group mean relative reaction time to lane-deviation events averaged across 80 90-min simulated driving experiments. (**b**) Relative reaction times to lane-deviation events across five example subjects, showing the differing variability of individual behavior. (**c**) Correlation between EEG spectral power and reaction time for five individual simulated driving sessions. Correlations are computed separately for 120 EEG frequencies between 0.98 and 30 Hz. (**d**) Reaction time, theta band power, and alpha band power over the course of one experimental session for subject S07

informative data for labeling and incorporation into a model and can be used to more rapidly calibrate systems to individual users (Settles 2010). Online and adaptive learning methods continually update as the system is used, attempting to adapt to the user. Deep Learning methods attempt to identify the most meaningful representation of the data and develop models based on that representation. Collaborative filtering approaches attempt to identify the data in a data set that is most relevant to a specific individual based on that individual's similarities and differences from a set of other known individuals.

One key aspect that unites all of these distinct methodologies is that their performance tends to increase as the amount of data available to train and infer upon increases. To date, however, the amount of data that is available from individuals using BCI technologies is highly limited, particularly with regard to serious, nonmedical BCI technologies. We argue that using these approaches to develop robust and reliable BCIs will require large amounts of data from large numbers of individuals using BCIs over long periods of time. In order to collect this large-scale and long-term data, it will be necessary to embed BCI paradigms and technologies into systems that people either need to use or will want to use. As a result, we argue that it should be possible to embed BCI paradigms into a video game in order to obtain the required data. Games have previously been utilized for both long-term and large-scale data collection. Often, these results are achieved through the use of *games with a purpose* or GWAP (Box 2).

Box 2: The Use of Games for Large-Scale and Long-Term Data Collection

Perhaps the earliest example of a successful large-scale GWAP is the ESP game, also known as the Google Image Labeler game (von Ahn 2006). In this game, two randomly selected players would be shown the same set of images. Their task was to provide as many labels for each image as possible within a short period of time, without communicating with the other player. If the labels provided by the two players matched, then both players received points. Players had profiles with persistent scores, and Google maintained high-score lists for the game. The image labels provided by these players served as the ground truth labels for searching the images with Google Image Search.

Another early example of large-scale data collection using a game was the deployment of the tracking real-time user experience or TRUE system (Kim et al. 2008), by Bungie Studios during the development of Halo 3. The TRUE system logged large quantities of data, including user interactions, contextual information, and attitudinal data from users who playtested Halo 3. While not technically a game with a purpose, the data collected using the TRUE system allowed identification of problematic and frustrating areas in the game, as well as improved fairness in multiplayer matches, resulting in an overall more enjoyable gaming experience (Thompson 2007).

A more recent game with a purpose is the Airport Scanner game for Android and iPhone, which simulates the task of searching for contraband using an airport X-ray scanner. To date, anonymous data from thousands of users performing millions of scans have been collected from players of the game. Researchers have used this data to explore questions about how well the airport baggage screening task can be performed. In particular, they found that the less often a target appeared, the less likely it was to be seen by the players. This was true even when looking at the high scoring subset of players (Mitroff and Biggs 2013). Interestingly, this low-probability, high-consequence detection problem is exactly what Transportation Security Administration (TSA) screeners encounter on a daily basis but is very hard to replicate in a relatively short laboratory experiment.

Games have also been used to study performance at cognitive tasks. In particular, Lumosity provides a web-based platform for simple games which are intended to be used for training various aspects of cognition. Recently, Lumosity has made the data collected by individuals playing these games available to researchers. This data set is the largest existing set of cognitive performance data, consisting of over 35 million users playing 600 million game trials. Researchers hope to use this data to obtain new insights into human cognition and how people think. Initial analyses reveal that aging and lifestyle factors such as amount of sleep taken and alcohol consumed significantly correlate with task performance (Sternberg et al. 2013).

(continued)

The final example here is the Foldit GWAP. Foldit is a collaborative online multiplayer protein-folding game (Cooper et al. 2010). Foldit players generate complex protein folds through directly manipulating protein structures and working with automated protein structure prediction technologies. Protein folding is an extremely complex problem, with even simple proteins having over 1000 degrees of freedom. It is also a problem that can have potentially high real-world payoff in the understanding of disease and in the development of novel drugs. One of the key successes of Foldit was in identifying the crystal structure of the retroviral protease (a key component of a retrovirus that influences viral maturation and proliferation) of the Mason-Pfizer monkey virus, which causes AIDS in monkeys. This structure had previously proved to be elusive to uncover through standard methods, and it is hoped that understanding this structure will lead to new antiviral treatments (Khatib et al. 2011).

Challenges in Developing GWAP for BCI

However, there is one primary challenge that must be overcome to develop a BCI-based GWAP. This is the challenge of developing an entertaining and engaging game such that large numbers of people will choose to play the game and use the BCI paradigms within the game in the face of the fact that the technical capability for robust and reliable BCIs is currently lacking. As a result, any BCI-based GWAP run the risk of either overemphasizing the BCI paradigms and being too frustrating and unpleasant to play or overemphasizing the gameplay and not providing viable data. Developing viable GWAP that incorporate BCI will require carefully balancing three key factors: the game must be *engaging* such that people will play for significant periods of time; BCI paradigms must be seamlessly *integrated* with the game while keeping gameplay entertaining and ensuring that sufficient experimental control is maintained; and sufficient information must be recorded to *contextualize* the resulting data so that it can be usefully interpreted. A hypothetical game that attempts to address all three of these challenges is shown in Fig. 8.

Developing Engaging Gameplay

In order to obtain the amount of data required to successfully address the issue of robustness and reliability in BCI, we will require the truly large-scale and long-term usage that only comes from intrinsic motivation, i.e., subjects play the game because they enjoy it. While there are an increasing number of BCI games available, none of them have obtained a large and dedicated audience sufficient to provide the needed large-scale and long-term data. One of the key reasons for

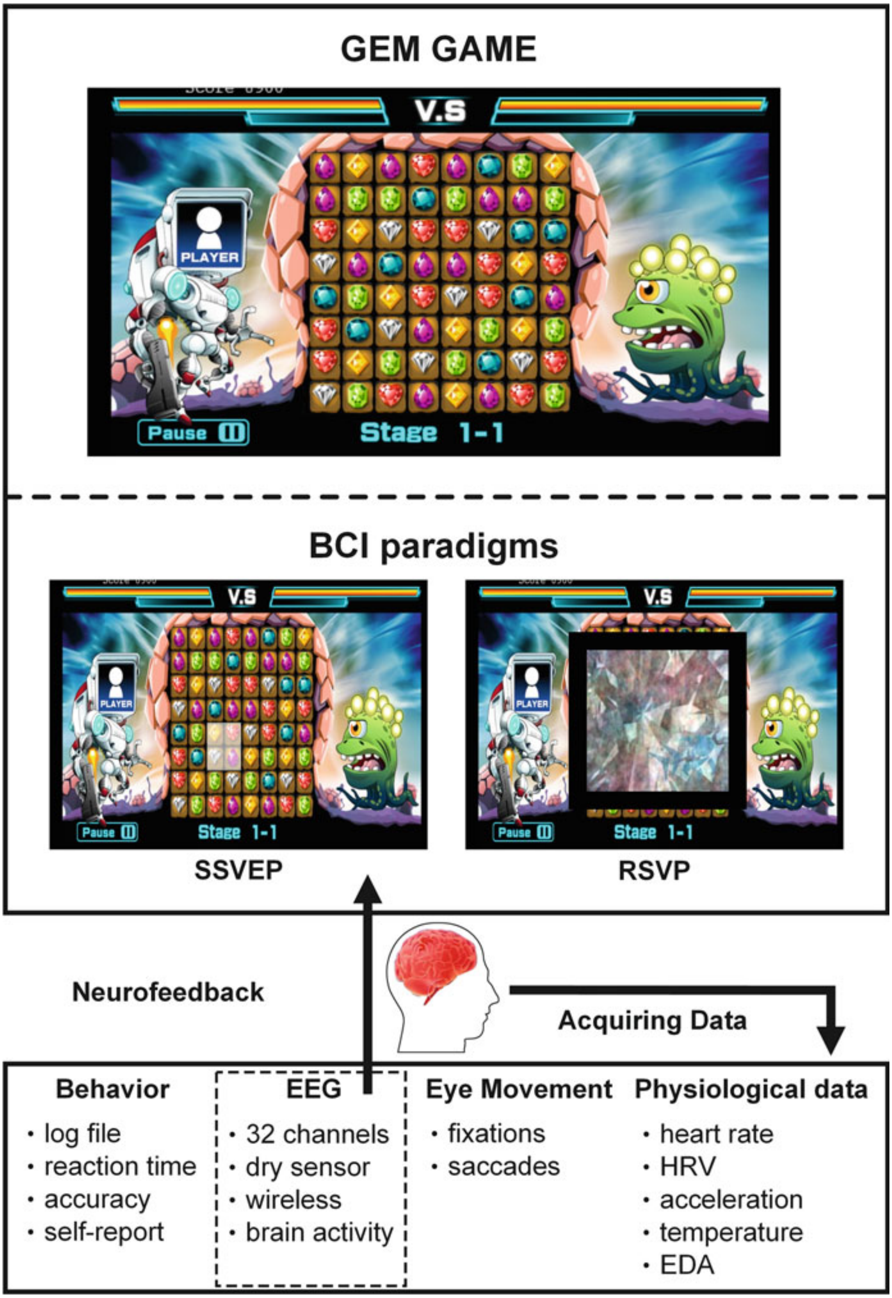


Fig. 8 A hypothetical game with a purpose for collecting long-term and large-scale BCI data. This hypothetical game is based upon an existing gameplay model which is known to be engaging, that of an opposed gem-matching game, similar to the Puzzle Quest series. The game includes multiple BCI paradigms that are integrated both mechanically and artistically with the game and records multimodal physiological and behavioral data for contextualization

this is that BCI games often focus on the BCI usage as the central aspect of the experience (Marshall et al. 2013).

Many of these games are focused on directly controlling the movement of objects on the screen through neural signals, either through the player directly imagining the movement of the object or through modulation of the spectral power within the alpha frequency band by attempting to feel more "relaxed" or "stressed." As the BCI paradigm is usually the central focus of the game, further aspects of gameplay and game design tend to be very limited. While these games can initially be highly entertaining, the novelty of the BCI wears off in a matter of minutes. Once this occurs, it becomes clear that BCI technologies currently lack the robustness and reliability required to be a primary human interface device, leading to gameplay which is frustrating and unenjoyable (Dakan 2010). In fact, if the user gets frustrated by a BCI with subpar functional performance, this frustration itself affects their neural signals and can further degrade functionality.

For BCI-based GWAP to be successful, the focus of the design should be on ensuring that the game is sufficiently engaging and enjoyable that players will be intrinsically motivated to play the game. Understanding the user experience is a crucial aspect of game design, and previous research shows a key relationship between engagement, user experiences, and increased play time (Hassenzahl and Tractinsky 2006). We argue that focusing the game design on engagement, entertainment, and immersion first gains several advantages. It will make it possible to create a game that can be fun for hours, making the long-term data collection much less painful on the subjects and thereby decreasing the likelihood of subjects dropping out of the experiment. It will also be much more likely to obtain a large player base that will willingly play for free, enabling large-scale data collection. One way to ease the development of a sufficiently engaging game is to utilize an existing model of gameplay.

Integrating BCIs with Gameplay: The Free-To-Play (F2P) Model

There is a large body of previous research focused on increasing engagement and immersion during gameplay, and there are many aspects of designing a game that is engaging or even addictive. One example of the criticality of engaging experiences is the development of free-to-play (F2P) games. F2P games are distributed online for free and make money by providing in-game items, areas, or capabilities which can be purchased. As a result, the game only makes money by drawing large numbers of people in and getting them engaged over the long term. There are lessons to be taken from the design of these F2P games on how to embed BCI usage within a game in such a way that both BCI usage and long-term gameplay are encouraged (Luban 2011; Luton 2013).

Under the F2P model, the use of the BCI can be viewed as the "price" that is paid to play the game. As a result, it should be treated similarly to purchasing upgrades in a standard F2P game. Thus, unlike the design of the core gameplay, which should emphasize engagement, the integration of the BCI paradigms should be focused on making the experience as painless and as rewarding for the player as possible. By this, we mean that the player should not have to go out of their way in order to activate the BCI paradigm and should receive a significant gameplay benefit for doing so. However, due to the robustness of current state-of-the-art BCI technologies, it is likely that the BCI paradigm will fail to work a significant percentage of the time. Therefore, the game must be designed such that it is possible for the player to circumvent the use of the BCI in the game, while still providing the required research data.

Consider how this could be done with the three BCI paradigms described in Box 1. For example, the RSVP paradigm is very similar to reward "slot machines" that are commonly used in F2P games. In these "slot machines," a small number of highly valuable in-game rewards are interspersed with a large number of low-value rewards that are rapidly displayed to the player, similar to a wheel on a slot machine. The player attempts to obtain a high-value reward by selecting it, which is difficult to accomplish due to the speed of the display. By simply adding a behavioral selection to the RSVP task and providing an in-game reward for selecting a target, it is possible to mirror a successful mechanism that commonly appears in F2P games. However, while this addresses the issue of integrating the RSVP paradigm into the game, it may not provide sufficient experimental control, depending on the goals of the experiment. For example, if the goal is to study the effects of changes in states such as fatigue on BCI performance, then one key problem that can arise is that of performance saturation. If the task is too easy, then all players will always perform it with near-perfect accuracy. When this happens, the fluctuations in behavior that would indicate low-performing states will simply not appear in the data, making it impossible to address the intended question. One possibility is to mask the targets and nontargets to make them more difficult to distinguish. Figure 8 shows a faceted masking, which fits artistically with the overall theme of the game, being used for this purpose.

There are many potential ways in which an SSVEP paradigm could be implemented in a game. Several targets could be displayed on-screen simultaneously, all flickering at different frequencies. The player could then use the SSVEP paradigm to select a desired reward or a target for an attack. However, in order to minimize the potential for frustration, the player must have the capability to circumvent a potentially nonfunctioning BCI. For example, the player could be allowed to click on the desired target. Unfortunately, while this integrates the SSVEP paradigm into the game, it does not provide sufficient experimental control. Instead, what is likely to happen is that the player will rapidly click the desired target without attending to it for the 1–2 s window that most SSVEP algorithms require. As a result, the neural data recorded will be insufficient for later analysis. Alternatively, the SSVEP paradigm could be interacted with through a mechanic where the player receives an award based on how quickly they click on the target but only once the target stops flickering. This would incentivize the player to pay close attention to the flickering target in order to maximize reward and ensure that the desired neural data of the player performing the SSVEP task would still be recorded.

Integrating the third BCI paradigm in Box 1, a passive behavioral performance estimation paradigm, is perhaps the simplest of the three BCI paradigms. As a passive BCI paradigm, no additional interaction with the player is needed, although if a specific task is being studied then that task could be embedded. Instead, what is required is extensive logging of the player's interaction with the game, as well as metrics and heuristics for estimating the quality of the player's moves. This data can then be used to study the relationship between neural activity and player behavior.

It is important to emphasize that the specific issues of maintaining experimental control when integrating the BCI paradigms with gameplay will vary based on the overall research goals being addressed. In addition, many of these issues, such as performance saturation, are not unique to any single paradigm but are relevant across many potential paradigms.

Finally, in order to encourage long-term gameplay, successful games provide for the player to experience task-reward feedback loops with multiple avenues for incremental progression. Example avenues for progression include difficulty, reward, capabilities, game mechanics, virtual space, and narrative space. Each of these also presents potential avenues for BCI integration. For example, there are usually various levels or difficulties in a game, and players will generally complete challenges in sequentially increasing difficulty (Brown and Cairns 2004). BCI paradigms can similarly be scaled in terms of difficulty. For example, both SSVEP and RSVP can be made more difficult by increasing the number of targets and distracters. This can also be tied to narrative progression, with difficulty or types of targets in a BCI also connected to sequential narrative episodes.

Contextualizing Data

Data recording and analysis is a key issue for F2P games, GWAP, and BCI implementation. In F2P games, games are often launched in a minimally playable state. Data from the game is then continuously monitored to determine key pieces of information such as what aspects of the game keep players engaged, what causes players to quit, and what drives players to perform in-game purchases. The game is then continually refined and updated based on this information (Luton 2013). When considering GWAP, in addition to monitoring this information, the experimental questions at the center of the GWAP must also be answered. As a result, additional data relevant to the experimental questions must also be recorded and analyzed (von Ahn and Dabbish 2008). Similarly, the analysis of the EEG data used for current BCI technologies requires a large amount of contextualization (Zander and Jatzev 2012). In order to estimate what the recorded neural signals may mean, it is necessary to know as much as possible about what the player is doing and how they are doing it.

As such, EEG data is not the only type of data that would need to be recorded. One key type of data needed is the information from the game itself, which includes player actions, game events, the times these occur, and contextual information from the state of the game that will serve as a basis for interpreting the other streams of data. Alternative sources of physiological data can also be used to contextualize the

EEG data and game data as well. For example, eye tracking technology provides the gaze direction in monitor coordinates and can be used to identify what aspect of the game the player is focusing on at any given time. While this and other physiological recording methods such as heart rate or electrodermal activity generally require laboratory-grade hardware, it is becoming increasingly possible to approximate these methods using common digital cameras. For example, pupil diameter, which can serve as an approximation for physiological arousal (Partala and Surakka 2003), can be detected using OpenCV, an open-source computer vision library (Schwarz et al. 2012). Similarly, heart rate can be detected using a webcam (Poh et al. 2011).

In order to uncover the correlates of longer-term state change, it will also be necessary to record self-report data from game players. Many of the correlates of this longer-term variation, such as sleep history and stress, cannot be derived from behavioral or physiological data while the subject is playing the game. However, some of these can be approximated using self-report (Buysse et al. 1989; Brantley et al. 1987).

One key challenge for recording this multimodal data is that the synchronization across the data streams must be highly accurate. As the quality of the synchronization decreases, it becomes increasingly difficult to associate the recorded neural signals with any external stimuli or behavior. For example, the amplitude of an averaged ERP is rapidly attenuated as synchronization becomes increasingly noisy. While other data streams may not degrade as rapidly with poor synchronization as EEG does, it will still become extremely difficult to make sense of inaccurately synchronized data during analysis.

Usable EEG Technologies

In addition to the challenges of developing an engaging game and integrating BCI paradigms into that game, there are also clear engineering challenges that need to be met in order to develop BCI-based GWAP which can obtain a large player base. The primary challenge is in the development of electroencephalography (EEG) recording devices which are low-cost, easy to use, and provide high signal quality (McDowell et al. 2013; Liao et al. 2012). While a BCI-based GWAP may still be of value without these devices, it will be limited in focus to single-session and longitudinal experiments with small subject populations.

One key usability issue is that current EEG technologies require a conductive NaCl-based gel or saline solution between the electrodes and skull to decrease the impedance and improve signal-to-noise. These gels and solutions serve as a bottleneck to using the BCI applications by requiring a time-consuming pregame application and postgame cleanup. Several research groups are developing dry EEG sensor technologies that would eliminate the need for this conductive gel while maintaining the reliable measurement of EEG signals (Sullivan et al. 2008; Grozea et al. 2011; Liao et al. 2014).

Likewise, traditional EEG systems transmit the measured brain activity through a cable which is connected between the EEG cap and computer. This greatly limits usability, and the swaying of the wires when the player moves their head causes additional artifact in the EEG signal. Wireless transmission using a protocol such as Bluetooth or Wi-Fi can overcome this limitation (Liao et al. 2014) and is easily extendable to tablet and mobile devices. While greatly adding to the usability of the EEG technology, using these wireless protocols may require additional power, raise bandwidth and packet loss issues, and increase the difficulty of properly synchronizing the EEG data.

While there are existing examples of commercially available dry-electrode, wireless EEG technologies (Sullivan et al. 2008), these systems tend to be too uncomfortable for long-term wear and to lack sufficient sensor density over a broad enough scalp distribution to examine different brain areas associated with a variety of cognitive functions. However, distinct types of BCI applications may relate to these cognitive functions and require collecting data from several different brain areas. For example, a motor imagery paradigm, where a BCI user attempts to control an object by imagining either its movement or their own, is usually based on the brain activity recorded from the motor cortex (Billinger et al. 2013). On the other hand, the data acquired from motor areas may be not fully appropriate for either SSVEP or RSVP paradigms. Instead, for these two types of BCI applications, the activity in the visual areas is usually measured and analyzed (Lin et al. 2014). For most research applications, the trend is to record the brain activity from the entire brain. While this may not be needed for BCI applications, it is unlikely that only one or two electrodes would be sufficient, although it is an open question as to how many electrodes are needed.

Conclusion

Performance issues in BCI technologies arise in a large part from human variability, including both individual differences and intraindividual variability, although there are other causes. In fact, in nearly every area of BCI, both individual differences and intraindividual variability have strong effects on the BCI users' neural data. This is not unique to BCI. There are existing approaches to deal with individual differences and intraindividual variability which may drastically improve the performance of BCI technologies by detecting and adapting to the individual user and their current mental state. However, using these approaches will require large amounts of data from people using BCI technologies over long periods of time.

We propose that one method for obtaining this data is through the development of BCI-based games with a purpose (GWAP) that not only are focused on obtaining the required BCI data but are also engaging and entertaining enough that large numbers of users will willingly play them. As a result, these BCI games which are expressly designed for entertainment and engagement have the potential to revolutionize research into a broad range of serious BCI applications.

There are critical challenges that must be overcome to develop BCI-based GWAP. The most pressing challenge is that of developing engaging BCI-based games when BCI technologies have such low levels of performance. As a result, developing these GWAP will require carefully balancing three key factors: developing an engaging and entertaining game, seamlessly integrating BCI paradigms with the game, and properly contextualizing the resulting data to enable useful interpretation and analysis. However, in order for data from a large number of users to be obtained, EEG systems that are comfortable, easy-to-use, low-cost, and that provide high-quality data will be required. A BCI-based GWAP would still be of value without this EEG technology, as currently available EEG systems could be used to obtain data from fewer participants over long periods of time. As a result, whether these EEG technologies come in the near future or are longer-term technologies, the future development of serious BCI applications for nonmedical applications would benefit greatly from large-scale and long-term collection of data from individuals performing BCI-relevant tasks in GWAP.

Recommended Reading

V. Balasubramanian, K. Adalarasu, A. Gupta, EEG based analysis of cognitive fatigue during simulated driving. Int. J. Ind. Syst. Eng. **7**(2), 135–149 (2011). doi:10.1504/IJISE.2011.038563

M. Billinger, C. Brunner, G.R. Müller-Putz, Single-trial connectivity estimation for classification of motor imagery data. J. Neural Eng. **10**(4), 046006 (2013). doi:10.1088/1741-2560/10/4/046006

G. Bin, X. Gao, Y. Zheng, B. Hong, S. Gao, An online multi-channel SSVEP-based brain–computer interface using a canonical correlation analysis method. J. Neural Eng. **6**(4), 046002 (2009). doi:10.1088/1741-2560/6/4/046002

D.P.-O. Bos, B. Reuderink, B. van de Laar, H. Gürkök, C. Mühl, M. Poel, A. Nijholt, D. Heylen, Brain-computer interfacing and games, in *Brain-Computer Interfaces*, ed. by D.S. Tan, A. Nijholt. Human-computer interaction series (Springer, London, 2010), pp. 149–178. doi:10.1007/978-1-84996-272-8_10

P.J. Brantley, C.D. Waggoner, G.N. Jones, N.B. Rappaport, A daily stress inventory: development, reliability, and validity. J. Behav. Med. **10**(1), 61–73 (1987)

E. Brown, P. Cairns, A grounded investigation of game immersion, in *CHI'04 Extended Abstracts on Human Factors in Computing Systems (CHI EA'04)* (ACM, New York, 2004), pp. 1297–1300. doi:10.1145/985921.986048

D.J. Buysse, C.F. Reynolds III, T.H. Monk, S.R. Berman, D.J. Kupfer, The Pittsburgh sleep quality index: a new instrument for psychiatric practice and research. Psychiatry Res. **28**(2), 193–213 (1989)

S. Cooper, F. Khatib, A. Treuille, J. Barbero, J. Lee, M. Beenen, A. Leaver-Fay, D. Baker, Z. Popović, F. Players, Predicting protein structures with a multiplayer online game. Nature **466**(7307), 756–760 (2010). doi:10.1038/nature09304

R. Dakan, Review: Emotiv EPOC, tough thoughts on the new mind-reading controller. Joystiq (2010). 27 Jan. http://www.joystiq.com/2010/01/27/review-emotiv-epoc-tough-thoughts-on-the-new-mind-reading-cont/

C.D. Geisler, L.S. Frishkopf, W.A. Rosenblith, Extracranial responses to acoustic clicks in man. Science **128**(3333), 1210–1211 (1958). doi:10.1126/science.128.3333.1210

A.S. Gevins, G.M. Zeitlin, S. Ancoli, C.L. Yeager, Computer rejection of EEG artifact. II. Contamination by drowsiness. Electroencephalogr. Clin. Neurophysiol. **43**(1), 31–42 (1977)

C. Grozea, C.D. Voinescu, S. Fazli, Bristle-sensors – low-cost flexible passive dry EEG electrodes for neurofeedback and BCI applications. J. Neural Eng. **8**(2), 025008 (2011)

M. Hassenzahl, N. Tractinsky, User experience – a research agenda. Behav. Inform. Technol. **25**(2), 91–97 (2006). doi:10.1080/01449290500330331

L.R. Hochberg, M.D. Serruya, G.M. Friehs, J.A. Mukand, M. Saleh, A.H. Caplan, A. Branner, D. Chen, R.D. Penn, J.P. Donoghue, Neuronal ensemble control of prosthetic devices by a human with tetraplegia. Nature **442**(7099), 164–171 (2006). doi:10.1038/nature04970

D.C. Jangraw, J. Wang, B.J. Lance, S.-F. Chang, P. Sajda, Neurally and ocularly informed graph-based models for searching 3D environments. J. Neural Eng. **11**(4), 046003 (2014). doi:10.1088/1741-2560/11/4/046003

T.-P. Jung, S. Makeig, M. Stensmo, T.J. Sejnowski, Estimating alertness from the EEG power spectrum. IEEE Trans. Biomed. Eng. **44**(1), 60–69 (1997). doi:10.1109/10.553713

F. Khatib, F. DiMaio, S. Cooper, M. Kazmierczyk, M. Gilski, S. Krzywda, H. Zabranska et al., Crystal structure of a monomeric retroviral protease solved by protein folding game players. Nat. Struct. Mol. Biol. **18**(10), 1175–1177 (2011)

J.H. Kim, D.V. Gunn, E. Schuh, B. Phillips, R.J. Pagulayan, D. Wixon, Tracking real-time user experience (TRUE): a comprehensive instrumentation solution for complex systems, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (ACM, 2008), pp. 443–452. http://dl.acm.org/citation.cfm?id=1357126

D.J. Krusienski, E.W. Sellers, F. Cabestaing, S. Bayoudh, D.J. McFarland, T.M. Vaughan, J.R. Wolpaw, A comparison of classification techniques for the P300 speller. J. Neural Eng. **3**, 299–305 (2006). doi:10.1088/1741-2560/3/4/007

B.J. Lance, S.E. Kerick, A.J. Ries, K.S. Oie, K. McDowell, Brain-computer interface technologies in the coming decades. Proc. IEEE **100**(Special Centennial Issue), 1585–1599 (2012). doi:10.1109/JPROC.2012.2184830

L.-D. Liao, C.-T. Lin, K. McDowell, A.E. Wickenden, K. Gramann, T.-P. Jung, L.-W. Ko, J.-Y. Chang, Biosensor technologies for augmented brain–computer interfaces in the next decades. Proc. IEEE **100**(Special Centennial Issue), 1553–1566 (2012). doi:10.1109/JPROC.2012.2184829

L.-D. Liao, S.-L. Wu, C.-H. Liou, S.-W. Lu, S.-A. Chen, S.-F. Chen, L.-W. Ko, C.-T. Lin, A novel 16-channel wireless system for electroencephalography measurements with dry spring-loaded sensors. IEEE Trans. Instrum. Meas. **63**(6), 1545–1555 (2014). doi:10.1109/TIM.2013.2293222

C.-T. Lin, R.-C. Wu, S.-F. Liang, W.-H. Chao, Y.-J. Chen, T.-P. Jung, EEG-based drowsiness estimation for safety driving using independent component analysis. IEEE Trans. Circuits Syst. Regul. Pap. **52**(12), 2726–2738 (2005). doi:10.1109/TCSI.2005.857555

Y.-P. Lin, Y. Wang, T.-P. Jung, Assessing the feasibility of online SSVEP decoding in human walking using a consumer EEG headset. J. Neuroeng. Rehabil. **11**(1), 119 (2014). doi:10.1186/1743-0003-11-119

B. Lou, Y. Li, M.G. Philiastides, P. Sajda, Prestimulus alpha power predicts fidelity of sensory encoding in perceptual decision making. Neuroimage **87**, 242–251 (2014). doi:10.1016/j.neuroimage.2013.10.041

P. Luban. The design of free-to-play games. Gamasutra (2011). 22 Nov. http://www.gamasutra.com/view/feature/134920/the_design_of_freetoplay_games_.php

W. Luton, *Free-to-Play: Making Money From Games You Give Away*, 1st edn. (New Riders, 2013)

D. Marshall, D. Coyle, S. Wilson, M. Callaghan, Games, gameplay, and BCI: the state of the art. IEEE Trans. Comput. Intell. AI Games **5**(2), 82–99 (2013). doi:10.1109/TCIAIG.2013.2263555

G. Matthews, P.A. Desmond, Task-induced fatigue states and simulated driving performance. Q. J. Exp. Psychol. A Hum. Exp. Psychol. **55**(2), 659–686 (2002). doi:10.1080/02724980143000505

K. McDowell, C.-T. Lin, K.S. Oie, J. Tzyy-Ping, S. Gordon, K.W. Whitaker, S.-Y. Li, S.-W. Lu, W.D. Hairston, Real-world neuroimaging technologies. IEEE Access **1**, 131–149 (2013). doi:10.1109/ACCESS.2013.2260791

D.J. McFarland, W.A. Sarnacki, J.R. Wolpaw, Electroencephalographic (EEG) control of three-dimensional movement. J. Neural Eng. **7**(3), 036007 (2010). doi:10.1088/1741-2560/7/3/036007

S.R. Mitroff, A.T. Biggs, Mitroff, Biggs, The ultra-rare-item effect visual search for exceedingly rare items is highly susceptible to error. Psychol. Sci. (2013). doi:10.1177/0956797613504221

C. Mühl, C. Jeunet, F. Lotte, EEG-based workload estimation across affective contexts. Neuroprosthetics **8**, 114 (2014). doi:10.3389/fnins.2014.00114

S.J. Pan, Q. Yang, A survey on transfer learning. IEEE Trans. Knowl. Data Eng. **22**(10), 1345–1359 (2010)

T. Partala, V. Surakka, Pupil size variation as an indication of affective processing. Int. J. Hum. Comput. Stud. **59**(1), 185–198 (2003)

G. Pfurtscheller, B.Z. Allison, C. Brunner, G. Bauernfeind, T. Solis-Escalante, R. Scherer, T.O. Zander, G. Mueller-Putz, C. Neuper, N. Birbaumer, The hybrid BCI. Front. Neurosci. **4** (2010). doi:10.3389/fnpro.2010.00003

M.-Z. Poh, D.J. McDuff, R.W. Picard, Advancements in noncontact, multiparameter physiological measurements using a webcam. IEEE Trans. Biomed. Eng. **58**(1), 7–11 (2011)

E.A. Pohlmeyer, J. Wang, D.C. Jangraw, B. Lou, S.-F. Chang, P. Sajda, Closing the loop in cortically-coupled computer vision: a brain–computer interface for searching image databases. J. Neural Eng. **8**(3), 036025 (2011)

P. Sajda, E. Pohlmeyer, J. Wang, L.C. Parra, C. Christoforou, J. Dmochowski, B. Hanna, C. Bahlmann, M.K. Singh, S.-F. Chang, In a blink of an Eye and a switch of a transistor: cortically coupled computer vision. Proc. IEEE **98**(3), 462–478 (2010). doi:10.1109/JPROC.2009.2038406

L. Schwarz, Humberto Remigio Gamba, Fabio Cabral Pacheco, Rodrigo Belisario Ramos, Miguel Antonio Sovierzoski, Pupil and iris detection in dynamic pupillometry using the OpenCV library, in *2012 5th International Congress on Image and Signal Processing* (*CISP*) (2012), pp. 211–215. doi: 10.1109/CISP.2012.6469846

B. Settles, in *Active Learning Literature Survey,* vol. 52 (University of Wisconsin, Madison, 2010), pp. 55–66

D.A. Sternberg, K. Ballard, J.L. Hardy, K. Benjamin, P.M. Doraiswamy, M. Scanlon, The largest human cognitive performance dataset reveals insights into the effects of lifestyle factors and aging. Front. Hum. Neurosci. **7**, 292 (2013). doi:10.3389/fnhum.2013.00292

T.J. Sullivan, S.R. Deiss, T.-P. Jung, G. Cauwenberghs, A brain-machine interface using dry-contact, low-noise EEG sensors, in *Circuits and Systems, 2008. ISCAS 2008. IEEE International Symposium on* (IEEE, 2008), pp. 1986–1989. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4541835

C. Thompson. Halo 3: how Microsoft labs invented a new science of play. http://archive.wired.com/gaming/virtualworlds/magazine/15-09/ff_halo?currentPage=all Mag. **15**(9) (2007). http://josquin.cti.depaul.edu/~rburke/courses/f07/gam224/doc/science_of_play.pdf

J. Touryan, G. Apker, B.J. Lance, S.E. Kerick, A.J. Ries, K. McDowell, Estimating endogenous changes in task performance from EEG. Neuroprosthetics **8**, 155 (2014). doi:10.3389/fnins.2014.00155

J. Van Erp, F. Lotte, M. Tangermann, Brain-computer interfaces: beyond medical applications. Computer **45**(4), 26–34 (2012). doi:10.1109/MC.2012.107

G. Vanacker, R. del José, E.L. Millán, P.W. Ferrez, F.G. Moles, J. Philips, H. Van Brussel, M. Nuttin, Context-based filtering for assisted brain-actuated wheelchair driving. Comput. Intell. Neurosci. **2007**, 3 (2007). doi:10.1155/2007/25130

L. Von Ahn, Games with a purpose. Computer **39**(6), 92–94 (2006). doi:10.1109/MC.2006.196

L. Von Ahn, L. Dabbish, Designing games with a purpose. Commun. ACM **51**(8), 58–67 (2008). doi:10.1145/1378704.1378719

J.R. Wolpaw, N. Birbaumer, D.J. McFarland, G. Pfurtscheller, T.M. Vaughan, Brain–computer interfaces for communication and control. Clin. Neurophysiol. **113**(6), 767–791 (2002). doi:10.1016/S1388-2457(02)00057-3

D.D. Wu, C.G. Courtney, B.J. Lance, S.S. Narayanan, M.E. Dawson, K.S. Oie, T.D. Parsons, Optimal arousal identification and classification for affective computing using physiological signals: virtual reality stroop task. IEEE Trans. Affect. Comput. **1**(2), 109–118 (2010). doi:10.1109/T-AFFC.2010.12

T.O. Zander, S. Jatzev, Context-aware brain–computer interfaces: exploring the information space of user, technical system and environment. J. Neural Eng. **9**(1), 016003 (2012). doi:10.1088/1741-2560/9/1/016003

T.O. Zander, C. Kothe, Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general. J. Neural Eng. **8**(2), 025005 (2011). doi:10.1088/1741-2560/8/2/025005

User-Centered BCI Videogame Design 10

Emilie Loup-Escande, Fabien Lotte, Guillaume Loup, and Anatole Lécuyer

Contents

E. Loup-Escande (✉)
CRP-CPO (EA7273), Université de Picardie Jules Verne, Amiens, France
e-mail: emilie.loup-escande@u-picardie.fr

F. Lotte
INRIA, Centre de Recherche Bordeaux Sud-Ouest, Talence Cedex, France
e-mail: fabien.lotte@inria.fr

G. Loup
LIUM, Université du Maine, Laval Cedex 9, France
e-mail: guillaume.loup@univ-lemans.fr

A. Lécuyer
INRIA, Centre de Recherche Rennes Bretagne-Atlantique, Campus de Beaulieu, Rennes Cedex, France
e-mail: anatole.lecuyer@inria.fr

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_3

Abstract

This chapter aims to offer a user-centered methodological framework to guide the design and evaluation of Brain-Computer Interface videogames. This framework is based on the contributions of ergonomics to ensure these games are well suited for their users (i.e., players). It provides methods, criteria, and metrics to complete the different phases required by a human-centered design process. This aims to understand the context of use, specify the user needs, and evaluate the solutions in order to define design choices. Several ergonomic methods (e.g., interviews, longitudinal studies, user-based testing), objective metrics (e.g., task success, number of errors), and subjective metrics (e.g., mark assigned to an item) are suggested to define and measure the usefulness, usability, acceptability, hedonic qualities, appealingness, emotions related to user experience, immersion, and presence to be respected. The benefits and contributions of the user-centered framework for the ergonomic design of these Brain-Computer Interface videogames are discussed.

Keywords

Brain-Computer Interface • Videogame • Ergonomics • User-Centred Design • Evaluation

Introduction

A BCI videogame combines a Brain-Computer Interface (BCI) and videogame content. It mainly consists in using a BCI as one of the input devices that can interact with the game. This BCI can be used in an active way to send control commands (e.g., Lotte et al. 2008), in a passive way to adapt the game content based on the user's brain activity (e.g., George and Lécuyer 2010), or both. In a BCI videogame, the user can play by means of brain activity, and the available output device (mainly visual feedback) is used to provide a meaningful feedback to the user (Lotte et al. 2013). As explained in Scherer et al. (2013), a game, including a BCI game, is a problem-solving activity performed with a playful attitude. The objectives of BCI videogames are multiple. Indeed, they can be used in the medical area: for example, Holz et al. (2013) suggested the BCI videogame "Connect-Four," which aims to entertain the severely motor-restricted end users in their daily life in order to increase their quality of life and decrease their depression. Connect-Four is "a strategic videogame with two competitive players in which coins have to be placed in rows and columns with the goal to connect four coins" (Holz et al. 2013, p. 113). A BCI videogame can also be a learning tool: for example, as a motivating and engaging way to learn to use a BCI, to later use it to control neuroprosthetics, as in Müller-Putz et al. (2007). Obviously, a BCI videogame can be an entertainment app: for example, “Mind the Sheep!” is a single- or multiplayer videogame which consists to fence sheep in as quickly as possible by herding them with dogs (Gürkök et al. 2013).

To design and evaluate these BCI videogames, two complementary approaches coexist: a technocentric design and an anthropocentric design.

The technocentric orientation aims to design and optimize an innovative application by testing its technological possibilities and solving technical challenges. This research path hardly considers the characteristics of human activity. Typically, these studies aim to improve the performance of the BCI at detecting mental states, in terms of information transfer rate (Yang et al. 2010), classification accuracy, or error rates, for instance.

The anthropocentric orientation aims to design an application which will be useful and usable by their end users. This approach focuses on the characteristics, capabilities, and resources of end users; the context of use of the designed application; and the users’ activity (Rabardel and Béguin 2005). In the context of BCI applications, the user-centered approach has already been introduced (e.g., Kübler et al. 2013; Schreuder et al. 2013). However, its use is more recent for the design of BCI videogames (e.g., Plass-Oude Bos et al. 2011; Van de Laar et al. 2011). This approach remains minor compared to the technocentric approach. One reason could be the absence of a methodological framework dealing with the ergonomic design of BCI videogames, compared to the research conducted in technocentric design.

The purpose of this chapter is to propose a user-centered methodological framework to guide the design and evaluation of BCI videogames. This methodological framework deals with the ergonomic design of these applications and suggests methods, criteria, and metrics.

The remainder of this chapter is organized as follows. In the following section, the relevance of the user-centered design for BCI videogames is explained. The third section details the ergonomic methods which are used or advocated in the context of user-centered design of BCI videogames. The fourth section describes which ergonomic criteria could be introduced to user-centered design BCI videogames, and proposes metrics for each criterion. The contribution of the proposed methodological framework to the user-centered design of BCI videogames is discussed in the fifth section. The sixth section provides a general summary and conclusion.

User-Centered Design and Design of BCI Videogames

Designing emerging applications suited to end users requires design models from the ergonomic literature. Among these models, a user-centered design is particularly promising for BCI videogames. Indeed, a BCI videogame is an interactive system that could benefit from the “user-centred design” model as any interactive system. This model is formalized in the ISO 9241-210 standard through four iterative phases: (1) understand and specify the context of use, (2) specify the user needs and the other stakeholders' requirements, (3) produce design solutions (e.g., scenario, mock-up, prototype), and (4) evaluate the solutions at all stages in the project from early concept design to long-term use to specify design choices. Consequently, it can be used to guide designers toward an anthropocentric design of these specific videogames.

Moreover, the term "user-centred design" is already known in the BCI community to evoke a concept which is assessed empirically (i.e., evaluated by a user during the accomplishment of an experimental task). For example, in their study, Plass-Oude Bos et al. (2011) tried to evaluate users' preference for three mental tasks (inner speech, association, stressed or relaxed mental state) which are more adapted than traditional paradigms (e.g., Motor Imagery, P300, and Steady-State Visually Evoked Potentials) considered too slow, nonintuitive, cumbersome, or just annoying. They measured the relationship between recognition performance and the preferences of users. To do that, 14 participants participated in 5 experiments consisting of playing World of Warcraft (WoW) for 2 h and for 5 weeks. They were divided into two groups. A real-BCI group controlled their shape-shifting action with their mental tasks, at least insofar as the system could detect it: i.e., the users received feedback on the recognition of their mental tasks in the form of an orange bar in the videogame (the bar is small if the system had detected the mental task related to elf form, while it is large if the system had interpreted the mental task related to bear form). A utopia-BCI group pressed a button to shape shift when they decided by themselves that they performed the mental task correctly (in this group, a BCI system with 100 % detection accuracy was simulated). After user test sessions, the participants filled out a user experience questionnaire. The results show that the users preferred the association tasks (i.e., the user had to feel like a bear to change into a bear and to feel like an elf to change into an elf) whereas the relaxed/stressed mental state seems to be disliked the most. However, in practice, such a BCI paradigm with which a user could imagine being a bear to turn its avatar into a bear is not necessarily realistic and robust, as such mental tasks cannot be reliably detected in EEG. Furthermore, it appears that recognition performance has a strong influence on users' preferences.

Ergonomics Methods Used or Advocated in the Context of User-Centered Design of BCI Videogames

It is now known that the combination of ergonomics methods produces complementary data enriching three of these four phases of user-centered design (e.g., Anastassova et al. 2005): specify the context of use (step 1) and the functionalities of the future videogame (step 2) and evaluate the designed solutions (step 3). Defining the future context of use and specifying the functionalities of the BCI videogames are closely related as both steps involve identifying existing and latent needs (section "Ergonomics Methods to Identify Existing Needs and Latent Needs in Order to Specify the Future Context of Use and the Functionalities of the BCI Videogame"). Evaluating solutions at all stages in the project from early concept design to long-term use to specify design choices implies using prototype evaluation methods (section "Evaluation Methods of Prototypes") and methods to analyze system appropriation (section "Analysis Methods of BCI Videogame Appropriation").

Ergonomics Methods to Identify Existing Needs and Latent Needs in Order to Specify the Future Context of Use and the Functionalities of the BCI Videogame

Ergonomics uses methods which allow the designers to understand the future context of use on the one hand and to identify existing needs and anticipate latent ones in order to specify the functionalities of the BCI videogame on the other hand. According to Robertson (2001), the existing needs correspond to

- Conscious needs: i.e., those clearly formulated by the (future) end users
- Unconscious needs: i.e., those which exist but are not clearly formulated by users, because they are not aware of the potential of the chosen technology. The latent needs are characterized by their not-proven-yet or undreamed nature (Robertson 2001). These needs are an important issue for emerging technologies, like BCI videogames, which are still in development in laboratories and whose uses are still sought.

Identification Methods of Existing Needs

Interview is a currently used or recommended method to identify relevant needs in order to suggest realistic expectations with BCI. For example, to design a suitable BCI system for domestic use by people with acquired brain injury in order to facilitate control of their environment, Mulvenna et al. (2012) have conducted interviews with disabled participants with acquired brain injury and participants without movement disabilities. Before imagining a videogame, designers can interview potential users about points such as

- The context of use, e.g., who are the users of the videogame (children, adults, etc.)? Can you explain the context in which the videogame will be used (in a classroom, at home, etc.)?
- The videogame, e.g., what is the goal of the videogame (entertainment, learning, rehabilitation, etc.)? Which actions can be done in the videogame? What could the virtual environment be composed of?
- The interaction, e.g., do you prefer to imagine that your hand is moving or to focus in order to move the avatar? Which other interaction device can be used in addition to the BCI?

Anticipation Methods of Latent Needs

Anticipating latent needs (i.e., not "existing" for users at a specific moment) involves the use of creativity methods to widen potential uses, the functionalities, and properties of the videogame (Burkhardt and Lubart 2010). This kind of methods is particularly relevant for the design of all emerging technologies, specifically those designed for different users' profiles like a BCI videogame. Among these creativity methods, focus groups are often used. This method aims to create small groups of potential users of an interactive system and to interact on a relevant topic for the study to achieve. As output, opinions, expectations, motivations, and attitudes from customs, practices, and experiences of participants are usually collected. Indeed, focus groups have recently been used for the design of BCI applications in a study conducted by Blain-Moraes et al. (2012). The purpose was to determine the barriers and mediators of BCI acceptance in a population with amyotrophic lateral sclerosis. The authors conducted a focus group which involved eight individuals with amyotrophic lateral sclerosis (having previously used a P300-based speller with a visual display) and their nine carers. The focus group consisted of open-ended questions asking participants about their desires and concerns with the BCI. It was transcribed in full, and data was thematically analyzed. Focus group analysis yielded two categories of mediators and barriers to user acceptance of this technology: personal factors (i.e., physical, physiological, and psychological concerns) and relational factors (i.e., corporeal, technological, and social relationships with the BCI). This study showed that a focus group is a relevant method to evoke needs in terms of mediators and barriers concerning the use of BCI by people with amyotrophic lateral sclerosis. This result is encouraging for the use of a focus group to anticipate latent needs of potential users of BCI videogames to stimulate them to discuss together and imagine new ideas on the usage of the technology (i.e., a person is able to imagine new needs through creativity generated within a group). For example, future users are able to cocreate together a videogame scenario and to decide their preferred interaction paradigm (e.g., association task). In addition to these methods, ergonomics uses and recommends methods to accomplish the fourth phase of the user-centered design process which corresponds to the evaluation of prototypes (3.2) and the appropriation of the BCI videogame by users (3.3).

Evaluation Methods of Prototypes

Prototypes are a means to immerse users in the context in which the BCI videogame will be integrated. Users are able to evoke functionalities which were previously latent and unconscious. A simulation of the future situation is an opportunity for the designer to explore the real impact of the videogame on the users' activities, in particular for the BCI videogames in learning and medical areas. This simulation allows them to identify the improvements to be implemented so that the users have a real benefit from its use (Burkhardt and Lubart 2010).

In a user-centered design of BCI videogames, prototypes can be evaluated through user-based testing (i.e., a simulation to study the users' behavior in front of the application) followed by questionnaires. This observation is illustrated with some studies detailed below. These studies provide an overview of the way the evaluation is carried out in the field of BCI videogames on three aspects: the number of participants, the location of the test, and the evaluation objectives (objective evaluation of performance, subjective evaluation, etc.).

Badia et al. (2011) conducted a study aiming to explore the synergies of a hybrid BCI and a BCI videogame for neurorehabilitation systems. This study involved 18 participants and consisted of 4 phases based on a training task called "Spheroids" which is a videogame-like task in which the user has to intercept incoming spheres by moving the arms of the virtual avatar. First, the BCI classifier was trained. Then, the "Spheroids" calibration phase was used to assess the level of control of the participants by asking them to drive the virtual arms to specific locations. Subsequently, participants played the "Spheroids" training videogame. Finally, all participants answered a 5-point Likert scale (1 lowest, 5 highest) of 23 questions covering different aspects: enjoyment of the experience, perceived performance learning during task execution, level of task ease, level of control of the virtual avatar, and appropriateness of the system configuration (for instance, if arms were too fast or too slow).

Van de Laar et al. (2011) have recently elaborated a standardized questionnaire, inspired by the Game Experience Questionnaire and the Engagement Questionnaire. This questionnaire can be used to evaluate the user experience in a BCI-based interaction for entertainment purposes. It offers optional items depending on the category of BCI (e.g., passive BCI, active BCI). For passive BCI, items measure the degree of comfort and of distraction from the main task due to the BCI hardware. For active BCI, items measure the applicability of the mental tasks and perceived speed of the BCI on the users' actions.

In their study, Zickler et al. (2013) evaluated the usability of Brain Painting, an application for entertainment offering the user creative expression by painting pictures. For that, they compared Brain Painting with a P300 spelling application in terms of user satisfaction through three questionnaires (i.e., Quebec User Evaluation of Satisfaction with assistive Technology 2.0, Assistive Technology Device Predisposition Assessment (ATD PA), and Device Form), the effectiveness (accuracy), the efficiency as measured by the information transfer rate, and the subjective workload using the National Aeronautics and Space Administration Task Load Index. Even if the results globally showed a good usability, they revealed that the system operability and the EEG cap are the main obstacles for use in daily life. Concerning the last observation, Nijholt and Gürkök (2013) suggested reducing the complexity of gel-based electrode installation by using wireless headsets and dry electrodes.

These studies show three elements. First, BCI videogames, even relatively immature, can be evaluated with a large number of users with highly varying skills and potentially very heterogeneous needs. This is a key element in the context of BCI videogames which must be adapted to users other than those who have been involved in the design. Second, these studies stressed the diversity of the implementations of user-based testing that can take place in laboratory conditions (i.e., controlled situations) and in conditions similar to real situations of interaction (typically, playing videogames at home where the variables are not controlled). This suggests that user-centered evaluations of BCI videogames can take place under experimental conditions (in the laboratory) and in more ecological conditions relative to the future situation of use. These two ergonomic approaches are complementary and coexist for the design of these emerging videogames. This opens perspectives for user-based evaluations of BCI videogames for several types of applications (e.g., medical, learning, or entertainment applications). Third, it provides information about the question properties: these questions are generally measured on Likert scales (5- or 7-point) and less frequently open-ended questions that allow the questioned person to express freely. The items covered in these questionnaires concern the system (e.g., appropriateness of the system configuration, application responsiveness to initiated actions) and subjective elements perceived by the user (e.g., enjoyment of the experience, level of task ease). Even if the questionnaire is a relevant method, it should not be used alone. From an ergonomics point of view, it must be complemented, for instance, by interviews that will justify the scores obtained in the questionnaire.

Analysis Methods of BCI Videogame Appropriation

To study system appropriation, longitudinal studies can be used. They are usually implemented through observations, interviews, and self-confrontation made at different periods (e.g., at 0 month, 3 months, and 6 months after the integration of the BCI application in a situation).

Longitudinal studies have already been conducted in the context of BCI to measure the influence of psychological state and motivation on the BCI performance (e.g., Nijboer et al. 2010). In this study, six participants were trained for several months either with a BCI based on sensorimotor rhythms, on event-related potentials (i.e., P300), or both. Questionnaires assessing quality of life, severity of depressive symptoms, mood, and motivation were filled out by the participant before each training session. Results suggest that P300-based BCI must be a first choice for allowing severely paralyzed patients to control a communication program based on a binary spelling system. Results also suggest that motivational factors may be related to the BCI performance of individual subjects and suggest that motivational factors and well-being should be assessed in standard BCI protocols.

This study illustrated that longitudinal studies have been conducted to measure the evolution of some parameters' effects on the use of a BCI-based speller and thus to study the appropriation of the BCI application by users. This is necessary because the lack of training on these new tools and the lack of support are factors that may lead end users to abandon their use. Typically in the context of BCI videogames, a longitudinal study may intend to measure the motivation and game engagement for several months. It can also be used to study how the user skills at BCI control evolve with time, in order to design adequate levels and content in the game that can match such BCI skill progress.

These methods can be used to design and evaluate usefulness, usability, acceptability, hedonic quality, sense of presence, immersion, and user experience. In the next section, the ways in which these ergonomic criteria can be taken account in the context of a BCI videogame are presented.

Ergonomic Criteria

Usefulness

In ergonomics, the term "usefulness" oscillates between two meanings: *purpose-usefulness* and *value-usefulness* (Loup-Escande et al. 2013).

Purpose-usefulness is the description of system features and its uses. This description can take many forms depending on the phase of the design process: concept (e.g., a BCI videogame to entertain the motor-restricted users in their daily life), specifications (e.g., this BCI videogame must be multiplayer and could consist in a virtual ping pong activity, the interaction must be based on relaxed/stressed mental state), prototype (e.g., the virtual environment is modeled, the ball-passing task between two users is implemented, the BCI interaction uses the Neurosky MindWave), final application (e.g., the prototype is improved with a new point-counting functionality and the integration of the Emotiv Epoc instead of the Neurosky MindWave). Purpose-usefulness corresponds to the functionalities of a videogame. These features are determined at a given time, even if they can be subsequently modified by players.

Several metrics can be used to assess the BCI videogames' usefulness. A well-known metric, and common to all emerging systems, is the adequacies versus inadequacies between the features implemented in a system on the one hand and those desired at a time T by the user on the other hand (Blandford et al. 2008). These adequacies or inadequacies are collected with the number of responses Yes versus No to the question: Do you think this functionality/information is useful? They are collected with metrics resulting from requirement prioritization: nominal scale, ordinal scale, ratio scale methods (Loup-Escande and Christmann 2013). In nominal scale methods, requirements are assigned to different priority groups. An example is the MoScoW method, which consists of grouping all requirements into four priority groups, such groups corresponding to the requirements that the project must/should/could/will not have. All requirements listed in a category are of equal priority, which does not allow a finer prioritization. Ordinal scale methods produce an ordered list of requirements; for example, the simple ranking where the most important requirement is ranked "1" and the least important is ranked "n." Another known method called analytic hierarchy process asks users to compare all pairs of requirements. Ratio scale methods provide the relative difference between requirements (e.g., the hundred dollar method asks users to allocate a sum of money to each requirement). In addition to an ordered list of requirements, this method also helps us to discover the relative importance of each requirement in relation to the others.

Value-usefulness is defined as a significant advantage of the videogame for players in activities mediated by a computer system; this advantage is always relative to the objectives of the user, the existing tools, the use environment, and the dependencies on other activities. Value-usefulness refers to improvements and benefits that the videogame provides to the players. In the field of BCI, these benefits are evaluated in the short, medium, or long term. For example, Yan et al. (2008) conducted a study to evaluate the usefulness of a BCI videogame for treating children with attention deficit hyperactivity disorder (ADHD). In other words, they wanted to show that the virtual reality neuro-feedback training could improve sustained attention. This study involves 12 subjects with ADHD aged 8 to 12 years, who were trained at least twice per week with 5 virtual environment games for 25–35 min. During the training period, subjects were requested to stop taking any medication and behavioral therapy. To measure children's attention, authors used the integrated visual auditory – continuous performance test (IVA-CPT) at the beginning of the treatment and after each of the 20 training sessions. Results show that subjects' attention had strengthened after 20 training sessions.

In order to evaluate the benefits and advantages for the user with respect to his goals, existing tools, the environment of use, and dependencies with other activities, the used metrics are more specific to the domain for which the application was designed. For example, to evaluate the usefulness of a BCI videogame to support learning, the metrics are specific to the knowledge that can be learned with the system.

In a BCI videogame for entertainment, the value-usefulness can be measured by the degree of engagement in the game if we consider that the more engaged the users, the more useful the videogame for them (Brockmyer et al. 2009).

Both for the medical area, for the learning domain, and for the entertainment field, the benefits provided by a BCI videogame can be measured using indicators of well-being, since the objective of the BCI videogame is to increase quality of life. Quality of life can be measured by the Schedule for the Evaluation of Individual Quality of Life (Holz et al. 2013). To measure game engagement, Brockmyer et al. (2009) have designed a questionnaire based on four dimensions:

- The psychological absorption describes the "total engagement in the present experience" (Brockmyer et al. 2009, p. 625): e.g., time seems to stand or stop.
- The flow describes "the feelings of enjoyment that occur when a balance between skill and challenge is achieved in the process of performing an intrinsically rewarding activity" (Brockmyer et al. 2009, p. 625): e.g., I play without thinking how to play.
- The presence is defined in terms of "being in a normal state of consciousness and having the experience of being inside a virtual environment" (Brockmyer et al. 2009, p. 624): e.g., I play longer than I meant to.

- Immersion describes “the experience of becoming engaged in the game-playing experience while retaining some awareness of one’s surroundings” (Brockmyer et al. 2009, p. 624): e.g., I really get into the videogame.

Usability

In 1998, the International Organization for Standards published a definition of usability (ISO 9241-11): “The extent to which a product can be used by specified players to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use.” Some authors add learnability and memorability to define usability in BCI context (e.g., Plass-Oude Bos et al. 2010). Usability is therefore a combination of five elements: effectiveness, efficiency, satisfaction, learnability, and memorability (Plass-Oude Bos et al. 2010):

- Effectiveness concerns the fact that the software allows the user to achieve the specified goal.
- Efficiency is the capacity to achieve a task with the minimum of resources for users, i.e., efforts. For instance, can the user achieve multitasking with the BCI and control another input at the same time?
- Satisfaction is influenced by ease of use (e.g., EEG sensor setup time, machine calibration time, etc.) and by noninstrumental qualities (e.g., aesthetics aspects).
- Learnability is what allows a novice user to devote himself quickly to a task, reducing the time needed to learn how to use the application. The second aspect of learnability is how efficiently can the game train the user to reliably perform mental tasks to use the BCI (Plass-Oude Bos et al. 2010; Lotte et al. 2013).
- Memorability is what allows the user to perform the tasks after a period of nonuse without having to relearn the functioning of the application (e.g., how well can a user perform the BCI mental commands from 1 day of gaming to the next).

Plass-Oude Bos et al. (2010) adapted the usability characteristics commonly used in human-computer interaction (i.e., learnability, memorability, efficiency, effectiveness, error handling, and satisfaction) to the design of BCI videogames. The usability of BCI applications has already been evaluated in some studies. For example, Ekandem et al. (2012) conducted a study to evaluate the usability of two BCI devices: the Emotiv EPOC and the Neurosky MindWave. Authors compared user comfort, experiment preparation time (i.e., total time from initial placement to final adjustment), signal reliability, and ease of use of each BCI. This study involved 13 participants. Each participant completed a training phase before playing several simple videogames included with each system like Pong, Tetris, and SpadeA. After having worn the BCI device for 15 min, the participant completed a postexperiment questionnaire. Results show that the preparation time for the Emotiv EPOC is longer than for the Neurosky MindWave and that the majority of participants indicated that the Emotiv EPOC was comfortable whereas the

Fig. 1 The “Use-the-force” application. The user had to lift up a virtual spaceship using foot motor imagery (©CNRS Photothèque/Hubert Raguet)

Neurosky MindWave was not. Moreover, the MindWave provided an easier signal acquisition whereas the EPOC clearly had contact issues due to participants’ hair. However, the signal was maintained and even improved during the session once the EPOC was connected and calibrated, while the MindWave experienced more signal fluctuations.

The previous study focused on the direct evaluation of a BCI device. Interestingly enough some of the studies concerned the usability evaluation of BCI-based interaction with another system. For example, Iturrate et al. (2009) aimed to evaluate a new brain-actuated wheelchair concept that relies on a synchronous P300 brain-computer interface integrated with an autonomous navigation system. The evaluation concerned the effectiveness and the efficiency of the BCI-based interaction and of the graphical interface.

In Lotte et al. (2008), the usability of a BCI videogame in 3D (Fig. 1) was assessed using questionnaires that measured the feeling of control, fatigue, comfort, and frustration induced by the BCI game, both for control commands based on real and imagined foot movements.

These studies suggested that the usability evaluation of BCI videogames can have different objectives. Indeed, it can concern either the BCI hardware (EEG cap), the BCI software (e.g., the reliability of mental state detection), or the BCI-controlled application (e.g., a robot).

A BCI-controlled application is often evaluated in terms of effectiveness using, for example, the degree of accomplishment of the task, the distance traveled to accomplish the task, the task success, the total time taken to accomplish the task (in seconds), and the number of missions to complete the task (Iturrate et al. 2009; Escolano et al. 2009).

To evaluate the efficiency of a BCI videogame, Holz et al. (2013) used the NASA-TLX to estimate the subjective workload experienced in a specific task and its main sources (i.e., mental, physical, and temporal demand, effort, performance and frustration). In addition, the efficiency of a BCI system is often measured using the information transfer rate (ITR), which measures how much information can be conveyed by the system in a given time (e.g., in bits per second) (Wolpaw et al. 2002). Thomas et al. (2013) listed several metrics used to quantify the performance of a BCI and evoked the evaluation strategies employed to compare the performance of two or more BCIs. They suggested the following metrics to measure the performance of BCI based on synchronized control (i.e., the system is periodically available to the user when it is on but does not support the noncontrol): accuracy (i.e., classification error), kappa, bit–rate (Wolpaw or Nykopp), confusion matrix, and task specificity. They recommended those to estimate the performance of BCI based on self-paced system (i.e., the system is continuously available to the user when it is on and supports noncontrol): hit-false difference, confusion matrix, task specificity, utility, efficiency, true positive (i.e., it quantifies the chance of correctly identifying intentional control states), and false positive (i.e., it quantifies the chance of incorrectly identifying the noncontrol state). To evaluate the satisfaction evoked by a BCI videogame, the semantic differential scales corresponding to the following items, provided by (Hassenzahl 2001), can be used: Understandable versus Incomprehensible, Supporting versus Obstructing, Simple versus Complex, Predictable versus Unpredictable, Clear versus Confusing, Trustworthy versus Shady, Controllable versus Uncontrollable, and Familiar versus Strange.

To the authors' best knowledge, there are very few studies on the evaluation of learnability and memorability in the BCI field. An exception is the study of Jeunet et al. (2014), which assessed the subjective learnability and memorability (among other measures) of standard BCI training procedures, using questionnaires. However, this was not in a gaming context. In technological systems though, learnability is better assessed by comparing task execution time by a novice group (who has never used the application before) and task execution time by an expert group (who already used the application). If the time taken by the group of novices is inferior or equal to that taken by the expert group, the learnability of the application is good.

The learnability is linked with a specific phenomenon called "BCI illiteracy/BCI deficiency" (Allison and Neuper 2010). Indeed, a given BCI system does not work for all users because several users cannot produce detectable patterns of brain activity necessary to a particular BCI approach (Guger et al. 2003).

Memorability is assessed by measuring task execution time at different times over a period of weeks. If the time obtained during the first week is superior or equal to the time obtained during the following weeks, then the memorability of application is good.

Acceptability

According to (Venkatesh et al. 2003), acceptability refers to an individual's perception of the system's value. To assess the acceptability, authors try to identify the intentions of individuals to use a system through questions. Thus, models of acceptability identify the variables that contribute significantly to the determination of intentions to use a technology. Among these models, the most complete is the UTAUT (Unified Theory of Acceptance and Use of Technology) by Venkatesh et al. (2003). According to this model, behavioral intention is influenced by the performance expectancy, effort expectancy, and social influence, and use behavior is influenced by behavioral intention and facilitating conditions. Previous experience with the system, the voluntariness or not of use, gender, and age moderate the effects of direct determinants like performance expectancy, effort expectancy, social influence, and facilitating conditions.

According to Venkatesh et al. (2003), acceptability is measured by eight factors:

- Performance expectancy is defined as the degree to which an individual believes that using the system will allow him/her to gain in task performance (e.g., gain speed in the achievement of the task, score a lot of points, be able to get in a relaxed/stressed mental state).
- Effort expectancy corresponds to the degree of ease associated with the use of the system.
- Attitude toward using technology corresponds to an overall affective reaction of an individual using a system.
- Social influence is defined as the degree to which an individual perceives that important others believe he/she should use the new system.
- Facilitating conditions are defined as the degree to which an individual believes that an organizational and technical infrastructure exists to support the use of the system (e.g., BCI patients association or BCI companies which provide technical support).
- Self-efficacy corresponds to "*an individual's belief in one's capability to organise and execute the courses of action required to produce given attainments*" (Bandura 1997, p. 3) (e.g., does the player believe he/she can control a game efficiently by mental activities alone).
- Behavioral intention to use the system corresponds to the intention to use the system in the next months.
- Anxiety. (e.g., is the player anxious that the BCI may read his/her mind without his/her consent).

Table 1 describes metrics (or items) for BCI videogame acceptability inspired and adapted from Venkatesh et al. (2003).

These items formalized in the form of positive affirmation are associated with the Likert scale on which the user notes the degree of agreement or disagreement.

Instead of questionnaires, some authors perform interviews with players after a test session. Gürkök et al. (2014) conducted semistructured interviews with

Table 1 Items used to assess factors influencing acceptability

Factors influencing acceptability	Items
Performance expectancy	I would find the BCI videogame useful Using the BCI videogame entertains me Using the BCI videogame enables me to learn/to treat/to rehabilitate
Effort expectancy	My interaction with the BCI videogame would be clear and understandable I would find the BCI videogame easy to use Learning to operate the BCI videogame is easy for me
Attitude toward using technology	Using the BCI videogame is a good idea Using the BCI videogame is fun I like using the BCI videogame
Social influence	People who influence my behavior think that I should use the BCI videogame People who are important to me think that I should use the BCI videogame
Facilitating conditions	I have the knowledge necessary to use the BCI videogame The BCI videogame is not compatible with other systems I use
Behavioral intention to use the BCI videogame	I intend to use the BCI videogame in the next <n> months I predict I would use the BCI videogame in the next <n> months I plan to use the BCI videogame in the next <n> months
Anxiety	I feel apprehensive about using the BCI videogame The BCI videogame is somewhat intimidating to me

42 players to investigate their opinions on control and playability of the BCI videogame based on the famous multiplayer online role-playing game "WoW." They used the answers to provide some general guidelines for the design and the evaluation of BCI videogame acceptability. For example, they observed that "the experience of fun resulting from playing a BCI videogame once does not reliably represent the experience of pleasure that unfolds by playing the videogame." Consequently, the authors suggested that the "BCI videogames should be developed and evaluated for the pleasure rather than the fun they provide." To do that, literature in ergonomics and psychology proposes to design and measure hedonic qualities and appeal.

Hedonic Qualities and Appeal

Hedonic qualities refer to the aspects of a BCI videogame that are related to a person's pleasure. The pleasure derived from the use of a BCI videogame is associated with its appealing characteristics and aesthetic.

To evaluate the hedonic quality of all applications including videogame, Hassenzahl (2001) suggested the following items using the semantic differential scales: Interesting versus Boring, Costly versus Cheap, Exciting versus Dull, Exclusive versus Standard, Impressive versus Nondescript, Original versus Ordinary, and Innovative versus Conservative. To evaluate the comfort and discomfort of BCI devices, Ekandem et al. (2012) used three indicators: the comfort of the device (very uncomfortable, uncomfortable, indifferent, comfortable, and very comfortable), the time the participants felt they could comfortably wear the device (0–5 min, 5–20 min, 20–60 min, 60–120 min, and more than 120 min), and the type of discomfort perceived (sharp, dull, itchy, heavy, throbbing, awkward, burning, or other).

To evaluate the appeal of these systems, Hassenzahl (2001) suggested the following items: Pleasant versus Unpleasant, Good versus Bad, Aesthetic versus Unaesthetic, Inviting versus Rejecting, Attractive versus Unattractive, Sympathetic versus Unsympathetic, Motivating versus Discouraging, and Desirable versus Undesirable.

Immersion and Presence

Two additional dimensions can be specified to evaluate the BCI videogames: immersion and presence. Each of these dimensions has led to many definitions (e.g., Witmer and Singer 1998; Brockmyer et al. 2009). The following definition remains the most comprehensive: immersion corresponds to the degree with which the system interface controls the sensory inputs for each modality of perception and action. So, immersion can be described (but not only) in terms of specific devices: a common dichotomy derives from the opposition of “immersive” systems (Head Mounted Display, Cave Automatic Virtual Environment) and “non immersive” systems (desktop, mouse).

According to Sanchez-Vives and Slater (2005), immersion determines the sense of presence perceived by users through display parameters, visual realism, sound, haptics, virtual body representation, and body engagement. The International Society for Presence Research (2000) ruled that presence is “a psychological state or subjective perception in which even though part or all of an individual’s current experience is generated by and/or filtered through human-made technology, part or all of the individual’s perception fails to accurately acknowledge the role of the technology in the experience.” In other words, the sense of presence is experienced when the place illusion (i.e., the sensation of being in a real environment) and the plausibility illusion (i.e., the feeling that the virtual scenario is actually occurring) occur (Slater 2009).

These links between immersion and presence appeared in the literature: immersion has been studied in the context of many studies on BCI and virtual environment. For example, an immersive environment can improve the sense of presence while carrying out navigational tasks through imaginary movements (Lotte et al. 2013). Donnerer and Steed (2010) observed that P300 can be used successfully in immersive virtual environments, and Groenegress et al. (2010) proved that P300-based navigation lowered the sense of presence compared to gaze-based navigation.

In the field of BCI, this dimension was evaluated in some studies. For example, Hakvoort et al. (2011) conducted a study with 17 participants aiming to measure immersion in a BCI videogame named “Mind the Sheep!.” Their experiment consisted in two different sessions: a BCI session and a non-BCI session. Each session was divided into three trials: a familiarity trial (i.e., participants had to collect 10 objects which were placed across the playground), an easy trial (i.e., participants had to park a small flock of 5 sheep using the dogs), and a difficult trial (i.e., participants had to gather 10 sheep, which were more scattered across the playground, into one pen that was placed in the center of the playground). After each session participants filled in a questionnaire on their perceived immersion in the videogame. The 31 questions dealt with cognitive involvement, emotional involvement, real-world dissociation, challenge, and control. Results showed that the BCI selection method was more immersive than the non-BCI selection method (i.e., a mouse).

To measure immersion and presence during a videogame, Witmer and Singer (1998) designed a questionnaire which was used in numerous studies. In their approach, these authors evaluate the presence according to four categories that we adapted below for BCI videogames:

- The control factors correspond to the degree of control that a person has in playing with a BCI videogame, the immediacy of control (i.e., the delay between the action and the result), the anticipation concerning what will happen next, whether or not it is under personal control, the mode of control (i.e., the manner in which one interacts with the environment is a natural or well-practiced method), and the physical environmental modifiability (i.e., the ability to modify physical objects in an environment).
- The sensory factors are the sensory modality (i.e., visual information and other sensory channels), the environmental richness in terms of information, the multimodal presentation to simulate completely and coherently all the senses, the consistency of multimodal information, the degree of movement perception (i.e., the observer must perceive self-movements through the virtual environment), and the active search (e.g., the observers can modify their viewpoints to change what they see).
- The distraction factors correspond to the isolation from the used devices (e.g., head-mounted display), the selective attention (i.e., the observer’s willingness or ability to focus on the videogame stimuli and to ignore distractions), and the interface awareness.
- Realism factors are the scene realism governed by scene content, texture, resolution, light sources, field of view and dimensionality, the consistency of information with the objective world, the meaningfulness of experience for the person, and the anxiety/disorientation when users return from the videogame to the real world.

User Experience

User experience is a consequence of an interaction between a user (with his characteristics) and a product (with its features and qualities) appearing after an evaluation process (Hassenzahl 2001). The user experience is defined by the perceived usefulness, usability, hedonic quality, appeal and the sense of presence, and the emotional reactions. The emotions arise from subjective feelings, physiological reactions, motor expressions, and cognitive appraisals (Mahlke et al. 2006):

- Subjective feelings can be evaluated and measured by the "Self-Assessment Manikin" defined as "a non-verbal pictorial assessment technique that measures the pleasure, arousal, and dominance associated with a person's affective reaction to a wide variety of stimuli." Pleasure is measured by the following items: Unhappy versus Happy, Annoyed versus Pleased, Unsatisfied versus Satisfied, Melancholic versus Contented, Despairing versus Hopeful, and Bored versus Relaxed. Arousal is measured with these adjectives: Relaxed versus Stimulated, Calm versus Excited, Sluggish versus Frenzied, Dull versus Jittery, Sleepy versus Wide-awake, and Unaroused versus Aroused. Dominance is measured with these items: Controlled versus Controlling, Influenced versus Influential, Cared for versus In control, Awed versus Important, Submissive versus Dominant, and Guided versus Autonomous.
- Physiological reactions can be estimated from peripheral (e.g., muscle tension, heart rate, electrodermal activity) and central nervous system signals (e.g., electroencephalogram) (Mühl et al. 2009).
- Motor expressions can be measured with electromyography of the two facial muscles associated with positive emotions (*zygomaticus major*) and negative emotions (*corrugator supercili*).
- Cognitive appraisals can be measured by the "Geneva Appraisal Questionnaire," which measures five appraisal dimensions: intrinsic pleasantness (i.e., a stimulus event is likely to result in a positive or negative emotion), novelty (i.e., a measure of familiarity and predictability of the occurrence of a stimulus), goal/need conductivity (i.e., the importance of a stimulus for the current goals or needs), coping potential (i.e., the extent to which an event can be controlled or influenced), and norm/self-compatibility (i.e., the extent to which a stimulus satisfies external/internal standards).

The user experience has been often evaluated in the context of BCI videogames. The user experience resulting from these technologies has been measured by its perceived usability, its hedonic quality, its appeal, and the sense of presence. Some studies on user experience in BCI and videogames are conducted to compare the user experience resulting from the use of different controllers (e.g., BCI, mouse) and to understand the added value relating exclusively to BCI control (Gürkök et al. 2013), or to find the differences between real and imagined movement in a BCI videogame in relation to user experience and performance (Plass-Oude Bos et al. 2010) using a user experience based on the Game Experience Questionnaire (Ijsselsteijn et al. 2008).

These studies are focused on the comparison of user experience according to paradigms and interaction devices, only in the videogame field. Thus, user experience was evaluated with a specific questionnaire dedicated to game experience.

Fig. 2 Two users playing BrainArena in a competitive trial

Competition and Collaboration, Two Relevant Criteria for Multibrain Videogames

In the entertainment area, some multibrain videogames begin to appear (Fig. 2): each player uses a BCI (e.g., Bonnet et al. 2013).

For these specific BCI videogames, it is important to provide and measure competition or collaboration (Nijholt and Gürkök 2013). In multibrain videogames providing competition, two players compete using their brain until one of them is the winner and the other is the loser (e.g., Hjelm and Browall 2000). For example, in the "Brainball" game (Hjelm and Browall 2000), two players have to compete by relaxing, their performance is measured by EEG, their brain activity is compared, and the difference determines the direction of the ball (i.e., moving into the direction of the player who is less relaxed).

In multibrain videogames promoting collaboration, Pope and Stevens (2012) suggest to distribute the control of the input devices among the two players, e.g., one player acts physically while the other uses his/her brain. According to Nijholt and Gürkök (2013), the evaluation of the collaboration level requires several metrics such as the tasks' repartition and the number of social interactions between the two players and has to integrate the environmental context, e.g., are players colocated or distributed? Is there an audience? What does the audience see and is there any interaction between the audience and the players?

Discussion

Our framework suggested specific methods to be used to define and measure specific criteria (Fig. 3). Overall our framework highlighted that usefulness, usability, and acceptability are criteria that need to be considered both to identify and to

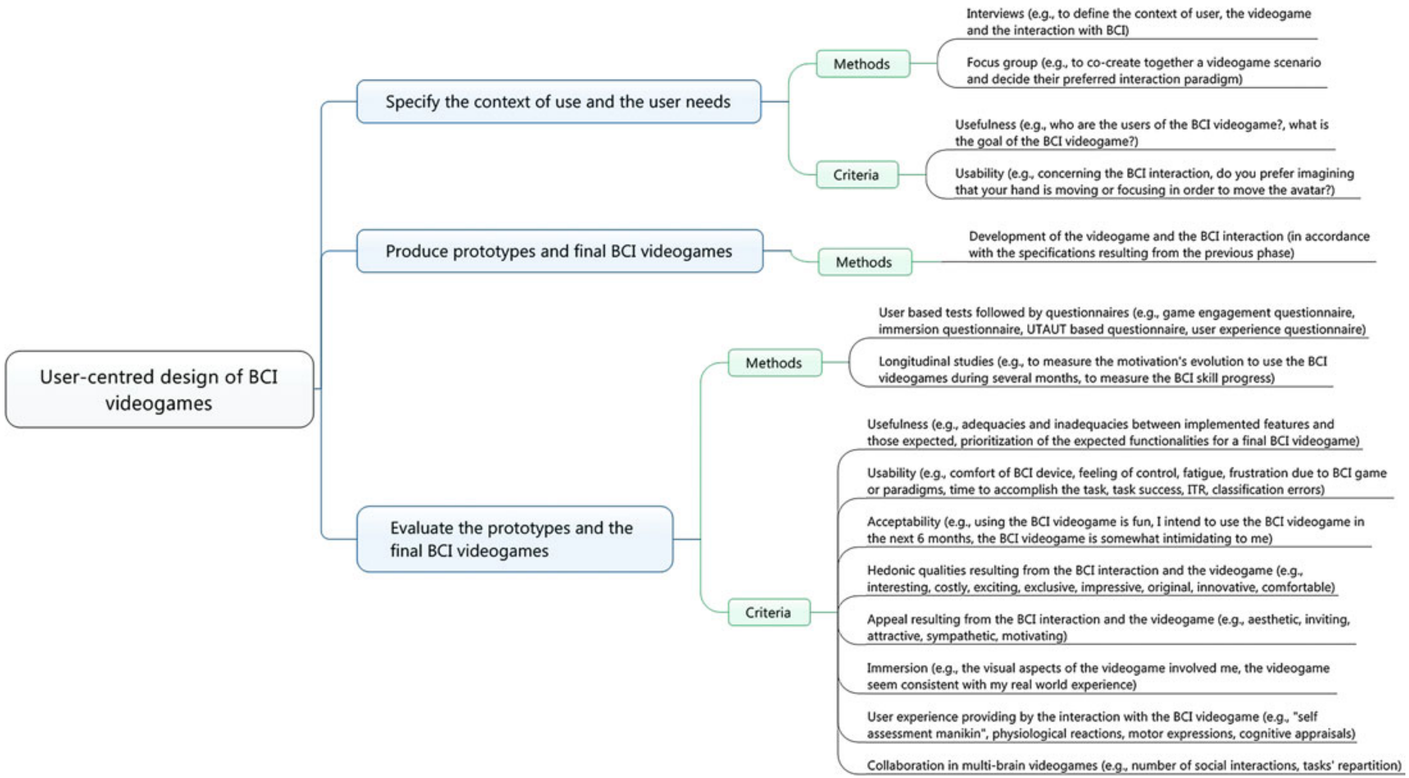


Fig. 3 Methodological framework for the user-centered design of BCI videogames

anticipate user needs (i.e., contexts of use and features) through interviews and focus groups. These criteria have also been integrated in the evaluation phase to assess both intermediate solutions (e.g., prototypes) with user-based tests associated with questionnaires and the system's appropriateness by conducting longitudinal studies. In the same way, our framework stressed that hedonic qualities, appeal, immersion and presence, emotions, and more generally user experience also need to be measured in the evaluation phase through user-based tests and questionnaires. Moreover, this framework recommended specific metrics associated with each ergonomic criterion. For example, in a questionnaire or in an interview's grid, some questions can allow to evaluate usability (e.g., the degree of complexity), usefulness (e.g., the order of priority of the suggested functions), acceptability (e.g., the intention to use the system in the next 4 months), hedonic quality (e.g., the degree of boredom), appeal (e.g., the degree of appeal), emotion (e.g., the degree of arousal), and immersion and presence (e.g., the degree of provision of visual aspects).

The studies concerning the design of user-centered BCI videogames are uncommon compared to studies aimed to optimize the performance of these applications. One possible reason may be that research on BCI videogames requires technological developments before players can be confronted with the system in an actual context of use. However, integrating players from the early design phase to the use of application is necessary to design applications which will be suitable. Despite this, very few papers have focused on the methods, criteria, and metrics for the ergonomic design of BCI videogames.

Concerning the studies on user-centered design of BCI (and not specifically BCI videogames), this chapter highlighted three observations. First, these studies concerned mainly the evaluation phase, i.e., the evaluation of prototypes and of the appropriateness of applications (e.g., Nijboer et al. 2010), to the detriment of the early design phase aiming to understand the context of use or to specify the user needs (e.g., Blain-Moraes et al. 2012). Second, the majority of studies used only one method: the user-based test is mainly used to evaluate the usability (e.g., Valsan et al. 2009) and the usefulness (e.g., Yan et al. 2008), the focus group to characterize the acceptability (e.g., Blain-Moraes et al. 2012), and the questionnaire to measure the user experience (e.g., Van de Laar et al. 2011) and, in longitudinal studies, to assess the application's appropriateness (e.g., Nijboer et al. 2010). Third, the assessments of BCI videogames focused on a single criterion which was mainly the usability (e.g., Holz et al. 2013; Gürkök et al. 2014) at the expense of other criteria such as immersion, presence (e.g., Hakvoort et al. 2011), or usefulness (e.g., Yan et al. 2008).

To overcome these observations, our framework suggests three guidelines. The first guideline insists on all phases of the user-centered design process for which ergonomics is equipped: understanding and specifying the context of use (phase 1), specifying the user needs (phase 2), evaluating the prototypes (phase 3), and assessing the appropriateness of applications (phase 4). Indeed, involving players from the identification of needs before the implementation of the application to the application's use allows designers to integrate the evolution of characteristics, needs, and expectations of players. Our methodological framework is comprehensive in that it deals with every phase of the process and not only the evaluation phase partially covered in the literature of human-computer interaction. The second guideline concerns the necessity to use several methods: interview and focus group to define specifications based on players' needs, user-based testing, questionnaire, and longitudinal studies to evaluate the BCI videogames. Indeed, using several ergonomic methods to study a criterion is advisable because they provide complementary data. Our methodological framework suggests implementing at least two methods to evaluate the same criterion. For example, to evaluate the usefulness of a BCI videogame, it is desirable to achieve (1) quasi experimentation-based longitudinal studies (i.e., comparison between a group that uses the application and another group that does not use it, for a period of several months), and (2) after-use questionnaire tests with questions on the intention to use the BCI videogame in the future, etc. The third guideline recommends defining and measuring several ergonomic criteria and underlines the importance of evaluating one criterion using several metrics. Indeed, designing a technological system suited to the end user implies integrating several criteria and not just one: typically, a usable BCI videogame cannot be used if it is not useful for its players. The methodological framework incorporates several ergonomic criteria such as usefulness, usability, acceptability, hedonic qualities, appeal, immersion and presence, emotions and user experience, and several metrics for each of them.

Conclusion

The analysis of the BCI applications literature suggested that current evaluations of BCI videogames are mainly technocentric in order to technically improve the BCI videogames before using them in real contexts of use. However, some recent works tend to show the need to include characteristics, expectations, and requirements of users in early design phases. To do so, methodological guidelines, from literature on ergonomics and human-computer interaction, are necessary.

In this paper, a user-centered methodological framework was proposed to guide the designers of BCI videogames so that the games are adapted to human characteristics, to the players' needs, and to the context in which these applications will be integrated (e.g., at home, in a classroom, during a medical consultation). For each user-centered design phase, our framework discussed methods, criteria, and metrics to guide designers in the design and evaluation of the usefulness, usability, acceptability, hedonic qualities, appeal, immersion and presence, emotions, and more general user experience associated with BCI videogames for therapy, learning, or entertainment.

The implementation of this methodological framework to design and evaluate numerous BCI videogames could suggest several research perspectives. A first perspective could be to assess empirically the potential of a user-centered methodological framework, comparing one BCI videogame resulting from a technocentric design process (i.e., as traditionally implemented to develop these applications) and an application resulting from a design process based on the user-centered framework. A second perspective could be to improve the framework with the empirical results using it to design applications for different scopes (e.g., BCI videogames for learning). An example of results could be the enhancement of the metrics' database for the usefulness criterion that is highly dependent on the applications' scope. Because one conclusion of the massive use of this framework could be a partial use depending on the application, a third perspective would be to develop a tool to assist in the decision-making concerning the selection of methods, criteria, and metrics according to the scope of the application.

Cross-References

▶ Brain-Computer Interface Games: Towards a Framework
▶ Brain-Computer Interfacing and Virtual Reality
▶ Ethics, Privacy, and Trust in Serious Games
▶ Introduction to the Ethics of New and Emerging Science and Technology

Recommended Reading

B.Z. Allison, C. Neuper, Could anyone use a BCI? in *Brain-Computer Interfaces* (Springer, London, 2010), pp. 35–54

M. Anastassova, J.-M. Burkhardt, C. Mégard, A. Leservot, User-centred design of mixed reality for vehicle maintenance training: an empirical comparison of two techniques for user needs analysis, in *HCI International* (2005)

B.S. Badia, H. Samaha, A.G. Morgade, P.F.M.J. Verschure, Exploring the synergies of a hybrid BCI – VR neurorehabilitation system, in *International Conference on Virtual Rehabilitation (ICVR)* (2011). doi: 10.1109/ICVR.2011.5971813

A. Bandura, *Self Efficacy: The Exercise of Control* (Freeman, New York, 1997)

S. Blain-Moraes, R. Schaff, K.L. Gruis, J.E. Huggins, P.A. Wren, Barriers to and mediators of brain-computer interface user acceptance: focus group findings. Ergonomics **55**, 516–525 (2012). doi:10.1080/00140139.2012.661082

A. Blandford, T.R.G. Green, D. Furniss, S. Makri, Evaluating system utility and conceptual fit using CASSM. Int. J. Hum. Comput. Stud. **66**, 393–409 (2008). doi:10.1016/j.ijhcs.2007.11.005

L. Bonnet, F. Lotte, A. Lécuyer, Two brains, one game: design and evaluation of a multi-user BCI video game based on motor imagery, in *IEEE Transactions on Computational Intelligence and Artificial Intelligence in Games* (*IEEE TCIAIG*), vol. 5 (2013), pp. 185–198. doi:10.1109/TCIAIG.2012.2237173

J.H. Brockmyer, C.M. Fox, K.A. Curtiss, E. McBroom, K.M. Burkhart, J.N. Pidruzny, The development of the Game Engagement Questionnaire: a measure of engagement in video game-playing. J. Exp. Soc. Psychol. **45**, 624–634 (2009). doi:10.1016/j.jesp.2009.02.016

J.-M. Burkhardt, T. Lubart, Creativity in the age of emerging technology: some issues and perspectives in 2010. Creativity Innov. Manag. **19**, 160–166 (2010). doi:10.1111/j.1467-8691.2010.00559.x

M. Donnerer, A. Steed, Using a P300 brain-computer interface in an immersive virtual environment. Presence: Teleoperators Virtual Environ. **19**, 12–24 (2010). doi:10.1162/pres.19.1.12

J.I. Ekandem, T.A. Davis, I. Alvarez, M.T. James, J.E. Gilbert, Evaluating the ergonomics of BCI devices for research and experimentation. Ergonomics **55**, 592–598 (2012). doi:10.1080/00140139.2012.662527

C. Escolano, J. Antelis, J. Minguez, Human brain-teleoperated robot between remote places, in *IEEE International Conference on Robotics and Automation* (*ICRA*) (2009), pp. 4430–4437. doi:10.1109/ROBOT.2009.5152639

L. George, A. Lécuyer, An overview of research on "passive" brain-computer interfaces for implicit human-computer interaction, in *International Conference on Applied Bionics and Biomechanics* (2010)

C. Groenegress, C. Holzner, C. Guger, M. Slater, Effects of p300-based BCI use on reported presence in a virtual environment. Presence: Teleoperators Virtual Environ. **19**, 1–11 (2010). doi:10.1162/pres.19.1.1

C. Guger, G. Edlinger, W. Harkam, I. Niedermayer, G. Pfurtscheller, How many people are able to operate an EEG-based brain-computer interface (BCI)? IEEE Trans. Neural Syst. Rehabil. Eng. **11**, 145–147 (2003)

H. Gürkök, A. Nijholt, M. Poel, M. Obbink, Evaluating a multi-player brain-computer interface game: challenge versus co-experience. Entertain. Comput. **4**, 195–203 (2013). doi:10.1016/j.entcom.2012.11.001

H. Gürkök, B. van de Laar, D. Plass-Oude Bos, M. Poel, A. Nijholt, Players' opinions on control and playability of a BCI game, in *International Conference on Universal Access in Human-Computer Interaction* (*UAHCI*) (2014), pp. 549–560. doi:10.1007/978-3-319-07440-5_50

G. Hakvoort, H. Gürkök, D. Plass-Oude Bos, M. Obbink, M. Poel, Measuring immersion and affect in a brain-computer interface game, in *Human-Computer Interaction – INTERACT 2011*, ed. by P. Campos, N. Graham, J. Jorge, N. Nunes, P. Palanque, M. Winckler. Lecture Notes in Computer Science (Springer, Berlin/Heidelberg, 2011), pp. 115–128

M. Hassenzahl, The effect of perceived hedonic quality on product appealingness. Int. J. Hum. Comput. Interact. **13**, 481–499 (2001). doi:10.1207/S15327590IJHC1304_07

S.I. Hjelm, C. Browall, Brainball – using brain activity for cool competition, in *NordiCHI* (2000)

E.M. Holz, J. Höhne, P. Staiger-Sälzer, M. Tangermann, A. Kübler, Brain–computer interface controlled gaming: evaluation of usability by severely motor restricted end-users. Artif. Intell. Med. **59**, 111–120 (2013). doi:10.1016/j.artmed.2013.08.001

W. IJsselsteijn, W. van den Hoogen, C. Klimmt, Y. de Kort, C. Lindley, K. Mathiak, K. Poels, N. Ravaja, M. Turpeinen, and P. Vorderer. Measuring the experience of digital game enjoyment. In A.J. Spink, M.R. Ballintijn, N.D. Bogers, F. Grieco, L.W.S. Loijens, L.P.J.J. Noldus, G. Smit, and P.H. Zimmerman (Eds.), Proceedings of Measuring Behavior (2008), Maastricht, The Netherland, August 26–29, 2008

I. Iturrate, J.M. Antelis, A. Kübler, J. Minguez, A noninvasive brain-actuated wheelchair based on a P300 neurophysiological protocol and automated navigation. IEEE Trans. Robot. **25**, 614–627 (2009). doi:10.1109/TRO.2009.2020347

C. Jeunet, A. Cellard, S. Subramanian, M. Hachet, B. N'Kaoua, F. Lotte, How well can we learn with standard BCI training approaches? A pilot study, in *International Brain-Computer Interface Conference* (2014)

A. Kübler, E.M. Holz, C. Zickler, T. Kaufmann, A user centred approach for bringing BCI controlled applications to end-users, in *Brain-Computer Interface Systems – Recent Progress and Future Prospects,* ed. by R. Fazel-Rezai (InTech, Croatia 2013). doi:10.5772/55802

F. Lotte, Y. Renard, A. Lécuyer, Self-paced brain-computer interaction with virtual worlds: a qualitative and quantitative study "out-of-the-lab," in *International Brain-Computer Interface Workshop and Training Course* (2008)

F. Lotte, J. Faller, C. Guger, Y. Renard, G. Pfurtscheller, A. Lécuyer, R. Leeb, Combining BCI with virtual reality: towards new applications and improved BCI, in *Towards Practical Brain-Computer Interfaces: Bridging the Gap from Research to Real-World Applications*, ed. by B.Z. Allison, S. Dunne, R. Leeb, J.R. Millán, A. Nijholt (Springer, Berlin/Heidelberg, 2013). doi:10.1007/978-3-642-29746-5_10

E. Loup-Escande, O. Christmann, Requirements prioritization by end-users and consequences on design of a virtual reality software: an exploratory study, in *International Conference on Evaluation of Novel Approaches to Software Engineering* (*ENASE*) (2013), pp. 5–14

E. Loup-Escande, J.-M. Burkhardt, S. Richir, Anticipating and evaluating the usefulness of emerging technologies in ergonomic design: a review of usefulness in design. Le travail Humain **76**, 25–55 (2013b). doi:10.3917/th.761.0027

S. Mahlke, M. Minge, M. Thüring, Measuring multiple components of emotions in interactive contexts, in *CHI Extended Abstracts on Human Factors in Computing Systems* (2006), pp.1061–1066. doi:10.1145/1125451.1125653

C. Mühl, Neurophysiological assessment of affective experience, in *Affective Computing and Intelligent Interaction* (2009)

G. Müller-Putz, R. Scherer, G. Pfurtscheller, Game-like training to learn single switch operated neuroprosthetic control, in *Internationational Conference on Advances in Computer Entertainment Technology* (2007)

M. Mulvenna, G. Lightbody, E. Thomson, P.J. McCullagh, M. Ware, S. Martin, Realistic expectations with brain computer interfaces. J. Assist. Technol. **6**, 233–245 (2012). doi:10.1108/17549451211285735

F. Nijboer, N. Birbaumer, A. Kubler, The influence of psychological state and motivation on brain-computer interface performance in patients with amyotrophic lateral sclerosis – a longitudinal study. Front. Neurosci. **4** (2010). doi:10.3389/fnins.2010.00055

A. Nijholt, H. Gürkök, Multi-brain games: cooperation and competition, in *Universal Access in Human-Computer Interaction. Design Methods, Tools, and Interaction Techniques for Inclusion*, ed. by C. Stephanidis, M. Antona. Lecture Notes in Computer Science (Springer, Berlin/Heidelberg, 2013), pp. 652–661

D. Plass-Oude Bos, B. Reuderink, B. Laar, H. Gürkök, C. Mühl, M. Poel, A. Nijholt, D. Heylen, Brain-computer interfacing and games, in *Brain-Computer Interfaces, Human-Computer Interaction Series*, ed. by D.S. Tan, A. Nijholt (Springer, London, 2010), pp. 149–178

D. Plass-Oude Bos, M. Poel, A. Nijholt, A study in user-centered design and evaluation of mental tasks for BCI, in *International Conference on Advances in multimedia modeling* (2011), pp. 122–134. doi:10.1007/978-3-642-17829-0_12

A.T. Pope, C.L. Stevens, Interpersonal biocybernetics: connecting through social psychophysiology, in *ACM International Conference on Multimodal Interaction* (2012), pp. 561–566. doi:10.1145/2388676.2388795

P. Rabardel, P. Beguin, Instrument mediated activity: from subject development to anthropocentric design. Theor. Iss. Ergon. Sci. **6**, 429–461 (2005). doi:10.1080/14639220500078179

S. Robertson, Requirements trawling: techniques for discovering requirements. Int. J. Hum. Comput. Stud. **55**, 405–421 (2001). doi:10.1006/ijhc.2001.0481

M.V. Sanchez-Vives, M. Slater, From presence to consciousness through virtual reality. Nat. Rev. Neurosci. **6**(4), 332–339 (2005)

R. Scherer, G. Moitzi, I. Daly, G.R. Muller-Putz, On the use of games for noninvasive EEG-based functional brain mapping. IEEE Trans. Comput. Intell. AI Games **5**, 155–163 (2013). doi:10.1109/tciaig.2013.2250287

M. Schreuder, A. Riccio, M. Risetti, S. Dähne, A. Ramsav, J. Williamson, D. Mattia, M. Tangermann, User-centered design in brain–computer interfaces – a case study. Artif. Intell. Med. **59**, 71–80 (2013). doi:10.1016/j.artmed.2013.07.005

M. Slater, Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. Philos. Trans. R. Soc. B **364**(1535), 3549–3557 (2009)

E. Thomas, M. Dyson, M. Clerc, An analysis of performance evaluation for motor-imagery based BCI. J. Neural Eng. (2013). doi:10.1088/1741-2560/10/3/031001

G. Valsan, B. Grychtol, H. Lakany, B.A. Conway, The Strathclyde brain computer interface, in *IEEE Engineering in Medicine and Biology Society* (2009), pp. 606–609. doi:10.1109/IEMBS.2009.5333506

B. Van de Laar, H. Gurkok, D. Plass-Oude Bos, F. Nijboer, A. Nijholt, Perspectives on user experience evaluation of brain-computer interfaces, in *Universal Access in Human-Computer Interaction – Users Diversity*, ed. by C. Stephanidis. Lecture Notes in Computer Science (Springer, Berlin/Heidelberg, 2011), pp. 600–609

V. Venkatesh, M.G. Morris, G.B. Davis, F.D. Davis, User acceptance of information technology: toward a unified view. MIS Q. **27**, 425–478 (2003)

B.G. Witmer, M.J. Singer, Measuring presence in virtual environments: a presence questionnaire. Presence: Teleoperators Virtual Environ. **7**, 225–240 (1998). doi:10.1162/105474698565686

J. Wolpaw, N. Birbaumer, D. McFarland, G. Pfurtscheller, T. Vaughan, Brain-computer interfaces for communication and control. Clin. Neurophysiol. **113**, 767–791 (2002)

N. Yan, J. Wang, M. Liu, L. Zong, Y. Jiao, J. Yue, Y. Lv, Q. Yang, H. Lan, Z. Liu, Designing a brain-computer interface device for neurofeedback using virtual environments. J. Med. Biol. Eng. **28**, 167–172 (2008)

F. Yang, W. Chen, B. Wu, Y. Qi, J. Luo, Y. Su, J. Dai, X. Zheng, An adaptive BCI system for virtual navigation, in *International Conference on Information Science and Engineering* (*ICISE*) (2010), pp. 64–68. doi:10.1109/ICISE.2010.5688650

C. Zickler, S. Halder, S.C. Kleih, C. Herbert, A. Kübler, Brain painting: usability testing according to the user-centered design in end users with severe motor paralysis. Artif. Intell. Med. **59**, 99–110 (2013). doi:10.1016/j.artmed.2013.08.003

Emilie Loup-Escande is assistant professor at the CRP-CPO (EA7273) from Université de Picardie Jules Verne. She is interested in the ergonomic design of emerging technologies. She has been involved in the design and evaluation of Virtual Reality applications for the design of children's' products, vocational learning by the disabled person and learning in education

Fabien Lotte is a research scientist at Inria Bordeaux Sud-Ouest in France, a research Institute dedicated to Computer Science and Applied Mathematics, which he joined in 2011. He has been working for about 10 years on Brain-Computer Interface design, study and applications. He notably focused on brain signal processing, Virtual Reality and gaming applications of BCI as well as BCI user training

Guillaume Loup is a phd student at the LIUM Lab from the University of Maine. He is interested in the realization of interactions 2D/3D real-time, the multiplatform implementation and the integration of devices for interacting with virtual environment. His expertise area concerns the design of mixed reality-based games

Anatole Lécuyer is a senior researcher at Inria, the French National Institute for Research in Computer Science and Control, which he joined in 2002. His main research interests are in Virtual Reality, 3D User Interface, Haptic Feedback, 3D Visual Displays, and Brain-Computer Interfaces

Part III

Entertainment Games

Current Status of Applying Artificial Intelligence in Digital Games

11

Youichiro Miyake

Contents

Y. Miyake (✉)
Square Enix Co., Ltd., Tokyo, Japan
e-mail: miyakey@square-enix.com; y.m.4160@gmail.com

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_70

Abstract
Artificial intelligence in digital games has developed in the last 40 years. It has a long and deep history with digital games. AI techniques in digital games evolved independently and differently from the academic AI research of science and engineering which require functionality in the real world. Digital games have complex and large-scale virtual 2D/3D worlds where game characters live in, recognize, make decisions, and design their motions to fitting their environment. The digital world of games is larger, more complex, and more detailed than any other virtual world. It is the most suitable experimental field to study and evaluate AI technologies in the virtual world. The AI for a game character is called "Character AI" and its function is for characters to make decisions. This is much different from functional AI in academic research, and making a character means to create one whole intelligence. The other unique AIs are "Navigation AI" which analyzes and recognizes the environment of game world and "Meta AI" which dynamically controls and changes the progress, situation, and drama of the game. These three AIs, namely, Character AI, Navigation AI, and Meta AI, cooperate with each other and develop one unified system to form a dynamic user experience. In addition, recently learning and evolution approaches have been introduced into AI for digital games. In this chapter such current status of AI in digital games is described.

Keywords
Character AI • Navigation AI • Meta AI • Agent Architecture • Behavior tree • Goal-Oriented Action Planning • Evolutionary game design

Introduction

One of the goals in digital games is to produce good user experiences. Digital game technologies can be called "Art and Science" as they create various kinds of experiences, and they produce their own unique experiences for each game by combining graphics, sound, AI, and other components. User experience has many variations and different intensity of emotions such as joy, fear, moving, excitement, etc. Game AI technology is one of game technologies to make user experience and belong to "Entertainment AI" (Miyake 2008a).

Both AI technologies in conventional games and digital games have progressed together. In the 1980s, a digital game was still small software, and there were little number of games to which AI technologies were applied. Games with AI technologies had to spend much processing power for AI, so that they look like “AI Game” (Morikawa 1999). But recently, digital games have become large-scale software with complex architecture, and it has become consisting of many modules. In these days, AI technologies are implemented in many modules of game software and give different intelligence to each module. The history of digital game is nearly 40 years long, and in the last 10 years digital games have become more structured. A rapid evolution of computer hardware produces more and more processing power and storage resources, and these combinations give both game and AI progress. In this chapter, a current perspective of game AI is described.

The most distinguished feature of digital games from the other AI is real-time processing which updates own processing status within 1/30 s or 1/60 s like robotic control. AI in digital games is also real-time system and restricted by the limit of update time. This limitation gives special feature to digital game AI.

Three Types of AI for Digital Games

AI technologies had been applied to digital games in the 1980s and 1990s. But in the latter half of the 1990s, game AI became large scale in size and well structured. They have specific roles and are divided into the three specific AIs: Character AI, Meta AI, and Navigation AI (Miyake 2010a).

Character AI

A virtual human or animal in digital games is called “character.” A character controlled by a player is called “player character,” and a character who is not controlled by a player is called “non-player character” (NPC) such as enemy or buddy. Sometimes, more simply, NPC is called “AI” (Fig. 1).

Character AI, which is evolving now, is the most important goal in digital game AI. There are two directions to evolve. One is autonomous AI who makes a decision by itself, and the other is entertainment AI who speaks previously prepared lines and acts as directed by scenario to make players enjoy (Miyake 2013).

Meta AI

Meta AI is the AI of game system itself. Meta AI has no physical body in game world. Game system is a module to control game sequence. The basic role of game system is to prepare and open game contents step by step. But Meta AI has further role to generate contents dynamically by observing users’ skills and game status. Furthermore, Meta AI changes game difficulty according to a user’s skill. Meta AI can distribute enemies at any place by predicting user’s route, generate a dungeon and terrain, and control timing and probability to spawn enemies.

Fig. 1 Final Fantasy XIV (2013, Square Enix). Player character (*front*) and treelike enemy character (*back*) (Final Fantasy XIV © 2010–2016 Square Enix Co., Ltd. All Rights Reserved)

Navigation AI

In digital games, a "level" means game world including terrains, objects, and characters. Navigation AI has a role to transform static and dynamic level data depending on character's specific data obtained by Character AI. For example, Navigation AI has navigation data for character's pathfinding and modifies them dynamically when the terrain and situation of the level change in game. Also Navigation AI has no physical body and supports Character AI and Meta AI by providing them level information.

These three types of distributed AI cooperate with each other to take an action as one game AI system (Fig. 2). A Character AI works as a brain of each NPC, Meta AI modulates game design by distributing and generating characters as game proceeds, and Navigation AI analyzes a game level and provides Character AI and Meta AI with information about terrain and game status. Navigation AI makes it possible for Meta AI and Character AI to make decisions by using information of the current level state dynamically.

Internal Structure of Digital Game

To explain how AI works for digital games, internal structure of modern digital game needs to be described first. "Objects" are nonliving thing on a level such as armor, rocks, items, and houses. Ground of level, called "terrain," includes forests, glasses, walls, and bridges which are impossible to move. A component of level is called "entity" such as a character, an object, and a terrain. Level design is a collective system of all components.

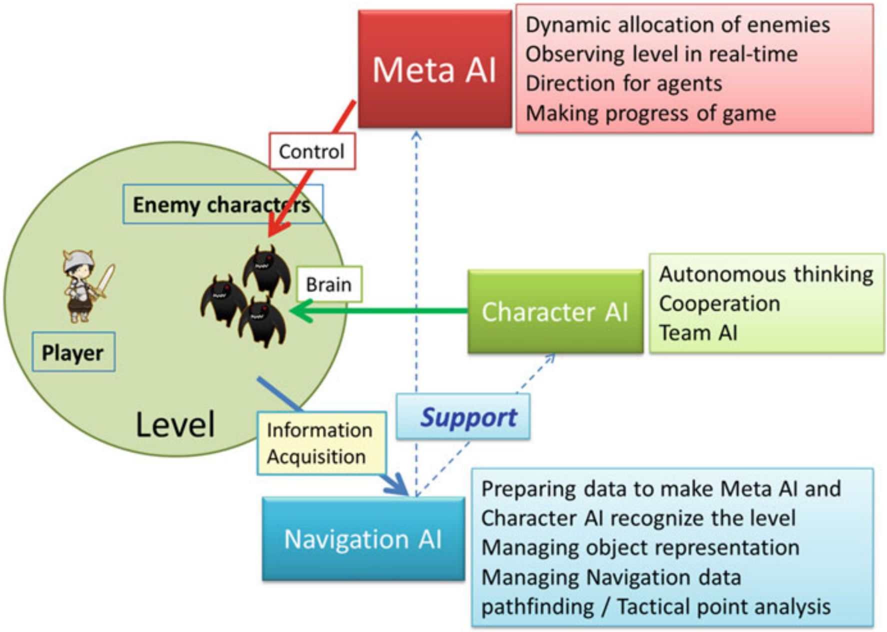


Fig. 2 Cooperation of three types of AI: Meta AI, Character AI, and Navigation AI

Physical interaction, such as collision, rotation, gravity, and flow, is defined for entities. For example, when an entity drops into a river, water flow pushes and moves it. An explosion quiver leaves, and a door is broken by strong collision into broken pieces.

An entity consists of a code and data. Data of entity is called “asset” such as 3D model data, texture data, shader data, and collision model. Assets are created by artists and codes are written by engineers, and scripts are written by level designers (Miyake 2007a).

Digital game system has multiple-layer structure and each layer is independent from others (Gregory 2009a). There are four layers related with character: graphics, collision, AI, and animation. The data of each layer is processed in each layer, but the processing order between layers is controlled by strict rules. The graphics pipeline should be updated in one frame in game, but AI pipeline needs not be updated within one frame. It is possible for AI pipeline to be updated asynchronously (Fig. 3).

Graphics Layer

Graphics layer creates a user’s vision. Until the end of the 1980s, game screen consists of 2D dots, pixels, and sprites. From the beginning of the 1990s, 3D graphics is introduced and graphics layer has become to have a huge number of assets such as 3D model data, textures, and shader data. Asset data which belongs to graphics layer is called graphics data.

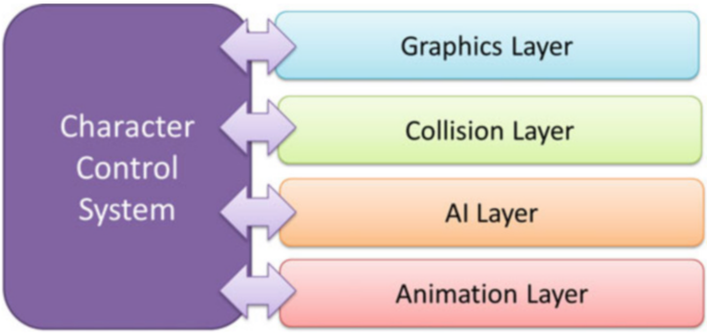


Fig. 3 Four layers for character control

Collision Layer

Collision layer controls interactions between entities. Data belonging to the collision layer is called "collision model." A collision model consists of rectangular parallelepipeds, spheres, and planes. Their combination covers the entire shape of an object. All physical interactions of entities are calculated by using collision models. For example, interactions between one character and objects are calculated by using their collision models.

AI Layer

AI layer has a role to make it possible for a character to recognize its environment in a game world. Huge data of knowledge representation of terrains, objects, and situations belongs to AI layer. These data form the recognition of game level. For example, data of a rock in AI layer should include attributes (soft, hard), affordance (movable, breakable), and rules (what will happen when a rock is moved). Each entity has its knowledge representation and forms recognition of object for AI.

Animation Layer

An animation layer has a role to control a motion for entities with bones and mutual restrictions to move autonomously. A character has bones like an animal and moves by replaying character animation data or makes a different motion calculating its character animation data dynamically.

Each layer calculates each data in the layer, and an order of calculation is restrictedly defined between layers.

Making Character AI

While technical works in modern game AI development have been focused on building AI layer, game design works have been directed to make characters by integrating graphics, collision, AI, and animation module. Each module has each respective data, and each respective data is processed in each pipeline. A character in game is required to move in natural motion and to realize natural motion even in a complex environment. It needs synchronous cooperation of three layers: AI, collision, and animation. AI module makes its decision and selects one animation. When it moves a character's body, collisions with objects and ground occur. Furthermore, it causes physical feedback to the character's body, and the character animation is deformed and changed rapidly in response. Basically an animation change is executed in an animation layer, but for a drastic change of motion, AI layer issues a command to animation layer. In such a way, AI, collision, and animation layers are closely related with each other. The main topic to develop character system is to build up interactions between layers.

Agent Architecture

A character has internal intelligence and a body, and the outer environment in game has natural structure to it. Environment is on the outside of a character and intelligence is on the inside of a character. A "sensor" obtains information from outer environment and sends them into the inside a character, and "effector" has an influence to the environment from the inside. The model that the inside of character is connected to the outside environment via sensor and effector is called "agent architecture" (Russell and Norvig 2009). In the area dividing the inside and outside of a character, there is a body of a character which has sensor and effector within it.

Agent architecture consists of four components connected in a circle which consists of sensor, intelligence, effector, and environments. AI develops knowledge and recognition from information in memory which sensors gather from the environment, makes its decision, and has an influence to the environment via effector. The loop of components keeps the activity of getting information from the environment (Fig. 4).

Agent architecture in digital game has the feature of a dynamic one-directed cyclic loop of information flow. The information loop is called "information flow" (Griesemer 2002). Once information flow is formed, it has hysteresis to keep its dynamic connection between environment and intelligence. Furthermore, information flow consists of many smaller loops (Miyake 2013, 2014).

There is also internal information flow in intelligence module called "internal cyclic information flow," which achieves evoking and inner thinking (Miyake 2014).

An introduction of agent architecture to game character development begins from the "C4 architecture model" for a creature's brain in virtual environment proposed by Synthetic Character Group in MIT Media Lab (MIT Synthetic Character Group 2003) in the late 1990s (Fig. 5) (Isla et al. 2001; Burke et al. 2001). C4 architecture was lectured by Damian Isla at GDC (Game Developers Conference) 2000. Later agent architecture was introduced into character's brain model of "Halo" (2001, Bungie) (Griesemer 2002). It was followed by many large-scale FPS (first-person shooter) games such as "F.E.A.R." (2005, Monolith Production), "Killzone 2" (2009, Guerrilla Games), and other games (Miyake 2008a; Orkin 2005, 2006; Straatman et al. 2009).

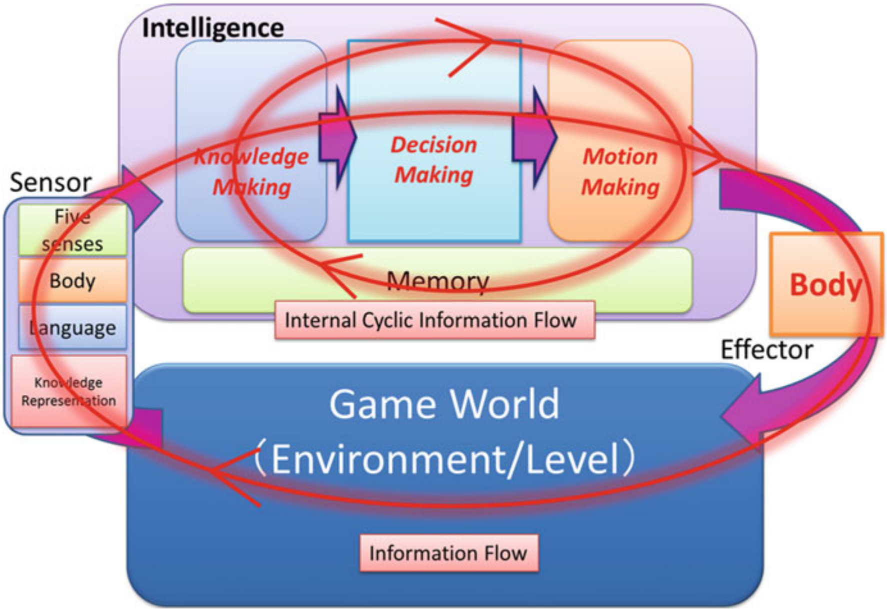


Fig. 4 Agent architecture and information flow

The main distinguishing feature of C4 architecture is blackboard architecture. Blackboard architecture is a widely used system that helps several distributed small AI software cooperate with each other. There are many use cases (Penny Nii 1986a, b; Ishida et al. 1996). Although in the 1990s it had become less used, in the 2000s it came to be used for building the internal model of agent architecture. Blackboard architecture consists of three components such as some numbers of KS (Knowledge Source) with specific function, a blackboard where KS writes and reads information, and one arbiter to control all of KS. All of KS are connected via a blackboard (Fig. 6). C4 architecture has the same structure as blackboard such that a group of modules (Fig. 5, left) are connected via memory regions (Fig. 5, right) (Isla and Blumberg 2002).

Sensor

"Sensor" refers to not only five primary senses but also sensors to get all information from the environment such as physical force to a body, language communication, and abstract information. Sensors form a big pipe to the environment to acquire concrete and abstract information and connect intelligence.

Sense of Vision

A character's vision is implemented to judge whether the object can be seen or not. Usually, this judgment is executed by two steps. The first step is to search whether an

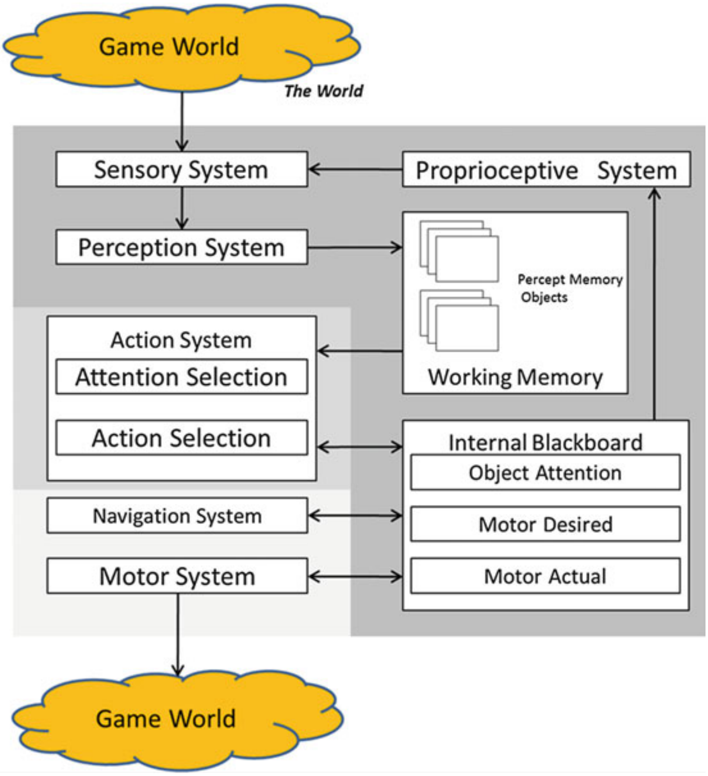


Fig. 5 C4 architecture (Burke et al. 2001)

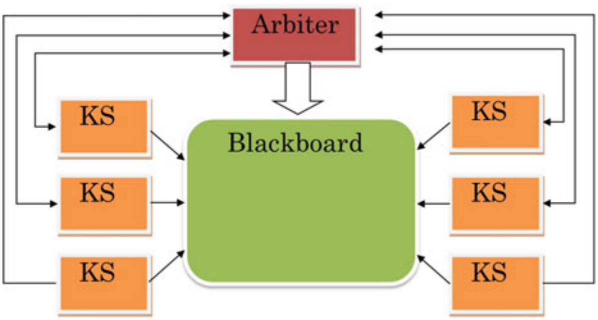


Fig. 6 Blackboard architecture (Isla and Blumberg 2002)

object is included in a region in front of the character (Walsh 2014). The second step is to judge whether a line between the character and the object is interrupted by other objects or not. This condition is called "LOS" (line of sight). Usually, two or three different LOS lines are used. AI does not use rendering data, but it gets knowledge representation (KR) data from an object which can be seen (in the AI layer) to make the recognition. KR data defines how an object should be recognized for a character.

The process of LOS judgments becomes heavy for a lot of objects. Several different ways to make it lighter have been researched. For example, in "Killzone 2" all of LOS to 384 directions from one point are pre-calculated for all directions in game development process, so that LOS process in game runtime becomes much lighter by using these data (Leeuw 2009). In Saints Row IV, by using multi-thread, it becomes lighter (Canary 2014).

Sense of Hearing

The implementation of hearing sensor is sound propagation simulation similar to LOS. A different point from LOS is sound diffraction. It is the judgment of whether or not a sound will arrive to a character from a sound origin. One way is to judge whether a character is within a propagation region (ordinarily a circle with a constant radius). Furthermore, in Splinter Cell: Conviction (2010, Ubisoft), a sound propagation simulation is executed rightly by pathfinding technology (Walsh 2014; Guay 2012). Like a sense of sight, a character gets KR data from an object from which it can hear the sound.

In this way, five sensors are implemented, instead of carrying out simulation, by easy alternative way to acquire information of KR data of objects. The other senses such as taste, smell, and touch are rarely simulated.

Physical Force

A character can change a position, posture, and motion by external force from object collision and recognizes a fact that it receives a force. There are a few reactive actions prepared for external forces: one case is to response automatically to an external force (like falling back animation), the second case is to keep balance automatically (e.g., a character keeps walking motion against weak external force), and the last case is to change the animation when the force is strong enough. There are multiple ways to response to external force in multiple layers of intelligence and the body. It is also the same for vector force field such as gravity, explosion, water flow, and crowd flow (Anguelov et al. 2012; Fauerby 2012).

Language Communication

When characters communicate in a digital game, or when a commander gives an order to a member, they will exchange symbols not by language but by predetermined symbols. Symbols such as "Gather," "Advance," and "Retreat" are directly sent to each other's recognition modules, not via sensors between characters (Isla 2008; Miyake 2010b).

Effector

In many cases, an "effector" is a collision model of character body or tools and weapons the character equips. But in the case of commander AI, it does not have a physical body, but gives indirect force to the environment by using a battle unit, soldiers, and cannons. Furthermore, there is an abstract effector such as languages and orders that give nonphysical influence instead of physical influence. The information needed for a character to take actions is acquired via sensors. Conversely, after an action is taken, some specified sensors become activated to check the result of action. The relation between sensor and effector is called "coordination" (Bernstein 1996). For example, to attack an enemy, a character's eyes need to follow the movement of the enemy, or after the attack action, it should be confirmed whether or not the attack was successful. Once an information flow to connect sensors, decisions, and effectors is formed, it connects also the environment and intelligence dynamically (Gibson 1986). Furthermore, the relation between actions and results becomes a base to introduce learning algorithms.

Intelligence

Intelligence of game character has three main parts: recognition, decision-making, and action-making. It is the most fundamental part for intelligence that a character can recognize which actions are possible to take in the environment. That is called "affordance" (Gibson 1986). For example, when a character recognizes that "Punch," "Kick," "Guard," and "Release magic weapon" are possible to achieve, it selects one of them in decision-making process and coordinates the action to the situation in the action-making process. Information which goes through the knowledge-making module, decision-making module, and action-making module is stacked in memory (Miyake 2013, 2014).

Recognition (Knowledge Making)

A role of recognition is to reconstruct a self-centered image of environment. A character gathers knowledge representation data in the environment via sensors, reconstructs self-centered image of the world, and acquires more information needed to make its decision and complete the missing information for reasoning. More accurate recognition leads to more accurate decision-making. Recognition relies on what kind of form of information is required for decision-making. As memories are stacked by recognition data, the changes of the world are calculated by them.

A future prediction is possible by mining time-series data of world state. For a simplest case, a prediction of ball position is possible by collecting the ball's position in every frame and completing an orbit of the ball (Burke et al. 2001). In the same way, fire warning or ambush is possible.

A memory region where knowledge is stacked through recognizing a world is called "working memory." In the case of "F.E.A.R.," all memories are represented by a unified form of memory called WMF (working memory fact), and they are stacked and grouped for each type of object (Orkin 2005, 2006). Each WMF is represented as a list of a set of variable and confidence such as position and its confidence, direction and its confidence, and so on (Miyake 2010a). For AI in F.E.A.R., the world is represented by the 20 symbols extracted from stacked WMF memory data. For example, knowledge of enemy characters is represented by some symbols: position, possibility to fight, whether it is armed or not, what vehicle to ride, and so on. The symbolic information is also used in decision planning as action representation described in section "Goal-Based AI."

The other role of recognition is to form the recognition of its own body. As all living things recognize their own body state at any time, game characters should recognize their own body and understand what actions are possible now to carry out in the next frame. A representation of character's body consists of dynamic and static parameters. The dynamic parameters consist of position, acceleration, attitude, affordance, stamina, magical power, and so on. The static parameters consist of maximum speed, maximum acceleration, maximum stamina, attack strength, weapons and its effects, tools and its effect, and so on. Decision-making needs information of both outside information of the environment and internal information of body.

Decision-Making

In the case of turn-based games such as Shogi and Go, the act of decision-making is to select one move from all possible moves. In this case, decision-making is an algorithm of searching for the best move in a game tree.

On the other hand, decision-making problem in real-time game cannot be considered in limited number of states. Real-time game has unlimited variations of state. In the case of the game with continuous time and space, the approach of selecting the best move is not effective. A better approach for a character is to make the better situation by dynamic decision-making. There are two type of decision-making: one is called reactive decision-making, and the other is nonreactive decision-making (Miyake 2012a).

There are typical seven types of decision-making algorithm represented as "-based" where "-based" indicates which component is used for a base unit. For example, "state-based" means thinking by a unit of state, and "goal-based" means that thinking begins from setting a first goal. There are reactive algorithms such as rule-based, state-based, and behavior-based. In the next chapter, seven types of algorithms are explained.

Motion Making

A character's decision-making process generates an action for the next moment as an output in a game. In the case of board games, an action is represented as a coordinate of a move. But in action game, a character has a complex structured body, and a decided action must be adjusted to the environment through complex process.

For example, when a character is assigned to the state "Walk," it needs to generate a motion to walk such as bending its knees and fit its feet to the bumpy road along a path. So it needs to combine pathfinding, IK (inverse kinematics), animation blending, collisions, and so on. In this way, the process to generate a motion from decision-making is necessary, which is called "motion making."

Decision-Making Algorithms

The most popular seven algorithms used for character decision-making are explained in this chapter (Miyake 2012a).

Rule-Based AI

Rule-based AI is a method to think using a collection of rules. Ordinarily the system has several numbers of rules, selects the best one to fit the situation, and takes an action by following the selected rule. In some cases, a routine to select one rule is implemented (Fig. 7). In many cases, a process to select one rule out of more than two rules is activated under the condition that priority of all rules is predefined.

In some of RPGs (role-playing games) including Dragon Age (2009, BioWare), it is possible for a user to set priority into predefined rules for character control. For example, a rule might be "if stamina is less than half-full, cast a revival magic" or "attack the strongest enemy in enemy group." In "The Sims 3" (2009, Electronic Arts), many production rules such as "if time becomes evening, watch TV" or "if you wake up in the morning, go out for jogging" are stored in character's memory, and they make up the personality of the character (Evans 2010).

State-Based AI

A "state" describes a character's action. When a character is in some state, it takes an action defined in that state. For example, when a character is in the "Walk" state, it keeps walking. When a state changes to the other "Attack" state, the character keeps attacking as described in the state. Each character has many states, and state transition happens by a specific condition. A condition from one state to the other is called "transition condition." A system of states connected via transition conditions is called a "finite-state machine" (FSM). Finite-state machine has been popularly used from the 1990s to the present day. It is used by thousands of game developers in games such as Quake (1993, id Software) and Uncharted 2 (2009, Naughty Dog) (Miyake 2010a; Evans 2010). For large-scale games, "hierarchical

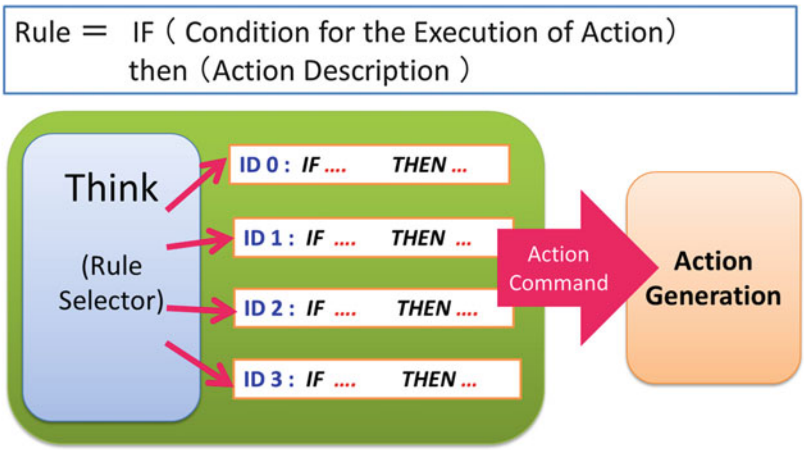


Fig. 7 Rule-based AI

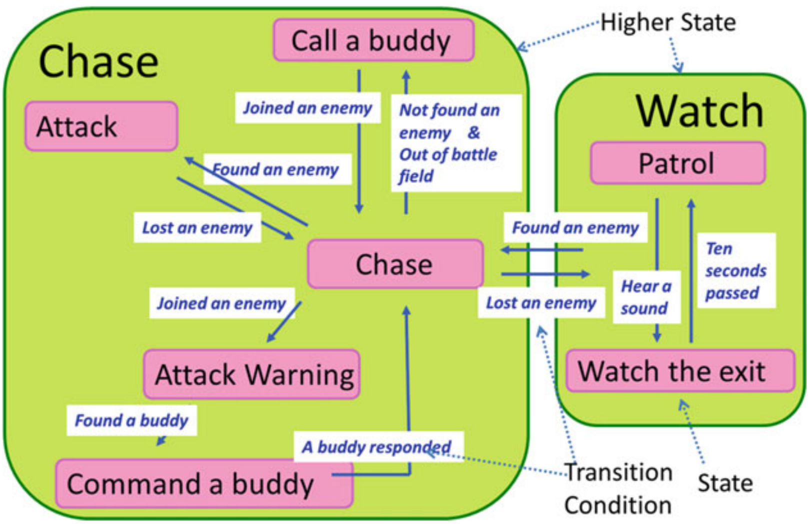


Fig. 8 Hierarchical finite state machine

finite state machine" (HFSM) is used. HFSM has a structure such that a state in a high layer has a lower state machine (Fig. 8).

Behavior-Based AI

A behavior describes a character's physical motion. Behavior-based AI is the system that decides a character's physical behavior. In these years, "behavior tree" system has become the most popular system in the game industry, and state machine has been replaced by behavior tree. Behavior tree was invented by Damian Isla to develop "Halo2" (2004, Bungie). Behavior tree has no cyclic structure like a state machine, but instead has a hierarchical tree structure which activates nodes from a root node to a leaf node (Fig. 9) (Isla 2005a). Only a leaf node describes the physical motion of a character. Other middle nodes describe abstract actions. When a leaf node becomes active, a behavior is executed. After a leaf node process finishes, the root node becomes active again (Miyake 2012a).

A behavior tree selects one node in each layer by child competitive principle, not by parent node selection. Each layer in behavior tree has several number of child nodes, and these nodes compete with each other to become active. The model is called "child competitive model," as a parent node does not select a child one, but child nodes compete with each other under "selection rule" defined in each layer. First of all, each behavior judges whether it is possible to become active or not under the current situation, and the nodes impossible to become active are removed from the selection. Then, nodes are selected based on rules applied to those remaining. Under the selection rule, a selection process goes. For example, "sequence rule" executes nodes in the order of priority in a layer, or "priority rule" executes only one

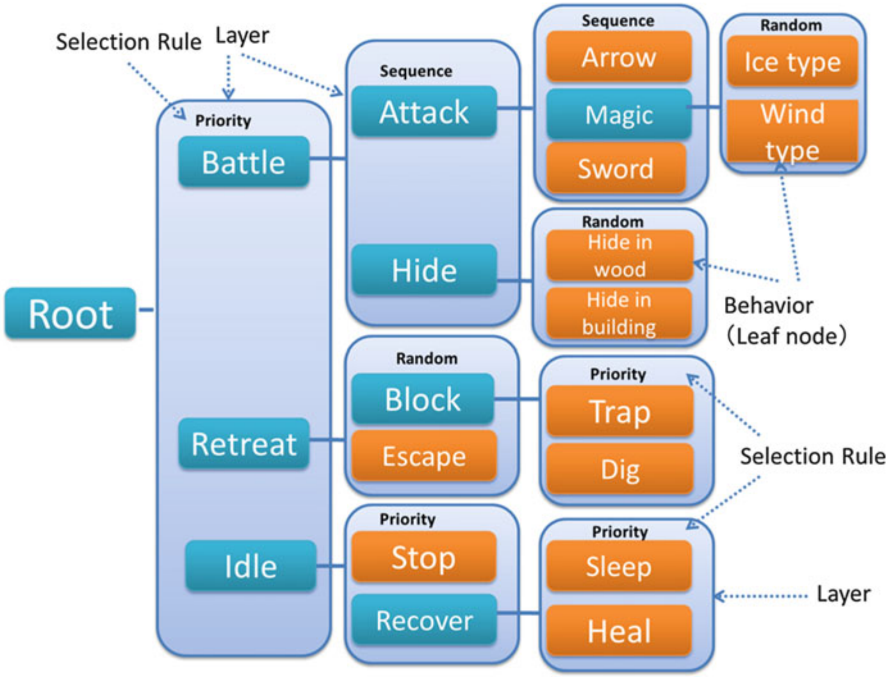


Fig. 9 Behavior tree

node with max priority in all nodes possible to be active. "Random rule" executes one node picked up at random.

Behavior tree has many restrictions, but it also has several good features. The first is that game designers can build behavior tree by themselves using behavior tree tool. The second is scalability to make AI from small-scale one to large-scale one on the behavior tree. The third is that it is easy to debug on visual tools. For these reasons, behavior tree has become a default tool to make AI. It is used in hundreds of game titles such as Spore (2008, Maxis) and CRYSIS (2012, Crytek) (Hecker 2009; Pillosu 2009).

Goal-Based AI

The best feature of reactive algorithms such as rule-/state-/behavior-based algorithms is to response to the situation immediately. For example, when a character avoids a flying arrow, or recovers itself when its stamina becomes a half of full, reactive algorithms make rapid decisions for a character. But it is impossible for reactive algorithms to have a long-term goal for a character. A goal-based algorithm has the function that when a character sets a long goal, it designs an action sequence to achieve the goal. For a game development, there are two types of goal-based algorithms. One is GOAP (Goal-Oriented Action Planning), and the other is "Hierarchical Goal-Oriented Planning."

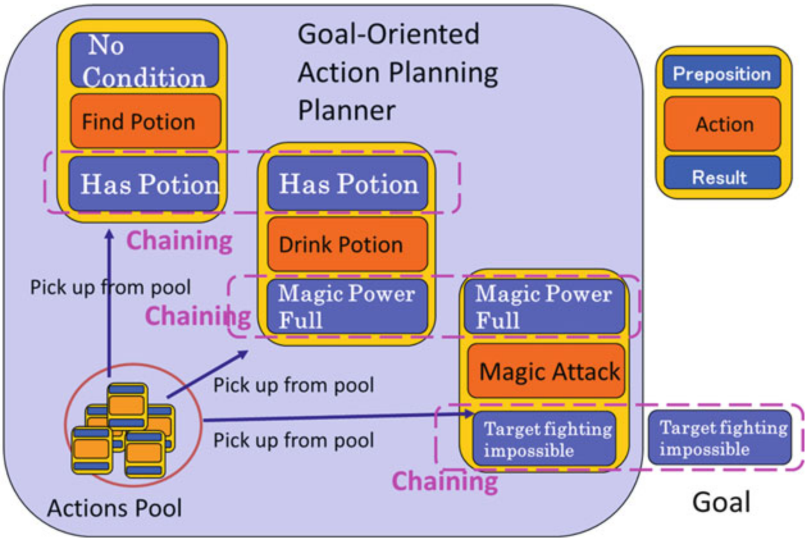


Fig. 10 Goal-Oriented Action Planning (*GOAP*)

GOAP is the method of chaining actions from a goal state to an initial state. In "F. E.A.R.," by using STRIPS-like algorithm (Russell and Norvig 2009), each action is represented by three parts such as precondition part, action part, and effect part. A precondition and effect parts are represented by using simple symbols. A lot of actions described by the form represented by these three are stored in an action pool. GOAP chains two actions by symbol matching so that one action has the same effect symbol with the precondition symbol the other action has. This operation is called "chaining." By repeating backward chaining from the goal to initial state, one action plan is composed (Fig. 10) (Orkin 2003, 2005, 2006).

On the other hand, Hierarchical Goal-Oriented Planning is, first of all, one big goal to set, and it is decomposed into smaller goals repeatedly until the end of decomposing where action commands are generated (Buckland 2004a).

This method was used for NPC's decision-making of "Chromehounds" (2006, From Software). By changing a way to dynamically decompose by NPC's ability, state, and goal, an action plan is generated and adjusted according to the battle situation (Miyake 2006, 2008b, 2010a; Miyake et al. 2008; Okamura 2011a). GOAP is also used in "Sakatsuku DS" (2008, SEGA) to make dramatic situations in the game (Ando 2010).

Utility-Based AI

Utility means "effectiveness value." For example, "utility of attack" is the damage value which a character could give to an enemy, or "utility of spell recovery" is the recovery value the spell gives. When there are four action choices such as "Attack," "Guard," "Spell," and "Recovery," by scoring each action utility with utility function, a character can select the best one with maximum score. This method is called decision-making by utility. By setting different utility functions to different characters, it can give a different personality to different characters (Namiki 2011).

Utility method is the oldest decision-making method for game characters, and it has been used from the 1980s to the present. It is used not only for action selection but also for any selection such as goal selection, weapon selection, magic spell selection, path selection, and so on (Buckland 2004a). For example, when an enemy has multiple weapons, utility method could automatically select the best one which has maximum utility (damage). The Sims (2000, Maxis) uses the utility method for the fundamental part of AI. A character has eight physiological parameters such as "want to sleep," "want to talk," and so on, and a utility curve is defined for each parameter. A utility curve is a graph which shows the relation between a parameter and its utility, and the shape is not a straight line such that when a parameter increases, at some stage utility value becomes saturated. This saturation phenomenon is called the "law of diminishing marginal utility." In The Sims, the characters' physiological parameters are under this law.

For example, when the utility value for the "Hunger" parameter changes from −80 to 10, and when it changes from 10 to 100, the change is calculated in this way.

$$\begin{aligned}
&\Delta(-80 \rightarrow 10)\text{Utility} \\
&= \text{W_Hunger}(10) * (10) - \text{W_Hunger}(-80) * (-80) \\
&\Delta(10 \rightarrow 100)\text{Utility} \\
&= \text{W_Hunger}(100) * (100) - \text{W_Hunger}(10) * (10)
\end{aligned}$$

W_Hunger is a weight value for the "Hunger" parameter and a function of the "Hunger" parameter. In The Sims, the utility curve is not prepared, but the utility weight curve is prepared such that the x-axis is "Hunger," and y-axis is W_Hunger (Fig. 11) (Miyake 2010a; Forbus 2002).

The value of the former is much bigger than the one of the latter. This result shows the law of diminishing marginal utility, and it means a character becomes more pleased by the change from a hungry state to a full state than the change from a satiety state to an even fuller state. Following the law, after the characters of The Sims are satisfied with one physiological need, it goes to satisfy another physiological need. For example, when a character is full, it might go to other place to talk with other characters. After it talks in much time, it goes to an arcade center to play games. After it becomes hungry while playing games, it might go to have dinner somewhere. In this way, The Sims character AI is implemented to select one action automatically which gives maximum utility from all actions (Miyake 2012a; Forbus 2002).

Task-Based AI

A task-based algorithm is the idea that a character action consists of several tasks, and any action can be decomposed into tasks. A task has its own definition of what a character should do. And tasks have the rules of ordering what tasks can follow other tasks.

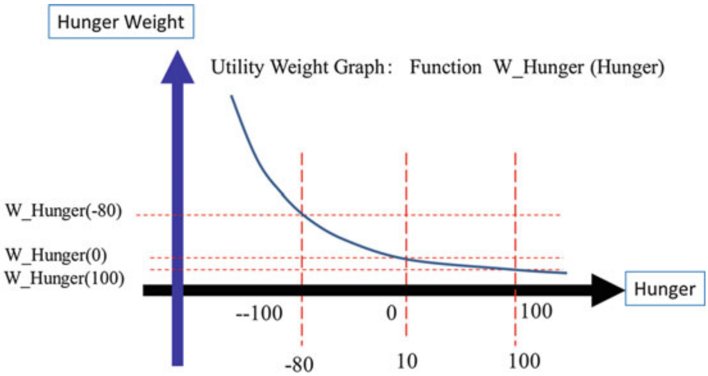


Fig. 11 Utility weight curve of "Hunger" parameter of "The Sims" (Forbus 2002)

HTN (hierarchical task network) (Nau et al. 2013) is applied to a decision-making system for Killzone 2. HTN has a long-term goal which is decomposed into tasks in a way called "method" repeatedly until all tasks are decomposed into primitive tasks which cannot be decomposed further. A group of tasks has an order of execution, so the tasks form a unidirectional networked structure under the order restriction (Miyake 2012a). In Killzone 2, the methods to decompose tasks are described by script language, so a unique task network is generated for the situations of a character and game (Straatman et al. 2009).

Simulation-Based AI

In a complicated terrain, to move a character with high degree of freedom of movement performance, making a motion plan is difficult by simple logic alone. In this case, a simulation-based algorithm is required to find a right motion plan, by repeatedly simulating a character orbit with a random combination of acceleration and timing. To simulate an orbit, precise terrain models are not needed, but simple collision models or navigation data are enough for simulating. In some cases, dynamic programming is used to find an ideal orbit in all combinations of possible orbits (Okamura 2011a, b).

Hierarchical Structure: Layered Structure

A character's decision-making system for a large-scale game usually has more than three layers. The structure "subsumption architecture" gives a higher-layer permission to interrupt the decision of the lower layers (Brooks 1986). A higher layer performs more abstract decision-making. Information in recognition modules becomes more abstract in higher layer, and an action decided in decision-making module becomes more concrete and precise in lower layers (Buckland 2004a). This

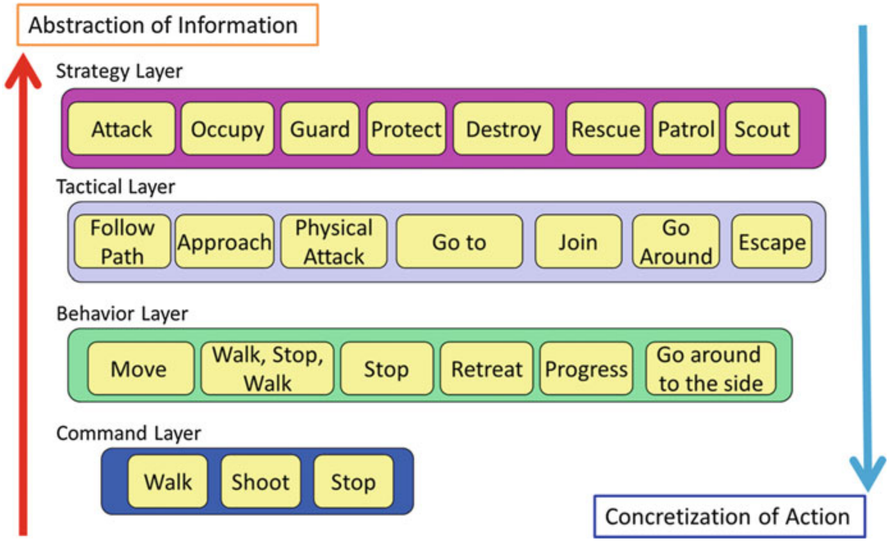


Fig. 12 Hierarchical structure and layered structure of NPC decision-making system in Chromehounds (Miyake 2006)

hierarchical structure is called a "multilayered structure." This structure is used for the whole AI system in the WCCF (2004, World Club Champion Football, SEGA) series (Tanabe 2009).

In Chromehounds, for example, when an "Attack Enemy" goal in the highest layer is decided, the goal is repeatedly decomposed into smaller and more detailed goals in the lower layer (Fig. 12) (Miyake 2006; Miyake et al. 2008). Specific modules modify the goal with more details such as direction, position, and path to the target while resolving the problem. Finally a list of simple action commands is generated, and animation data will be played following the action command sequence. When a factor interrupts the process of playing a list, a lower layer modifies its action. The goal in the highest layer can be changed by a very strong interrupting factor (Miyake 2010a).

For example, when a character is attacked by an enemy on the way to approach to an attack point, the character avoids the enemy's attack reactively. This reactive decision-making should be done in the lowest layer. When an enemy changes its position, the goal in the middle layer should update the position to attack. But when the character loses against the enemy, or the enemy was beaten by other agents, the character should change the highest goal in the highest layer.

Connection Between Decision-making Layer and Animation Layer

As described in the previous chapter, decision-making algorithms are a discrete symbolic system. On the other hand, animation system is a numerically continuous system. For example, when decision-making system can cast two symbolic commands such as “Throw” and “Run” to animation system, the animation system needs to execute them within the environmental restrictions, blend them smoothly, and deform them continuously for external force. This phenomenon is caused by the connection of discrete decision-making layer and continuous animation layer.

If there are only two layers such as highest decision-making and body, a highest decision-making layer must know a lot of restrictions of environment and body, and it becomes difficult for the highest layer to think abstract things. So a few middle layers are set between the highest decision-making layer and the body layer to connect them as indirectly as possible. In Hitman: Absolution (2012, IO Interactive), there are six layers between a highest decision-making layer and a body layer and further post process to adapt a motion smoothly and immediately for external forces (Anguelov 2013).

There are two reasons why multilayered architecture should be prepared for intelligence. All layers are connected each other via commands and messages. One reason is that each layer’s independence and modularity make it possible to develop and expand each layer independently. The other reason is that a middle layer must be interrupted to make a character play a specific role. The multilayered structure is also used to control each member in a team by Meta AI or team AI (Miyake 2010b). Even now, optimal multilayered model between a decision-making layer and a body layer is an unsolved problem and has been studied, and many multilayered models for each game title have been suggested.

Cooperation Among Characters

Cooperation among characters is a multi-agent problem in general, but there are three approaches in digital games. One way is to make it possible for characters to communicate with each other by sending messages based on predefined protocol. The second way is that team AI gives commands to all characters and controls them as a commander. The last way is a special cooperation system described later.

Messaging

This method is based on messaging using predefined protocol among NPCs to make a character group unified. But it is difficult for this method to manage all characters collectively and adjust them to the current game status that changes rapidly. So the message content should be just simple and temporal. For example, a message shares some kinds of information such as targeting enemies, requirements to rescue and guard, commands to move, and so on. In the previous chapter, seven decision-making algorithms were explained. When all characters are state-based AI, they can communicate with each other by assigning other’s state directly. Also when they are goal-based AI, they can communicate with each other by assigning the other’s goal (Miyake et al. 2008).

Team AI

Team AI is a facilitator for cooperation of characters. Sometimes, one of the characters becomes a commander, and sometimes team AI, that has no body, is set as a commander (Tanabe 2009; van der Sterren 2013). Team AI has agent architecture, but sensors and effectors are very specific. A sensor of team AI is to gather messages from each team members. Team AI recognizes the current team situation by gathering information from team members. The effector of team AI gives an order to each character. Team AI communicates with team members via messages and commands to control them. Team AI can reduce the complexities of messaging between all team members.

In Killzone 3 (2011, Guerrilla Games), team AI has three layered structures such as a commander, squad leaders, and individual members. The squad leaders and members report their current local status in their environment and battle situation around themselves to their commander. Meanwhile, the commander gathers a lot of local information, reconstructs the global situation of the battle field, and orders each squad leader by issuing a command. When the squad leaders receive the order, they order their members to accomplish the commander's order.

Waypoints are prepared for character pathfinding, and they are also used as an influence map which increases the value of the points where enemies are approaching.

Each member watches the influence map and makes moves by taking their enemies' degree of threat into consideration. A tactical global map consists of regions made by clustering hundreds of waypoints and is also used as an influence map of global enemies' distribution. An AI squad executes pathfinding by analyzing the data of the global influence map (Fig. 13) (Straatman et al. 2013; Champandard 2008).

Specific Cooperation System

In digital games, AI technique is needed not only to control the team but to achieve partial cooperation. The "targeting" technique is to resolve "who will attack which enemy" when there are many enemy characters. It's not efficient for NPCs to randomly attack their nearest enemies. Such strategy looks like very foolish to players. So a targeting system should distribute NPCs' attacking power intelligently. To implement a targeting system, a blackboard system can be used very effective in making it possible for characters to cooperate. First of all, an "arbiter" of blackboard observes how strong enemies are and prepares slots for each enemy on the blackboard. The number of slots for each enemy is proportional to its strength. For example, three slots for the strongest enemy and one slot for the weakest one. A slot means a token to attack the enemy. Each NPC applies for one slot, an arbiter accepts the application, and the slot is occupied by the NPC. When all slots of the enemies are occupied by NPCs, further applications are refused. So other NPCs must apply for other empty slots. In this way, the targeting system distributes NPCs to enemies.

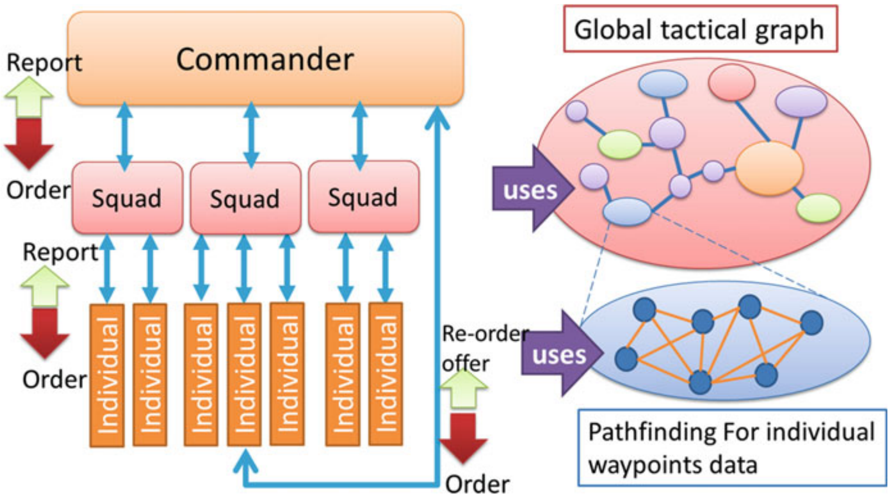


Fig. 13 Organization chart of NPCs in Killzone 2 and navigation data (Straatman et al. 2013; Champandard 2008)

When a lot of enemy characters surround a player, they should avoid synchronous attacks to the player. So one token to attack is prepared to make only one enemy to attack at the same time (Dawe 2013).

How to Make Navigation AI

Decision-making methods are used in both game industry and in robotics, but Navigation AI is a system unique for game development. Navigation AI was originally a character navigation system from one point to another. But as modern game level design became more complicated, Navigation AI has extended its role to analyze static and dynamic level, extract terrain information, and provide useful information to character AI and Meta AI. This means that in one sense, Navigation AI lives in the environment. Navigation AI can extract dynamic global terrain information by real-time terrain analysis and situation analysis.

In the real world, it is impossible to embed information into the environment, but it is possible to do it in digital world. It is especially possible to embed information to an object for a character AI to understand it. It means that Character AI in digital game gets enough information without high-performance sensors from the terrain and objects, in which such information is embedded, which human and animal can get only via five sensors and perception. Navigation AI has the heavy task to analysis environment, and Character AI and Meta AI get the information which Navigation AI provides.

Basic information to recognize environment is spatial-representation data and object data which are embedded as additional information necessary to make environmental recognition. Information for static environment and terrain is prepared in game development.

A role of Navigation AI is to provide environmental information to Character AI just like animals get environmental information from the world. Of course, the type of necessary environmental information depends on the type of character. It means that the first role of Navigation AI is to provide information for recognition and to make the AI's subjective world. There are two kinds of information. One is information of objects via sensors, and the other is information of objects required to take an action in the world. The former is called the sensory world, and the latter is called action world (von Uexküll 1956). The concept explained above belongs to ecology and cognitive science. But in artificial intelligence, it is called knowledge representation of objects. There are many different types of knowledge representation of objects and terrains.

World Representation

The representation of whole terrains and space in game level design is called "world representation (WR)" (Straatman et al. 2005). World representation consists of two kinds of data. One is navigation data, and the other is data which relates to the location on navigation data. Navigation data has two styles: waypoint and navigation mesh (Buckland 2004b, c).

In the early 1990s, there was no navigation data in most digital games. To navigate a character, fixed points where a character moves on or fixed regions like a room and a limited field where a character walks around were used. In many cases, a character is bound in narrow region by script control. For example, the control like "when a character is 5 m far from a center point of a room, turn back" is scripted.

But, in the middle of the 1990s, 3D games appeared. Navigation data and A* algorithms to find a path dynamically between any two points on a level had been introduced rapidly. Now it has become the standard method (Buckland 2004b; Snook 2000). Waypoints are connected points distributed over the terrain. In the latter half of the 1990s, waypoint systems are widely to ease data making. But in the early 2000s, navigation mesh systems had become widely used (Miyake 2007b). Pathfinding technology makes it possible to move characters between any two points. In Armored Core V (2012, From Software), pathfinding in 3D space is realized (Okamura 2011b).

Navigation meshes are networked mesh data which cover the whole terrain with convex polygons connected to each other with the neighboring mesh. Usually navigation meshes are rougher than collision models of the terrain and objects. In general, the problem with navigation mesh system is generating navigation data automatically. There are many middleware to generate navigation meshes automatically from collision models.

In Final Fantasy XIV (MMORPG), navigation meshes are generated automatically from a large map of collision models (Fig. 14), and the mesh data are stored on the online server. To reduce the load of the pathfinding execution process, instead of A* algorithm lookup tables with overhead regions are used for pathfinding (Fig. 15) (Gravot et al. 2012a, 2013; Yoneda 2012).

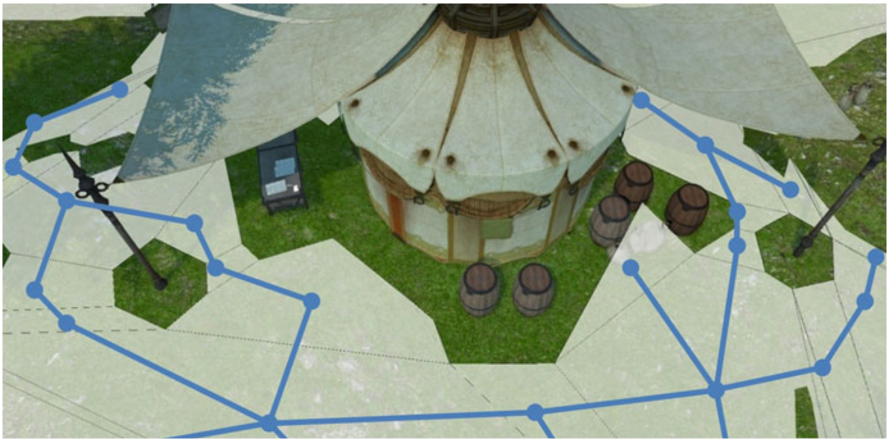


Fig. 14 Navigation meshes and linked graph structure of Final Fantasy XIV (Final Fantasy XIV © 2010–2016 Square Enix Co., Ltd. All Rights Reserved)

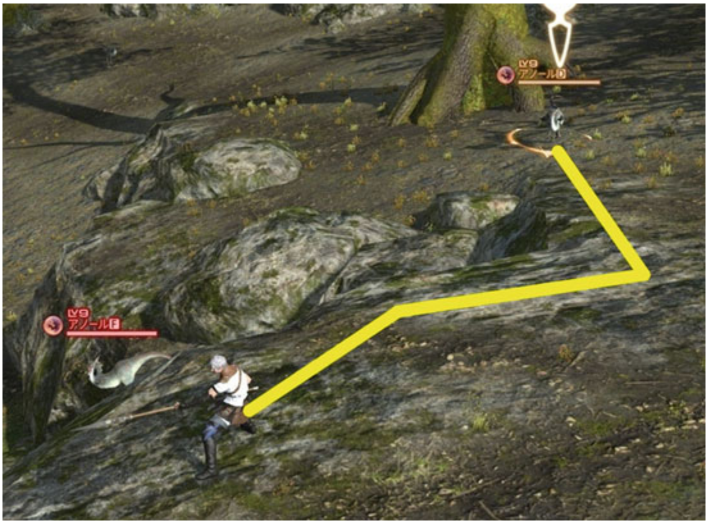


Fig. 15 Pathfinding on navigation mesh in Final Fantasy XIV (Final Fantasy XIV © 2010–2016 Square Enix Co., Ltd. All Rights Reserved)

Navigation mesh is a polygon plane which contains terrain surface information. Pathfinding by using attribute information of a terrain surface embedded in navigation such as snow, water, glass, concrete, and soil can generate a different path for a different character corresponding to the movement performance (Miyake et al. 2008; Buckland 2004b). For example, if the cost of water is set high for a character who cannot swim, the character will not go through water region. If the cost of snow is set low, for a character who can go on snow with high speed, it goes through snow region. Furthermore, by setting one-way direction as mesh attribute, a character's behavior such as dropping or going around is realized (Gravot et al. 2012b, 2013; Yoneda 2012). As a character knows additional information of environment including cover points, or good place to attack enemies from a distance, a character can understand deeply and take actions by using features of the environment (Miyake 2006; Straatman et al. 2005). In Splinter Cell: Conviction, navigation mesh models are dynamically modified by changes of game level such as explosion or destroying objects. It makes possible for a character to go through a mesh where the character could not go through before. Or on the contrary it prevents a character from going through a mesh where a character could go through (Walsh 2010).

Object Representation

Every object in an environment of a game world must have a knowledge representation which a character uses to recognize objects. Usually, knowledge representation means knowledge format of inner intelligence. But in digital games, knowledge representation means information each object has. There are many types of information for each object corresponding to different actions a character carries out to the object. Typical types of knowledge representation are explained below.

Symbol Representation

Symbols are used to categorize objects into different types such as natural objects, artifacts, characters, enemies, buddies, tribes, and so on. In Halo 2, all enemies are categorized and organized by a family tree (Isla 2005a).

Affordance Representation

Affordance is information attached to an object and represents an action a character can take with the object. For example, an action for a rock is "to move to a specific direction," and an action for a lever is "to pull." Furthermore, affordance of a rock is which direction it can be moved to, and affordance of a car is which position a character can ride the car from. By using more precise information of an afforded action, a character takes an action more exactly. In Halo 2, cars and objects have affordance information (Isla 2005b). In Monster Hunter Diary: Poka Poka Airu Village (2010, Publisher: Capcom, Developer: FromSoftware), all objects on the map have an affordance value for each potential action (Namiki 2011).

Animation Auxiliary Data

To realize natural animation, various animation auxiliary data is necessary. For example, to realize a character's motion to sit down on a chair naturally, information of the direction and height of chair are required to support modifying "sitting down" animation to the chair. But there are many chairs in the game world, and it is impossible to make a proper animation for each chair. So information to adjust an original animation to each chair is necessary for the object. For example, when a character cut down a tree, the tree must have the information for the cutting position.

When some of characters form a line in front of a shop, the shop must have information on where the line should be formed and the place to get food in the shop. In BioShock Infinite (2013, Irrational Games), there are many auxiliary data for each object such as a place to sit, a lookup point, and a place to stop. Furthermore, an object controls a character to take such auxiliary data close to the object. An object to control a character is called a smart object or smart terrain (Abercrombie 2014).

Enemy Representation

Enemy is an object to which a character should pay the most attention. Objective information such as position, type, stamina, magic power, velocity, and species are important, but the subjective image on how an enemy looks from a character's point of view is also important for decision-making. Subjective information is calculated from objective information. For example, threat is one of the subjective information that means how dangerous an enemy is for a character, and it's calculated by an estimate function from enemies' position, ability, and velocity (Isla and Gorniak 2009). By using threat values of enemies, a character selects a target enemy. Furthermore, an enemy's intention such as "Attack," "Retreat," and "Wait" can be evaluated from the records of enemy's position, velocity, and acceleration. A character makes a decision to select an action by predicting an enemy's action. In general, recognition of an enemy is called "awareness" (Rabin 2013).

Event Representation

Event representation is the information of what happens in game. The information format of representation falls under the 5W1H style "Who, When, Where, What, How, Why," a method to get information such as seeing or hearing and a confidence value of the information.

In Gunslinger (2002, not released, Surreal), a Western game, an event manager is watching all events in the village, and one NPC who sees an event shares the information with other NPCs. And the event propagates via conversation between NPCs over the village. It is called "dynamic reputation system" (Miyake 2010b; Alt and King 2002).

Navigation AI always observes the world and prepares the world representation data for the static structure and dynamic change of the world. Character AI and Meta AI recognize the environment and make a decision by using the data Navigation AI provides.

How to Make Meta AI

Meta AI is one of the AIs that dwell in game system. Meta AI observes the status of the game in real time and modulates and changes the game dynamically. In the 1980s, Meta AI's role was to adjust the difficulty level of the game automatically to the user's playing skill as implemented in Xevious (1983, Namco)

(Iwatani 2005a, b). Conversely, the modern Meta AI has an active role in generating game contents dynamically (Miyake 2010a, b, c).

In the 1980s, it was difficult to distinguish the role of Meta AI from Character AI. The AI system watches over the game world and controls not only characters but also all game gimmicks. But through the 1990s, the AI system differentiates into three independent AIs. The first AI is Navigation AI to recognize the environment, the second AI is Character AI to make it possible for a character to think autonomously, and the last AI is Meta AI. Meta AI is free from controlling characters and focuses on controlling the whole game. After Meta AI controls the distribution of NPCs, each NPC moves autonomously in the level (Miyake 2011, 2012a).

Meta AI recognizes the situation in the same way as Character AI by using the information that Navigation AI provides. Meta AI recognizes the game status by getting information from Navigation AI. What Meta AI needs to recognize in the game world is the global situation of the game and user's game-play records. Meta AI analyzes a user's game records to extract user skill and game progress and analyzes the global game situation to extract the battle situation and terrain around the user. User play records are gathered from game data and the analysis results. The global information of the game including enemies, buddies, and gimmicks is gathered via navigation mesh information. Navigation mesh has the role to detect what is happening on navigation mesh in complicated terrain and game situation. Navigation AI supports Meta AI.

Cases of Meta AI

Left 4 Dead is an online game in which a four-player team fights against enemies' AI group. In the game, Meta AI predicts a user's path by using navigation mesh and user play records (past paths, battle log, number of enemies the user defeated), calculates a user's tension from controller input and battle histories, and spawns enemies dynamically at the location unseen from a user on a predicted path. Usually, positions and number of enemies are defined as fixed data, but in this game, Meta AI generates the game situation dynamically corresponding to the user's different skill and behavior patterns (Booth 2009a, b).

Meta AI in Left 4 Dead was developed under the policy that the fun in a game comes from an iteration of tension and relaxation. The goal of Meta AI is to generate iterations of tension and relaxation. The equation to calculate user's tension consists of user's controller input and battle logs. There are four phases of Meta AI process. The first phase is the interval when Meta AI spawns monsters around the predicted user path and lets them fight with users. After the monsters increase the user's tension (buildup phase), the tension is kept constant (sustain phase). After Meta AI stops spawning monsters, they diminish and the intensity of the tension curve decreases (fade phase). When the curve becomes almost 0, relaxation time begins so that nothing happens (relax phase). After this phase has run for a while, Meta AI begins to reproduce monsters (buildup phase). In this four-phase cycle, iterations of tension and relaxation can be generated repeatedly.

Meta AI in Warframe (2013, Digital Extremes), which has additional functions from the Meta AI in Left 4 Dead, has the ability to generate a dungeon. It analyzes the dungeon topology, extracts a structure to distribute enemies automatically in it, and observes a user's movement on a heat map (Brewer 2013, 2014).

Agent Architecture for Meta AI

The inner structure of Meta AI is the same agent architecture as character AI or team AI has. The sensors of Meta AI are able to access all game information. On the other hand, the effectors are able to control all entities except a player and regions in the game. As Meta AI in Left 4 Dead, Meta AI estimates a user's tension, which becomes the basis for decision-making of Meta AI.

The Future of Meta AI

Compared to other game technologies, one specific feature of game AI technology is the fact that game AI is deeply connected with game design and level design. Furthermore, Meta AI works to generate game design and level design.

One distinguished feature of traditional game system is to use mostly fixed data which game designers prepare in game development process. But Meta AI changes the game system such that Meta AI generates and modifies the game and level design dynamically. Therefore, Meta AI is AI embedding game designer's intelligence. The implementation cost of Meta AI is low because the Meta AI development is later in the game development and it can use all the structure of Character AI and Navigation AI. By replacing Meta AI with a new one, a new game system can be generated with the same components. For example, a new different game with the same components can be distributed only by downloading additional lightweight Meta AI.

In general, the process to transform fixed data contents to dynamically changing contents is called "procedural." In the last 20 years, as games progressed, much fixed data are transformed to procedural contents such as intelligence logic, character path, terrain, scenario, and so on. Meta AI uses game procedural contents and also generates game procedural content. This is the natural evolution of digital game (Fig. 16).

Integration of Three AIs

A purpose of digital games is to give users unique and proper experience depending on game title. First, local and short-time interactions around a player character are formed by Character AI. Then, Navigation AI makes it possible for all characters to walk on all stages freely and to intelligently use objects in the level. Characters can

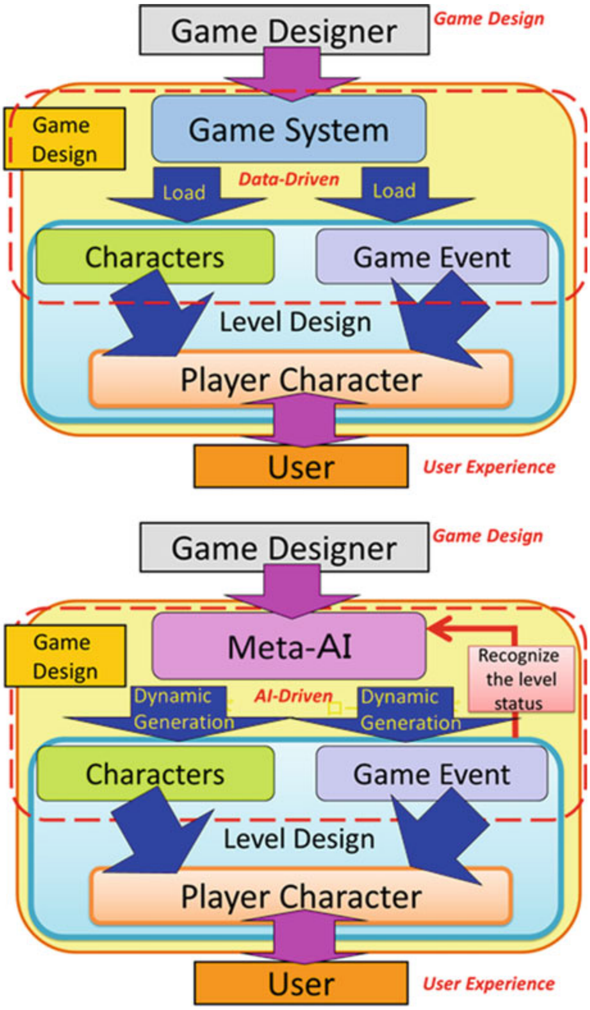


Fig. 16 The structure of conventional game system (*up*) and new game system with Meta AI (*below*)

move along a long path, go around, and use objects using Navigation AI. Navigation AI expands the space scale of a user's experience. Meta AI observes the status of users, characters, and level design, and modifies the game, and changes the game's progress dynamically. It also expands the time scale of user experience.

In this way, while AI technology in digital game handles characters and games, it can, to some degree, freely create a multi-scale user experience in time and space (Fig. 17).

In the 1980s and the 1990s, there was one immature AI system. Later it was divided into three distributed AIs: Character AI, Navigation AI, and Meta AI. In the present era, these three AIs cooperate and this integrated AI system can control user experience artificially to some degree.

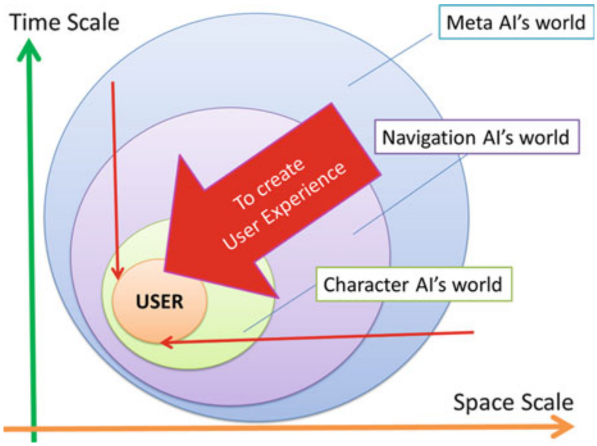


Fig. 17 Three AIs and user experience

Learning, Evolution, and Auto-generation Method

For about 40 years since the down of digital games, basically, digital games have worked by using fixed and prepared data. But such a static game system changed in the beginning of the 2000s. It is called a "procedural approach" to replace fixed data with dynamic content generation. It has changed also game system. In procedural approach techniques such as learning, evolution, and generation methods happen as the big challenge to replace fixed data with dynamic calculation in the core of digital games. Also it was a difficult challenge to introduce and harmonize the approach with the other parts of game. In this chapter, a few recent cases to introduce these procedural techniques are described.

Character Learning and Evolution

There are three techniques for character learning and evolution: data-driven approach, pseudo-learning, and learning algorithms. They are explained here one by one. The first two approaches are not ordinarily called "learning," but they are categorized in "learning" for a character in digital game.

Data-Driven Approach

As the game goes on, character data are replaced repeatedly by other data. This is the data-driven approach. By introducing this approach, a character looks like they are learning and evolving step by step. This method is the most popular approach for digital games because game designers can make and control the data as much as they want to customize.

Pseudo-learning

Pseudo-learning is a method by which all abilities are implemented in a character in game development and are released one by one on each step in game. For example, all special actions are set in a NPC, but they are hidden at first. When the NPC fights with a player character, a skill, which a player used, becomes possible to use for the NPC. From a user's view, the NPC looks like it's learning the user's skill.

Learning and Evolutionary Algorithms

Recently a number of applications of learning and evolutionary algorithm to digital games increase but still the number is not so big. So far the total number of games that use learning and evolution is about 200 titles, and even at present less than 20 games that adopt learning system are released per year. There are mainly four reasons why only a small number of games adopt learning algorithms, such as short of processing power, anxiety of unstableness, difficulty to harmonize the algorithm with game design, and difficulty to construct learning environment in game.

In the last 14 years since 2000, rendering required most of the processing power, and it was difficult to assign huge processing power to AI learning, but recently the problem has become resolved in some sense because of growing processing power. The problem of unstableness of learning and evolutionary algorithms is not avoidable because learning itself is unstable and not deterministic in its principle. It has always the possibility of success and failure. If it is required that learning should always be successful, data-driven approach is better.

The number of games that use learning algorithms reveals the role of learning in game playing. An affinity of learning algorithm with game design depends on the situation of the game world. Learning algorithm requires that in a game play simulated situation should be repeated. But usually a character has the specific role in each game stage, and the same situation does not repeat again. Also constructing environment in a game that is adequate to learning has become the most difficult problem. Learning and evolutionary algorithms work within a simulation framework with set parameters. In the development of digital game, game itself keeps changing. Because of this it is difficult to construct the learning framework. Even though it is possible to try again and again to select a parameter set, consider how to normalize parameters, and determine which algorithm should be used, and how to tune it, once the game design changes, all trials are set back. It may be possible to introduce them after game development, but there is no room to apply learning system in the completed game. Therefore, introducing learning algorithm should be considered in game design before game development begins. Also it is the central problem that the parameter range is not large enough for learning and evolutionary algorithm. If the parameter range is not large enough, it happens that the algorithm conversely limits game design. As AI development makes a progress with game development progress, the problem is so serious. But some games can clear all of these problems, and some games use learning algorithm only in game development process. The examples of games which use learning algorithms are explained in the next chapter.

Game Titles with Learning and Evolutionary Algorithm in the 1990s

There is small number of games that adopt learning and evolutionary algorithms in the 1990s, and most of these games are AI-centric game. AI-centric game means that the AI is the core of the game design because most of processing power is used for AI processing. In Creatures (1986, Millennium Interactive), a user breeds a monster and teaches it how to use many objects in the level such as a ball, a box, and so on. A monster has dozens of lobe in the brain which is a perceptron-type neural network with hundreds of nodes. By using these neural networks, a character learns a teacher signal from a player and uses different objects in the progress (Grand and Cliff 1997). Creatures Online is also developed by the same approach (Hamaide 2014).

In Astronoka (1998, muumuu), monsters evolve by genetic algorithms. A player disrupts a monster in the field setting many traps, but the monster learns the traps and becomes to be able to avoid them by genetic algorithm. A character has a set of gene parameters and takes an action to a trap by following the parameter values. An evaluation value for a gene depends on how near the character reaches to the goal. And the crossover and mutation for two different genes in higher rank can make new two genes. Furthermore, characters learn the environment by the genetic system and can go near the goal. Behind the game screen, 20 characters are simulated to try the same trap field seen on the screen. It enables a rapid evolution such that a player feels the constant speed of evolution in game playing (Morikawa 1999, 2008).

Game Titles with Learning and Evolutionary Algorithm in the 2000s

In Black & White (2001, Lionhead Studios), a character uses a simple perceptron neural network with a few nodes (Evans 2002). In Supreme Commander 2 (2010, Gas Powered Games), an enemy character has a perceptron with 35 nodes as an input layer which get input signals from the game world, 98 nodes as a middle layer, and 15 nodes as an output layer to select one target. But learning is used only for the game development, and the configuration of the neural net is fixed in the game (Robbins 2013).

In Forza Motorsport (2004, Microsoft Game Studios, Turn 10 Studios), Drivatar (Microsoft Research 2014) is an AI to learn user's driving techniques by machine learning. Its race course consists of blocks, and every block has an ideal course line. A car in Forza Motorsports has a real driving simulation model of a handle, an accelerator, and a brake connected to tires. Drivatar detects the difference between an ideal course line and a player's course line which a player's car runs on. The difference is reduced to the handle's motion, timing of acceleration, and brake strength. Drivatar learns player's driving features. Drivatar learns these driver's techniques in different courses and stocks many player's driving features. A player can fight with Drivatar the player has bred. Furthermore, any player can upload player's own Drivatar and download another's Drivatar. A player can also have a race with other player's Drivatar (Herbrich et al. 2008).

In Hitman: Absolution, a character learns animation transitions by reinforcement learning (Büttner 2012). In Hitman: Absolution, the animation system is a motion-graph system such that each motion has transition points to other motions. When a character changes velocity and posture, some of the transition routes are considered. But it is difficult to find the best route which makes the change looks natural. Then by calculating the cost from the change of velocity and posture, a character learns the route with the lowest cost by reinforcement learning.

In NERO (2005, Texas University, Neural Networks Research Group), which is the game developed through the research of Texas University, the character has "neuro-evolution" system for decision-making which is a combination of neural network and genetic algorithm (Stanley et al. 2005). The neural network topology is represented by gene information. By the crossover of two genes, one neural network with a new topology is generated (Buckland 2002). In NERO, a character who wins can evolve its neural network by genetic algorithm. This method is now wildly used for academic research (Neuro-evolving robotic operatives 2005).

Procedural Techniques

Basically in game development, asset data is prepared and used to construct a game world in runtime game system. On the other hand, procedural techniques generate content data by program in runtime (Miyake 2010b, c).

Procedural terrain generation is generating terrain by using fractal and random function. Procedural characters and asset distribution allocate characters dynamically at adequate positions by calculation. Procedural story generation is to make a branch story dynamically by calculation.

Procedural technique reduces asset data and development cost and makes data size smaller. But quality of procedural contents is not good as what human designers create. So, procedural content generation is used to make many asset data for background models and side stories. Pure procedural generation cannot avoid showing specific patterns of algorithms. On the other hand, semi-procedural generation is the combination of generation and original data. For example, semi-procedural animation is generated by retargeting an original animation (Hecker et al. 2008; Kurachi 2009a).

Procedural technique is often categorized into AI technology and has deep relationship with Meta AI. The origin of procedural technique is automatic dungeon generation in Rogue (1980) (UNIX 1980), followed by ELITE (1984, BBC), the level generator of Mystery Dungeon: Shiren the Wanderer (1995, CHUNSOFT), and many game titles (Braben 2011). In the Age of Empire (1997, Ensemble Studios) series, each terrain is auto-generated for every battle (Pottinger 2000). In Far Cry 2 (2008, Ubisoft), vegetation on the island is all procedurally generated (Guay 2008). In Spore (2008, Maxis), many procedural techniques are used to generate terrain, animation, texture, sound, and so on (Miyake 2008b; Kurachi 2009b). In recent years, procedural techniques such as terrain generation, vegetation generation, and distribution are built in game engines such as Dunia Engine (2011, Ubisoft) and Frostbite engine (2011, EA DICE) (Widmark 2012).

As a game development advances, modifying or adding game data becomes more difficult for the complex relations among a lot of data. To lessen the phenomena, procedural techniques are effective for large-scale games.

Procedural techniques keep it easy to modify and improve contents data even in the latter parts of game development by procedural generation system. Game development keeps its own flexibility until the end of development by using procedural approach, and the procedural system can also be reused also for the other titles.

Game Evolution Technology

Procedural techniques make it possible to generate game itself by generating levels, sounds, and distribution of enemies. Automatic generating content is called "procedural content generation" (PCG). Game development by procedurally generating game and testing the game automatically by AI is called automatic (automated) game design (Nelson and Mateas 2007; Nelson 2012; Hom and Marks 2007; Togelius and Schmidhuber 2008).

Evolutionary Game Design

The problem of automatic game design is that the quality is lower than human creators generate. Nevertheless, there is a technique to make the quality higher. The method is called "evolutionary game design" to evolve a game automatically by evolutionary algorithm by AI's playing test repeatedly (Nelson 2012).

Usually, genetic algorithms or genetic programming is used as an evolutionary algorithm. In Age of Empires, after the terrain is procedurally generated, a connectivity check between generated terrains is executed. It ensures that any character can travel from a start point to an end point by pathfinding. It is a minimum assurance for generated games (Pottinger 2000).

Furthermore, AI plays and evaluates the procedurally generated game and evolves the game by evolutionary algorithm which uses the estimation values. The idea is called "evolutionary game design." For example, level design in a side-scrolled action game is procedurally generated, and AI plays the game repeatedly and researches whether or not the game is possible to clear, the score, the clear time, number of jumps, speed changes, and the number of defeated enemies. By using this information, the AI calculates the total score of the game. The AI changes the procedural algorithm and parameters and evolves the game to a direction to increase the estimation value (Shaker et al. 2012, 2013; Dahlskog and Togelius 2014).

Automatic Evolutionary Board Game

"Representation of game" is the most important element to implement evolutionary game design. If the representation is completed, game design can be changed by modifying each component. For example, evolutionary game design for board games is categorized into two representations. One is board representation and the other is rule representation. Each representation can be evolved in the crossover by generic algorithm or genetic programming, while an AI plays the game and returns the feedback evaluation value. This method makes board game design evolve automatically. As a result, new game rules on new boards are generated (Browne and Maire 2010; Pell 1992).

Summary

In digital games, users spend much time interacting with characters. As it has been explained, AI can generate user experiences more and more precisely. Evolution of AI makes for a new user experience. The progress of AI technology means the evolution of digital game design.

AI technologies in digital games must be explained with game design. In this chapter not only AI technologies but also their relationships with game design are explained. Also, AI technologies are a part of the game development process. There is no completed game like Shogi or Go at the start of game AI development. While game development advances, AI in digital games meets the requirements of games, makes them better, and becomes a core of the game. Sometimes AI technology changes game designs, and sometimes game design evolves AI technology.

It is also the important problem in game industry how to build AI development into the game development pipeline. For example, navigation mesh is such a big amount of data, and to make it, the development flow must be equipped with a tool pipeline, data flows, and building system.

In game industry in the last 20 years, AI makes a big progress and forms foundations. This evolution is almost independent from the academic progress of AI. But AI in games has built up a base to introduce academic research, or AI in the game industry now can provide academic research with new problems. New cooperation with the game industry and academic research makes them advance each other with increasing speed.

Information Resource of AI of Game Industry

Usually techniques and cases of AI in the game industry are not written in papers, but presented in game industry conferences such as GDC in the United States, CEDEC in Japan, and Game AI Conference in Europe, and the lecture materials are public now on public library servers such as the game developer's publication websites and game AI websites (AiGameDev). Furthermore, some books of compiled articles written by game developers have been published. But it is not easy to gather all information, and some of the information resources are listed up below.

GDC (Game Developers Conference) Vault
http://www.gdcvault.com/

CEDEC Library
http://cedil.cesa.or.jp/
Game AI Conference Recordings
http://gameaiconf.com/recordings/
AAAI AIIDE (Artificial Intelligence for Interactive Digital Entertainment)
http://www.aiide.org
AiGameDev
http://aigamedev.com/
Valve Software
http://www.valvesoftware.com/company/publications.html
Guerrilla Games
http://www.guerrilla-games.com/publications.html
Naimad Games
http://naimadgames.com/publications.html
Bungie
http://halo.bungie.net/inside/publications.aspx
Square Enix
http://connect.jp.square-enix.com
AI Game Programming Wisdom 1–4
Edited by Steve Rabin, 1-2, 4 Charles River Media
3 Cengage Learning (2002–2008)
Game AI Pro
Edited by Steve Rabin, A K Peters/CRC Press (2013)
Game Programming Gems 1–8
Born Digital (2001–2011)

And there is the summary of game AI themes with references (Miyake 2012b).

Recommended Reading

J. Abercrombie, Bringing bioShock infinite's Elizabeth to life: an AI development postmortem, in *GDC 2014* (2014), http://www.gdcvault.com/play/1015387/Animation-Driven-Locomotion-for-Smoother

G. Alt, K. King, Chapter 8.6, A dynamic reputation system based on event knowledge, in *AI Game Programming Wisdom*, vol. 1, Charles River Media, Boston (2002), pp. 426–435

T. Ando, Soccer game AI system for "Sakatsuku DS", in *CEDEC 2010* (2010), http://cedil.cesa.or.jp/session/detail/379

B. Anguelov, Managing the movement, in *GDC 2013* (2013), http://www.gdcvault.com/play/1018145/Managing-the-Movement-Getting-Your

B. Anguelov, G. Leblanc, S. Harris, Animation-driven locomotion for smoother navigation, in *GDC 2012* (2012), http://www.gdcvault.com/play/1015387/Animation-Driven-Locomotion-for-Smoother

N.A. Bernstein, *Dexterity and Its Development* (Psychology Press, New York, 1996)

M. Booth, Replayable cooperative game design: Left 4 Dead, in *GDC 2009* (2009a), http://www.valvesoftware.com/company/publications.html

M. Booth, The AI systems of Left 4 Dead, in *AIIDE 2009* (2009b), http://www.valvesoftware.com/company/publications.html
D. Braben, Classic game postmortem – ELITE, in *GDC 2011* (2011), http://www.gdcvault.com/play/1014807/Classic-Game-Postmortem
D. Brewer, AI postmortems: Assassin's Creed III, XCOM: Enemy Unknown, and Warframe, in *GDC 2013* (2013), http://www.gdcvault.com/play/1018223/AI-Postmortems-Assassin-s-Creed
D. Brewer, The living AI in Warframe's procedural space ships, in *Game AI Conference* (2014), http://gameaiconf.com/recording/living-ai-of-warframe/
R. Brooks, A robust layered control system for a mobile robot. IEEE J. Robot. Autom. **2**(1), 14–23 (1986)
C. Browne, F. Maire, Evolutionary game design. IEEE Trans. Comput. Intell. AI Games **2**(1), 1–16 (2010)
M. Buckland, Evolving neural net topologies, in *AI Techniques for Game Programming* (Cengage Learning PTR, 2002)
M. Buckland, Chapter 9, Goal-driven agent behavior, in *Programming Game AI by Example* (Wordware/Jones & Bartlett Learning, Burlington, 2004a)
M. Buckland, Chapter 8, Practical path planning, in *Programming Game AI by Example* (Wordware/Jones & Bartlett Learning, Burlington, 2004b)
M. Buckland, Chapter 5, The secret life of graphs, in *Programming Game AI by Example* (Wordware/Jones & Bartlett Learning, Burlington, 2004c)
R. Burke, D. Isla, M. Downie, Y. Ivanov, B. Blumberg, Creature smarts: the art and architecture of a virtual brain, in *Proceedings of the Game Developers Conference*, San Jose (2001), pp. 147–166
M. Büttner, Reinforcement learning-based character locomotion in Hitman: absolution, in *Game AI Conference* (2012), http://gameaiconf.com/recording/hitman-reinforcement/
A. Canary, Free-range AI, in *GDC 2014* (2014), http://www.gdcvault.com/play/1020110/Free-Range-AI-Creating-Compelling
A. Champandard, On the AI strategy for KILLZONE 2's multiplayer bots, AiGameDev (2008), http://aigamedev.com/open/coverage/killzone2/
S. Dahlskog, J. Togelius, A multi-level level generator, in *Proceedings of the IEEE Conference on Computational Intelligence and Games (CIG)*, Dortmund (2014)
M. Dawe, Chapter 28, Beyond the Kung-Fu circle, A K Peters/CRC Press, Boca Raton, in *Game AI Pro* (2013), pp. 369–375
R. Evans, 11.2. Varieties of learning, AI Wisdom, vol. 1 (2002), pp. 567–578
R. Evans, Modeling individual personalities in The Sims 3, in *GDC 2010* (2010), http://www.gdcvault.com/play/1012450/Modeling-Individual-Personalities-in-The
K. Fauerby, Crowds in Hitman: absolution, in *GDC 2012* (2012), http://www.gdcvault.com/play/1015315/Crowds-in-Hitman
K. Forbus, Simulation and modeling: under the hood of The Sims, Northwestern University, Lecture note (2002), http://www.cs.northwestern.edu/%7Eforbus/c95-gd/lectures/The_Sims_Under_the_Hood_files/frame.htm
J.J. Gibson, *The Ecological Approach to Visual Perception* (Routledge, Florence, 1986)
S. Grand, D. Cliff, Creatures: entertainment software agents with artificial life. Auton. Agent. Multi-Agent Syst. **1**, 39–57 (1997)
F. Gravot, T. Yokoyama, Y. Miyake, Following a path(movie) (2012a), https://www.youtube.com/watch?v=9_vdpndn4jI
F. Gravot, T. Yokoyama, Y. Miyake, Dropping(movie) (2012b), https://www.youtube.com/watch?v=o3ihN2kSzMs
F. Gravot, T. Yokoyama, Y. Miyake, Chapter 20, Precomputed pathfinding for large and detailed worlds on MMO servers, A K Peters/CRC Press, Boca Raton, in *Game AI Pro* (2013), pp. 269–287
J. Gregory, *Game Engine Architecture* (A K Peters/CRC Press, Boca Raton, 2009a)
J. Gregory, State-based scripting in uncharted 2: among Thieves, in *GDC 2009* (2009b), http://www.gdcvault.com/play/1730/State-Based-Scripting-in-UNCHARTED

J. Griesemer, The illusion of intelligence: the integration of AI and level design in Halo, in *GDC 2002 Proceedings Archiv* (2002), http://halo.bungie.net/inside/publications.aspx
D. Guay, Procedural data generation in FAR CRY 2, in *GDC 2008*, San Francisco (2008)
J. Guay, Real-time sound propagation in video games, in *GDC 2012* (2012), http://gdcvault.com/play/1015492/Real-time-Sound-Propagation-in
J. Hamaide, Creatures 3 online, in *Game AI Conference* (2014), http://gameaiconf.com/recording/2014-creatures/
C. Hecker, Spore behavior tree docs (2009), http://chrishecker.com/My_liner_notes_for_spore/Spore_Behavior_Tree_Docs
C. Hecker et al., Real-time motion retargeting to highly varied user-created morphologies. Proc. ACM SIGGRAPH **27**(3), 27:1–27:12 (2008)
R. Herbrich, T. Graepel, J. Quiñonero Candela, Halo, Xbox Live the magic of research in microsoft products, Microsoft Research (2008), http://research.microsoft.com/en-us/projects/drivatar/ukstudentday.pptx
V. Hom, J. Marks, Automatic design of balanced board games, in *Proceedings of the 3rd Artificial Intelligence and Interactive Digital Entertainment Conference*, Stanford (2007), pp. 25–30
R. Ishida, K. Kuwabara, Y. Katagiri, Chapter 4, Blackboard model, in *Distributed Artificial Intelligence* (Corona Publishing, Tokyo, 1996)
D. Isla, Managing complexity in the Halo2 AI, in *Game Developer's Conference Proceedings 2005* (2005a), http://www.gamasutra.com/view/feature/130663/gdc_2005_proceeding_handling_.php
D. Isla, Dude, where's my warthog? From pathfinding to general spatial competence, in *AIIDE, 2005* (2005b), http://naimadgames.com/publications.html
D. Isla, Halo 3 objective trees: a declarative approach to multi-agent coordination, in *Artificial Intelligence and Interactive Digital Entertainment (AIIDE)* (2008), http://naimadgames.com/publications.html
D. Isla, B. Blumberg, Chapter 7.1, Blackboard architectures, in *AI Game Programming Wisdom*, vol. 1, Charles River Media, Boston (2002), pp. 333–344
D. Isla, P. Gorniak, Beyond behavior: an introduction to knowledge representation, in *AI Summit, GDC 2009* (2009), http://naimadgames.com/publications.html
D. Isla, R. Burke, M. Downie, B. Blumberg, A layered brain architecture for synthetic creatures, in *Proceedings of IJCAI*, Seattle (2001)
T. Iwatani, *Introduction to Academic Research of PAC-MAN* (Enterbrain, Tokyo, 2005a), p. 101
T. Iwatani, Self-game-control-system, in *International Game Designers Panel* (2005b), http://game.watch.impress.co.jp/docs/20050312/gdc_int.htm
N. Kurachi, *Real-Time Motion Retargeting Pipeline in Spore*, vol. 125 (CGWORLD, Tokyo, 2009a), pp. 80–85
N. Kurachi, *A Perspective of Procedural Technologie*, vol. 125(CGWORLD, Tokyo, 2009b), pp. 44–47
M. Leeuw, The PlayStation®3's SPUs in the real world: a KILLZONE 2 case study, in *GDC 2009* (2009), http://www.guerrilla-games.com/publications.html
Microsoft Research: Drivatar™ in Forza Motorsport (2014), http://research.microsoft.com/en-us/projects/drivatar/forza.aspx
MIT Synthetic Character Group (2003), http://characters.media.mit.edu/
Y. Miyake, A perspective of game AI in"Chrome Hounds", in *CEDEC 2006* (2006), http://cedil.cesa.or.jp/session/detail/50
Y. Miyake, Chapter 5, Managing and controlling contents technology, in *Research Report of Advanced Technology Application for Digital Contents Creation* (Digital Content Association of Japan, Tokyo, 2007a)
Y. Miyake, Chapter 3, Game AI technologies, in *Research Report of Advanced Technology Application for Digital Contents Creation* (Digital Content Association of Japan, 2007b), http://www.dcaj.or.jp/project/report/pdf/2007/dc08_07.pdf

Y. Miyake, How to use AI technologies to develop digital games. J. Jpn. Soc. Artif. Intell. **23**(1), 44–51 (2008a)
Y. Miyake, Game AI technologies and procedural techniques in Spore, in *Digital Game Research Association Japan, 14th Monthly Seminar*, Chiba (2008b), http://digrajapan.org/?wpdmact=process&did=Ni5ob3RsaW5r
Y. Miyake, Chapter 23, Digital game AI, in *Digital Game Textbook* (Softbank Publishing, Tokyo, 2010a), pp. 214–218
Y. Miyake, Applied AI and procedural technologies for online game. J. Jpn. Soc. Fuzzy Theory Intell. Inform. **22**(6), 745–756 (2010b)
Y. Miyake, Chapter 22, Procedural technology, in *Digital Game Textbook* (Softbank Publishing, Tokyo, 2010c), pp. 395–420
Y. Miyake, Chapter 5, The progress and important things of digital game AI, in *Core Technologies in Japanese Video Games* (Softbank Creative, Tokyo, 2011), pp. 127–214
Y. Miyake, *Introduction to Game AI*, vol. 68 (WEB + DB Press/Gijutsu-Hyohron, Tokyo, 2012a), pp. 87–120
Y. Miyake, Recent research topics summaries for next-generation digital game AI, in *Game Programming Workshop 2012 Proceedings*, Hakone 2012(6) (2012b), pp. 108–113
Y. Miyake, The fundamental theory of digital game AI. J. Virtual Real. Soc. Jpn. **18**(3), 28–33 (2013)
Y. Miyake, AI engine in digital game. J. Inst. Image Inf. Telev. Eng. **68**(2), 125–130 (2014)
Y. Miyake, T. Yokoyama, T. Kitazaki, Developing character AI based on agent architecture in digital games, in *4th Digital Content Symposium, Proceedings*, 2-2 (2008)
Y. Morikawa, Use of artificial intelligence for video game. J. Jpn. Soc. Artif. Intell. **14**(2), 214–218 (1999)
Y. Morikawa, Affinity of game and AI, in *CEDEC 2008* (2008), http://cedil.cesa.or.jp/session/detail/156
K. Namiki, The case study of affordance oriented AI in "Monster Hunter Diary: Poka Poka Airou Village", in *CEDEC 2011* (2011), http://cedil.cesa.or.jp/session/detail/697
D. Nau et al., SHOP2: an HTN planning system. J. Artif. Intell. Res. **20**, 379–404 (2013), http://www.guerrilla-games.com/publications.html
M. Nelson, Encoding and generating videogame mechanics, in *Computational Intelligence and Games 2012 Tutorial* (2012), http://www.kmjn.org/notes/generating_mechanics_bibliography.html
M. Nelson, M. Mateas, Towards automated game design, in *AI*IA 2007: Artificial Intelligence and Human-Oriented Computing*, Lecture Notes in Computer Science, Springer, Berlin 4733 (2007), pp. 626–637
Neuro-evolving robotic operatives (2005), http://nerogame.org
N. Okamura, Hierarchical goal oriented planning and machine control in ARMORED CORE V, in *CEDEC 2011* (2011a), http://cedil.cesa.or.jp/session/detail/591
N. Okamura, Pathfinding in ARMORED CORE V, in *CEDEC 2011* (2011b), http://cedil.cesa.or.jp/session/detail/593
J. Orkin, Chapter 3.4, Applying goal-oriented planning for games, in *AI Game Programming Wisdom*, vol. 2, Charles River Media, Boston (2003), pp. 217–227
J. Orkin, Agent architecture considerations for real-time planning in games, in *AIIDE 2005* (2005), http://web.media.mit.edu/~jorkin/
J. Orkin, 3 states & a plan: the AI of F.E.A.R., in *Game Developer's Conference Proceedings* (2006), http://web.media.mit.edu/~jorkin/
B. Pell, METAGAME in symmetric chess-like games, in *Heuristic Programming in Artificial Intelligence 3 – The Third Computer Olympiad* (Ellis Horwood, Hemel Hemstead, 1992)
H. Penny Nii, The blackboard model of problem solving and the evolution of blackboard architectures. AI Mag. **7**(2), 38–53 (1986a), http://www.aaai.org/ojs/index.php/aimagazine/article/view/537

H. Penny Nii, Blackboard application systems, blackboard systems and a knowledge engineering perspective. AI Mag. **7**(3), 82–107 (1986b), http://www.aaai.org/ojs/index.php/aimagazine/article/view/550

R. Pillosu, Coordinating agents with behavior trees, in *Paris Game Conference* (2009), http://aigamedev.com/open/coverage/paris09-report/#session3

D. Pottinger, Terrain analysis for realtime strategy games, in *GDC 2000*, San Jose (2000)

S. Rabin, Agent awareness and knowledge representation, in *Game AI Pro* (CRC Press, Boca Raton, 2013)

M. Robbins, Chapter 30, Using neural networks to control agent threat response, A K Peters/CRC Press, Boca Raton, in *Game AI Pro* (2013), pp. 391–399

S. Russell, P. Norvig, *Artificial Intelligence, A Modern Approach*, 3rd edn. (Prentice Hall, Upper Saddle River, 2009)

N. Shaker, G. Yannakakis, J. Togelius, Digging deeper into platform game level design, in *Proceedings of the European Conference on Applications of Evolutionary Computation (EvoGames)*, Malaga (2012)

M. Shaker, M. Sarhan, O. Naameh, N. Shaker, J. Togelius, Automatic generation and analysis of physics-based puzzle games, in *Proceedings of the IEEE Conference on Computational Intelligence and Games*, Niagara Falls (2013)

G. Snook, Chapter 3.6, Simplified 3D movement and pathfinding using navigation meshes, in *Game Programming Gems*, vol. 1 (Charles River Media, Boston, 2000), pp. 288–304

K. Stanley, B. Bryant, R. Miikkulainen, Evolving neural network agents in the NERO video game, in *Proceedings of the IEEE 2005 Symposium on Computational Intelligence and Games*, Colchester (2005)

R. Straatman, A. Beij, W.V.D. Sterren, *Killzone's AI: Dynamic Procedural Combat Tactics* (2005), http://www.cgf-ai.com/docs/straatman_remco_killzone_ai.pdf

R. Straatman, T. Verweij, A. Champandard, Killzone 2 multiplayer bots, in *Paris Game/AI Conference* (2009), http://www.guerrilla-games.com/publications.html

R. Straatman, T. Verweij, A. Champandard, R. Morcus, H. Kleve, Chapter 29, Hierarchical AI for multiplayer bots in Killzone 3, A K Peters/CRC Press, Boca Raton, in *Game AI Pro* (2013), pp. 377–390

M. Tanabe, Engineering and design of soccer AI in WCCF, in *Digital Game Research Association of Japan, Regular Open Seminar of May in 2009* (2009), http://digrajapan.org/?wpdmact=process&did=OS5ob3RsaW5r

J. Togelius, J. Schmidhuber, An experiment in automatic game design, in *Proceedings of the 2008 IEEE. Conference on Computational Intelligence and Games*, Perth (2008), pp. 111–118

UNIX free software developed by Michael Toy, Glenn Wichman, Ken Arnold (1980)

W. van der Sterren, Chapter 13, Hierarchical plan-space planning for multi-unit combat maneuvers, in *Game AI Pro*, vol. 1, A K Peters/CRC Press, Boca Raton (2013), pp. 169–183

J. von Uexküll, *Streifzüge durch die Umwelten von Tieren und Menschen: Ein Bilderbuch unsichtbarer Welten* (Rowohlt, Reinbek, 1956)

M. Walsh, Dynamic navmesh – AI in the dynamic environment of splinter cell: conviction, in *GDC 2010* (2010), http://www.gdcvault.com/play/1012651/Dynamic-Navmesh-AI-in-the

M. Walsh, Modeling AI perception and awareness in splinter cell: blacklist, in *GDC 2014* (2014), http://www.gdcvault.com/play/1020195/Modeling-AI-Perception-and-Awareness

M. Widmark, Terrain in Battlefield 3: modern, complete and scalable system, in *GDC 2012* (2012), http://www.gdcvault.com/play/1015414/Terrain-in-Battlefield-3-A

S. Yoneda, [SQEXOC 2012]AI technologies in FFXIV, 4gamers (2012), http://www.4gamer.net/games/032/G003263/20121205079/

Hand Gesture Interface for Entertainment Games

12

Kiyoshi Hoshino

Contents

Abstract
The author proposes "games that can be played in a virtual world through the same movements in the real world without installation of sensors and a special controller." In particular, this chapter introduces a virtual 3D block-building play system that can control the movements with human gestures. This system requires only two small inexpensive RGB high-speed cameras as the peripheral devices. It is not necessary to attach reflective markers on the body of the user.

K. Hoshino (✉)
University of Tsukuba, Tsukuba, Ibaraki, Japan
e-mail: hoshino@esys.tsukuba.ac.jp

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_47

The core of this technology is a hand pose estimation that restores 3D postures from 2D hand images. The technology, which is a so-called data glove without glove, can determine the poses of human hands and fingers or joint angles seamlessly. It is not a pointing device where specific movements are assigned to specific functions of a game. A user therefore does not have to master the instruction operations by the hands in advance to use the functions of the game. He can operate a virtual game by behaving in a manner similar to the daily routine movements as in the physical world. In this chapter, the author also introduces the 3D modeling system based on hand gesture (hand gesture-based CAD system) using the depth sensor for recognizing hand poses.

Keywords
Estimation of full articulation of a human hand from markerless visual observations • Hand gesture-based CAD • Hand pose estimation • Motion capture • Physical simulation of the linkage between human 3D gestures and virtual CG/ robots • RGB high-speed camera • Virtual block-building play system

Introduction

Most of the technological and engineering specialists eagerly desire to achieve an ambition. It is to develop a practical technology to enable the users/operators to operate all the artificial systems, such as information and communication equipment, CAD, digital signage displays, home appliances, remote-controlled robots, and medical devices, by means only of gesture including sensitive hand motions with no need for sensors equipped and for prior learning of manipulation.

It should be specially mentioned that pointing devices have been commonly used for conventional human interfaces, most of which are based on the motion vector analysis technique, namely, the hand tracking technique. In fact, it may be assumed with respect to these devices that rock and scissors of the rock-paper-scissors game are assigned to the F1 key and the F2 key on the keypad, respectively. However, a human interface based on such a technique as hand motion capture, namely, “hand pose estimation,” is required for the users/operators to operate artificial systems and remote-controlled robots with no sensors equipped simply using movements in daily living. This technique enables estimation of full articulation of a human hand from markerless visual observations at every frame of video; more specifically, it enables physical computation (physical simulation) of the linkage between 3D gestures including sensitive finger movements and artificial structures/virtual CG/robots, entertainment games, digital signage, etc.

The approaches adopted in the existing hand pose estimation systems equipped with an RGB camera are largely classified into two categories. One of them is a 3D model-based approach, which extracts a local feature, e.g., silhouette, from an image acquired by the camera and assigns the 3D pose model previously built within a computer to the extracted feature (Rehg et al. 1994; Kameda et al. 1996; Lu et al. 2003; Gumpp et al. 2006). The other is a 2D appearance-based approach, which directly compares the appearances between input images and the images stored in a database without extracting the feature (Athitos et al. 2002; Hoshino and Tanimoto 2005; Wu et al. 2005). The former allows for high-accuracy hand pose estimation but has disadvantages of being prone to self-occlusion and requiring longer processing time. In contrast, the latter reduces computation time. However, a large scale of matching database is required to correspond to changes in 3D appearance of the hand including the wrist and forearm movements. The author's research group had assumed that if any difficulty in hand pose estimation arose essentially due to the complicated hand pose and self-occlusion, high-accuracy hand pose estimation could be theoretically achieved simply by building a large-scale database containing a huge amount of information on all the appearances of hand pose even though these negative factors remain unsolved and that the feasibility of this approach might depend on a search algorism for finding out about the similar images from the large-scale database at high speed. On this assumption, we proposed some high-speed and high-accuracy hand pose estimation systems (Hoshino et al. 2007; Hoshino et al. 2012; Tomida 2015). These systems have achieved a process of hand pose estimation, including image acquisition, estimation, and robot control, on a single notebook computer at 150 fps of maximum speed and about 5° of mean error in joint angle estimation.

Meanwhile, recently, inexpensive, smaller depth sensors, which allow for acquisition of depth information in addition to RGB 2D images, have been commercially available. Accordingly, 3D information may be also directly applied to input information and models when the aforementioned 3D model-based approach is used. For example, the method has been proposed to calculate the model parameters, which minimizes any inconsistency between hand depth information given by the depth camera and the 3D model by particle swarm optimization (Oikonomidis et al. 2011). This approach, however, could only just achieve low processing performance of about 15 fps even on a high-performance computer due to high computation cost.

Therefore, the author first introduces "games that can be played in a virtual world through the same movements (gestures) in the real world without installation of sensors and a special controller" in the second section. In particular, the chapter introduces a virtual 3D block-building play system that can control the movements with human gestures. This system requires only two small inexpensive RGB high-speed cameras as the peripheral devices. It is not necessary to attach reflective markers on the body of the user. The author second introduces hand pose estimation at high speed with high accuracy, with 3D images acquired by a depth sensor in the third section. And finally, the author introduces the 3D modeling system based on hand gesture (gesture CAD) using the depth sensor for recognizing human hand poses in the fourth section, and wearable hand Motion capture system with an ultraminiature camera in the fifth section.

Virtual Block-Building Play System

Hand and Arm Motion Capture with Two Cameras

The core of this technology that is introduced in the second section is a hand pose or hand posture estimation technology that restores 3D postures from 2D hand images. This technology, which is a so-called data glove without glove, can determine the poses of hands and fingers or joint angles seamlessly. Similar technologies that are compared with it are a hand-movement vector analysis technology and a hand tracking technology that detect the hand region and the direction and the speed of movement. These technologies are designed to identify movements to limited types of states and are not designed to comprehend the poses of the hands themselves. These technologies are used as pointing devices where specific movements are assigned to specific functions of a game. Therefore, the user must master the instruction operations by the hands in advance to use the functions of the game. On the contrary, in the system that uses the hand pose estimation technology as the core as described in this manuscript, a user can operate a virtual game by behaving in a manner similar to the daily routine movements as in the physical world.

Hand pose estimation, estimation of full articulation of hand, and capturing of articulated hand motion are realized as follows in our system. A cross-checking database of an enormous size is prepared by storing the following three types of information items as a set for hand images of various poses: (i) “image pose features” that indicate the degree of bending of four fingers and the thumb and the degree of tilt of the wrist (Hoshino et al. 2012); (ii) low-level “image values” that present the positions and amounts of the dots, the line segments, the inclinations of the line segments, the edges, and the 45-degree lines that exist in each hand image; and (iii) “joint angles” of the hand pose. Since this system uses two RGB high-speed cameras that are installed in a loose orthogonal relationship, two information items that are captured in advance by those cameras are used for (i) and (ii). The joint angle information of (iii) is used for output of the estimation result.

When a hand of a user is captured by two cameras, as the first step, similar hand poses are selected from the hand data in the database at high speed. As the second step, the most similar hand pose is selected by the detailed similarity calculation based on the image feature values of (ii). However, if an upper body including a hand is captured by two cameras, two image pose ratios and image feature values are obtained for the hand pose. Since the two cameras are installed in the positions in a mutually loose orthogonal relationship and the four fingers and the thumb can be considered to be bent toward the palm, one camera captures an image viewed from the back or palm of the hand, and the other comparatively clearly captures the opposing state of the four fingers and the thumb according to the rotation angle of the forearm. Firstly, the sizes of the hand regions that were captured by the two cameras were compared and the larger one was selected. Approximate narrowing was performed by using the pose ratio of the hand of the selected images only. To put it simply, approximate poses viewed from the back of the hand were determined. Secondly, the degrees of finger bends were confirmed using the images in the orthogonal relationship among the selected candidates (Hoshino et al. 2012).

For the upper body motion capture, the two cameras in an orthogonal relationship were used. The region that is comparatively slimmer near the tip of the arm was assumed as a wrist, and the 3D coordinates of the wrist and the elbow and the wrist vector and the elbow vector were obtained.

Virtual Block-Building Play System

The virtual block-building play system that is performed by hand and arm gestures comprises the following components. 3D virtual space is prepared in advance. Multiple objects such as blocks, a floor, and a hand are prepared in the virtual space. As shown in Table 1, each object has two types of information, drawing information and physical information, each of which has information items of a posture, a surface material, and a constraint condition.

The real-time physics engine (physical simulation software), Bullet, is used. Force information including speed, acceleration, gravity, and center of gravity is provided to the physical simulator as parameters regarding the physical posture of the object, and the physics engine calculates which motions and posture changes occurred to the object at each clock time including the case of the occurrence of collision.

Among the posture information of hands and arms that were obtained from the motion capture that was described above, the physical parameters are updated from the outside of the engine by using position information only.

Table 1 Information attributed to each object

Object	Drawing Information	Posture
		Surface material
		Constraint condition (joints of the hand)
	Physical Information	Posture
		Surface material (friction coefficient, reflection coefficient, density, and mass)
		Constraint condition (For instance, the link structure of the hand to simplify calculation according to the condition or the relationship between the block and the hand to improve the stability when a block is picked up by fingers.)

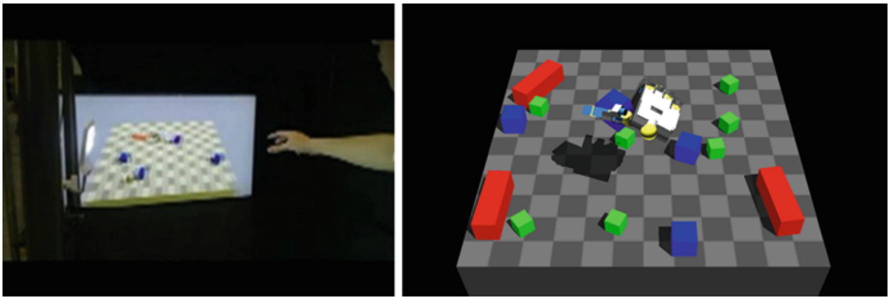


Fig. 1 Appearance of the virtual block-building play system

Evaluation Experiment

A high-speed camera (manufactured by ViewPlus, with wide-angle lens installed in Firefly MV) was installed at the top and side of the user. These cameras were positioned to ensure that the hand and the forearm are captured when the user moves the dominant arm freely at the front of the body.

Figure 1 shows an example of a snapshot using this system. You may intuitively understand that the user can enjoy virtual block-building play by moving the hand and the arm in the same way as the real block-building play. The 3D virtual space that is constructed enables the drawing of the system in a stereoscopic view of 3D. The stereoscopic display assists the user to understand the correspondence between the real space and virtual space, thereby enhancing the operability of the block-building play.

Turning Pages on a 3D Book Icon

Change of the icon from virtual 3D building block to 3D book causes the shape of the book icon to be changed as if it was an actual book simply by turning pages in the same manner as an activity of daily living. Being not a clear image, the following figure shows the changing process of the shape of the book icon when the user turns pages on the 3D book icon by fingers. In the lower part of each of the images, the user's pinching motion, which is made opposite the whitish display in the upper part, is reflected. It can be seen that the target page is appropriately pinched by the user's fingertip pinching motion, page turning CG being successfully created in the same manner as in the physical world (Fig. 2).

Hand Pose Estimation Based on 3D Images

Hand Pose Estimation Using a Depth Sensor

The author introduces the system in the third section, which has fully adopted the knowledge learned in the process of building the monocular camera-based 2D hand pose estimation system to enable hand pose to be estimated on 3D images acquired

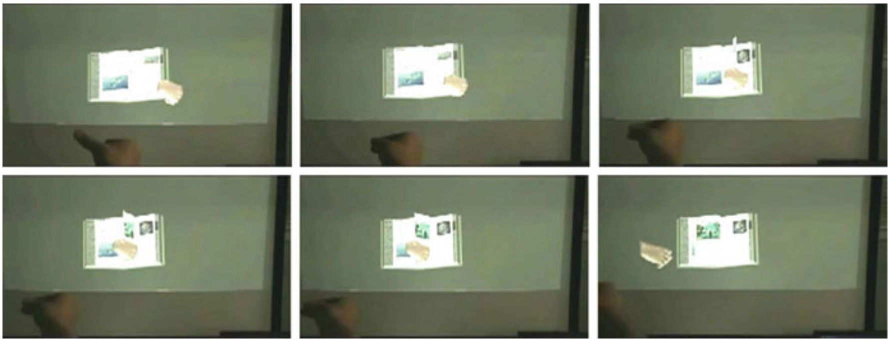


Fig. 2 3D digital book allowing the shape of the digital book icon to be changed by page turning in the same manner as in the physical world

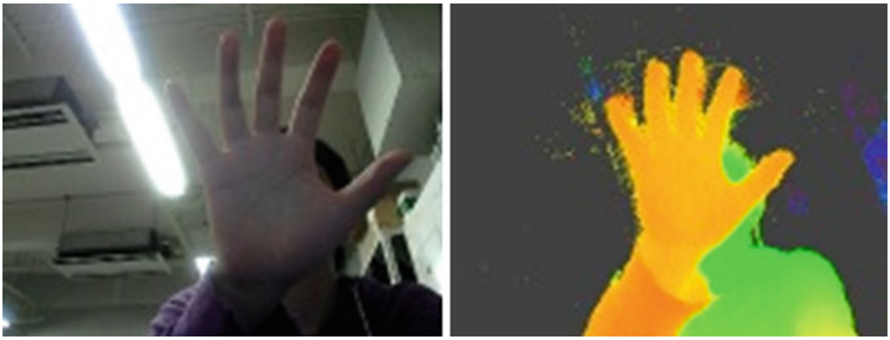


Fig. 3 A palm placed over the sensor (*left*) and an example of the observed hand depth image (*right*)

by a camera equipped with a depth sensor at high speed. In this system, the small-sized TOF depth sensor (DepthSense 325 from SoftKinetic) is installed under the display. As illustrated in Fig. 3, color images and depth images may be acquired with resolutions of HD 720p and QVGA and maximum speeds of 30 fps and 60 fps, respectively. The focus distance is short, falling within a range of 15 cm to 1 m. Only depth images are used in this system.

System Configuration and Estimated Result

The matching database for collating with depth images including those of the hand, which is similar to the database for hand pose estimation based on RGB 2D images, contains data sets, each consisting of three types of data, the image pose ratio and image feature amount (HLAC) of the hand cut off from the image and the angles formed by finger joints to output estimated results. The matching database contains thousands to tens of thousands of data sets; thereby, first, the data sets are narrowed based on an image pose ratio with the less number of dimensions, and second, similarity is precisely collated with the narrowed candidate data set. In hand pose estimation using the RGB camera, the image pose ratio is represented by the values for three parameters: the ratio of longitudinal length (rough longitudinal length of the hand pose, or rough rotating state of the forearm), the ratio of upper length (degree of upper-part length, or rough bending state of four fingers), and the ratio of right length (the degree of deviation in right and left balance, rough bending state of the thumb) (Hoshino et al. 2012). In hand pose estimation based on depth information from the depth sensor, the image pose ratio, which reflects to what extent the hand area has deviated forward or backward (rough degree of forward projection of the thumb and other four fingers), is introduced. On the other hand, the image pose ratio is calculated for the depth image of each frame with four degrees of freedom (DOF) during estimation to narrow the data sets in the matching database. With four DOF, which is lower degree of freedom, data sets can be narrowed within the large-scale database containing thousands to tens of thousands of data sets in an extremely short time period.

Figure 4 shows an example of hand pose estimation based on the depth image from the depth sensor. The input image is shown on the left side and the estimated result (the searched most similar data set) is shown on the right side. Although in rare cases, a little bit estimation error may occur (as seen at the bottom of the figure), it can be understood that the angles formed by finger joints were estimated at high accuracy even when the palm did not face toward the camera because of the rotating forearm. When the matching database containing 20,000 data sets was used, the mean processing time required to estimate one image was 70 fps on the desktop PC with general specification (Intel® Core™ i5-2500 CPU 3.30GHz, 8.00GB RAM).

Hand Gesture-Based CAD System

System Configuration

Similarly, the 3D modeling system based on hand gesture (gesture CAD) uses the depth sensor for recognizing hand poses (Hamamatsu et al. 2013; Hoshino et al. 2015). Specifically, 3D CAD and 3D modeling software are available for operating and creating 3D models on the computer. These devices, however, are 2D input devices such as mouse and keyboard and 3D devices such as at best 3D mouse and 3D joystick. They have a disadvantage that they require special knowledge and skill to operate or they are difficult to operate intuitively. To solve these problems, focusing on "pinching a thing by fingers," the movement in daily life, as an input interface based on 3D gesture recognition, we made it possible to give single DOF corresponding to mouse click input based on whether the fingers are in contact with each other, independently of six DOF input for the position and pose.

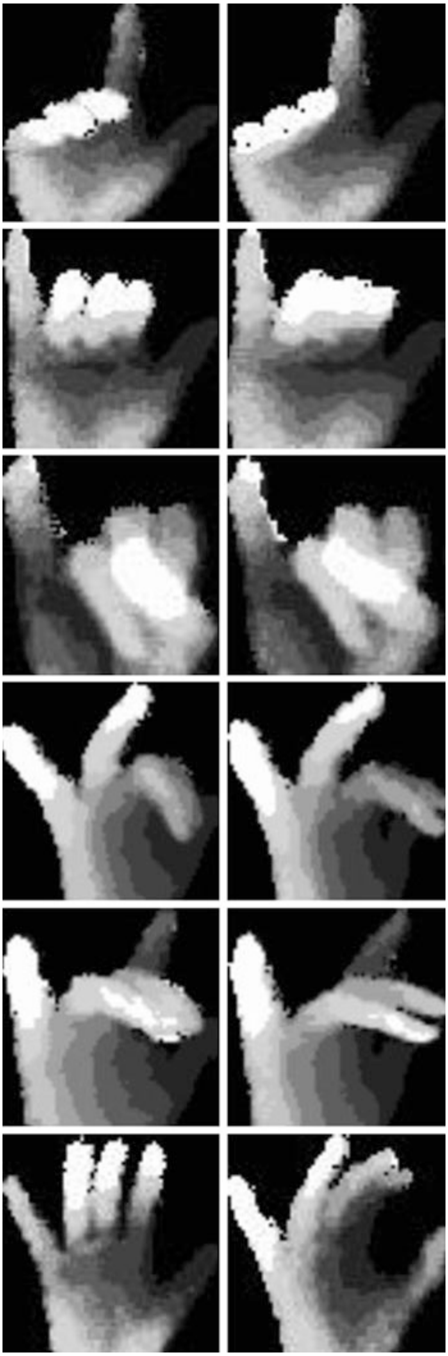


Fig. 4 Example of hand pose estimation based on the depth image from the depth sensor. Input image is on the left side and the estimated result is on the right side

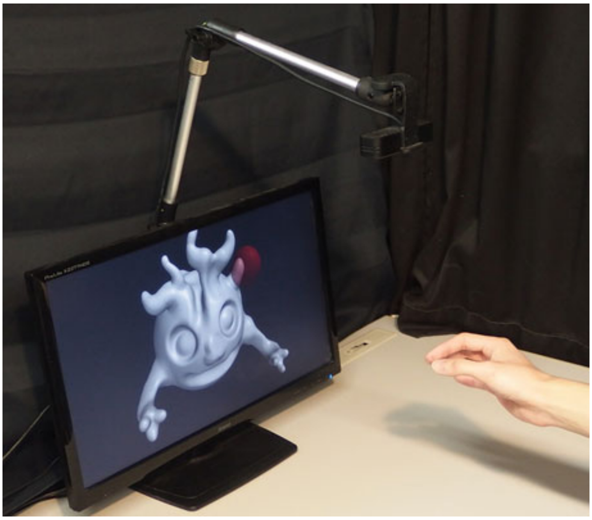


Fig. 5 Appearance of the gesture CAD system

In this system, as illustrated in Fig. 5, a single depth sensor (DepthSense 325 from SoftKinetic) is installed on the top of the monitor. This TOF sensor is able to acquire a depth image of 320 $\times$ 240 pixels at the maximum speed of 60 fps. For the depth image, which serves as an input to the system, information on the pixel positions and the depth are represented by the coordinates and pixel values, respectively.

The depth value of the image obtained directly from the depth sensor contains a large amount of noise. To eliminate this noise for smoothing, three types of filters are overlaid on the sensor in advance. First, a 5 $\times$ 5 pixel median filter is applied to the obtained image. The median filter has a profound noise reduction effect on shot noise specific to the image, which is a local outlier in a space. Second, an adaptive moving average filter is applied. This filter varies the weights on the result of moving average calculation at the previous time point and on the result at the earlier time points for each frame. Usually, the moving average of sufficiently long time period is applied. For example, when the depth value largely varies, the value for the weight is decreased to weaken smoothing to enhance responsiveness, resulting in prevention of output delay. Third, a 5 $\times$ 5 pixel Gaussian filter is applied twice. This smooths fine unevenness, on which the median filter has lower reduction effect, into consecutive values. After three types of filtering processes, the depth image is obtained as shown in Fig. 6.

To extract the hand area, an image (hereafter, referred to as an intensely smoothed image (Fig. 7b), which is intensely smoothed by applying an 11 $\times$ 11 pixel Gaussian filter eight times, is created in addition to the image (Fig. 7a) obtained after three types of filtering processes. Next, the gradients in the x and y directions are calculated on these two images using a 3 $\times$ 3 Scharr filter. The pixels, for which depth values and gradient strengths are within a given range of values, are binarized to create a silhouette image of the hand area (Fig. 7c). The

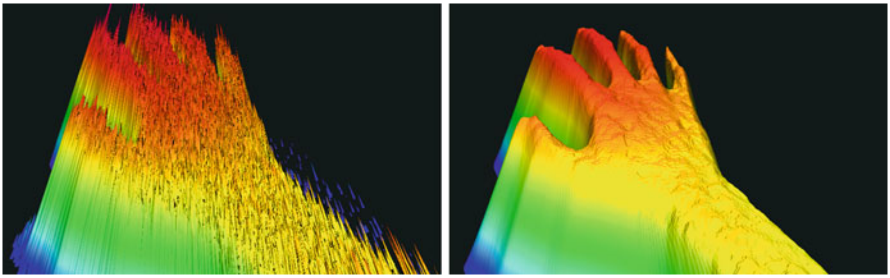


Fig. 6 The original image obtained from the depth sensor (left) and the image obtained after three types of filtering processes (*right*)

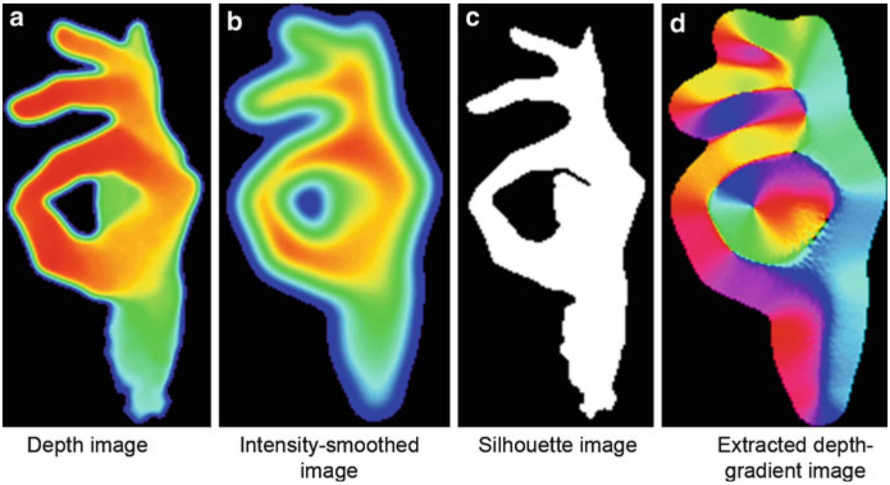


Fig. 7 Various types of created images

gradient direction is calculated based on vectors in two different directions on the intensely smoothed image to define the image containing this information as an abstracted depth gradient image (Fig. 7d).

Other than our research, a detection method of pinching was proposed using the RGB camera as a GUI operational method in the desktop environment (Wilson 2006). Moreover, pinching gesture recognition was proposed for tabletop entertainment system (Fukuchi et al. 2010). Both the researches have three big problems with pinching motion recognition as an input interface. The first problem lies in that positions are defined based on the centroid of the inner space formed by making two fingers into contact with each other. Since it is difficult to reflect sensitive movements of the hand on the input device, the user needs to move the hand excessively more than necessary. The second problem is no consideration of the hand state taken when no pinching motion is made. The user cannot grasp what position should be recognized as input until making fingers in contact with each other.

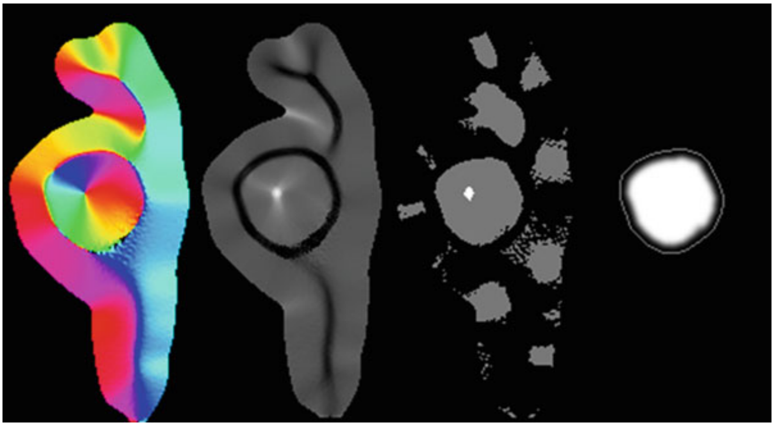


Fig. 8 Extraction of a concave part

The third problem is that only positions and rotation state can be recognized two dimensionally. In short, the hand movement cannot be recognized when the hand pose changes three dimensionally. To solve these problems, the author's research group had paid attention to the correlation between the hand pose during pinching motion and the depth value on the image (Hoshino et al. 2015). Specifically, when one makes the fingers into contact with each other for pinching a thing, a closed space is formed by the fingers. It, however, is difficult to detect the closed space by digitizing the hand area and the background area because how to pinch varies between individuals and the other finger or palm is reflected in the closed space when the hand moves. To address this problem, multiple regression analysis was used with the depth values around the finger areas to eliminate the hand poses other than those of the fingers used for input and the positions of the wrist and forearm.

The aforementioned abstracted depth gradient image has information on the abstracted rough depth gradient of the hand, instead of the precise hand pose. Based on this gradient information, a concave (凹) part, which occurs in the hand area during pinching motion, is extracted. On the abstracted depth gradient image illustrated (Fig. 8, at the far left), it is determined whether eight pixels around the pixel of interest form a radial gradient. After matching is performed with all the pixels, the pixels with higher similarity are ternarized based on two thresholds. In the same figure, the similar pixels are shown in gray and more similar pixels are shown in white. The gray similar pixel area is segmented by labeling and only the areas containing white similar pixels are extracted. The area covering the largest amount of extracted pixels is determined to be the concave (凹) part (Fig. 8, at the far right) formed by pinching motion. Since this concave part is detected based on the abstracted depth gradient image, it may be detected even before the fingers come into complete contact with each other. It is detected whether pinching motion is made based on the touching state of the fingers in this concave part. To estimate the hand pose, the silhouette around the fingers is approximated on the 3D plane based on depth information on the hand around the outline of the concave part as shown in Fig. 9.

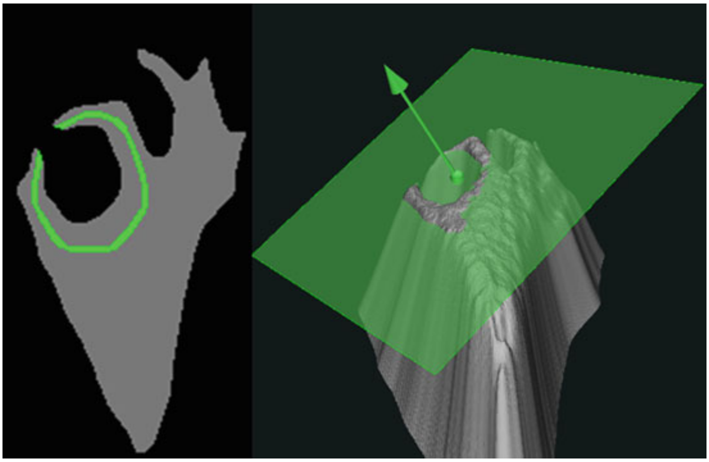


Fig. 9 Plane approximation of a group of points by multiple regression analysis using the least-square method

Example of Hand Gesture-Based Modeling

The aforementioned procedure using the depth sensor can be used to create 3D structures by means of pinching motion of fingers as illustrated in Fig. 10. Moreover, as illustrated in Fig. 11, both the right and left hands may be used to make modeling gesture. For example, for the right-handed user, the right hand is used for modeling gesture, while the left hand is used to move the viewpoint or the operational object, achieving efficient 3D modeling (Fig. 11, upper). It was demonstrated that a single depth sensor might recognize the motions of up to 25 hands at the same time on the PC with Intel Core i7-3770 K (3.50 GHz) mounted, at the processing speed of 35 fps. In the case of a single hand, the maximum processing speed was 106 fps. Since the image acquiring speed of the sensor is 60 fps, the pinching motions of up to nine hands can be recognized at the same time with no reduction in recognizing speed (Fig. 11, lower). This means several users may make gesture modeling together.

Wearable Hand Motion Capture System

Ultraminiature Camera-Based Hand Pose Estimation

Motion capture system has been widely used to reproduce the computer animation of human bodies in movies and the actions of characters imitating human motions in games. Moreover, a system, which enables a facial expression to be drawn in real time as a computer animation by the combined use of human motion capture, has been also practically used. The ultimate technique, which is critical for editing of

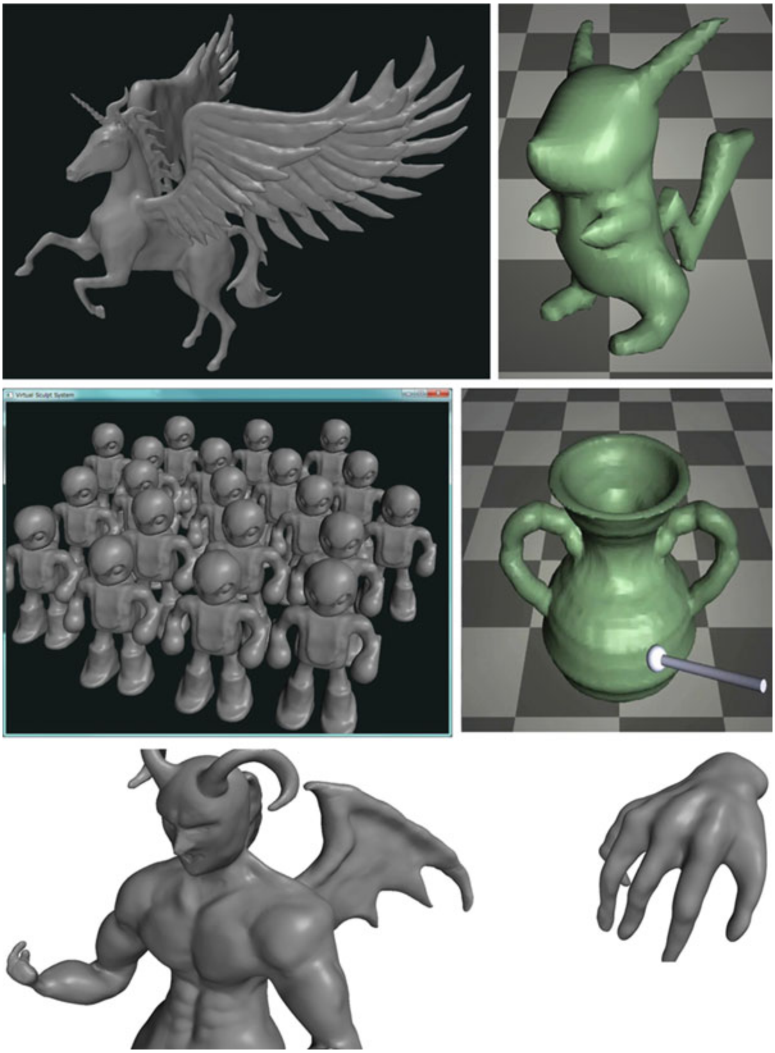


Fig. 10 Examples of modeling with a hand gesture CAD

human motion images but has not yet reached the practical level, is “motion capture system for fine hand and finger motions,” namely, 3D hand pose estimation (Fig. 12).

The technical specifications, which do not affect motion capture, are required to use the hand pose estimation with the motion capture system. Such technical specifications include: the ability to estimate the pose of a relatively small hand even while a subject is moving freely around a wide studio for optical motion capture; when the subject grasps a thing or his/her palm comes into contact with any of other objects such as the floor, wall, desk, or the other person, the ability to

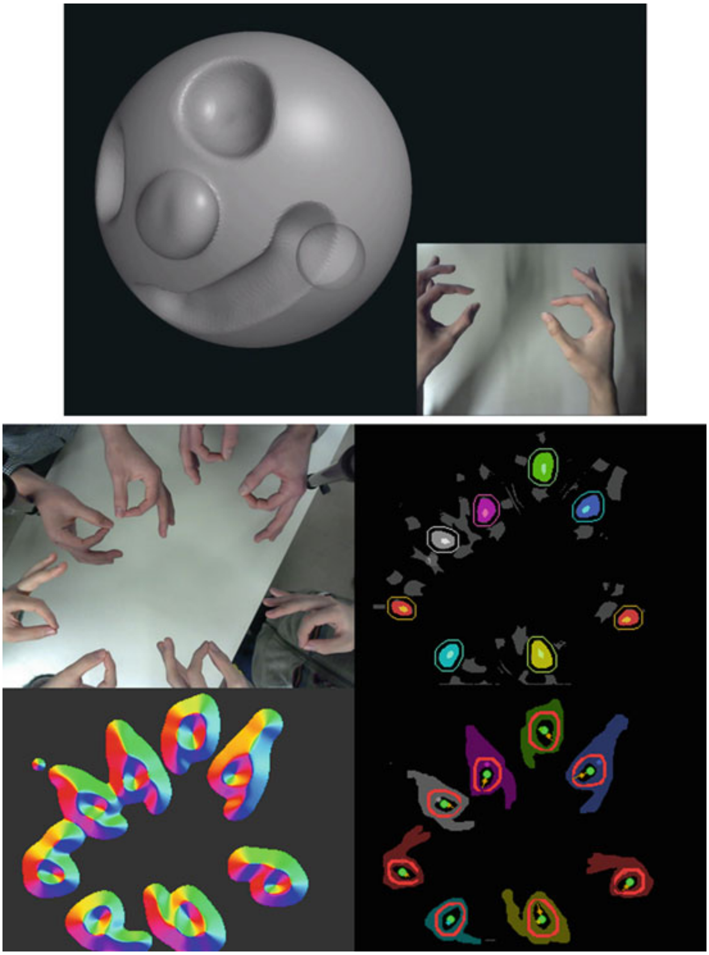


Fig. 11 Examples of gesture modeling made by several hands (*upper*) and various types of generated images (*lower*)

estimate his/her hand pose; and the equipment having a form and mass weight which does not cause the subject to feel uncomfortable. The authors have proposed a non-contact system, which enables hand pose estimation to be performed immediately and accurately using one or two RGB high-speed cameras (Hoshino and Tanimoto 2005; Hoshino et al. 2007; Hoshino et al. 2012). However, this non-contact hand pose estimation system has a disadvantage: when the subject moves freely around a wide studio as with motion capture, the images of the hand away from the camera(s) are very small, making it difficult to carry out hand pose estimation. To solve this problem, we have conceived an idea that the use of a wearable camera, rather than completely non-contact type, may achieve compact hand pose estimation equipment.

Fig. 12 Performance of the craft of drawing out the sword and CG Samurai drawn with the motion data (Endo 2015)

One of the techniques we figured out is to use a wireless data glove. However, the data glove has the following disadvantages: being expensive, feeling of tight, its strain indicator (a sensor) and wiring easy to be broken, and inability to be used for quick motion capture. Moreover, this technique is not suitable for the motions such as object manipulation with the subject's palm and fist making. Digits (Kim et al. 2012) are one of the examples of hand pose estimation equipment using a wireless wearable camera. Digits capture the images of the subject's hand, who has attached an infrared sensor or infrared camera on the wrist, from the palm side to measure the distance between the sensor and the fingers, making hand pose estimation. Note that this system, which requires a sensor attached on the wrist, has a disadvantage of difficulty in feasible response to flexing and extending motions. In addition, in this system, the sensor must be attached on the palm side of the hand; thereby, it may disturb the subject's motions such as daily actions (e.g., walking) and grasping/manipulating or coming into contact with other objects, leading to impossibility of measurement. Other techniques have been reported by a study on hand pose estimation using an RGB camera and an AR marker (Fernandes et al. 2008) and a study on hand pose estimation based on the convex and concave shapes of the wrist using large quantities of photo reflectors attached on the wrist (Fukui et al. 2011). Both of them also have disadvantages of preventing the subject from doing work, insufficient accuracy, etc.

To address these problems, the author proposed an originally developed compact system (Tomida 2015; Hoshino 2014), which allows for estimation of the hand pose of the subject, who has attached an ultraminiature wireless RGB camera on the back side of the hand, rather than the palm side. Attachment of the camera on the back side of the hand may make it possible to minimize the restraint on the subject's motions during motion capture. The conventional techniques attach the camera on the palm side of the hand for the reason that the images of fingertips always need to be captured by the camera before any other parts of the hand. In contrast, the algorism, the author proposes here, is capable of estimating the hand pose with no need for capturing first the fingertips.

Camera Mounting Fixture

A camera fixing device consists of a wiring of 3 mm in diameter. One end of the wiring is fixed to the camera, and the other end is formed into a figure of eight (or ∞ shape) so as to be stably attached on the back of the hand. To fix the camera mounting fixture to the back of the user's hand, a hook and loop fastener is wound around the metacarpophalangeal (MP) joint on its wrist side (the palm). To relieve discomfort from attachment, soft-type Velcro is used. An ultraminiature wireless camera captures VGA images at 30 fps. The captured images are received at a specialized receiver and input into a PC via video capture cable. Figure 1 shows the appearance of the hand with the camera mounting fixture and the ultraminiature wireless camera attached on it. The weight of the mounting fixture (including the camera) is 26 g.

To enable the operator to estimate the flexing and extending motions of the fingers, especially the thumb, index, and third finger, as correctly as possible, by a single camera, the camera captures the back of the hand obliquely downward from the upper anterior side with a focus on the area between the index and third finger. With 90° in the angle of camera and 10 cm in the distance between the center of the back of the hand and the camera lens, the tips of the thumb and other four fingers are not captured (Fig. 13).

Construction of Matching Database

The present method uses 2D appearance-based hand pose estimation. However, 2D RGB images are used as input, while 3D information on the angles of the thumb and finger joints is necessary for output. Specifically, such a problem remains to be solved that 3D shapes must be restored from 2D images. To address this problem, information on the angles of these joints is added in a 2D hand pose database for matching to enable the 3D information on the angles of the hand-finger joint to be output from the 2D images.

An experimenter wears the abovementioned camera mounting fixture on one hand with the other hand on a data glove (CyberGlove, Virtual Technologies). Then, he/she makes a variety of same motions with both hands carefully. This data glove with 2° of freedom is capable of outputting 24 types of information on the thumb and finger angles of the joints. The information is composed of 24 components in total, including the angles of three joints (metacarpophalangeal (MP), proximal interphalangeal (PIP), and distal interphalangeal (DIP) joints), with two kinds of unused information, 4° of freedom (4-DOFs) opening/closing between four fingers excluding the thumb and 3° of freedom wrist motion. Note that the angles of

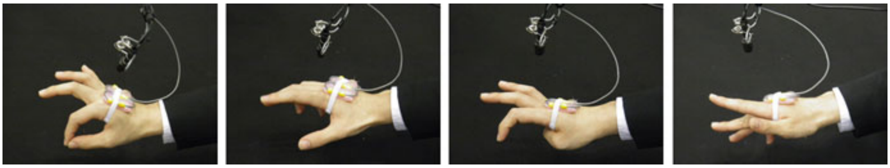


Fig. 13 Appearances of the ultraminiature camera-based hand pose estimation system

the DIP joint of the four fingers excluding the thumb have 18° of freedom because they are calculated using the angles of the PIP joint. Moreover, information on 3° of freedom wrist motion, which can be acquired only by combining this camera with a motion capture, was not used in this study.

Based on the data from the data glove and the hand images, both of which are captured at the same time, the amount of feature was found by following the procedure described below. First, only the hand region was cut out from the captured image and the image was reduced to the size of 64 × 64 pixels. Second, this image was split into smaller pieces to form an 8 × 8 matrix. Third, for each total of 64 (8 × 8) split pieces, the features of its high-order local autocorrelation were found. The high-order local autocorrelational function is useful in recognizing and measuring images; accordingly, the features of the reference point and its neighborhood were calculated using a high-order local autocorrelational function. This means that a hand image is described using the 1,600 (in total, 8 × 8 × 25) dimensional amounts of features.

In the matching database built in this manner, each data set has a combination of 1,600 dimensional amounts of features and 18° of freedom angles of the MP/PIP/DIP joint. In this study, subjects, whose hands were different in size and shape, were participated, the matching database containing about 30,000 data sets being built.

Similarity was calculated using the simple Euclidean distance method.

Evaluation Experiment

To study the effectiveness of this system, the author first conducted search of similar hand images from the database on sequential images of human hand motions. With an ultraminiature wireless RGB camera on the back side of the hand, the subjects freely moved his/her hand fingers. Figure 14 indicates an output image during the test captured on the screen, as example of estimated result. A hand image in the left and CG hand in the right indicate a captured image and generated CG image with estimated results, respectively. The sampling frequency of the camera is 30 fps. From this drawing, you can see that a processing speed of approximately 30 fps is obtained as well, even in the case where a personal computer of conventional functions (Pentium IV 2.8 GHz, main memory 1 GB) is used.

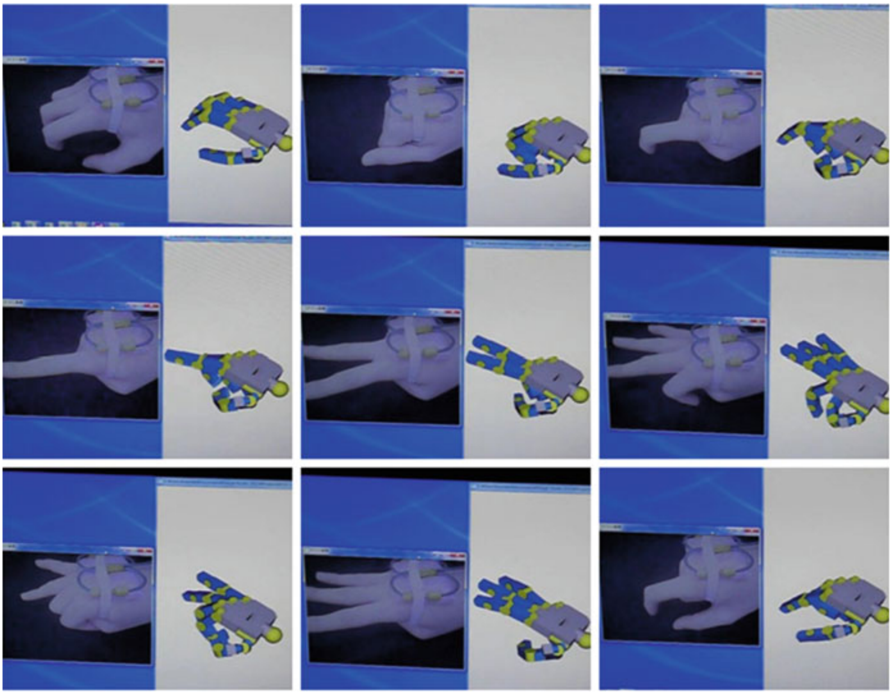


Fig. 14 Snapshots of estimated results drawn using CG hands

Recommended Reading

J.M. Rehg, T. Kanade, Visual tracking of high DOF articulated structures: an application to human hand tracking, *European Conference on Computer Vision*, 1994. pp. 35–46

Y. Kameda, M. Minoh, A human motion estimation method using 3-successive video frames, *Proceedings of the virtual systems and multimedia*, 1996. pp. 135–140

S. Lu, D. Metaxas, D. Samaras, J. Oliensis, Using multiple cues for hand tracking and model refinement, *Proceedings of the CVPR2003*, 2003. 2: pp. 443–450

T. Gumpp, P. Azad, K. Welke, E. Oztop, R. Dillmann, G. Cheng, Unconstrained real-time markerless hand tracking for humanoid interaction, *Proceedings of the IEEE-RAS international conference on humanoid robots*, 2006. CD-ROM

V. Athitos, S. Scarloff, An appearance-based framework for 3D hand shape classification and camera viewpoint estimation, *Proceedings automatic face and gesture recognition*, 2002. pp. 40–45

K. Hoshino, T. Tanimoto, Real time search for similar hand images from database for robotic hand control. IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences **E88-A**(10), 2514–2520 (2005)

Y. Wu, J. Lin, T.S. Huang, Analyzing and capturing articulated hand motion in image sequence. IEEE Trans. Pattern Anal. Machine Intell. **27**, 1910–1922 (2005)

K. Hoshino, E. Tamaki, T. Tanimoto, Copycat hand – robot hand imitating human motions at high speed and with high accuracy. Advanced Robotics **21**(15), 1743–1761 (2007)

K. Hoshino, T. Kasahara, M. Tomida, T. Tanimoto, Gesture-world environment technology for mobile manipulation – remote control system of a robot with hand pose estimation. Journal of Robotics and Mechatronics **24**(1), 180–190 (2012)

M. Tomida, K. Hoshino, Wearable device for high-speed hand pose estimation with a miniature camera. J. Robot. Mech. **27**(2), 167–173 (2015)

I. Oikonomidis, N. Kyriazis, A.A. Argyros, Efficient model-based 3D tracking of hand articulations using Kinect, *Proceedings of the 22nd British Machine Vision Conference* (*BMVC'2011*), 2011. 1: pp. 1–11

K. Hamamatsu, K. Hoshino, Detection of pinching gestures using a depth sensor and its application to 3D modeling, *2013 IEEE/SICE international symposium on system integration*, 2013. TA2-K(1): pp. 814–819

K. Hoshino, K. Hamamatsu, Gesture CAD, *Proceedings of the 38th international MATADOR conference on advanced manufacturing*, 2015. 2(1): pp. 17–22

A.D. Wilson, Robust computer vision-based detection of pinching for one and two-handed gesture input, *Proceedings of the UIST'06*, 2006. pp. 255–258

K. Fukuchi, T. Sato, H. Mamiya, H. Koike, Pac-pac: pinching gesture recognition for tabletop entertainment system, *Proceedings of the international conference on advanced visual interfaces*, 2010. pp. 267–273

M. Endo, The inseparable pen and sword – study in spirit and in practice. J. Japanese Soc. Biofeedback Res. **42**(2) 2015, printing

D. Kim, O. Hilliges, S. Izadi, A. Butler, J. Chen, I. Oikonomidis, P. Olivier, Digits: freehand 3D interactions anywhere using a wrist-worn gloveless sensor, *UIST'12 proceedings of the 25th annual ACM symposium on user interface and software technology*, 2012. pp. 167–176

L.A.F. Fernandes, V.F. Pamplona, J.L. Prauchner, L.P. Nedel, M.M. Oliveira, Conceptual image-based data glove for computer-human interaction. Rev. Inf. Teórica Aplicada (RITA) **15**(3), 75–94 (2008)

R. Fukui, M. Watanabe, T. Gyota, M. Shimosaka, T. Sato, Hand shape classification with a wrist contour sensor: development of a prototype device, *UbiComp '11*, 2011. pp. 311–314

K. Hoshino, Ultraminiature camera-based hand pose estimation. *Proceedings of the 18th International Conference on Mechatronics Technology* (*ICMT*), 2014, 18(77): pp. 1–8

Intelligent Character Technologies for Entertainment Games

13

Hiroshi Mori

Contents

Abstract

Human characters are essential to adding a sense of reality to virtual reality worlds in entertainment games. These characters require intelligence to give the virtual world a sense of reality or entertainment as a game. This chapter describes technologies for generating character behaviors and the application game system, using an immersive interactive game, the Versus Beat'em-up game, and an interactive storytelling game as examples.

Keywords

Intelligent character • Entertainment game

H. Mori (✉)
Graduate school of Engineering, Utsunomiya University, Utsunomiya-City, Tochigi, Japan
e-mail: hmori@is.utsunomiya-u.ac.jp

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_48

Introduction

Digital entertainment games provide many rich experiences. They can be used for education and training using immersive realities in video games as well as for entertainment. In many cases, the digital entertainment games use human characters as one of the essential elements for adding a sense of reality to the represented virtual reality world. The characters require intelligence to give the virtual world a sense of reality or entertainment as a game. For example, they can be lifelike or act in ways that will increase the user's enjoyment. The "intelligence" in this chapter not only describes character intelligence (AI) but also the intelligence that users are given when they see the character. In order to generate a character that seems and acts like a real human, several approaches have been proposed.

This chapter describes technologies for generating behavior for intelligent characters and the application game system, using as examples an immersive interactive game, the Versus Beat'em-up game, and an interactive storytelling game.

The Technologies for Interactive Conversational Character

Conversational characters are used in first-person viewpoint games or games where the viewpoint is seen through the user's avatar. These characters are required not only to respond properly to the user but also to exhibit natural-looking behavior when responding to the environment, the situation, and the users' actions.

This section describes the technologies used for interactive conversational characters and gives examples of applications in systems using this type of character.

Generating Conversational Eye and Head Movement

When a user talks to a life-size 3D character in an immersive environment (Laster et al. 1997; Isbister et al. 1997; Cassell et al. 1999), head and eye movements are key components in demonstrating a character's intelligence.

Several approaches have been tried to deal with such a task. Some examples are the modeling of eye movements on the actual measurement of eyes captured by motion sensors (Lee et al. 2002) and the generation of head gestures based on voices in conversations (Tsukahara and Nigel 1997).

It is difficult to generate movement requiring the coordination of the head and eyes, such as the shifting of the view line to an object indicated by the conversation.

It has been pointed out that nonverbal representations of the conversation between participants include various dependences, such as when the movement that is represented is affected by the state of the conversation (Argyle and Dean 1965).

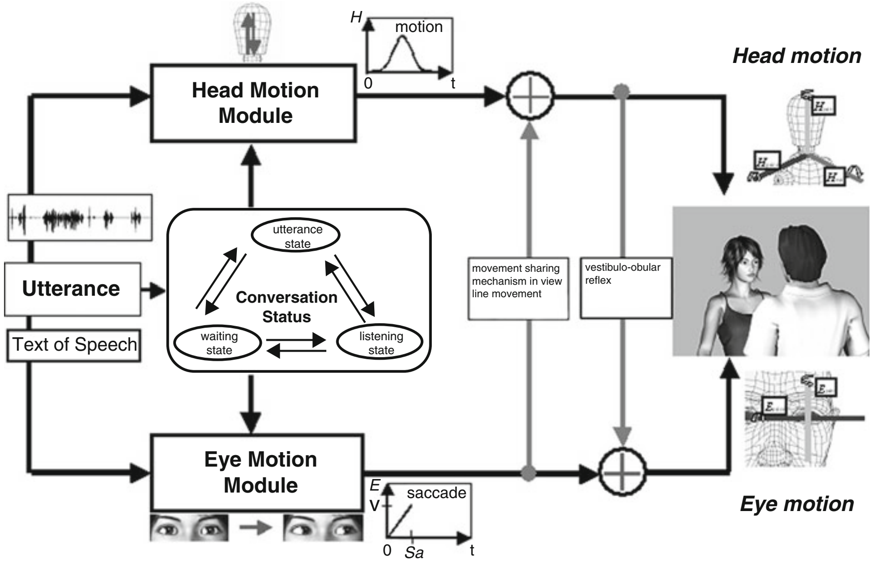


Fig. 1 Synchronized head-eye movement method

It is also desirable in conversations involving characters that the generated movement be adjusted according to the state of the conversation.

Masuko et al. proposed a method that generates composite head and eye movements of CG characters synchronized with the conversation in a simple way (Masuko and Hoshino 2006, 2007).

Their approach is not to handle the eye and head movements separately, as is done in conventional methods, but to consider a more complex animation generation, focusing on the coordinating mechanism of the eyes and head in view line movement. In the coordinating mechanism of the eyes and head, the vestibulo-ocular reflex and the mechanism of the sharing mechanism in view line movement between the head module (generating the head movement) and the eye module (generating the eye movement) are considered. In other words, the head and eye movements are generated as a composite mechanism (Fig. 1).

In an actual conversation, different movements are generated depending on the state of a conversation. For example, a listener might nod following the utterance of a counterpart. Consequently, three states are considered, depending on the utterance of the virtual actor. These are the utterance state, the listening state, and the waiting state without conversation. The conversation state is shared by the modules. This configuration makes it possible to generate head and eye movements in synchronization with the conversation state. Thus, the animation between the characters can be generated in a simple way in synchronization with the conversation (Fig. 2).

Fig. 2 Result of multiparty conversation

Generating for Conversational Gesture and Posture

Suitable expression of a character's gestures and posture is one of the important components in making a user feel that the character has some intelligence when interacting with it.

Characters perform various conversational motions that exhibit individuality. In most cases, however, these motions and their timing are built in by the creators. If highly expressive conversational motions similar to those built in by the author could be generated automatically, the technique would be applicable to a variety of content.

Cassell and colleagues proposed a method of generating gestures automatically from sentences in a conversation in which particular words and relationships between words are extracted from the input text and used to assign the character's gestures (Cassell et al. 1994, 2001). Kopp and Wachsmuth proposed a gesture generation method emphasizing the timing of gestures (Kopp and Wachsmuth 2000a, 2000b, 2001, 2002).

Nakano and colleagues proposed the layered behavior synthesis technique for integrating mental behaviors and conversational gestures (Nakano et al. 2005, 2006b; Nakano and Hoshino 2007). Figure 3 shows the concept for the behavior synthesis technique based on the mental status of the character. In this method, the psychological components of conversational gestures are added by combining poses, conversational gestures, and unconscious movements. The motion unit network controls the occurrence probability and continuity of motion units. Each motion units is indexed by mental parameters (Fig. 4). When utterances and mental parameters are given, evaluation values are calculated to select the most proper combination of postures, gestures, and unconscious movements across time. During these evaluations, the optimal behavior for the current state is selected from the

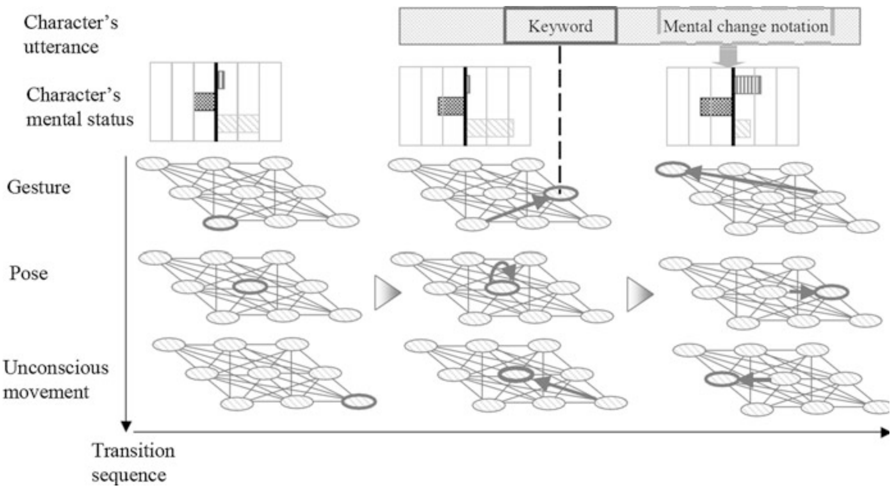


Fig. 3 Layered behavior network architecture

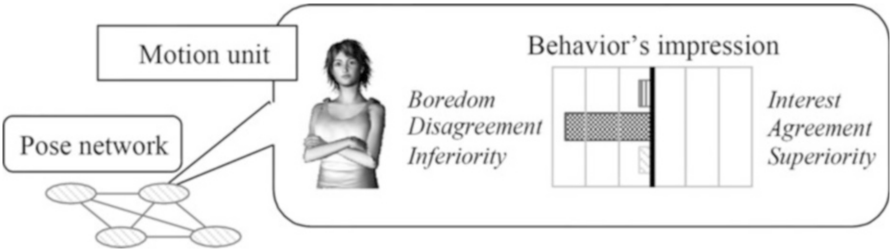


Fig. 4 Motion Unit. Motion data is associated with mental impression values

network of each attribute. The selected behaviors are accumulated hierarchically to generate the final behavior.

Figure 5 shows an example of the game. Figure 5a, b shows the differences between conversational gestures and composite mental behaviors.

Generating the Reactive Attention for Voice Communication Games

Characters' suitable reactive attention is a key component in indicating a character's intelligence to a user.

In human behavior, it has been pointed out that the contemplation and reflection processes (Damasio 1994) are necessary to establish smooth communication.

In the contemplation process, input from the outside affects emotions. Behavior is determined by directing the selective attention to the input and goals for the strongest emotions.

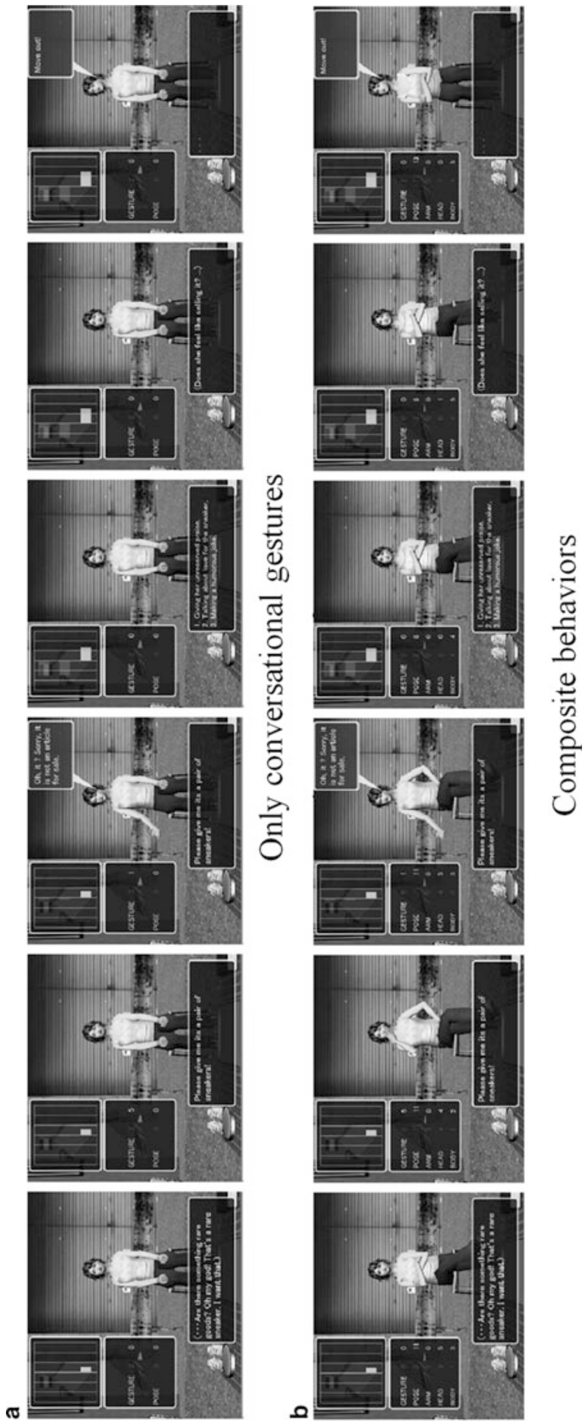


Fig. 5 Comparison of conversational gestures and composite behaviors

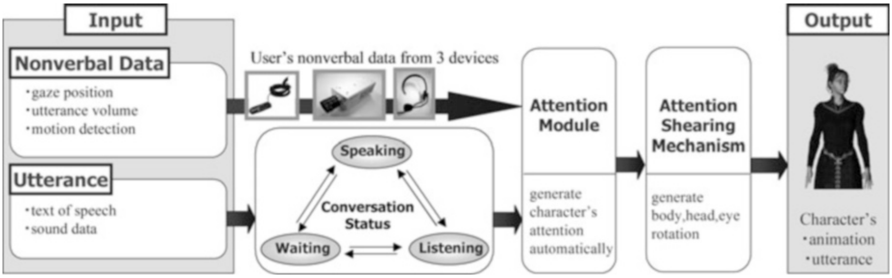


Fig. 6 Overview of generating reactive attention

The reflection process is one that reacts unconsciously against the contemplation process and cannot handle stimulus.

In the conventional method, the focus is on the contemplation processes; there are many approaches that generate a response to the contents of each user's speech.

However, since the reflective process is not sufficiently taken into account, characters' reaction time is often felt to be unnaturally slow.

The reason for this is that there is an operation delay while the speech recognition process is occurring.

Thus, in order to prevent this delay and in order to achieve a more natural realistic sense of interaction, the generation technique of the reaction attention of the character that reflects the nonverbal information of the user and the conversation state has been proposed (Hoshino and Mori 2011).

This approach is able to generate reflexive attention behaviors and attention operations linked to the conversation states using the quickly detected three pieces of nonverbal information: the user's gaze position, the user's operating body parts, and the user's speech volume. The conversation states in parallel with this take a long time to detect language information.

Figure 6 shows the overview of the method. The attention modules generate the attention according to the state of the conversation and the reactive attention according to the three pieces of nonverbal information from the devices.

Figure 7 shows the result of generating the reactive attention in voice communication games. It can be confirmed that the character reacts immediately with the user's perceived utterance. And the results of evaluating experiments show that this approach is effective in improving the way in which interaction is felt.

Locomotion Control for Immersive Conversational Space

Composing locomotion and gestures is also important for applications such as entertainment games. In the interactive story environment, characters do not simply stand in one position. Gestures and locomotion should be determined based on story locations and surrounding objects.

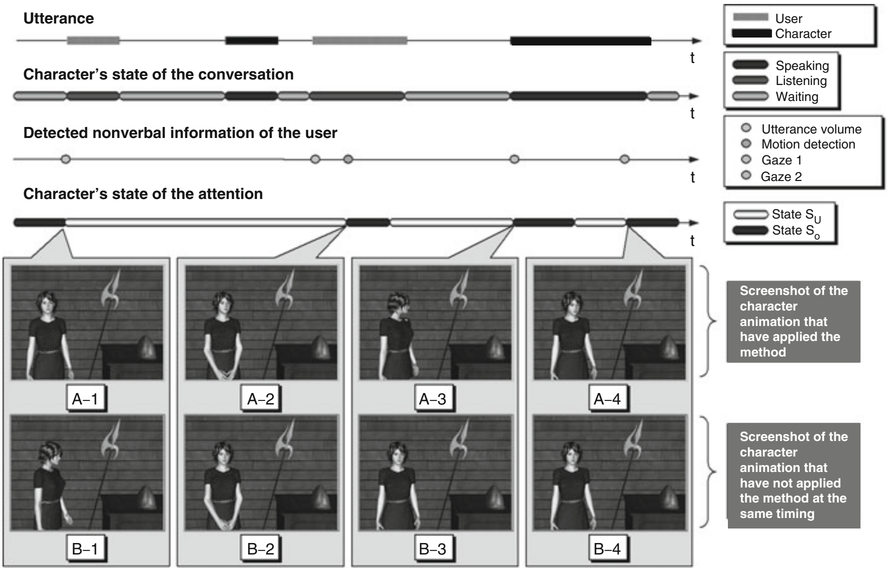


Fig. 7 Result of reactive attention in voice communication games

The proper location and timing of the character's movements is influenced by various elements, such as the connection between scene locations and the current environment. The apparent size of objects and the detail of explanations affect how much closer the character should move. Connection between scene locations also affects the character's current position.

When the character refers to particular objects during a conversation and the objects are far from the character's current position, it is time-consuming to make him or her approach the object every time. However, when the referenced object is close to the next scene location, it is more reasonable that the character should move closer to that object.

To deal with this problem, a mechanism for fluid conversational locomotion for virtual characters is proposed (Chan 2007). This is done by calculating the optimal locomotion path, which is influenced by the conversation and the story location, and then subsequently having the character generate composite walking and conversation actions. The character also locally adjusts his or her position so as not to limit the referenced object from the user's sight. Figure 8 shows a typical example of conversational locomotion. In this scenario, the character first explains that the vending machine cannot be approached in the disaster zone. In the next scene, the user asks about the other specific objects that cannot be approached and the character then moves closer to explain more about the background relating to the disaster.

Figure 9 shows the conversational locomotion architecture. The system has a locomotion module, conversation modules, and a story manager. A story consists of a set of scene units that control the discourse of the conversation. A scene unit has a precondition, scene location, and links to a collection of possible conversation modules.

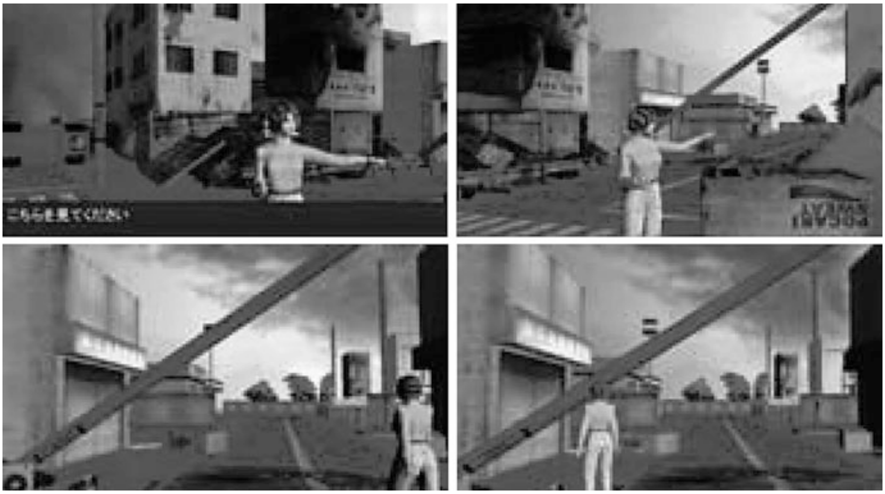


Fig. 8 Example of conversational locomotion

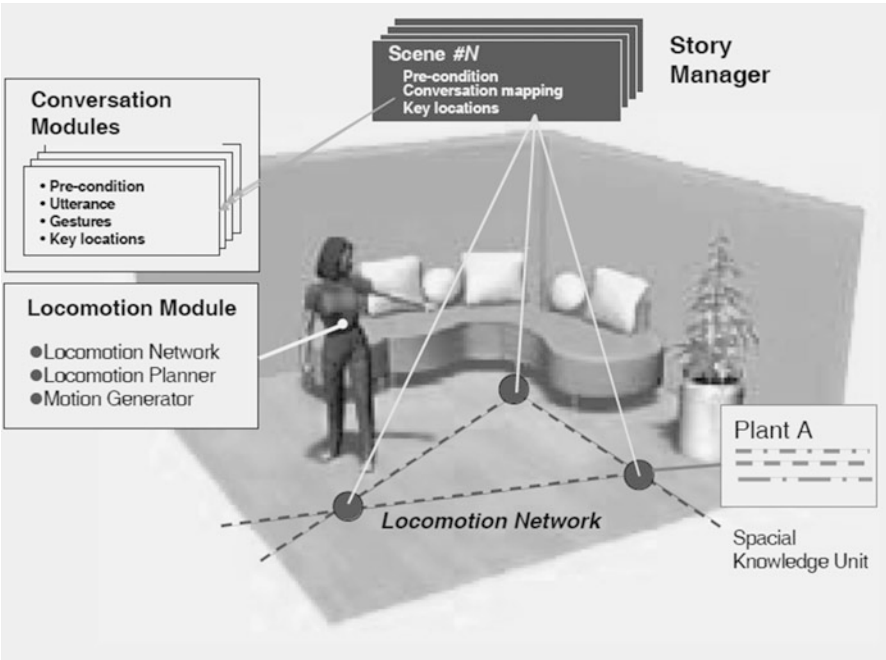


Fig. 9 Overview of the conversational locomotion architecture

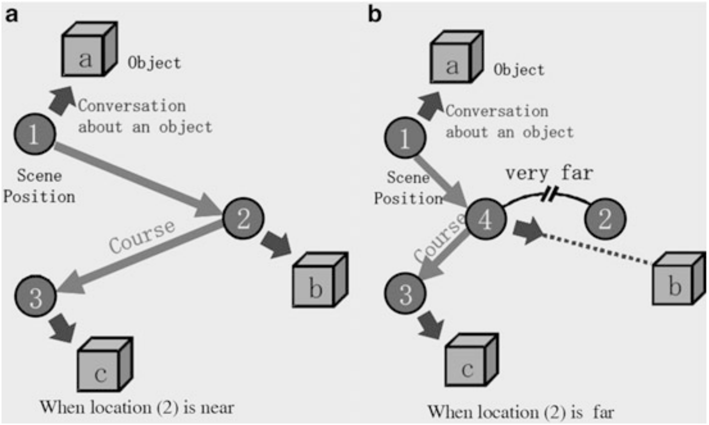


Fig. 10 Concept of conversational locomotion using a simple example

Figure 10 shows a typical example of locomotion planning during a conversation. Assuming that the character should move from node 1 to node 3, it is reasonable that he or she should stop at node 4 if the referenced object is visible enough.

Example of the Immersive Conversational System

Figure 11 shows a snapshot of the second language learning game system (Hoshino et al. 2009). The interface of the system consists of a large screen, a camera for user behavior analysis, and a voice recognition system. The game character recognizes users' voices and returns conversation sentences based on the negotiation of the meaning model. The learner experiences task-based language training in various daily situations such as shopping, cooking, and working. The task-based learning method enables learners to obtain communication skills while practicing particular missions using voice and gesture communications with life-size 3D game character.

Figure 12 shows the digital signage system using a character to attract the attention of a passerby, the shop character system (Mori and Hoshino 2011). The target user of this system is a customer interested in products in a specific area, who may not intend to make a purchase. The character, presented on a large display, gives an explanation according to mutual distances with the user to improve the user's attention and concern.

The Technologies for Versus Beat'em-up Game

In the entertainment game called "The Versus Beat'em-up," a player is able to play against a human player or the computer. This section describes intelligent character technologies for this game.

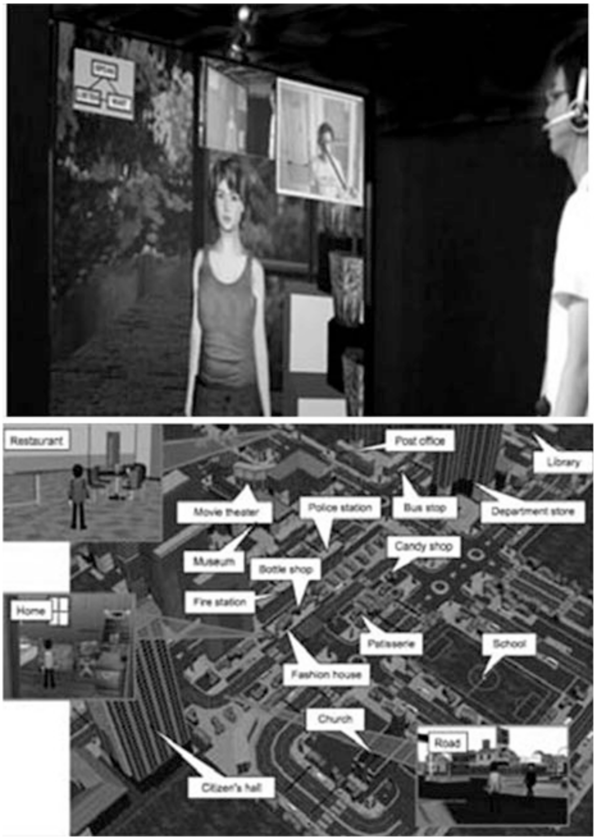


Fig. 11 Second language learning game

Fig. 12 Overview of the shop character system

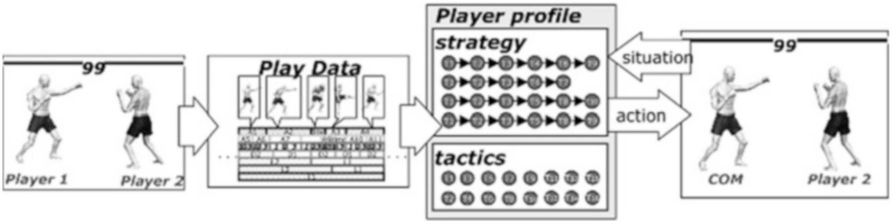


Fig. 13 Overview of imitating human player behavior in action games

NPC Imitating Player-Controlled Character

One of the limitations of computer opponents in action games is that the computer-generated characters are constructed in advance. This means that players may become bored quickly. A human player can play the same game repeatedly, learn the behavior of the computer-controlled game character, and win easily. This is one of the main reasons why human players soon grow tired of "fighting with a computer." Thus, many players prefer playing with another human instead.

Consequently, Hoshino and colleagues proposed a system that enables a computer character to imitate a human player (Nakano et al. 2006a, Hoshino and Mori 2009).

As shown in Figure 13, the proposed system records the moment-to-moment actions that a game player takes and the timing of interactions between game players as the play data. From this data, the system calculates correlations between situations and tactics and the actions taken by the player. In addition, the system acquires players' strategies from the play data, which are represented by a sequence of tactics. The acquired tactics and tactic sequences are stored as the player profile of the particular player to be imitated. Then, the system creates actions imitating the player based on his or her player profile. The computer evaluates the distance between the situation in the game and the strategy selection criteria in order to create the behavior imitating the player by choosing and carrying out the strategy in the player profile that the imitated player would be most likely to take.

Virtual Adversary Player by Generating Emotional Utterance

The approaches are proposed not only to imitate a human controlled-character (as described in section "NPC Imitating Player-Controlled Character") but also to imitate an adversary human player.

Shiratori et al. proposed a virtual fighting game player who would speak on various situations relating to a game's progress (Shiratori et al. 2008). The method of capturing a game status from a Life-point-gauge and skill-taking-patterns is read from special pixels on a game video's graphic frame. Emotional parameters are calculated from this value. Utterance sound is captured and classified at the same time. After those preparations, the game is played with this system. Appropriate

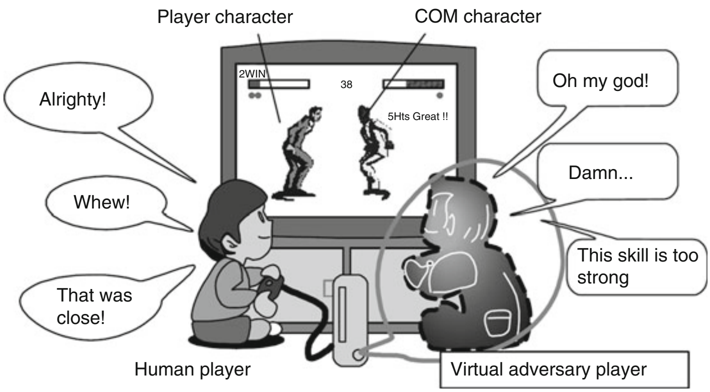


Fig. 14 Overview of the virtual adversary player

voices made from real speech from the player are selected depending on the situation. Based on information from a user's evaluation, this method adds to game enjoyment (Fig. 14).

The Technologies for Interactive Storytelling Game

Many interactive storytelling applications have relied on progress in animation technology to create characters that act autonomously.

To make these characters seem lifelike, it is important for them to perform massive actions, such as daily actions like greeting someone, taking a walk, or going to work; reflex actions, which require reacting to input from a user; perceiving actions, where the character perceives an object and reacts to it; active actions; and actions based on personalities or feelings.

This raises the problem of action planning. A character has to behave like a human being in the real world – carrying out the actions listed, keeping schedules and maintaining a personality, and reacting flexibly to user interaction, while still maintaining story flow.

For example, "The SIMS 2" produced by Electronic Arts (EA) and "Kenran Butoh Sai" produced by Sony Computer Entertainment Inc. (SCEI) are examples of this type of game.

The characters have parameters relating to their memory or feelings and show anger or laughter as per the player's selections.

These games also reflect the memory that is demonstrated through strong feelings in conversations that occur later. This process is humanlike.

The characters are controlled through the story arc using hierarchical task networks (HTN), i.e., tree of hierarchical task (Cavazza et al. 2001). A façade is proposed as an interactive drama in which the game progresses through natural-language conversation as a more reformative work (Mateas and Stern 2003).

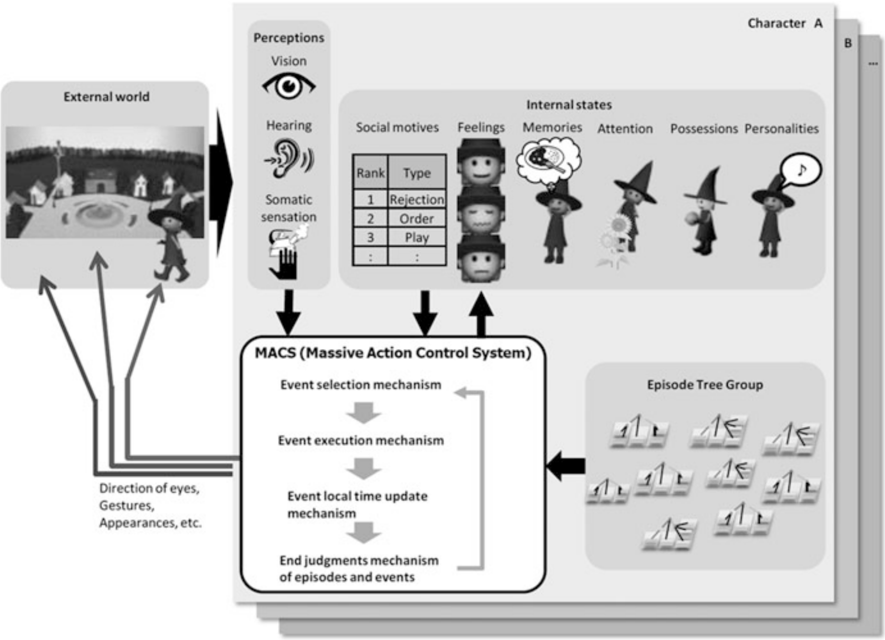


Fig. 15 Overview of the massive action control system (MACS)

Hamana et al. propose the massive action control system (MACS) (Hamana et al. 2008; Mori et al. 2010) which can execute massive actions in multiple characters (Fig. 15).

This system continuously selects an appropriate fragmentary behavior control module called an episode tree based on the character's inner states, such as motives, feelings, and personality, and the state of the external world, such as characters and objects surrounding the character. Thus, the character can continue with his or her actions and interference time is freed.

Figure 16 shows an application (Spilant World) that is used to verify the system. Multiple characters have motives behind their actions, feelings, and personalities and live daily in this application. When a user adds a new object to the virtual world or touches an object in the virtual world, the characters recognize the object or action and autonomously start new actions. This, in turn, affects the action selection of other characters, creating an opportunity for those characters to change their actions. The user can experience the story not only by tying the user's actions to the changes in the characters' actions but also by having the effect spread over the long term.

Conclusion

This chapter describes the intelligent character technologies for entertainment games. An immersive interactive game, the Versus Beat'em-up game, and an interactive storytelling game are used as examples.

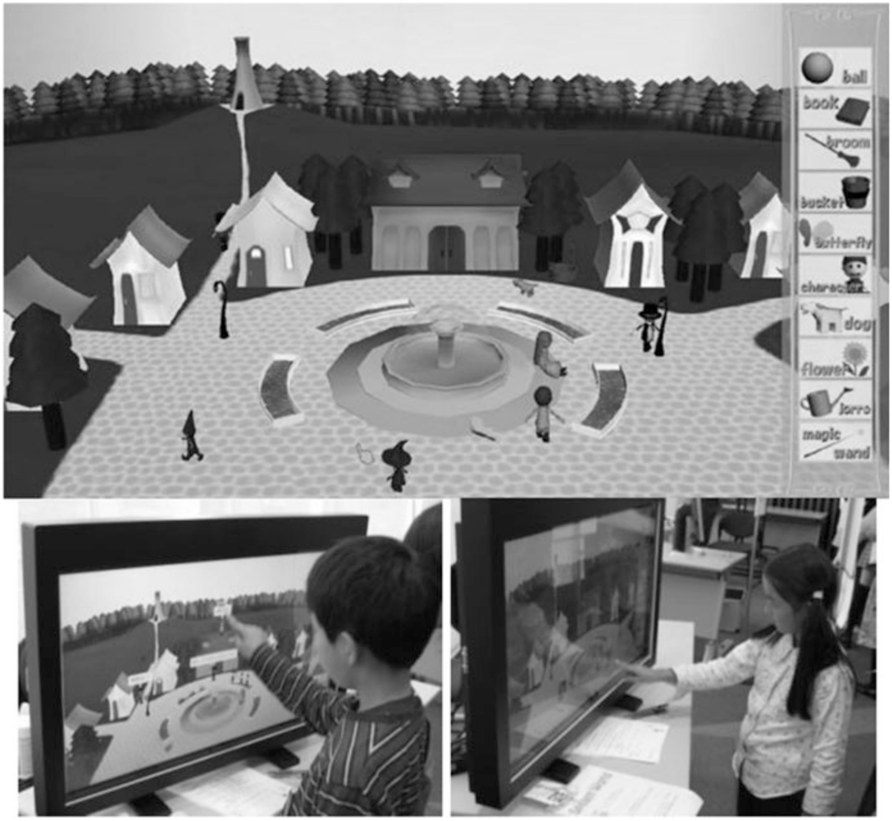


Fig. 16 Spilant World system

The majority of entertainment games have human characters to add a sense of reality to the represented virtual reality world.

The characters require a kind of intelligence to give the virtual world a sense of reality or entertainment as a game.

The high intelligent behaviors of the human characters are required to measure up to a specific level of quality of graphics in tandem with the advances in the technologies of graphics. Additionally, AI technologies have also steadily advanced with the progress of the performance of computers. For example, recently, some advanced voice assistants have been implemented through the introduction of AI using machine learning techniques. Therefore, the creation of intelligent characters that integrate a high-definition graphics and high-performance AI is expected. I want to pay attention to the advancement of intelligent character technology in the future.

Recommended Reading

M. Argyle, J. Dean, Eye-contact, distance and affiliation. Sociometry **28**, 289–304 (1965)

J. Cassell, C. Pelachaud, N.I. Badler, M. Steedman, B. Achorn, T. Becket, B. Douville, S. Prevost, M. Stone, Animated conversation: Rule-based generation of facial expression, gesture and spoken intonation for multiple conversational agents, in *Proceedings of the SIGGRAPH'94*, 1994, pp. 413–420

J. Cassell, T. Vickmore, M. Billinghurst, L. Campbell, K. Chang, H. Vilhjalmsson, H. Ya, Embodiment in conversational interfaces, in *Rea, CHI-99*, 1999, pp. 520–527

J. Cassell, H. Vilhjálmsson, T. Bickmore, BEAT: The behavior expression animation toolkit, in *Proceedings of the SIGGRAPH'01*, Los Angeles, 2001, pp. 477–486

M. Cavazza, F. Charles, S.J. Mead, AI-based animation for interactive storytelling, in *Proceedings of Computer Animation*. IEEE Computer Society Press, 2001, pp. 113–120

R. Chan, Building immersive conversation environment using locomotive interactive character. Computer **13**(2), 149–160 (2007)

A.R. Damasio, *Descartes' Error* (Quill, New York, 1994)

K. Hamana, A. Nakano, J. Hoshino, Massive action control system. in *SIGGRAPH ASIA 2008 Emerging Technologies*, 2008, p. 46

J. Hoshino, H. Mori, Incremental learning algorithm for online action game system, in Entertainment computing – ICEC 2009. Lecture notes in computer science, vol. 5709 (Berlin/Heidelberg, 2009), pp. 319–322

J. Hoshino, H. Mori, Generating reactive attention of CG character for voice communication games. J. Soc. Art Sci. **9**(1), 20–28 (2011) (in Japanese)

J. Hoshino, T. Saito, K. Shiratori, Task-based second language learning game system, in *Ifip International Federation For Information Processing*, 2009, pp. 323–324

K. Isbister, B. Hayes-Roth, Social implications of using synthetic characters, in *Workshop on Animated Interface Agents in IJCAI-97*, 1997, pp. 19–20

S. Kopp, I. Wachsmuth, A knowledge-based approach for lifelike gesture animation, in *ECAI 2000 – Proceedings of the 14th European Conference on Artificial Intelligence* (IOS Press, 2000a), pp. 663–667

S. Kopp, I. Wachsmuth, Planning and motion control in lifelike gesture: A refined approach, in *Postproceedings of Computer Animation 2000* (IEEE Computer Society Press, 2000b), pp. 92–97

S. Kopp, I. Wachsmuth, Model-based animation of coverbal gesture, in *Proceedings of the Computer Animation 2002* (IEEE Press, 2002), pp. 252–257

J. Laster, S. Barlow, S. Converse, B. Stone, S. Kahler, R. Bhogal, The persona effect: Affective impact of animated pedagogical agents, in *CHI-97*, 1997, pp. 359–366

S.P. Lee, J.B. Badler, N.I. Badler, Eyes alive. ACM Trans. Graphics **21**, 637–644 (2002)

M. Mateas, A. Stern, Façade: An experiment in building a fully-realized interactive drama, in *Game Developer's Conference: Game Design Track*

S. Masuko, J. Hoshino, Generating head-eye movement for virtual actor. Sys. Comput. Jpn. **37** (12), 33–44 (2006)

S. Masuko, J. Hoshino, Head-eye animation corresponding to a conversation for CG characters. Comput. Graph. Forum. **26**(3), 303–312 (2007)

H. Mori, J. Hoshino, The digital signage system using virtual human for getting attention of passersby. J. Info. Process. **52**(4), 1453–1464 (2011) (in Japanese)

H. Mori, K. Hamana, C. Feng, J. Hoshino, Narrative entertainment system with tabletop interface, in Entertainment computing – ICEC 2010. Lecture notes in computer science, vol. 6243 (Berlin/Heidelberg, 2010), pp. 422–424

A. Nakano, J. Hoshino, Composite conversation gesture synthesis using layered planning. Sys. Comput. Jpn. **38**(10), 58–68 (2007)

A. Nakano, K. Shioiri, J. Hoshino, Composite behavior synthesis technique for mental communication games, in *ACM SIGGRAPH 2005 Sketches and Applications*, 2005

A. Nakano, A. Tanaka, J. Hoshino, Imitating the behavior of human players in action games. in *International Conference on Entertainment Computing 2006*. Lecture notes in computer science, vol. 4161 (2006a), pp. 332–335

A. Nakano, K. Shioiri, J. Hoshino, Synthesizing pose, unconscious movement, and gesture for mental behavior expression of interactive characters, in *ACM ACE*, 2006b

K. Shiratori, K. Shioiri, J. Hoshino, Virtual adversary player by generating emotional utterance. J. Soc. Art Sci. **7**(2), 65–74 (2008) (in Japanese)

W. Tsukahara, W. Nigel, Nod as conversational phenomenon without understanding. Lang. Month. **26**, 90–97 (1997)

I. Wachsmuth, S. Kopp, Lifelike gesture synthesis and timing for conversational agents. Gesture and sign language in human–computer interaction, in *International Gesture Workshop (GW 2001)*, Revised Papers (2001). Lecture notes in artificial intelligence, vol. 2298, pp. 120–133

Real-World Game Platform for Lifelong Learning Society

14

Junichi Hoshino

Contents

Abstract

In a flexible society where individual diversification and individualization are advancing, due to the change of roles in the society and job content, areas that require individual judgment are increasing in daily life. In this chapter, I would like to overview the real-world game platform that can conduct systematic selection supports of various learning services provided in the area, as well as development, provision, and evaluation of new services which use game technologies, in an integrated manner, for the lifelong learning society.

Keywords

Real-world game • Lifelong society • Production support system • Learning games

J. Hoshino (✉)
Graduate School of Systems and Information Engineering, Entertainment Computing Laboratory, University of Tsukuba, Tsukuba-shi, Ibaraki, Japan
e-mail: jhoshino@esys.tsukuba.ac.jp

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_50

Introduction

In a rigid society, where individuals' roles are decided, the ways of problem solving that are necessary in daily life, the width of decision-making, and the human relations are limited because they require certain labor and are limited to specific fields of activities. However, in a flexible society where individual diversification and individualization are advancing, due to the change of roles in the society and job content, areas that require individual judgment are increasing in daily life. This tendency has become more remarkable due to the spread of networks and IT technology; at the same time, learning methods to enhance individual ability, such as learning new specialized knowledge and skills, have been increased.

The learning that can be required in daily environment follows:

- Obtaining basic knowledge for problem solving in daily life
- Obtaining specialized knowledge and selection of occupations that suit the purpose and characteristics of individuals
- Decision-making that takes conditions from various viewpoints into consideration
- Social skills such as building human capitals
- Understanding of different cultural backgrounds
- Understanding of the system of the society or nature, such as protection of natural environment, process of food distribution, and safety

Conventional school curriculum is not enough to cover these wide studies, so it is necessary to make use of various daily opportunities.

In this chapter, I would like to overview the real-world game platform that can conduct systematic selection supports of various learning services provided in the area, as well as development, provision, and evaluation of new services which use game technologies, in an integrated manner, for the lifelong learning society.

Continuous Development of Learning Services

In studies of learning support, many efforts have been conducted, such as informal learning which utilizes various experiences in daily environment comprehensively (Marsick et al. 1990; Foley 1999; McGivney 1999), gamification which is an edutainment accommodating entertainment property, games for change and so on. Analyses and evaluations of users' actions by game analytics (El-Nasr et al. 2013) also have been valued. In recent years, utilizing the network environment, the use of MOOCs (massive open online courses) (Daniel 2012; Yuan 2013) that shares several thousands of various online educational materials has also progressed.

These efforts are effective for purposes and users in each service; however, there are so many various services, and it is not easy for users to choose the proper service. For the system to support lifelong learning, the following frames are required:

1. It is necessary to be able to choose and build a service that suits individual purposes and characteristics from various learning services.
2. In the service development in these years, various specialties and skills are required.

In order to do that, not only individual service provisions by entities such as colleges, enterprises, research institutions, and specialists but also continuous frameworks to develop new services by collaborating, combining them to human resources cultivation, are necessary.

By introducing an information sharing system shown in Fig. 1, the selection of service that suits an individual from various services by the subject of service provision and feedback usage conditions becomes easy.

Possibly, this will enable the cocreation of values as follows:

- Users over many generations
  Choosing a service that suits one's purpose, preference, and characteristics for understanding from various courses. Also the developers can convey users' needs and issues.
- Specialists
  By recognition of specialized characteristics, matching and communication in service development become easier. Also, accumulation of production knowledge of learning services and sharing technologies developed by researchers enable the production of high-quality learning services.
- Enterprises
  Development of advanced services collaborated with specialists, production, and provision of suitable services for users and grasping usage conditions become easier.
- Administrative agencies
  They can consider the activation measures by being informed about the usage conditions of learning courses and the situations of service providers.
- Public institutions
  They can easily make public about their providing services to interested users and also can easily know the related services.

For the methods to develop the services that suit users' needs, there are several ways such as lean production which is the representative method, just-in-time system (in Toyota, Agile software development method that develops promptly in short cycles depends on user's needs), DevOps that combines software development and operation, and Lean Startup method that combines those.

In Fig. 1, process of the cocreation of service values can be described as follows: (a) analyses of social environment that analyzes the association between social environment and stakeholders, (b) service design, (c) system development, (d) service provision, and (e) learning needs from users by big data analysis.

Utilizing the systems that are actually under operations enables the use of big data analysis that clarifies our lives, human relations, the change of lifelong

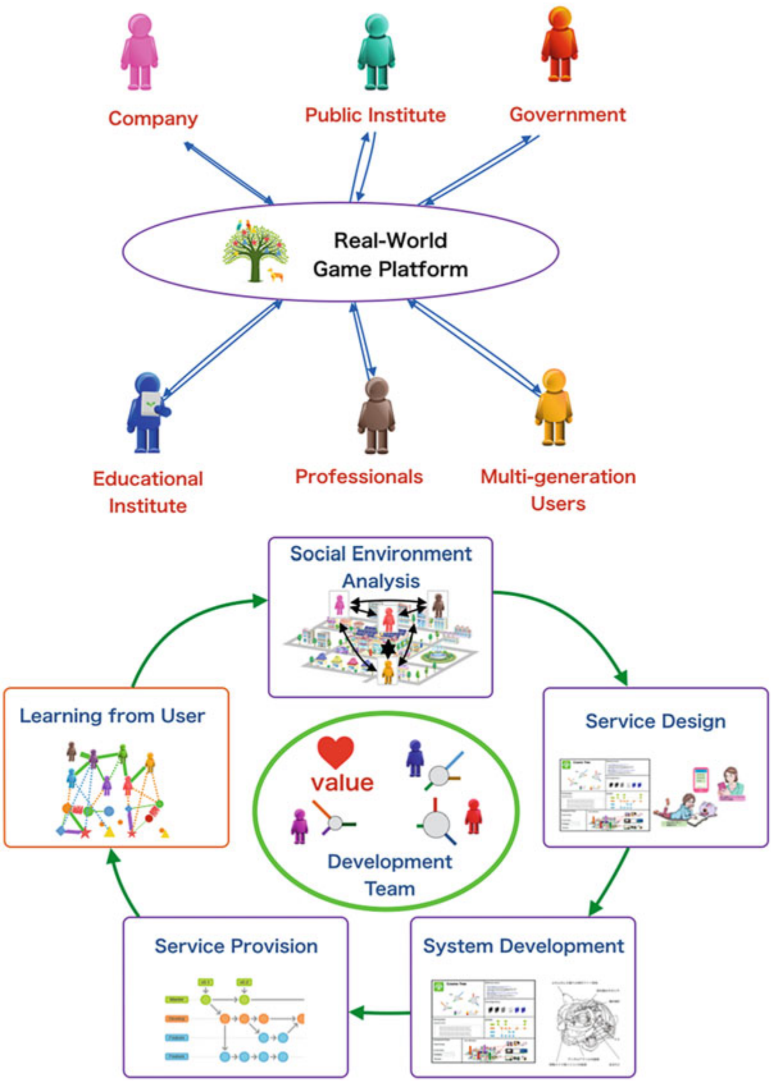


Fig. 1 Information sharing and continuous development process in lifelong learning society. Multi-generation users, companies, education institutes, and administrative agencies are collaborating to create new services

learning, impacts on industries, and social changes. For example, using the accumulation of usage data of the system, the developers can conduct pattern mining of specified user behaviors and services, which allows development and provision of services focusing on user group with certain patterns by service provider or friends finding support with the same preference. Utilizing the information academically, considering privacy protection as well, there is a possibility of open science by big data analysis that clarifies our lives, human relations, the change of lifelong learning, impacts on industries, and social changes and by opening linked data (Bizer 2009) to the public.

Real-World Game Platform

In the previous chapter, I reviewed the possibility of integrated provision of various lifelong learning services and continuous development of new services by introducing information sharing system. In this chapter, I would like to mention the realization methods using game technology as concrete approaches of these systems. In this system, it provides geographic space information and the basic functions to express and manage structural elements in the city such as local residents, specialists, enterprises, and government at various abstract levels. With this, the developers can consolidate various learning service information and provide it to users and also realize learning games that reflect local information.

Real-World Game

It is not easy to produce the game that uses modeling of actual world since it requires description of structural elements of people and society and also composes dynamic behaviors. Also, 3D graphics with high definition can effectively depend on the usage; however, in case the purpose is learning, it will be used on smartphone, and activities in an actual environment are important as well; thus the setting at proper abstract level is essential.

In this chapter, the author calls the game by abstract modeling of structural elements, such as people in actual world, specialists, enterprises, government, and geographic space information, as real-world game (RWG).

For example, a business game occurring at the actual urban space or a history learning game with reproduced historic city can be considered.

In real-world game, the following basic elements are important:

1. Multilevel abstraction of the real world
   In reality game, an abstract level along the purpose, such as modeling of geography and buildings in an actual environment, including 3D, setting a high value on specific knowledge like plants learning games, and dealing with historic scene, is necessary.
2. Realization of the game world activities
   When the developers create learning game, various learnings in the world reflecting the knowledge in the actual environment and also the mechanism to bring what is experienced in the game world into the actual world are required. For example, games that give points for performance in the actual world, experiencing the game with GPS in the actual environment, are included.

Fig. 2 Concept example of real-world game

For realization examples of RWG, modeling of urban environment and provision of various learning services such as history, geography, and economy. For example, in a historic urban space by 3D graphics and autonomous game characters, explain "structure of urban space," "structure of house," "historic events," "recycle," "various occupation," "public culture," etc. from various characters positions, experience of game event that is useful for understanding of associations, deep understanding linking with information in actual environment, etc. are considerable. Also it is possible to learn the second language in the daily environment situation and also experience a business game (Fig. 2).

Platform Architecture

I would like to mention a platform that provides basic function to produce RWG. This system is structured from development support system, user experience control system, and application framework, based on the geographic information, persona information, and information management system of service modules, as shown in Fig. 3.

Figure 4 shows samples of service development interface in service development support system. The provision support system that provides support function of analyses and visualization of specialized specifics, matching support, communication support, work process management, technology library group, usage condition, etc. suggests contents at optimized level guessing user's purposes and interests and allows following relations.

Figure 4 shows samples of service provision interface of user system. It supports the many versus many matching between users over many generations and service module group, research exploring using relations, and refined process of user models by communication with users.

To produce learning contents that utilized network technologies and digital medias, various technologies are required such as SNS, smartphone, tablet PC, Internet of Things (IoTs), machine learning, and recognition technology by mass data analyses.

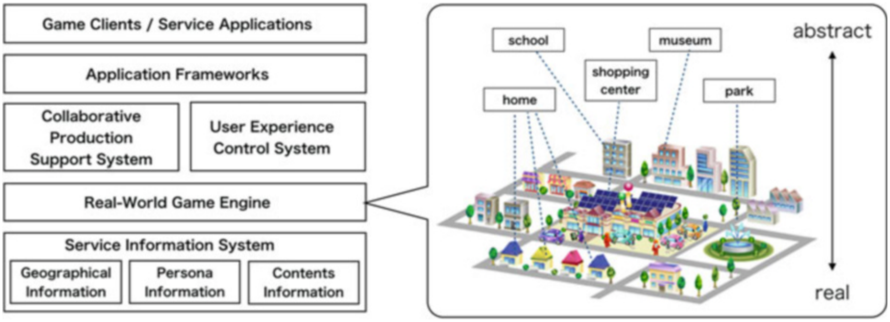


Fig. 3 System architecture of real-world game system

In this system, visualization of specialized specifics allows easier support by the system's ongoing production process, besides easier matching between service developers.

In this system, the developers can structure production tasks by structuring task unit of learning entertainment-producing process (problem analysis, design of concept, prototype development, analysis, and evaluation) with process analyses and schedule optimization. The task units are linked task unit with information that is necessary for individual task, design plan for the past, specification, report, paper, and library. The system conducts recommendation of proper task unit that depends on the production purpose and users' skill level and also task planning according to the deadline. Moreover, the developers can statistically analyze and visualize the usage condition of service modules and utilize it for service improvements and the next plan.

Service Examples

Task-Based Second-Language Learning in Virtual City

Task-based learning method enables learners to obtain communicative skills through the practice of particular "missions." As a result, the efficient improvement of communicative skill becomes achievable. In a 3D game environment, at home learners can practice their communicative skills in a simulated town, where they must converse with the local people, who in turn will help them accomplish missions.

Figure 5 shows a snapshot of using second-language learning game system. The user's environment of the system is a public classroom or a personal study room. The learners are assumed to be people who studied English in junior high school and high school. As a result, further study should build upon the skills one has learned, instead of wasting the knowledge acquired in school. The interface of the system consists of a large screen, a camera for the user behavior analysis, and a voice recognition

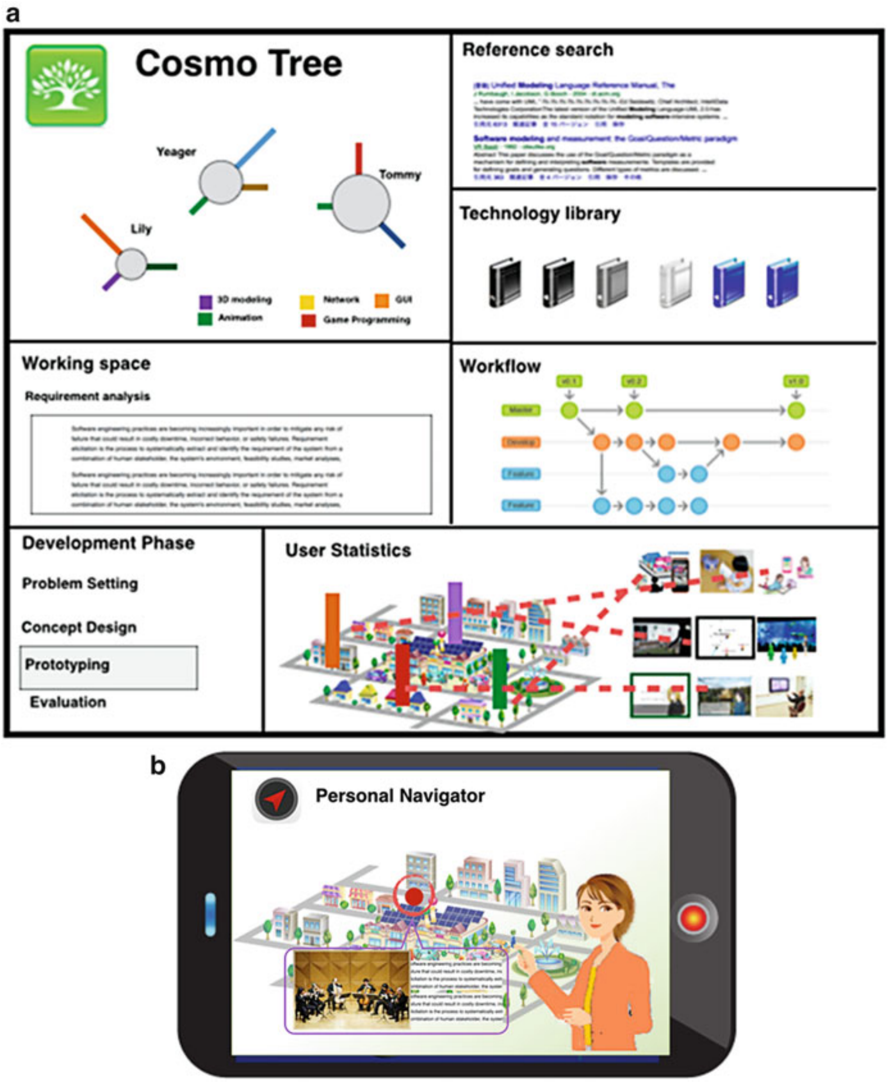


Fig. 4 Cooperative development support system and user experience navigator: (**a**) service development interface and (**b**) service selection interface

system. The game character recognizes user's voice and return conversation sentences based on the negotiation of meaning model.

The learner experiences task-based language training in various daily situations such as shopping, cooking, and job. The task-based learning method enables learners to obtain communicative skills through the practice of particular missions using voice and gesture communications with life-size 3D game character. Moreover, in this chapter, the focus is applied to the acquisition of necessary English knowledge and skill for use

Fig. 5 A snapshot of using second-language learning game system

in daily life. That is, the actions of the learner decide the outcome, which encourages learners to learn on their own and think for themselves in order to clear the task.

This system, using TBLT methodology, provides the negotiation of meaning function, and through its use, the imperfect understanding between the learner and the game character can be resolved. The learner advances through tasks by "talking" with English-speaking characters in a virtual city environment. The content of the conversation is displayed in a balloon above the heads of the characters, and the learner types their response.

Disaster Experience in Real World

Recognizing potential natural hazards and preparing for them are difficult, and people often fail to evacuate when warned about immanent disasters. Past research has suggested that residents' preparedness against natural disasters is not enough. The recognition of risks in a possible disaster situation is known as "risk perception." Appropriate decision-making and behavioral judgment are needed, but this is not common knowledge since disaster situations do not occur frequently. Public schools conduct evacuation drills to practice appropriate disaster response behaviors, but it is difficult to determine an escape route because participants are following instructions, not their instincts. Meanwhile, there are training lectures which teach effective decision-making skills using disaster maps and illustrations. In a manner this is an effective training method involving risk perception, but it has some feasibility problems to be a nationwide program because it requires many instructors with the necessary expertise and experience. It is also quite difficult to apply regionally specific risk perception into action.

In this section, a disaster experience game system which instructs the user about general knowledge and regionally specific disaster risks is described. The system integrates experiential learning and task-oriented learning approaches that are separately used in risk education. The system leads to recognize regionally specific risk and experience with a disaster with user's own two feet. Disaster drills are often messy and unorganized, but through frequent disaster drill performing, the authors were able to introduce a game concept that makes these drills feel a bit more fun.

There are various simulation tools for duplicating a disaster to obtain accurate information. This information can be used to minimize damage and help defend us from future disasters as "we cannot take appropriate measures in an unclear situation."

Meguro developed an evacuation simulation tool using the potential model. He enhanced calculated efficiency, simplified the algorithm, and designed easier simulations using a unified concept (e.g., space features, individual traits, and statues of disaster). The system projects evacuees in a VR space from various perspectives allowing users to get the whole picture and make clearer evaluations.

Kobayashi developed a "Tangible All Disaster Scenario and Simulator" which included a table UI system. The system takes into account various disaster prevention countermeasures that have been previously discussed in group situations.

Animations of the situation are displayed according to the disaster's class and extent, the time of the evacuation call, status of disaster, and refugee situation. Sakai proposed an MR technology system observing the results of simulations consisting of CG videos on a diorama. The diorama vividly transports viewers into the area. All these methods propose new visual-simulation systems and interfaces to enhance both risk perception and accuracy of predictions but do not consider behavior and decision-making.

NIED developed an AR technical application, "Disaster Risk Finder," which displays risk information of hazard maps, disaster records, etc., onto real footage using a smartphone camera. The system assists the user in recognizing surrounding disaster risks. Suezawa suggested a thinking type of muster drill, Komatta-Kun (Fig. 6).

The basic framework of disaster experience game (DEG) is to integrate experiential learning method in real-world and task-based learning methods that are usually separated in risk education. The game system uses actual user's current location such as home, office, school, and public transportations by GPS to select game events. The game events proceed in real time. By experiencing virtual natural hazards in user's daily living areas and solving problems, users can vividly image potential risks that would be helpful for future decision-making. Figure 1 shows the retargeting process of the real-world disaster situations to user's daily life environment. The author referred the reports of the Southern Hyogo Prefecture Earthquake from the Cabinet Office, Government of Japan, in this chapter. Figure 2 shows the game event categories.

According to the research of Smartphone Contents Report, 48.3 % of men and 42.5 % of women in their 20s have a smartphone in Japan. The author asked 11 people (men, 10; women, 1) who have a certain level of IT literacy to try playing the game. All five groups were able to succeed in reaching their final destination. They each walked an average of 1 km and experienced ten disaster events. Each group took between 30 and 40 min to complete the exercise (Figs. 7 and 8).

The author asked users to consider their risk perception before and after the experiment. The level of risk perception for all 11 users increased from their respective levels prior to the experiment. The results from the *t*-test showed a significant difference between before and after. This data shows that the use of our proposed system helps the user to gain better risk perception and learn how to react when confronted with disaster situations in the real world.

Next, the author looked at the user's statements related to risk perception in greater depth. Before the experience, there were similar answers: "to be buried under a collapsed building" or "to be damaged from fires and escape." It was thought that after the effects, this statement would become established as pragmatic information. But after the experience, you can see:

1. Parts of buildings could collapse; make sure to get as far away from them as possible and take notice of the area near the roof.
2. It is important to know first aid procedures. If you found someone collapsed, you should determine whether the person is conscious and the degree of the presented injuries.

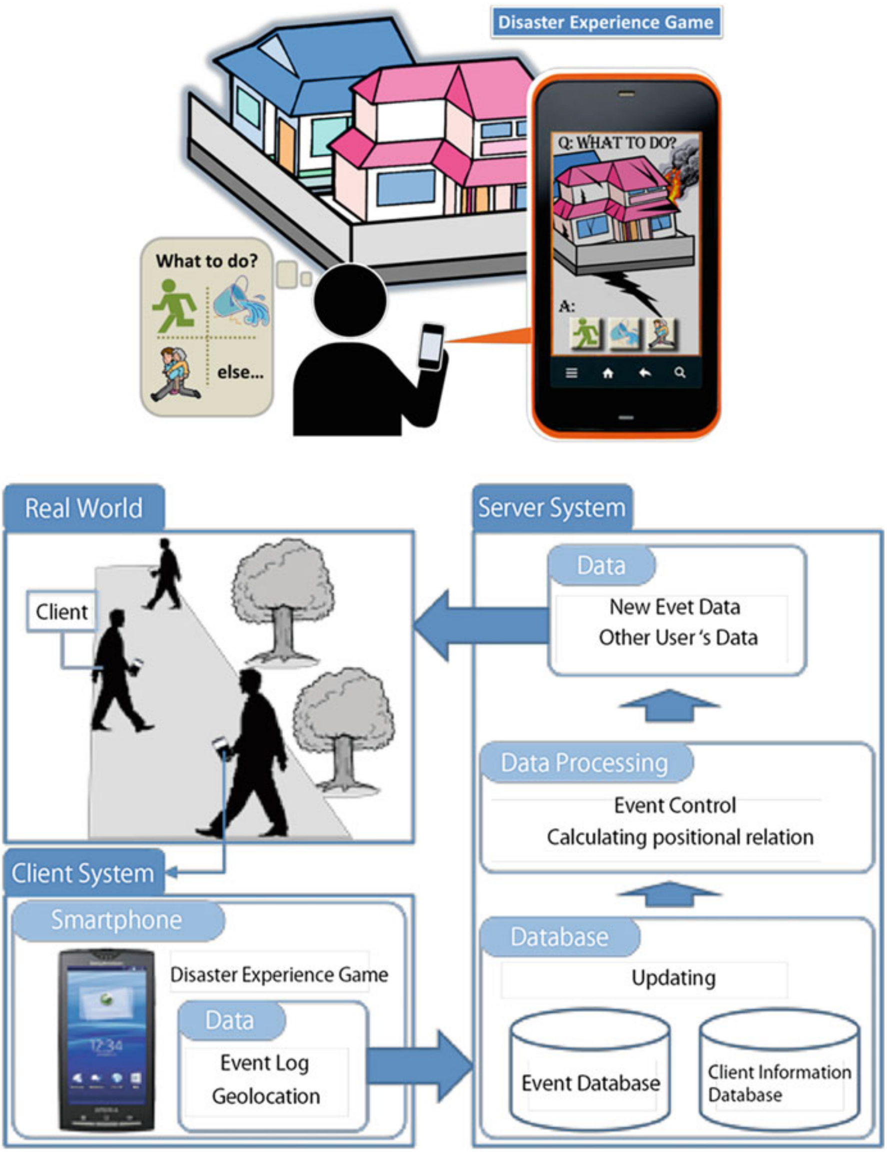


Fig. 6 The concept of DEG and the retargeting process of real-world disaster situations to user's daily life environment

3. If you find a damaged road, be careful and do not rush.
4. If you are caught in a crowd, your hands and feet may be bound. So stay away if possible. If you are caught in a stream, do not swim against the tide.
5. To avoid being isolated, accurately determine the location of a nearby evacuation shelter.

Fig. 7 Game event categories. The map shows the experimental fields of University of Tsukuba

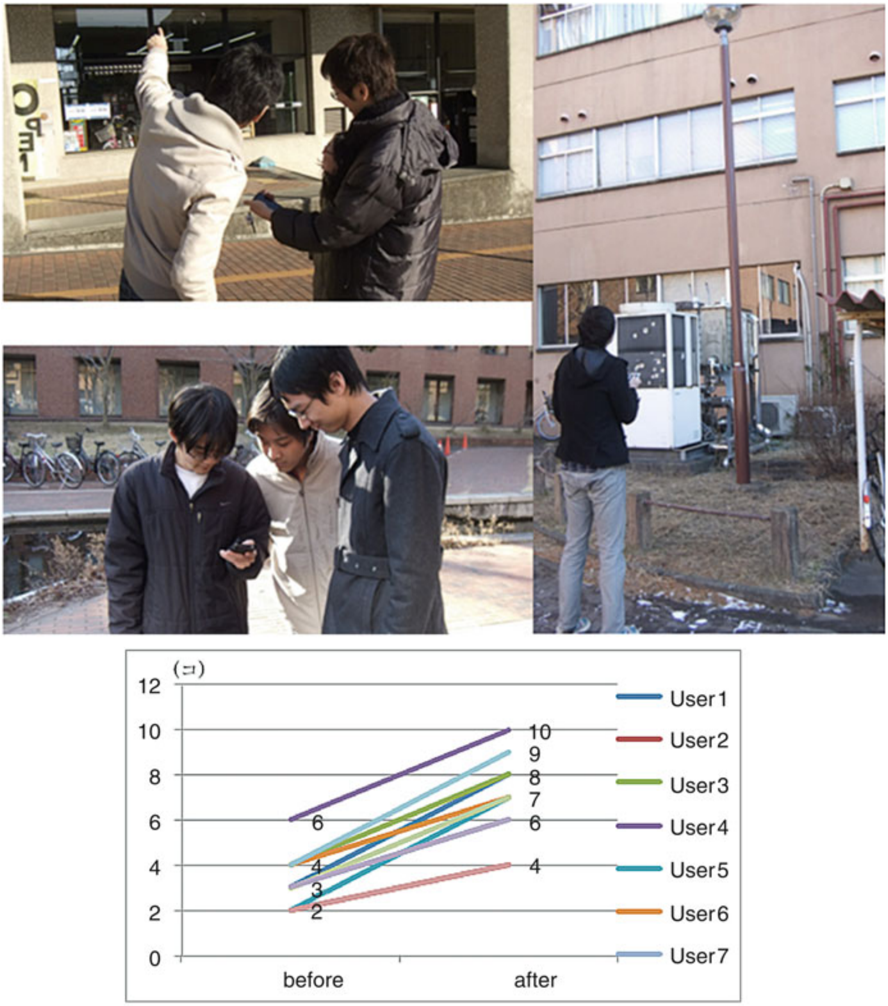


Fig. 8 The scene of experiment and results before/after the experiments

6. Cold weather might lead to a decrease in physical condition. You should prepare in advance for protection from cold weather.

Within the University of Tsukuba campus, there were many places for potential falling objects due to decrepit conditions from the disaster. If the system takes users through a series of building-related events, then we can get an answer like in no. 1. Also, any number of unexpected events could occur because of a disaster, so the game consistently advises locating a road and walking toward a nearby shelter, which then results in answer no. 3.

These answers suggest that incorporating important risks into the game, the user can perceive and consider these risks from the game experience in the real world. As for no. 5 and no. 6, they can be considered to be the characteristic answers of the disaster experience game in the real world, because of the season during which the experiment took place. The experiment was conducted in winter and therefore the weather was quite cold and shelter was necessary. As the author described, playing the proposed game is useful not only for teaching these important lessons but also for encouraging the user's initiative and evacuation consciousness.

Conclusion

In this chapter, we have summarized the real-world game platform that can conduct systematic selection supports of various learning services. The system supports creation of learning entertainment contents and selection by users. Future research will include constructing user experience framework, content engines, and evaluation by providing lifelong education services.

Recommended Reading

B. Anderson, S. Vongpanitlerd, *Network Analysis and Synthesis* (Dover, Mineola, 2013)

N. Azman et al., Public education in heritage conservation for geopark community. Proc. Soc. Behav. Sci. **7**, 504–511 (2010)

Z. Bauman, *Liquid Modernity* (Wiley, New York, 2013)

U. Beck, *Individualization: Institutionalized Individualism and Its Social and Political Consequences*, vol. 13 (Sage, Thousand Oaks, 2002)

C. Bizer, T. Heath, T. Berners-Lee. Linked data-the story so far, in *Semantic Services, Interoperability and Web Applications: Emerging Concepts*, P. S. Amit ed., pp.205–227 (Information Science Reference, Hershey/PA, 2009)

P.J. Carrington, J. Scott, S. Wasserman (eds.), *Models and Methods in Social Network Analysis*, vol. 28 (Cambridge University Press, Cambridge, 2005)

J. Daniel, Making sense of MOOCs: musings in a maze of myth, paradox and possibility. J. Interact. Media Educ. **2012**(3), Art-18 (2012)

R.H. Dave, *Lifelong Education and School Curriculum*. UIE Monographs 1 (Unesco Institute of Education, Hamburg, 1973)

R.H. Dave (ed.), *Foundations of Lifelong Education: Studies in Lifelong Education* (Elsevier, Amsterdam, 2014)

M.S. El-Nasr, A. Drachen, A. Canossa, *Game Analytics: Maximizing the Value of Player Data* (Springer, London, 2013)

G. Foley, *Learning in Social Action: A Contribution to Understanding Informal Education*. Global Perspectives on Adult Education and Training (St. Martins Press, New York, 1999)

N. Longworth, W.K. Davies, *Lifelong Learning: New Vision, New Implications, New Roles for People, Organizations, Nations and Communities in the 21st Century* (Kogan Page, London, 1996)

R.F. Lusch, S.L. Vargoeds, *The Service-Dominant Logic of Marketing: Dialog, Debate, and Directions* (M.E. Sharpe, Armonk, 2006)

V.J. Marsick, K.E. Watkins, *Informal and Incidental Learning in the Workplace* (Routledge Revivals, Abingdon-on-Thames, 1990), 270 p.

V. McGivney, *Informal Learning in the Community: A Trigger for Change and Development* (National Institute of Adult Continuing Education, London, 1999)

W. Mischel, Y. Shoda, A cognitive-affective system theory of personality: reconceptualizing situations, dispositions, dynamics, and invariance in personality structure. Psychol. Rev. **102** (2), 246 (1995)

M. Patzak, W. Eder, UNESCO GEOPARK. A new Programme-A new UNESCO label. Geologica Balcanica **28**, 33–36 (1998)

J.E. Pachankis, The psychological implications of concealing a stigma: a cognitive-affective-behavioral model. Psychol. Bull. **133**(2), 328 (2007)

A. Pentland, *Social Physics: How Good Ideas Spread-The Lessons from a New Science* (Penguin Press, New York, 2014)

S.L. Vargo, P.P. Maglio, M.A. Akaka, On value and value co-creation: a service systems and service logic perspective. Eur. Manag. J. **26**(3), 145–152 (2008)

S. Wasserman, K. Faust, *Social Network Analysis: Methods and Applications*, vol. 8 (Cambridge University Press, Cambridge, 1994)

L. Yuan, S. Powell, J. Cetis, MOOCs and Open Education: Implications for Higher Education, Centre for Educational Technology, Interoperability and Standards, CEITS White paper (CEITS, Bolton, 2013)

Part IV

Interactive Storytelling

Discourse and Camera Control in Interactive Narratives

15

Arnav Jhala

Contents

Abstract

This chapter provides a survey of research on the analysis and generation of narrative discourse that deals with effective presentation of story content through the visual medium. It starts with a theoretical grounding in narratology and cognitive science where the distinction between story and discourse is established. Theories of visual discourse that expand the notions of textual discourse to fit the analysis of visual narratives will be described. Finally, a discussion of automatic generation of coherent visual discourse in terms of viewpoint selection in virtual environments will be carried out.

Keywords

Narrative • interactive narrative • narratology • discourse • camera control

A. Jhala (✉)
University of California Santa Cruz, Santa Cruz, CA, USA
e-mail: jhala@soe.ucsc.edu

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_56

Introduction

Narratologists have long been in agreement that there are two or more levels in narratives. Structuralist narrative theorists such as Barthes, Genette, Todorov, Bremond, and Chatman, among others, generally consider the selection of content (*fabula: what is included*) and its presentation (*discourse: how it is shown*) (Barthes 1981; Bremond 1980; Chatman; Prince 1987) as separate layers in the analysis of texts. In the visual medium, film theorists such as Branigan also consider film analysis in terms of layers that include not only the *fabula* part of the story but also aspects of *performance* of the narrative. In visual media, such as film, there is a complex interaction between a set of performers that are overseen by a director. One such aspect is staging and lighting or the layout of physical space to enable actions in the story to be performed. Another aspect is the performance of actors to communicate both content of the story (dialogue) and mental states of the character. Finally, the composition of camera shots and transition between shots where decisions on content selection (what to include in the frame) and presentation (how to compose the objects within the frame) are taken. When this framework of discourse analysis is taken to interactive scenarios, we also need to include a human interactor with affordances provided by the underlying interactive narrative system to influence content selection and presentation at some or all levels of narration (Fig. 1).

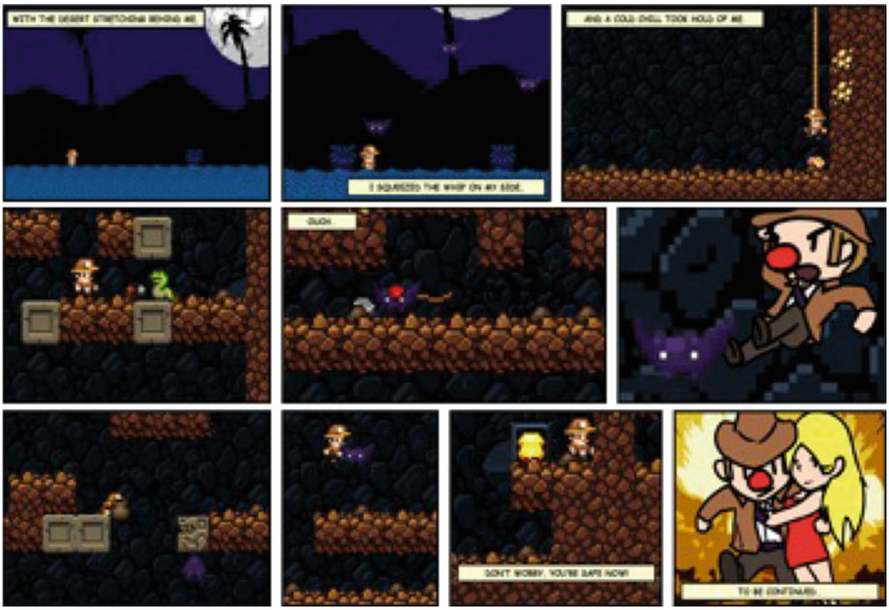


Fig. 1 Visual narrative automatically generated from replays in a procedurally generated dungeon crawler game. Interesting events are selected from game logs and frames are selected from a library of prerendered shots

Unlike narratology, that seeks to analyze how authors *create* narrative, cognitive scientists study how narratives are *perceived*. These investigations seek to answer questions about mental representations that are formed in readers/viewers during the process of comprehension of the narrative that is presented to them. There are study aspects of narratives that lead to cognitive phenomena such as perception of sequence, memory of events, casual dependencies of events, and overall coherence of the narrative. Narrative constructs a world and immerses a user in it, such that aspects of the narrative like order of events, causality, and emotive characteristics that are not overtly specified can still be inferred by readers or viewers (Bae and Young; Broek 1989; Cheong and Young 2008; Zwaan and Radvansky 1998).

Camera shots in a film communicate information about the story world to the viewer. Film theorists (Arijon 1976; Joseph 1970) have described film as being composed of sequences of *frames* that collectively form *scenes*. A *Frame* is a snapshot of the story world at any instant of time. A *scene* is a sequence of frames in which events in the story occur within the same spatial and temporal proximity. This is a nontraditional definition of the scene that we will use in our work as it readily lends itself to a computational representation. In textual discourse there is a similar, if not identical, hierarchical structure where sentences collectively form paragraphs of text (Fig. 2).

Each paragraph of well-written text communicates a particular topic through careful placement and presentation of sentences (Grosz and Sidner 1986) that individually and collectively contribute to the overall communicative intention of the paragraph. Similarly, well-crafted films have scenes comprising shots that

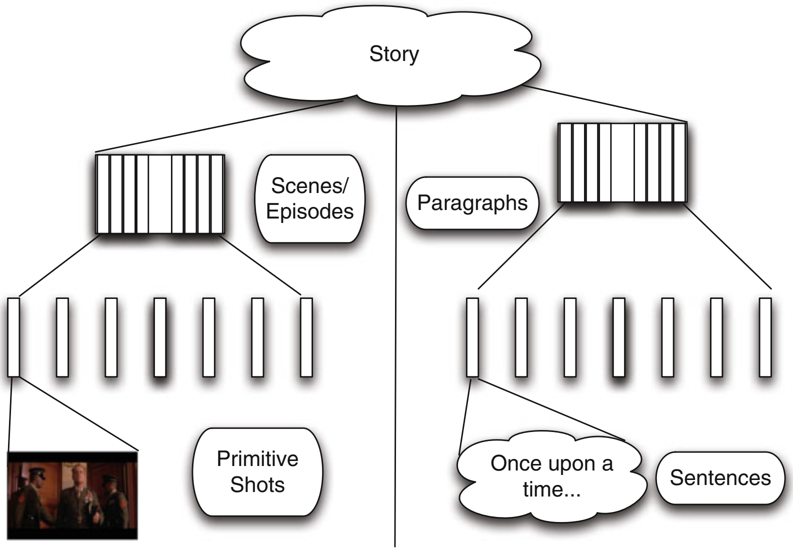


Fig. 2 Hierarchical representation of narrative structure in film and text. Camera shots are the surface/primitive actions, scenes are collection of spatially colocated and temporally contiguous primitive actions, and a collection of scenes together form the narrative

individually and collectively communicate the topic of the scene as intended by the director. For instance, consider the following sentences: *Bob fired his gun at Jane. John fell on the ground*. These sentences are ambiguous because it is not clear how they are related. Adding one more sentence makes the discourse coherent and unambiguous. *Bob fired his gun at Jane. John fell on the ground. He had just jumped in the way of the bullet and sacrificed himself for his friend*. Thus, the information contained within each sentence, when taken individually does not communicate the information that is inferred when they are presented together in a discourse. As described in the section "Theories of Visual Discourse," linguists describe text segments as collection of sentences that collectively convey communicative purpose. Each segment's purpose imposes constraints on individual sentences. For instance, according to one rule, anaphoric expressions are used under the constraints that they refer to the subject already introduced in previous sentences. Individually the sentence *He won* communicates little information about the context in which a character "won." Now consider: *Michael fired an ace on match point. He won*. Here it is clear that *He* refers to Michael because the anaphora can be resolved by looking at the relationship between the two sentences. Consider now a picture with a close shot of a man holding a trophy. This shot conveys that the man is holding a trophy and does not tell us anything about the context. If the same shot is followed by a wide shot of a tennis stadium with applauding audience then the two shots put together provide additional information about the context. It can be seen from this example that there are parallels between inter-sentence relationships seen in natural language and inter-shot relationships in the visual medium.

Theories of Visual Discourse

The discourse aspects of textual language are well-studied in linguistics and are successfully transitioning to computational models. Theories of linguistic discourse structure propose the nature and type of relationships that exist between sentences and between other segments of text. Recently, there have been several theories proposed for various aspects of visual discourse. McCloud (1993, 2006) proposes that transitions between adjacent visual panels in a comic book format provides more information about the semantic characteristics of the story in terms of temporal change and shifts between characters and scenes. Choice of frame composition of adjacent panels appeals to the visual perception system in a way that the inferential process fills in the gaps between the frames. There has been a long-standing tradition in film theory for studying *montage* and *mise-en-scene*. Eisenstein (1942), Metz (1974), Monaco (2000), and others have argued that shots combined together form an inferred meaning. (This is particularly highlighted by the Kuleshov effect where viewers ascribe emotion to the same expression of a character based on the image that is juxtaposed between two of such frames.)

A more recent theory from the cognitive sciences focuses on the local relationships in processing of verbal and visual discourse. The event-indexing model (Magliano et al. 2001; Zwaan et al. 1995; Zwaan and Radvansky 1998) identifies five domains that readers actively monitor when reading text: space, time, entities, motivation, and causality. If there are changes in any of these domains, it causes a shift in the cognitive processes as readers or viewers integrate this new information in working memory. In film, changes in space, time, number of characters, and causality also cue changes in scenes. Cohn reconciles prior theories of visual discourse and provides a structure that is informed by and complements the structure of textual narrative discourse. Cohn argues that discourse structure needs to take into account both the local changes between units of language and global impact of each unit to the overall narrative arc. This structure takes a detailed view of many important aspects of visual discourse structure including phenomena such as the one shown in Fig. 3. Here the effects of reordering the presentation sequence changes the beliefs of viewers about the original story based on inferences drawn from the different sequences.

Computational Analysis of Visual Discourse

Computational analysis of visual discourse deals with automated and semiautomated techniques for extracting boundaries of grammatical units (such as shots, frames, and scenes), identification of semantic elements, and an overall narrative structure from video. These techniques are still in their infancy due to the semantic gap that exists between low-level feature analysis from video and high-level intent mapping. Late 1990s and early 2000s saw a huge effort in the multimedia community towards automated video and film understanding motivated by a huge uptick in demand for various multimedia applications including compression, summarization, browsing, search, and retrieval. See (Truong and Venkatesh 2007) for a comprehensive survey. These efforts often motivated by principles from film theory used hand designed data representations and computational pipelines and used relatively small datasets for learning. Consequently, after initial rapid and promising advances, work in this area has largely plateaued. There are various elements of visual style explored in the multimedia analysis literature; these exist at various structural scales in the visual narrative from frames to shots to scenes to the entire movie (Fig. 4). Sanokho and Christie (2014) propose a model for visual balance in terms of focus of attention (actor's) body and head positions, their orientations, and gaze with respect to the camera as well as the empty space in the frame. Wang and Cheong (2009) propose a taxonomy of directing semantics (shot types) based on presence or absence of a focus of attention (FOA) actor(s) in their case, the camera and FOA motions, and the distance of the FOA to the camera. Adams et al. (2002) proposed a computational model for film rhythm by classifying the motion behavior as nonexistent, fluid, and staccato for a given shot. They showed that rhythm could be used to group sequences of shots (scenes) into seven different kinds of scene types based on their semantic content. In a subsequent work, they propose a model for film pacing or tempo based on the average motion magnitude in the shot and the length of the shot.

Merabti et al. (2014) are the first ones to propose to learn such a generative model from actual movies. They model the film-making process as a hidden Markov model

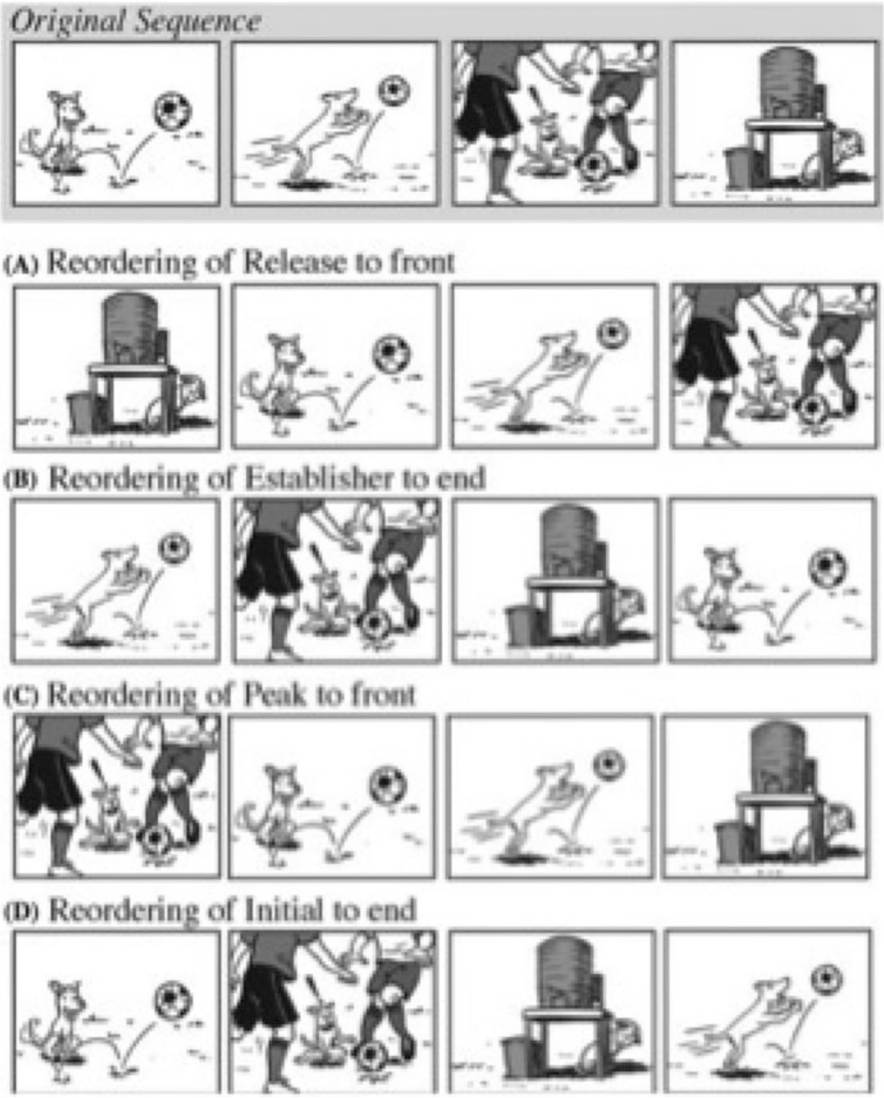


Fig. 3 Effect of reordering frames on the inferences made by viewers on the story events. Cohn

(HMM) where the hidden states denote different shot types while the observation nodes denote different kinds of dialog events between two characters, their relative importance, and presence in the fabula. Further, they make an important observation that film-making style corresponds to the parameters of the HMM model. Thus, two HMMs with different parameters encode different styles and given the same input (sequence of events) will rightly produce different outputs (sequence of shots). Their model can take as input an event sequence from an annotated script and produce a

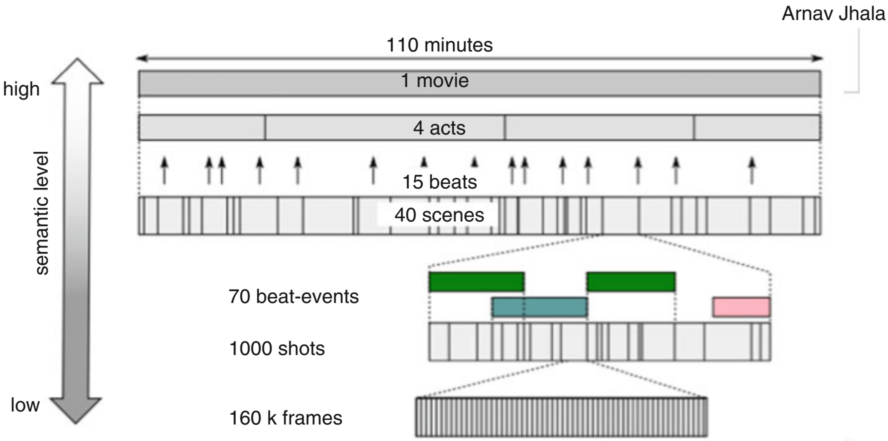


Fig. 4 Potapov et al.(Douze et al. 2014) claim reliable classification of scene boundaries and narrative beats based on changes detected in location, set of characters, or major narrative event movement in the video. Model learns and uncovers temporal structure using a conditional random field taking support vector machine classification scores as input features

sequence of shot types in the learnt style. Merabti et al. use a small hand-annotated dataset of movies to train their model. Further, they assume the availability of expert-annotated scripts and limit the model to rendering shots with dialogs between two people. Further, they are only useful in representing short-term temporal dependencies and cannot fully exploit the temporal relationships in films that exist at several temporal scales.

Generation of Visual Discourse: Camera Control

The computational cinematography community has proposed a number of approaches to automatically generate cinematographic sequences capturing activities in a 3D virtual world. The problem is formulated as that of selection of camera viewpoints (shot types) and shot transitions to optimally communicate the action being captured in the virtual 3D world. Computational approaches employed include estimation using finite-state machines (He et al. 1996), searching over film-tree representations (Christianson et al. 1996), dynamic programming over a collection of template shots (Elson and Riedl 2007), and hierarchical planning (Jhala and Young 2005).

Earliest approaches to camera control such as He et al. (1996) encode camera behavior as a finite-state machine (FSM). States in the FSM represent primitive single shots of the camera with fixed composition parameters, and transitions represent changes in story states that trigger corresponding camera composition changes in the FSM. Drucker and Zeltzer (1995) and Bares et al.(1998) improve on the variety of shots expressed in the FSM approach by modularizing and expanding the constraint language for controlling the camera. By relying on Arijons cookbook (Arijon 1976) of camera setups, these approaches encode stereotypical ways of shooting character configurations and actions by taking into account the physical layout of the scene. Other approaches rely on a combinatorial branching system such as the algorithm proposed by Christianson et al. (1996) that represents a movie as a hierarchical structure of sequences decomposed into scenes, candidate idioms, and candidate frames. Computing the best sequence of frames consists in searching the best frames, the best idioms, and the best transitions between idioms by exploring and evaluating all the possibilities in the film tree. A more general process that includes the tasks of blocking (placing the characters and the scene), shooting, and editing is implemented in the CamBot system (Elson and Riedl 2007). CamBot relies on a combination of search and dynamic programming to identify the best shots from a library of predefined stage layouts and narrative beats. The approach offers variations on directorial styles through manual weighting of parameters of the search heuristic. Jhala and Young (2005) operationalize the story/discourse distinction drawn in narratology in the Darshak system. Marshak encodes cinematic idioms as a hierarchy of plan operators that represent the presentation aspect. Plan operators in this approach query a content knowledge base that consists of fabula elements such as actions, events, participants, causality between actions, and temporal ordering. These operators are then sequenced through a partial order causal link planning system to generate a cinematographic sequence that fulfills directorial goals and conveys a sequence of actions. This approach closely follows the computational approaches to textual discourse generation in an attempt to clearly separate out the fabula part and the discourse part and make the knowledge representation at the fabula level uniform across modalities. The planner recursively explores the plan operators and adds and propagates constraints until reaching primitive camera actions that set visual constraints.

Cinematic discourse in the Darshak system is generated by a hierarchical partial order causal link planner whose abstract actions represent cinematic patterns and whose primitive actions correspond to individual camera shots. As shown in Fig. 2, the system takes as input an operator library, a story plan, and a set of communicative goals. The operator library contains a collection of action operators that represent camera placement actions, transitions, abstract cinematic idioms, and narrative patterns. Camera placement and transition actions, represented as primitive operators, affect the focus of visual they represent and have preconditions that encode continuity rules in cinematography. Operators representing abstract cinematic idioms and narrative patterns affect the beliefs of the viewers and encode recipes for sequencing primitive or abstract operators. A description of the story to be filmed is input as a plan data structure that contains the description of the initial state of the story world, a set of goals achieved by the story characters actions, and a totally ordered sequence of the story characters actions and the causal relationships between them. The input story plan is added to the knowledge base for the discourse planner in a declarative form using first-order predicates that describe the elements of the data structure. The communicative goals given to the system describe a set of belief states to be established in the mind of the viewer. The cinematic discourse planning algorithm performs both causal planning and temporal scheduling. To build a discourse plan, it selects camera operators from the operator library and adds them to the plan in order to satisfy specific communicative goals or preconditions of other communicative actions already in the plan. The algorithm binds variables in the camera operators, like start-time and end-time, relating the camera actions to corresponding actions in the story plan. The output of the planning algorithm is a plan data structure containing a temporally ordered hierarchical structure of camera operators with all operator variables bound. The resulting plan is merged with the story plan and sent to a game engine for execution (Videos of different visualizations can be found at: http://www.youtube.com/user/cyclonurb).

Discussion

This chapter presents an overview of the history and current state of research on analysis and generation of visual discourse. In this exploration, we notice the convergence of narratology, cognitive science, and computation as disciplines that are central to further advances in this field. Two aspects that are not directly addressed in this chapter but are worth mentioning are evaluation and adaptation. In generative and interactive systems, user evaluation of discourse elements poses interesting challenges. One aspect of discourse – comprehension – has been evaluated in the Darshak system. Comprehension is easier to measure in noninteractive visual narratives due to the availability of a fixed underlying story and indirect measures through questions about the content of the narrative that is independent of the presentation aspect. Other aspects of discourse such as ordering of events, affect, attention, style, and inclusion/exclusion of content have not been systematically evaluated in generative systems. In interactive systems, it is even more challenging to evaluate computational algorithms of discourse due to the changes effected by players in the underlying story. If interactive systems are seen as a dialogue between the human and the computer then research on dialogue acts, turn-taking protocols, and communicative acts of interaction can be brought in to further expand current models of visual discourse. Finally, most interaction in games is designed around a fixed viewport (first person, bird's eye, top-down, point-of-view) and clear demarcation between cut-scenes and interactive aspects of the game. With better understanding of the visual interaction processes and adaptive computational approaches, new forms of interaction and storytelling might emerge. Finally, interactive storytelling as a collaborative process as in the exquisite corpse game offers to provide challenging research questions when casted in visual medium.

Recommended Reading

B. Adams, C. Dorai, S. Venkatesh, Toward automatic extraction of expressive elements from motion pictures: tempo. IEEE Trans. Multimedia **4**(4), 472–481 (2002)

D. Arijon, *Grammar of the Film Language* (Silman-James Press, Los Angeles, 1976)

B.-C. Bae, R.M. Young, A use of flashback and foreshadowing for surprise arousal in narrative using a plan-based approach, in *International Conference on Interactive Digital Storytelling*, eds. by U. Spierling, N. Szilas. LNCS 5334, pp. 156–167 (Springer, Berlin, 2008)

W.H.Bares, J.P. Gregoire, J.C. Lester, Realtime constraint-based cinematography for complex interactive 3d worlds, in *AAAI/IAAI*, pp. 1101–1106 (AAAI Press/MIT Press, Boston, 1998)

R. Barthes, *Introduction to the Structural Analysis of the Narrative*. Communications (Centre for Contemporary Cultural Studies, University of Birmingham, 1981)

C. Bremond, The logic of narrative possibilities. New Lit. Hist. **11**(3), 387–411 (1980)

P.V. Broek, The effects of causal structure on the comprehension of narratives: implications for education. Read. Psychol. **10**(1), 19–44 (1989)

S. Chatman, *Story and Discourse: Narrative Structure in Fiction and Film* (Cornell University Press, New York, 1978)

Y.-G. Cheong, R. M.Young, Narrative generation for suspense: modeling and evaluation, in *ICIDS*, pp. 144–155 (Springer, Berlin, 2008)

D. Christianson, S. Anderson, L.-W. He, D. Salesin, D. Weld, M. Cohen, Declarative camera control for automatic cinematography. In *AAAI '96* (1996)

M. Douze, Z. Harchaoui, C. Schmid, D. Potapov, Category-specific video summarization. In *European Conference on Computer Vision (ECCV)*, Zurich, (Springer, Berlin, 2014)

S. M. Drucker, D. Zeltzer, Camdroid: a system for implementing intelligent camera control. In *Symposium on Interactive 3D Graphics*, Monterey, pp. 139–144 (ACM, New York, 1995).

S. Eisenstein, *The Film Sense* (Hartcourt, Orland, 1942)

D. Elson, M. Riedl, A lightweight intelligent virtual cinematography system for machinima production, in *Artificial Intelligence in Interactive Digital Entertainment (AIIDE)*, Palo Alto, 2007. Association for the Advancement of Artificial Intelligence (AAAI)

B.J. Grosz, C.L. Sidner, Attention, intentions, and the structure of discourse. Comput. Linguist. **12**(3), 175–204 (1986)

L.-W. He, C. Michael, S. David, The virtual cinematographer: a paradigm for real-time camera control and directing, in *Proceedings of SIGGRAPH*, New Orleans, pp. 217–224 (ACM, New York, 1996)

A. Jhala, R. M. Young, A discourse planning approach to cinematic camera control for narratives in virtual environments, in *AAAI*, pp. 307–312 (AAAI Press/MIT Press, Boston, 2005)

M. Joseph, *The Five C's of Cinematography* (Cine/Grafic Publications, Hollywood, 1970)

J. Magliano, J. Miller, R. Zwaan, Indexing space and time and film understanding. Appl. Cogn. Psychol. **15**, 533–545 (2001)

S. Mccloud, *Understanding Comics: The Invisible Art* (Tundra Publishing, Northampton, 1993)

S. Mccloud, *Making Comics: Storytelling Secrets of Comics, Manga and Graphic Novels* (Harper, New York, 2006)

B. Merabti, M. Christie, K. Bouatouch, A virtual director inspired by real directors, in *Workshops at the Twenty-Eighth AAAI Conference on Artificial Intelligence*, Quebec (ACM, New York, 2014).

C. Metz, *Film Language: A Semiotics of the Cinema* (Oxford University Press, Oxford, 1974)

J. Monaco, *How to Read a Film: The World of Movies, Media, and Multimedia: Language, History, Theory* (Oxford University Press, Oxford, 2000)

G. Prince, *A Dictionary of Narratology* (University of Nebraska Press, Lincoln, 1987)

C.B. Sanokho, M. Christie, On-screen visual balance inspired by real movies, in *Workshops at the Twenty-Eighth AAAI Conference on Artificial Intelligence*, Quebec (ACM, New York, 2014)

B.T. Truong, S. Venkatesh, Video abstraction: a systematic review and classification. ACM Trans. on Multimedia Computation, Vol.3, No.1, pp. 1–37 (ACM, New York, 2007)

H.L. Wang,L. F. Cheong, Taxonomy of directing semantics for film shot classification. IEEE Transactions on Circuits and Systems for Video Technology, Vol.19, No.10, pp.1529–1542 (Oct 2009)

R. A. Zwaan, G.A. Radvansky, Situation models in language comprehension and memory. Psychological Bulletin, Vol.123, No.2, pp.162–185 (American Psychological Association, Washington, 1998)

R.A. Zwaan, M.C. Langston, A.C. Graesser, The construction of sitiation models in narrative comprehension: An event-indexing model. Psychological Science, Vol.6, No.5, pp. 292–297 (Sage Publications, London, 1995)

Interactive Storytelling Paradigms and Representations: A Humanities-Based Perspective

16

Hartmut Koenitz

Contents

Abstract

Interactive Storytelling is an interdisciplinary field in which the humanities meet artificial intelligence. Collaborations between scholars rooted in the humanities and the computer sciences like the one between Brenda Laurel and the OZ group at Carnegie Mellon University have had a major influence on the field. At the same time, there are indications that the relationship is often tenuous, for example, between models of narrative in the humanities and their application in computational research projects. This chapter investigates the relationship, notes challenges, and identifies opportunities for an enhanced collaboration. Additional scrutiny in understanding context and scope of narrative models in the humanities would improve access for AI researchers to the vast space of

H. Koenitz (✉)
Department of Entertainment and Media Studies, University of Georgia, Athens, GA, USA
e-mail: hkoenitz@uga.edu

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_58

available models and allow for the codification and re-use of adaptation strategies. Simultaneously, the work of many AI researchers could be recast and recognized as contributions to narrative theory. In this regard, film theory can serve as a potential model for a narrative theory of Interactive Storytelling.

Keywords

Bordwell's view of film • Descriptive and prescriptive narrative models • Interactive Storytelling (IS) • Basic orientation • Computation and humanities • Descriptive and prescriptive narrative models • History of influence • Humanities in AI • Humanities vs. AI • Narrative models • Opportunities • Recent developments • Requirements for dialog • Roles of AI researcher • Narrative models • Descriptive and prescriptive • Narratology • Prescriptive and descriptive narrative models • Interactive digital narrative (IDN)

A Field Between Computation and Humanities

Interactive Storytelling can be described as a marriage of computation and narration that brings together perspectives origination in Computer Science and the Humanities. The interdisciplinary aspect is both a chance and a complication. A chance, in that the product of two distinct fields might combine the best of both worlds, and thus enable expressive artifacts that speak of the human condition in novel and compelling ways. And a complication, in that terminology and general approaches differ to a degree that a serious translation effort seems inevitable. Without that, there is a considerable risk of miscommunication. The situation is complicated further by semantic differences as the meaning of the same term varies across fields. For example, the understanding of "model" in computer science might diverge considerably from the understanding of the same word in the humanities. Conversely, "story" in a more general sense meaning "a narrative" clashes with the basic structuralist distinction in narratology between "story," meaning the "what" of a narrative, its content, in contrast to "discourse," denoting the "how," the presentation, its "telling." This example illustrates the potential for miscommunication on a basic level. Even though poststructuralist and postclassical narratology have refocused on other aspects of narrative, the story/discourse distinction is still a cornerstone in academic training in the humanities and no contemporary humanities scholar would use these terms without reference to their specific meaning. When an AI researcher talks about "interactive storytelling," or just "story" without providing a specific definition or acknowledging the etymology of these terms, a scholar grounded in the humanities might wonder whether the colleague considers the narratological context and in particular the distinction between story and its telling.

For this reason, I avoid the term "storytelling" in my own work and instead use the abbreviation IDN for interactive digital narrative (Koenitz 2015). The abbreviation is also intended to reduce conceptual problems inherent in the compound "interactive narrative." When we create the conceptual framing of narrative + interactivity, of interactivity as an addition to traditional narrative forms, we are prone to misunderstand the challenges and opportunities of interactive narrative as a specific form of narrative expression, in which interactivity is a fundamental ingredient, not an additional feature.

The editors of this section prevent these terminological problems by clearly defining their usage of terminology and distinguish between Interactive Storytelling for the Research field, Interactive Narrative for the individual "product," and Interactive Storytelling System for the technology. In this chapter, I will adhere to this nomenclature.

Mapping the opportunities afforded by the humanities for computer science requires a careful consideration of the etymology and semantics of both fields, but foremost an effort in a clear descriptions of concepts. I ask the reader to remember this aspect if some of the descriptions in this chapter might appear verbose.

In this chapter, I will start with some general remarks about the relationship between humanities and computer science before tracing some of the developments between humanities and AI in Interactive Storytelling. On this basis, I propose strategies for future research bringing humanities concepts to AI.

My goal with this chapter is threefold – first, to draw attention to the richness of available perspectives in the humanities, particularly outside of the confines of literary studies, which can serve as models for use in Interactive Storytelling. Second, to reframe the adaptation of humanities models by AI researchers in this field as genuine theoretical contributions and thus reconsider the relationship to the original models and their historical and epistemological context. Third, to raise awareness of the mutual benefits an enhanced collaboration between humanities and AI would bring to both fields.

Points of Orientation

Marc Cavazza and David Pizzi (2006) provide a good overview of narratology from the perspective of AI by considering a range of popular models of narrative, including Aristotle, Propp, and Barthes. Another valuable overview is in a paper by Pablo Gervas et al. (2006), originating in an interdisciplinary collaboration between humanities and AI scholars focused on narrative generation.

Of particular value as a basic orientation are summaries of the diverse narratological approaches in the humanities that also consider the historical development. Many good resources exist in this regard, a particularly easily accessible one is the *Living Handbook of Narratology* (Hühn n.d.). In addition to gaining an understanding of different approaches, it is also important to consider their underlying assumptions. Dominique Robert and Shaul Shenhav (2014) provide a current overview of this aspect of narrative analysis.

I like to suggest that having this kind of context is crucial in finding appropriate models, but also in avoiding fundamental assumptions potentially detrimental to an AI narrative project. The latter point pertains to the web of influences not readily apparent on the surface, but equally important. For example, a narrative model might embed the structuralist story/discourse dichotomy without explicit consideration of the structuralist assumptions governing this distinction.

A History of Influence

There is a long history of influence between the two fields, starting with one of the earliest scholarly works on Interactive Storytelling, Brenda Laurel's 1986 doctoral dissertation in theatre studies at Ohio State University, "Toward the design of a computer-based interactive fantasy system" (Laurel 1986). Her thesis caught the interest of the computer scientist Joseph Bates, who made her work a foundational element of his research in believable characters. Bates founded the OZ group at Carnegie Mellon University, which accumulated a number of researchers interested in the same subject, to work on "technology and art to help artists create high quality interactive drama, based in part on AI technologies" (description on the OZ project website). The group's particular focus was on building believable agents. This connection between OZ and Laurel's work was not merely conceptual, she actually worked with the group, as reported in her book *Computers as Theatre* (Laurel 1991).

Michael Mateas, a member of the group, went on to create the first fully realized interactive drama at Georgia Tech in 2005. *Façade* (Mateas and Stern 2005b) puts the player/interactor into the position of a friend trying to save the marriage between the two virtual characters Trip and Grace. Lauded as a milestone, the work is an exemplary result of the combination of concepts originally developed in the humanities and then realized with the capabilities of artificial intelligence in computer science.

Fundamental Differences Between Humanities and AI

While the trajectory from Laurel's PhD thesis to *Façade* is certainly a compelling success story, a closer look also reveals some underlying issues. A basic concern shared between the two fields is for models; however, the concrete approaches differ considerably. Broadly, they can be described as "concrete" versus "abstract" – while AI models tend to be formal descriptions of concrete technical implementations, most narratological frameworks in the humanities are abstract, aiming to identify the essence of narrative, rather than any particular case. Seymour Chatman describes the approach as follows:

> [. . .] literary theory is the study of the nature of literature. It is not concerned with the evaluation or description of any particular literary work for its own sake. It is not literary criticism but the study of the givens of criticism, the nature of literary objects and their parts (Chatman 1980)

This perspective, rejecting the evaluation of specific artifacts, contrasts with basic scientific practice that aims to generalize from the particular case. Where Newton starts with the observation of a falling apple (a particular event) to explain the mechanics of the universe, the narratologist Wladimir Propp starts with a collection of Russian folk tales to identify shared characteristics. This comparison also embeds another important difference in the status of the resulting categories and overall concepts. Whereas scientific definitions conceptually originate in empirical observation, definitions in the humanities are deduced, created by the researchers, as Chatman reminds us: "literary theory [...] should assume that definitions are to be made, not discovered" (Chatman 1980).

Such tendencies – different starting points of the academic inquiry, as well as the divergent nature and origin of rules and categories – are deeply engrained in the two disciplines and therefore enter, sometimes unconscientiously, scholarly work.

Another distinction pertains to the scope of a given framework. The humanities scholar works from the premise that the sought-after model needs to accommodate a wide range of already existing narratives, while AI researchers evaluate models based on their concrete output. Chatman makes quite clear that the detailed critical description of the individual works in the canon used to derive a narratological model from is a task outside the realm of narratology. He sees this task as best left to the field of literary criticism. In contrast, the AI researcher needs to scrutinize any particular output to judge the success of a given system.

In a similar vein to Chatman, Gerald Prince states unequivocally that aesthetic evaluations have no place in narratology:

> For narratology, there are no great, or beautiful, or profound narratives (there are only well-formed ones). (Prince 1982)

The perspectives of Chatman and Prince demonstrate the fundamental scope of narrative models in the humanities. The aim for universal descriptions has consequences if directly translated into AI models. An AI formalism built on narratology might create well-formed narratives. However, these narratives might still be boring, haphazard, and trivial. An additional layer of "aesthetic evaluation" is necessary to produce interesting and aesthetically pleasing pieces.

It is for this reason narratological frameworks that cover semantic and pragmatic aspects seem to be a more promising avenue for Interactive Storytelling. Prince (1982) does indeed propose such frameworks, but cautions that semantic and pragmatic interpretations should take care to remain generic, and thus be applicable to a wide range of works. Therefore, narratological frameworks are most likely insufficient for the creation of successful Interactive Narratives because of the limited coverage of necessary aspects.

In addition, most models in the humanities do not contain formalisms in a form that could be applied directly to interactive narrative. Gervás et al. in their discussion of this problem mention the lack of "logic language (e.g., predicate calculus) or other structured representation, including tables, graphs, etc." (Gervás et al. 2006).

A Multitude of Models

A further obstacle to the direct adoption of narrative models from the humanities in AI is their sheer number, which make the respective advantages and disadvantages for application in interactive storytelling difficult to judge, due to widely variable approaches. The identification of basic elements, structural relationships, and terminology vary considerably between models origination in schools of thought as far apart as Marxism and Psychology. Casting the net even wider to include East Asian or African traditions yields cyclical multi-climatic structures in some African oral narrative traditions or models of narrative forms without conflict in Japan (Kishotenketsu).

What is commonly referred to as narratology is but one of many available analytical perspectives in the humanities. In other words, narratology, as a field understood to contain scholars such as Propp (1928), Barthes (1975), Chatman (1980), Prince (1982; 2003), or Herman (2002), is a specific subfield of the different ways of analyzing narratives in the humanities. In addition, semiotic (Eco 1989; Kristeva 1980; Simpkins 2001), Marxist (Eagleton 2006, 2013; Lukács 1983), feminist (Butler 1999; Moi 2002; Ruthven 1990), and psychological (Lacan 2003; Schram and Steen 2001; Sugg 1992) frameworks have been used to analyze narratives. AI scholars approaching the humanities might not immediately be aware of this plethora of approaches. Frameworks outside of the narratological core provide a rich – and so far mostly untapped – source for AI research in Interactive Storytelling.

An example for a successful application of this approach is McCoy et al.'s "Comme il faut" engine (McCoy et al. 2009), the platform for their work *Prom Week*. These researchers use dramaturgy, and not literature, as a starting point and move beyond models directly concerned with narrative artifacts by drawing from knowledge about the authoring process. Finally, McCoy et al. base their approach on models from outside the humanities, namely the social sciences.

Descriptive and Prescriptive Models

Frameworks for narrative in the humanities are mostly descriptive in nature, conceived to analyze artifacts, but without any intention to provide the necessary detailed information an author would need. In contrast, prescriptive frameworks are also meant as models for artists to create works and thus are more appropriate as models of AI. Aristotle's Poetics can be seen as a rare example of a prescriptive framework, with detailed perspectives on practical aspects, including the poetic meter for language used and appropriate topics. The prescriptive aspect of the Aristotelian framework is one reason for the lasting appeal of this conception of narrative. However in practice, the distinction between prescriptive and descriptive is not be entirely clear-cut, as artists might appropriated descriptive frameworks, and in this way adapted them for use as prescriptive formalisms. Artists at times have also developed formal descriptions of their work. Examples include the Russian film director Sergei Eisenstein (1969) and the German playwright Bertolt Brecht (1960). Eisenstein critically reflected his work and thus is not only a pioneering film artist but also one of the first film theorists. Brecht, developed an activist theatre in theory and practice. His Japanese-influenced Epic Theatre provided a model for Arturo Boal's "Theatre of the Oppressed" which in turn has been proposed as a model for activist video games (Frasca 2001).

A further complication for the application of existing frameworks in AI lies in their media-specific basis. Aristotle's Poetics is originally concerned with stage drama (and the specific subcategory of tragedy), while many twentieth-century narratologists focus on literary narrative. Consequently, these frameworks do not automatically extend beyond the specific media they are conceived for.

Media-Specific Aspects: Bordwell

Film theory provides a valuable lesson in this regard. The film theorist David Bordwell considers the dominant basis of narratology in literary studies, predicated on verbal forms of narrative as particularly problematic. For him, examples of narrative in film, but also music or dance are testimony to the contrary, that narrative can exist outside the confines of (mostly written) language. Bordwell thus understands narrative as a "preverbal phenomenon" (Bordwell 2007) supported by a framework in human cognition that enables the understanding of narratives without requiring language. The insight that narrative is not exclusively linked to literature has important consequences for analytical perspectives and resulting models, Bordwell reminds us. When language is taken as the starting point, elements and structures identified in literary studies should be available in other forms of narrative as well. However, if other media forms are taken as having specific affordances and providing unique opportunities, we should not expect to successfully apply the same elements and structures:

> If language sets the agenda for all narrative, then we ought to expect all media to follow along. So in a film the analyst will look for equivalents of first-person point of view, or something analogous to the voice of a literary narrator. But if we think that language is on the same footing as other media, a vehicle for some but not all more fundamental narrative capacities, then we might not expect to find exact parallels between literary devices and filmic ones. Different media might activate distinct domains of storytelling. Perhaps, that is, filmic point of view might be quite different from literary point of view, and there may be no cinematic equivalent of a verbal narrator. (Bordwell 2007)

Bordwell's view of film as a specific form of narrative also applies to narrative forms in interactive digital media. While film narrative is setting up opportunities for interpretation as Bordwell observes, Interactive Storytelling adds a further dimension in providing opportunities for interaction. Seen in this light Interactive Narratives are even further away from literary form than film, by including the exploration of virtual worlds, choices and the experience of consequences, and the ability to replay.

On the basis of his observations, Bordwell defines film narrative as comprised of three-dimensions: *story world*, *plot structure*, and *narration*, in contrast to the structuralist dichotomy of fabula/syuzhet or story/discourse that is still prevalent in literary narratology.

The Humanities Turn in AI

In summary, AI researchers will most likely encounter frameworks in the humanities that are either descriptive in their intent or media-specific and therefore inappropriate and incomplete as the basis for practical implementations in digital media. This does not preclude the usage of such a framework for AI in principle, but necessary adaptation and additions of missing elements mean that the result is actually a different framework, with a complex relationship to the original. I like to suggest that this fact alone does not constitute a fundamental problem and rather should be understood as standard academic procedure. However, the tendency to foreground the original framework is problematic as the assumptions on which it was based are no longer valid in their entirety. The underlying problem here might very well be an attempt to gain additional "academic prestige" for interdisciplinary work by referencing major theoretical frameworks originating outside a researcher's home discipline.

In particular, Aristotle via the Poetics can be considered the "undying ghost in the machine," of the field, a frame of reference that is often taken as a given without proper contextualization of the relationship to the ancient text. In this vein, Cavazza and Pizzi consider the importance of Aristotle's Poetics as a model for Interactive Storytelling as having been overstated:

> [. . .] the Aristotelian model seems to have been primarily used as an inspiration, a theoretical framework in which to describe narrative concepts, rather than a source of narrative formalisms, let alone their implementations in IS [interactive storytelling] systems. (Cavazza and Pizzi 2006)

Further investigating the relationship between models and implementations, the authors classify the connection between the Poetics and many concrete implementations as tenuous, quoting Wilks's (1992) insight that "systems do not always work by means of the formalisms that decorate them."

Maybe one aspect that AI researchers could take from the humanities is that existing frameworks are not sacrosanct and narratological concepts are constantly evolving. Later additions are valuable and welcome. For example, instead of emphasizing the connection to ancient drama, Mateas' and Stern's *Façade* can be more aptly described as based on the theoretical framework developed by Mateas himself (2001), modifying Laurel's (1991) reinterpretation of Neo-Aristotelian concepts by implementing Janet Murray's (1997) analytical and phenomenological categories. The reference to the rather distant – several times removed – original notion is more distracting than helpful. Indeed, Aristotle's unmodified concept is inappropriate for Interactive Storytelling in its basic demand of a plot that is perfect if nothing needs to be added or can be removed, a perspective which is in fundamental disagreement with the idea of interactive narrative that needs to offer excess material for alternative paths and optional elements.

Seen in this light, Mateas' and Stern's claim of an "Aristotelian tension arc" (Mateas and Stern 2005a) in Façade might be misleading. Given the lineage outlined earlier, an unwieldy but more correct description would be of a "Mateas/Stern tension arc based on Laurel's reconceptualization of Neo-Aristotelian concepts modified to accommodate Murray's media-specific categories." This perspective does in no way lessen the appreciation of the work of the involved scholars. The crucial insight from this discussion is the realization that many – if not all – AI researchers working in Interactive Storytelling are no longer just that. Instead, by adapting and further developing existing theoretical frameworks for their needs, they engage in activities typical of humanities scholars. We might even say that AI researchers working in Interactive Storytelling – at least partially – become humanities scholars themselves, for pragmatic reasons, since few, if any, narratological models can be implemented as is. Yet, the effort that goes into making the necessary adjustments is more than just "technical preparation work" and has so far not been fully recognized. Indeed, when Szilas and Richle (2013) pinpoint the fuzziness of the concept of dramatic tension in narrative theory or Gervás et al. (2015) identify commonalities between the different categorizations of plot, they also work in a humanities capacity. Existing perspectives that emphasize common ground between AI and the humanities, for example, Gervás et al.'s notion of "shared abstract models" (Gervás et al. 2006) are therefore only a starting point. It is necessary to recognize the contribution of AI researchers to narrative theory.

At the same time, AI researchers are encouraged to more fully document their process in adapting and augmenting frameworks originating in the humanities. Once the assumptions underlying this translation process are available for research as well as the strategies in conducting the translation, they could be consciously re-applied, modified, and critiqued, and thus help end a practice of uncritical re-application, described by Gervás et al.:

> Often, the most significant reason behind a particular choice of theory is that AI researchers find it easier to work with models that a previous researcher has already translated to AI jargon and applied in some previous computer program. (Gervás et al. 2006)

An important aspect of the availability of such "translation rules" would be as tools for the adaptation of additional frameworks, by making the plethora of models in the humanities more accessible.

Requirements for a Dialog

AI scholars working in interactive narrative can only gain from an enhanced dialog with the humanities. Those who ignore the context of humanities frameworks do so at their own peril – instead of gaining additional insight by means of an interdisciplinary dialog that could help with the concrete implementation. Yet, such a dialog requires at least a minimal acknowledgement of narratological models and vocabulary.

From this perspective, McCoy et al.'s otherwise outstanding work on Prom Week is marred by the lack of a clear definition of "story." For example, in one of their papers, (McCoy et al. 2013) the authors variously refer to story as "personally meaningful" game experience, as "campaign," as "gameplay," as a "collection of levels [. . .] where the player can take social actions," or they talk about "player agency at the story level" and the "player's path through the story." Each of these descriptions depicts "story" as a slightly different concept. In a related perspective, Markku Eskelinen fundamentally questions the value of contributions missing a concrete definition of narrative and deems them "non-academic" (Eskelinen 2012) While I consider Eskelinens's critic exaggerated, the lack of a definition is still a serious omission. In any case, it represents an obstacle to interdisciplinary dialog. How a scholar working in the space of Interactive Storytelling can integrate a meaningful discussion of relevant humanities terminology and definitions is exemplified in Michael Nitsche's Video Game Spaces (Nitsche 2008). This author dedicates considerable space in his volume to the discussion of narratological concepts and their adaptation for interactive media.

Bordwell's approach towards film, outlined previously in this chapter, can serve as an example for AI researchers interested in developing models for interactive narrative. While mindful of the existing tradition in literary studies, his model of film narrative reflects the specific characteristics of the medium. Similar, Interactive Storytelling researchers might be best served by developing specific frameworks for narrative in interactive digital media that are mindful of the existing narratological tradition. Such an effort would have the added benefit of providing a common ground for discussion with scholars in the humanities.

The Many Roles of the AI Researcher

Earlier, I have pointed out some of the fundamental differences between the two fields. This discussion also embeds a perspective on the role of the AI researcher in interactive narrative in contrast to the humanities scholar. Interactive Storytelling research encompasses aspects of both narrative analysis and literary/aesthetic criticism. And while it is true that many leading figures in the humanities have also filled both roles – for example Roland Barthes – AI can still profit from an awareness of the differentiation, and its different dimensions. However, this perspective does not yet cover all the roles filled by many AI researchers. As I have argued earlier, there is also a creative dimension (Koenitz 2014). In many cases, the researcher also becomes the author using her/his system to create an artistic work. Thus, AI work in Interactive Storytelling combines aspects of several different roles from a humanities perspective, including narrative theorist, literary critic, and creative author.

In fact, given the technical nature of AI work, the creative role of the AI Interactive Storytelling researcher can be further broken down in creative and technical aspects, thus opening yet another dimension (Table 1). As the technical

Table 1 Roles of typical AI researcher as seen from the humanities

Roles in the humanities	Narratologist	Literary critic	Creative author	Printer
Dimension	Structural	Aesthetic	Creative	Technical
Roles in AI	Interactive Storytelling researcher			

programmer, the AI researcher is also the equivalent of camera technician doubling as movie director. In a literary setting, this would equate to a consideration of print technology and the materiality of books. While edge cases do exist in regard to the treatment of artists' books (however these are often seen as an entirely separate form from the perspective of fine arts (Drucker 2004)), such aspects are not normally considered in literary studies. Technical aspects are more likely to be covered in areas such as film studies, where they translate into particular aesthetic presentations and resulting effects on the audience. For example, the film scholar Robert L Carringer foregrounds technical aspects – such as wide angle lenses, specific lightning techniques, and camera movements – in his volume on the making of *Citizen Kane* (Carringer 1996).

In this case, the humanities can help the AI researcher to better understand the many different dimensions of work in this field. These can be used as research directions, or as the basis of defined roles in a research team.

Recent Developments

Since about 2005, AI researchers have increasingly investigated and employed models outside the literary domain and moved away from the focus on narrative structures. One such effort attempts to identify structures in improvisational theatre for use in Interactive Storytelling (e.g., Fuller and Magerko 2010). The goal here is to model the cognitive understanding shared between performers and apply the results in a system with virtual actors. In a similar vein, research into emergent narrative (Louchart and Aylett 2004) centers on virtual autonomous actors. Other efforts focus on modeling the author/director using AI planning methods (Porteous and Cavazza 2009; Riedl and Young 2006; Young et al. 2004). Another approach has instead focused on particular aspects of narrative, for example, the already mentioned research by Szilas and Richle on dramatic tension (Szilas and Richle 2013) or O'Neil and Riedl's work on modeling suspense (O'Neill and Riedl 2014).

As already mentioned, Michael Mateas, the creator of *Façade,* has focused on the notion of "social games" in the more recent work *Prom Week* (McCoy et al.2011). The model of social situations used is derived from research in the social sciences.

A trend in recent years is the development of specific computational models of narrative. Researchers in this vein seem to have decided that the differences in scope and methodology make the adaptation of traditional narratological models impractical. Their efforts amount to the creation of a specific field of computational narratology, often based on cognitive science's understanding of narrative – here they intersect with the humanities again in the form of narratologists like David

Herman, who sees narrative as a basic cognitive frame and understands cognitive sciences and narratology as two different perspectives on the same issue. Examples for work in this manner are, for example, George Lakoff and Srini Narayanan who propose a "Computational Model of Narrative" (Lakoff and Narayanan 2010) based on a shared "cognitive structure of human motivation, actions, goals, and events." Another example is Chris Martens et al. (2014) who explore a concept of "Proofs as Stories" based on ideas by Bosser et al. (2010) in an application of linear logic for narrative, while Tory Anderson (2015) explores the implications of the concept of Episodic Memory.

Opportunities

Finally, I like to offer suggestions for future work exploring humanities models and concepts for AI approaches to Interactive Storytelling. I have already mentioned the idea of documenting the process of adapting humanities models to AI implementation, with the hope to identify "translation rules."

Janet Murray (2015) reminds us that the humanities can also provide partial models – narrative concepts of particular dramatic situations, genre-specific language, or character pairings. Accordingly, she proposes a strategy to abstract the "women between two suitors" situation found in literature as a formalism for AI characters.

In addition, there are also many complete narrative models that present opportunities for AI treatment. For example, French neoclassical drama of Moliere and Corneille appears to be a particularly good candidate with its strict set of rules. On the other end of the spectrum, the theatre of the absurd (Beckett, Ionesco, Pinter) is testimony of successful narrative communication that purposefully undermines causal and temporal relations and conventional narrative development in order to expose the futility and absurdity of everyday life. Implementing the latter would mean a considerable shift from the causality-based models that have typically been used so far, but also from the goal of completeness; Clara Fernández-Vara's (2014) implementation of narrative dream puzzles is an important pointer in this direction.

Similarly on the trajectory of causal breaks, but well-specified through the detailed writings of its creator, Epic Theatre represents a particular opportunity for AI. In the 1930s and 1940s, the German playwright Bertolt Brecht developed a concept of non-Aristotelian drama (Brecht 1960). Influenced by the Japanese No theatre, Brecht was motivated by Marxist calls for the education of the masses. His form of drama was designed to make the audience reflect the depicted circumstances and ultimately turn to political action. With the experience of pity and fear and the resulting catharsis, Aristotelian drama arguably reinforces the status quo by demonstrating the futility of human effort in the light of fate. Brecht's concept wants the opposite, an audience educated about the need for action. And while Aristotle demands a complete action, a plot that needs nothing and cannot have anything taken away, Brecht's plot is episodic and open to continuation – ideally in the form of political action outside the theatre. Characters also have a different function in the epic theatre. Instead of being fully immersed in their role, they are supposed to "stand beside themselves", and comment on the character roles. Brecht likens this new kind of acting style to a street scene in which a motorist explains an accident to bystanders or the police. While she might act out certain parts of her report, she would also comment on them: "This is how I looked in the mirror and then this car came out of the side street..."

Several aspects make epic theatre an interesting model for Interactive Storytelling. The removal of the requirement for stringent causal connections between events is already helpful. However, a major change is in the treatment of characters. It is no longer the aim to be "in the role" to create a believable likeness, but instead to comment on the circumstances, both in and outside the role. As the creation of believable characters has been a stable of AI work in Interactive Storytelling, this change in perspective provides an interesting avenue for future work. While the emphasis on commentary would put higher demands on the character's understanding of the world, the effort for maintaining believability would be considerably reduced.

Conclusion

AI researchers in the field of Interactive Storytelling still have much to gain from the humanities – an enhanced understanding of the subject matter of narrative, models outside of literary studies, applications of additional fields such as psychology to narrative, but also considerations of authorship and knowledge in performance studies applicable to the understanding of interactive processes. However, the potential gains will be even greater when the relationship is defined as bidirectional. After all, both sides often work on the same key problems, as Gérvas et al. already pointed out in their pioneering collaboration of AI and narratology experts (Gervás et al. 2006). In this regard, the work of many AI researchers on aspects of narrative deserves proper recognition that can be framed as contributions towards a field of a media-specific narratology, "Interactive Storytelling Studies" in line with "Film Studies."

Recommended Reading

T.S. Anderson, From episodic memory to narrative in a cognitive architecture, in *6th Workshop on Computational Models of Narrative (CMN'15)*, eds. by M.A. Finlayson, B. Miller, A. Lieto, R. Ronfard. Atlanta, GA, USA, 2015 pp. 2–11. doi.10.4230/OASIcs.CMN.2015.2

R. Barthes, S/Z. (trans: Miller R). (Hill and Wang, New York, 1975)

D. Bordwell, *Poetics of Cinema* (Routledge, New York, 2007)

A.-G. Bosser, M. Cavazza, R. Champagnat, Linear logic for non-linear storytelling, in *Proceedings of the Nineteenth European Conference on Artificial Intelligence ECAI,* vol. 215, eds. by H. Coelho, R. Studer, M. Wooldridge. (ECAI, Lisbon, Portugal, 2010). doi.10.3233/978-1-60750-606-5-713, pp. 713–718

B. Brecht, *Kleines Organon für das Theater* (Suhrkamp, Frankfurt, 1960)

J. Butler, *Gender Trouble* (Routledge, London/New York, 1999)
R.L. Carringer, *The Making of Citizen Kane* (University of California Press, Oakland, 1996)
M. Cavazza, D. Pizzi, Narratology for interactive storytelling: a critical introduction, in *Technologies for Interactive Digital Storytelling and Entertainment* (Springer, Berlin/Heidelberg, 2006), pp. 72–83. doi:10.1007/11944577_7
S.B. Chatman, *Story and Discourse: Narrative Structure in Fiction and Film* (Cornell University Press, Ithaca/London, 1980)
J. Drucker, *The Century of Artists' Books* (Granary Books, New York, 2004)
T. Eagleton, *Criticism and Ideology: A Study in Marxist Literary Theory* (Verso, London/New York, 2006)
T. Eagleton, *Marxism and Literary Criticism* (Routledge, London, 2013)
U. Eco, *The Open Work*. (trans: Cancogni A). (Harvard University Press, Cambridge, MA, 1989).
M. Eskelinen, *Cybertext Poetics* (Bloomsbury Publishing USA, Indian Trail, New York, 2012)
S. Eisenstein (1969). Film Form: Essays in Film Theory. (J. Laeyda, Trans.). San Diego, CA: Harcourt Brace
G. Frasca, *Videogames of the Oppressed*. (2001)
D. Fuller, B. Magerko, Shared mental models in improvisational performance. Presented at the The Intelligent Narrative Technologies III Workshop, (ACM, New York, 2010). doi.10.1145/1822309.1822324
P. Gervás, B. Lönneker-Rodman, J.C. Meister, Narrative models: narratology meets artificial intelligence. Presented at the International Conference on Language Resources and Evaluation, Genoa, Italy (2006)
P. Gervas, C. Leon, G. Méndez, Schemas for narrative generation mined from existing descriptions of plot, in *6th Workshop on Computational Models of Narrative (CMN'15)*, ed. by M.A. Finlayson, B. Miller, A. Lieto, R. Ronfard (Dagstuhl Publishing, Atlanta, GA, USA, 2015), pp. 54–70. doi:10.4230/OASIcs.CMN'15.54
D. Herman, *Story Logic* (U of Nebraska Press, Lincoln, 2002)
P. Hühn et al., (ed.), *The Living Handbook of Narratology*. (n.d.). Retrieved 5 Oct 2015, from http://www.lhn.uni-hamburg.de
H. Koenitz, Five theses for interactive digital narrative, in *Interactive Storytelling: 7th International Conference on Interactive Digital Storytelling, ICIDS 2014, Singapore, Singapore, November 3–6, 2014, Proceedings,* vol. 8832, eds. by C. Fernández-Vara, A. Mitchell, D. Thue. (Springer International Publishing, Cham, 2014). doi.10.1007/978-3-319-12337-0_13, pp. 134–139
H. Koenitz, Towards a specific theory of interactive digital narrative, in *Interactive Digital Narrative*, ed. by H. Koenitz, G. Ferri, M. Haahr, D. Sezen, T.I. Sezen (Routledge, New York, 2015), pp. 91–105
J. Kristeva, *Desire in Language* (Columbia University Press, New York, 1980)
J. Lacan, *Ecrits* (W. W. Norton & Company, New York, 2003)
G. Lakoff, S. Narayanan, Toward a computational model of narrative, in *Computational Models of Narrative Papers from the AAAI Fall Symposium*. Palo Alto: . . . Symposium: Computational Models of Narrative. (2010)
B. Laurel, *Toward the Design of a Computer-Based Interactive Fantasy System* (Ohio State University, Columbus, 1986)
B. Laurel, *Computers as Theatre* (Addison-Wesley, Boston, 1991)
S. Louchart, R. Aylett, Narrative theory and emergent interactive narrative. Int. J. Cont. Eng. Edu. Life Long Learn. **14**(6), 506 (2004)
G. Lukács, *The Historical Novel. 1937* (University of Nebraska Press, Lincoln, 1983)
C. Martens, J.F. Ferreira, A.G. Bosser, M. Cavazza, Generative story worlds as linear logic programs, in *Intelligent Narrative Technologies 7: Papers From the Workshop*. (Intelligent Narrative Technologies 7, Palo Alto, 2014)
M. Mateas, A preliminary poetics for interactive drama and games. Digit. Creat. **12**(3), 140–152 (2001). doi:10.1076/digc.12.3.140.3224

M. Mateas, A. Stern, Procedural authorship: a case-study of the interactive drama Façade. Presented at the Digital Arts and Culture, (2005a)
M. Mateas, A. Stern, Structuring content in the Façade interactive drama architecture. Presented at the AIIDE, (2005b)
J. McCoy, M. Mateas, N. Wardrip-Fruin, Comme il Faut: a system for simulating social games between autonomous characters, in *Proceedings of the Digital Arts and Culture Conference DAC*. Digital Arts and Culture 2009 (Irvine, CA, USA, 2009)
J. McCoy, M. Treanor, B. Samuel, M. Mateas, N. Wardrip-Fruin, Prom week: social physics as gameplay, in *Foundations of Digital Games* (ACM, Bordeaux, 2011). doi.10.1145/2159365.2159425, pp. 319–321
J. McCoy, M. Treanor, B. Samuel, A.A. Reed, M. Mateas, Prom week: designing past the game/story dilemma. in *Foundations of Digital Games 2013*. (2013)
T. Moi, *Sexual/Textual Politics: Feminist Literary Theory* (Routledge, London/New York, 2002)
J.H. Murray, *Hamlet on the Holodeck: The Future of Narrative in Cyberspace* (Free Press, New York, 1997)
J.H. Murray, A tale of two boyfriends: a literary abstraction strategy for creating meaningful character variation, in *Interactive Digital Narrative*, ed. by H. Koenitz, G. Ferri, M. Haahr, D. Sezen, T.I. Sezen (Routledge, New York, 2015)
M. Nitsche, *Video Game Spaces* (MIT Press, Cambridge, MA, 2008)
B. O'Neill, M. Riedl, Dramatis: A computational model of suspense. Presented at the 28th AAAI Conference on Artificial Intelligence, Quebec City, 2014
J. Porteous, M. Cavazza, Controlling narrative generation with planning trajectories: the role of constraints, in *Interactive Storytelling* (Springer, Berlin/Heidelberg, 2009), pp. 234–245. doi:10.1007/978-3-642-10643-9_28
G. Prince, Narrative analysis and narratology. N. Lit. Hist., **13**(2), 179–188. (1982). Retrieved from http://www.jstor.org/stable/468908
G. Prince, *A Dictionary of Narratology* (U of Nebraska Press, Lincoln, 2003)
V. Propp, *Morphology of the Folktale* (University of Texas Press, Austin, 1928)
M.O. Riedl, R.M. Young, From linear story generation to branching story graphs. IEEE Comput. Graph. Appl. **26**(3), 23–31 (2006). doi:10.1109/MCG.2006.56
D. Robert, S. Shenhav, Fundamental assumptions in narrative analysis: mapping the field. Qual. Rep. **19**, 1–17 (2014)
K.K. Ruthven, *Feminist Literary Studies* (Cambridge University Press, Cambridge, 1990)
D.H. Schram, G. Steen, *The Psychology and Sociology of Literature* (John Benjamins Publishing, Amsterdam, 2001)
S. Simpkins, *Literary Semiotics* (Lexington Books, Lanham, 2001)
R.P. Sugg, *Jungian Literary Criticism* (Northwestern University Press, Evanston, 1992)
N. Szilas, U.M. Richle, Towards a computational model of dramatic tension. in *Workshop on Computational Models of Narrative 2013*. (Hamburg, 2013). doi.10.4230/OASIcs.CMN.2013.257
C.F. Vara, Creating Dreamlike Game Worlds Through Procedural Content Generation, in *Intelligent Narrative Technologies 7: Papers from the Workshop*. Association for the Advancement of Artificial Intelligence, (2014)
Y. Wilks, Form and content in semantics, in *Computational Linguistics and Formal Semantics*, ed. by R. Johnson, M. Rosner (Cambridge University Press, Cambridge, 1992)
R.M. Young, M.O. Riedl, M. Branly, A. Jhala, R.J. Martin, C.J. Saretto, An architecture for integrating plan-based behavior generation with interactive game environments. J. Game. Dev. **1**(1), 51–70 (2004)

Introduction to Interactive Storytelling 17

Marc Cavazza and R. Michael Young

Contents

Abstract

Interactive Storytelling is a new media endeavor, which aims at developing technology and content for a new kind of narrative experience, one in which a story's unfolding can be influenced by its audience. The interdisciplinary nature of Interactive Storytelling (IS) has led it to revisit many issues such as narrative structure, visual discourse, role of characters, and emotion elicitation, from the perspective of its underlying techniques, artificial intelligence, 3D animation, and user interfaces. In this introductory chapter, we discuss core concepts around which the discipline has organized itself, which are both technical challenges as well as dividing lines between competing approaches. After introducing a framework characterizing user experience through a combination of immersion and involvement, we introduce the main concepts of IS that will be revisited throughout this section. Among these are longstanding discussions on the respective roles of characters and plot in narrative structure, as well as the relationship between

M. Cavazza (✉)
School of Electronics and Digital Arts, University of Kent, Canterbury, UK
e-mail: M.O.Cavazza@kent.ac.uk

R.M. Young
North Carolina State University, Raleigh, NC, USA
e-mail: young@csc.ncsu.edu

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_55

story and discourse, both of which strongly determine the specific articulation of technology and content that constitutes an Interactive Narrative. Finally, we announce the various chapters composing this section and their objectives.

Keywords
Interactive Storytelling (IS) • Interactive narratives • Plot-centric approach • Narrative arc

Interactive Narratives are stories whose unfolding, pace, and outcome can be influenced by some intervention by a spectator or an entire audience. They can be defined by contrast with linear narratives, which, regardless of the medium in which they are presented, follow inflexibly the original design and author's intention. There has been a longstanding interest in producing nonlinear narratives, and our intention here is not to provide a historical account of the field, which can be traced back to Činčera et al. (1967) *Kinoautomat*.

Instead, we will introduce the main concepts of Interactive Storytelling, a recent field of research in Entertainment Computing, which has revisited the Interactive Narrative endeavor by taking advantage of progress in artificial intelligence (AI), as well as visualization techniques based on computer graphics. What distinguishes it from early attempts at developing branching videos or movies is to seek a principled approach to both the generation of the storyline and visual content. AI techniques, in particular planning, support the generation of narrative action sequences in a way that preserves narrative causality in the presence of user intervention. The production of alternative visual content that could account for all possible story variants has been for long a major bottleneck in the development of Interactive Narratives: the advent of computer graphics and animation has made it possible to generate alternative content in a much more efficient fashion. In particular, game engines have been widely adopted in Interactive Storytelling research for the integrated environment they provide and their support to character animation and staging.

In the literature, it is possible to find terms such as *Interactive Storytelling*, *Virtual Storytelling*, *Digital Storytelling*, *Interactive Drama*, *Interactive Narrative* to name but a few, being used apparently interchangeably. To avoid any confusion in this chapter, some brief terminological clarification is required. We shall restrict the term *Interactive Storytelling*, or *IS*, to refer to the research topic itself, and we shall use *IS system* to designate a complete prototype system. Furthermore, we reserve the term *Interactive Narrative*, or *IN*, to designate an end product, that is, the media object by which the user experiences the presentation of a narrative.

To provide a consistent presentation, we will largely base this chapter on our own work, together with current or former colleagues, rather than attempt a comprehensive review of the topic. This will inevitably come with a specific perspective, one we however think is the most relevant to the audience of this handbook, as it firmly grounded in the computing techniques that support Interactive Narratives, as well as their integration.

It is often argued that the objective of Interactive Storytelling is neither achievable nor desirable. These counterarguments generally build upon a traditional vision of the narrative experience as depending entirely on authorial talent, which by definition leaves no room to external intervention. The recent history of new media, in particular computer games, brings a more nuanced perspective, as well as the need to extend the traditional definition of narrative experience. To a large extent, computer games have been following an opposite, yet convergent path towards Interactive Narrative, starting from a highly interactive medium to progressively improve the narrative content and experience.

Throughout this section, we will adopt a more pragmatic stance considering that many of these theoretical issues are actually premature, until Interactive Narratives comparable in duration and richness of experience to traditional narratives become available. In this chapter, we will consider Interactive Narratives as an evolution of linear stories allowing user intervention: even though the visual medium used to stage this story may be reminiscent of computer games, we will exclude from our definition of Interactive Storytelling contemporary commercial games regardless of their claims to narrative experience obtained through cut scenes and other cinematic techniques including branching endings or "quick time events."

The arguments against interactive narratives stem from two different perspectives. The first one is the author's exclusive ability to craft a compelling story and sometimes their inalienable right to authorship, which cannot accommodate audience participation. A common concern is that user interference, if uncontrolled, might destroy an author's work. Consider for example, an IN based on the plays of Shakespeare, which allows for the possibility of different versions of *Romeo and Juliet* with endings that have the couple living happily together or of *The Merchant of Venice* where Shylock gives up his bond in return for Antonio's friendship. How far would this actually betray Shakespeare? The situation is not necessarily any better with contemporary authors: on what basis would they relinquish control over the unfolding of a plot that they have designed and that they see as their creation?

The second is the spectator's perspective: while it is sometimes argued that narrative experience is incompatible with interaction, there are certainly dissenting views emanating even from those less suspect of worshipping technology. For instance, in his classical account of suspense in Hitchcock films, Truffaut (1983) proposed one of the earliest rationales for interaction, suggesting that "[the] audience is longing to warn the characters on the screen."

Finally, some of this debate has taken a different form by prompting a theory of interactive narrative, again in our view premature in the absence of scalable systems. One recurrent theme is the specific status of the interactive narrative author, which we should address briefly as part of our list of core topics for Interactive Storytelling.

There is no unique or universally accepted definition of Interactive Storytelling, in part because it can be implemented through several different paradigms, each of which is characterized by different combinations of user involvement, narrative generation, story visualization and presentation, and interaction design. This is why we will not attempt to introduce another definition beyond the minimalistic one that opened this chapter and will concentrate instead on some key concepts. In our view, the best introduction is to discuss the typical components of an Interactive Storytelling system and the various paradigms for Interactive Storytelling, which provide different examples of how they can be integrated.

An example overview of the architecture of a typical IS system is shown in Fig. 1. It illustrates the core functionality required in an IS system, regardless of the technical solutions employed. Shown in the figure are: a narrative engine that is capable of dealing with generation (and regeneration) of plot instantiations; an interaction component that enables users to interact with the system and interfere with the unfolding plot (if and when they choose); and a real-time visualization engine where the plot is presented to the audience (or "executed"), for example, using 2D or 3D graphics but this could also use filmic video content (Piacenza et al. 2011; Ursu et al. 2007). Consequently, a key task of an IS system is to manage the real-time interleaving of these different components.

While all IS systems share the dynamic generation of a storyline taking into account potential user interaction, there exists multiple ways in which these two aspects can be integrated in practice. Beyond simple variation in the technical integration of the generative and the interactive components, it is possible to define paradigms of IS which can be characterized by how integration is perceived from a usability perspective.

To define these paradigms we introduce two dimensions along which systems can be categorized:

1. The degree of involvement of a user in an IN. This ranges from occasional intervention determined by the user to full-time involvement dictated by the narrative itself (e.g., if the user is playing the part of a member of the cast). This cannot be distinguished from the diegetic (An element is termed diegetic if it is part of the story world itself, and non-diegetic if it is added to the story presentation. This distinction has been popularized by the description of sounds constituting a film's soundtrack. For instance, in the film *Apocalypse Now*, the *Ride of the Valkyries* is a diegetic sound, since it is actually played on board the attack helicopters. The same piece used during the car chase of the *Blues Brothers* is part of the added soundtrack, hence non-diegetic.) nature of user interaction. However, it is important here to distinguish narrative concepts from interactive ones (e.g., gameplay). For instance, any user utterances that form part of story dialogues are diegetic; any motion of a user's avatar that forms part of the visual story is diegetic as well.
2. The level of immersion in the story world, which can be described as onstage or offstage, internal or external. It should be noted that since the earliest systems (Cavazza et al. 2001) there have been onstage modes which were not diegetic: these were "ghost modes" in which the user could interfere through an invisible avatar. It may be difficult to define precise boundaries of a story in which the user is immersed – the test in such cases should be from an external observer watching an IN without intervening.

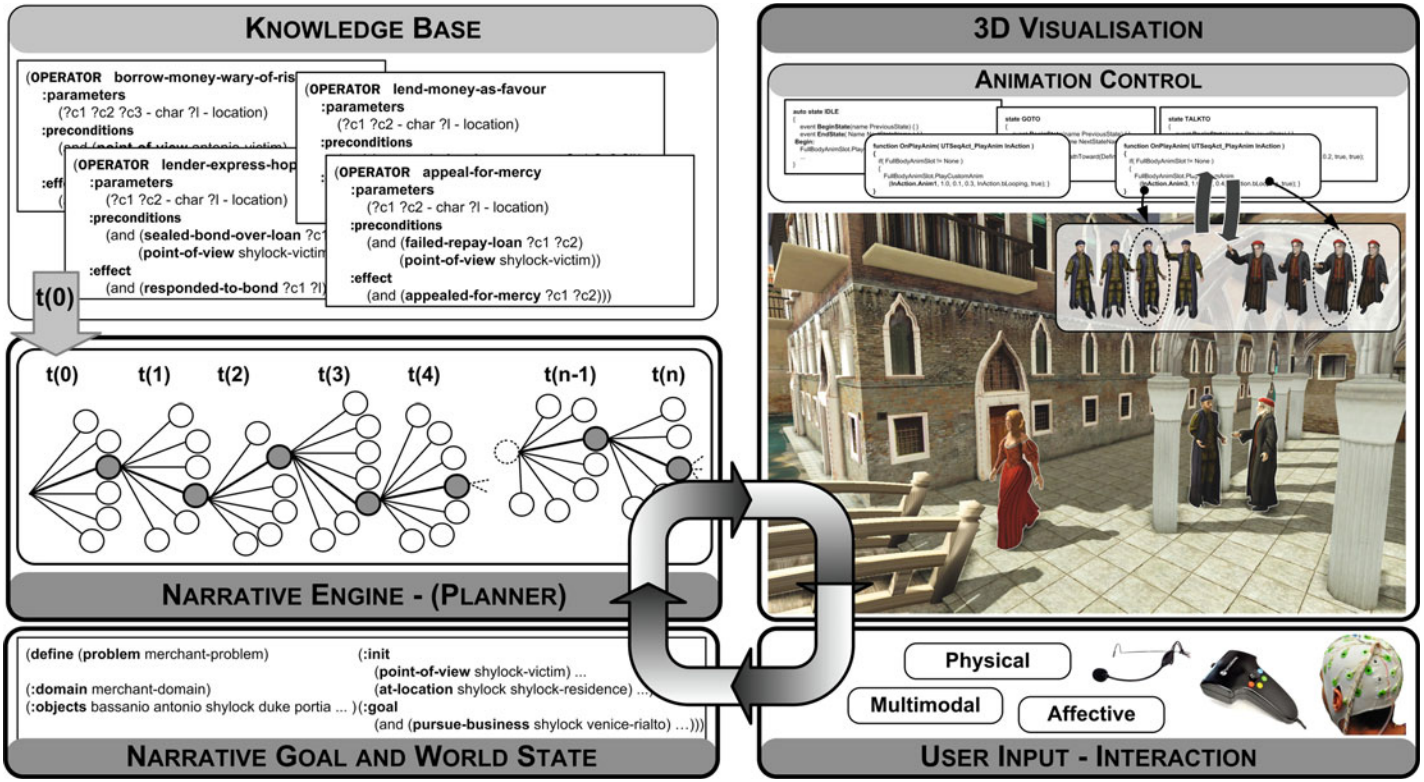


Fig. 1 Architecture of an Interactive Storytelling system. Narrative engine input consists of a knowledge base for the narrative world, current world state, and narrative goal; as the narrative is generated, actions are passed to animation control for 3D visualization; user interaction, along with changes in the 3D world, triggers narrative regeneration

Fig. 2 An illustration of different degrees of user involvement in an interactive narrative: (*1*) first person view, where the user impersonates a character in the narrative and is constantly involved in the action, and (*2*) "ghost" mode view, where the user is able to interfere with the narrative through an invisible avatar

Table 1 A classification of IS prototype systems by interaction paradigm

	Anytime interaction (non-diegetic)	Full-time involvement (diegetic)
First person (onstage)	VR Emma (ghost mode) (Lugrin et al. 2010) (Cavazza et al. 2007)	VR Emma (actor mode) (Lugrin et al. 2010) AR Façade (Dow et al. 2007)
First person (external)		Façade (Mateas and Stern 2005) MRE (Swartout et al. 2001) EmoEmma (Cavazza et al. 2009)
External (god mode)	I-storytelling ("Friends") (Cavazza et al. 2002) Prom Week (McCoy et al. 2013)	
Third person		MRIS (Cavazza et al. 2003)

These dimensions are of course not congruent: while a user impersonating one of the characters has to be onstage and interact regularly at a rhythm dictated by the immersive IS systems, the user can jump on stage but be invisible and interact at anytime but at their discretion. This is most often implemented via physical interaction, especially with virtual stage objects, but recently we have also extended it to affective interaction via emotional speech recognition (Lugrin et al. 2010).

As an illustration, Fig. 2 shows two different degrees of user involvement and immersion in an Interactive Narrative. The left hand side of the figure shows a first person view where the user is playing the part of a character in the narrative and is constantly involved in the action. The right hand side of the figure shows a "ghost" mode view where the user is able to interfere with the narrative through an invisible avatar. A summary of Interactive Storytelling paradigms adopted in previous systems is shown in Table 1.

Balancing Interaction and Narrative

In creating Interactive Narratives, there is always a dichotomy between interaction and narrative. This idea was discussed by Ryan (2001), who tried to reconcile immersion and interactivity by the description of the involvement of the body. She argued that interactivity and immersion support each other in virtual reality, though they conflict in literature because interactivity requires an awareness of semiotics, while immersion depends on their disappearance. For instance, the central idea underlying a subarea of IS sometimes referred to as interactive drama is to minimize the difference between author, spectator, actor, and character, as the user steps onto the stage, assumes the role of a character, and influences the development of the plot through their speech and actions. A typical example of such an immersive environment is a Holodeck™-like setting (Cavazza et al. 2007; Swartout et al. 2001; Murray 1997), where users can be immersed in a story, of which they become a character, while the narrative revolves around their actions.

In such systems, where users have freedom of interaction, it is possible for them to make choices that conflict with the story intended by the author: a situation referred to in IS as the "narrative paradox" (Aylett 2000). This tension between narrative and interaction exemplifies the original incompatibility between linear media such as film, in which the spectators' experience is entirely under the control of the authors and interactive media. At this point, let us first dispel a myth, according to which it is possible for complex narratives to emerge from the arbitrary interactions of autonomous characters gifted with intelligence, emotions, and generic motivations and that constitute the cast of an Interactive Narrative. This idea, which could be dubbed "The Sims™ fallacy," arises from confusion surrounding the attribution of narrative properties to event sequences that have been generated on the basis of local principles whereas true narrative properties relate to global principles, which are not considered in this scenario. To some extent, the approach known as "emergent storytelling" (Aylett et al. 2006) shares these limitations, although the definition of specific characters' feelings does account for some narrative plausibility.

As observed by Young (1999), when interaction results in changes to the narrative world that were not anticipated when the narrative was generated, then the system can respond in a number of ways. For instance, it might intervene in the virtual world by preventing the consequences of a user's actions (e.g., a gun misfires), or it may take some measures to get the story "back on track" by performing some form of repair. The decision of whether the system should repair a narrative by retaining as much of the current narrative as possible or by abandoning it in favor of regeneration depends on factors such as the interaction paradigm underlying the system and the extent to which plan failure is visible to an audience (e.g., in some story genres, such as quests, plan failure can be used to dramatic effect).

Clearly a key challenge in Interactive Storytelling is how best to propagate the consequences of user interaction, while maintaining shared control between user and system, as highlighted by Young (2000). An IS system's manipulation of the consequences of a user's actions may, in some cases, be masked by the lack of control a player has over the outcome of their own actions. For instance, when the system causes a crossbow bolt fired at an opponent to miss its target, a user may easily attribute the missed shot to their own error. But such an approach, when exposed to a user, can break their sense of agency, damaging the user experience. Alternatively, when a system decides to repair a narrative, careful consideration must be given to how much of the previously generated narrative to retain and what unexecuted portion, if any, should be disregarded. Regenerating a narrative may result in previously experienced portions of the story no longer playing a causal role in those new parts of the story yet to be experienced. Because global coherence is a critical contributor to an IS system's user experience, too many action sequences that appear to the user to be unmotivated can have a significant negative impact.

Additional means of manipulating a story world in response to a user's unanticipated action present different benefits and challenges. For example, by anticipating a user's likely future story actions, an IS system can manipulate system characters to alter the state of the story world in advance of the user's attempt, making errant actions impossible to execute (Harris and Young 2009). If a system models the user's knowledge of the story world's conditions, for instance, by tracking their observations as they move through the virtual environment, the system can manipulate conditions in the past which the user had not directly observed, effectively rewriting the historical record to prune off options for unwanted actions before they can be attempted (Robertson and Young 2014). Similarly, in story worlds where users are unfamiliar with the rules governing actions in the domain, the system can choose to manipulate those rules to alter the conditions necessary for an action's execution (Robertson and Young 2015).

Story and Discourse

Implicit in the design of many IS systems is the relation between a system-generated user experience and the structure of narrative itself. As indicated above, our intent here is not to advance a theory of interactive narratology; however, one significant way that work by IS researchers can be factored is by characterizing its relation to parallel narrative theoretic elements. As generally characterized by narrative theorists (e.g., (Chatman 1990)), a narrative can be viewed as being composed of two distinct parts: the story and the discourse. A narrative's story typically is defined as consisting of the elements of plot and setting – all the actions, locations, objects, and characters that populate the world of the narrative. The narrative discourse, then, is composed of the medium-specific communicative elements used by an author to convey the underlying story to a reader or viewer. As described by Young (2007), elements of IS systems can explicitly or implicitly address the generation of each or both of these categories.

IS systems that address story generation typically focus on the creation of action sequences – plot lines – that characterize actions for both the user as well as system-controlled characters. As we discuss here, AI planning approaches have been used to strong effect to automate the synthesis of a story's action sequences (Cavazza, et al 2000). Intuitively, there's a clear match between the functions computed by automatic planning systems, which generate valid plans to transform a world from its current state into a desired one, and the result of an author's design of a plot in which a collection of characters pursue their own goals. At the core of both lies the importance of causality, as we discuss below, and the role that causal validity (or its intentional lack) plays in the output of each.

Increasingly, researchers building plan-based story generation systems are introducing variations on or extensions to classical planning representations in order to address structural requirements unique to Interactive Storytelling. These algorithms reason about elements not typically modeled by AI planning systems, though their relevance to that body of work may have significance in terms of efficiency and/or expressivity. For example, Riedl and Young's work (Riedl and Young 2010) on the generation of plans that model characters' intentions towards the actions they take in a story produces narratives that are perceived as significantly more coherent. Ware and Young's work on planning algorithms that create and maintain harmful interactions between agents in a story produces IS experiences supporting users' perceptions of narrative conflict (Ware and Young 2014). Work by Porteous et al. (2011) clearly demonstrates the value of authorial control of plan structure (i.e., planning using landmarks) in both general purpose planning as well as story generation in IS systems.

In contrast to story's focus on the dynamics of a story world, narrative discourse focuses on the dynamics of a reader's experience. In the case of IS systems, this focus involves reasoning about the communicative actions presented to the user as a way of providing a window in to the unfolding storyline. Systems that address narrative discourse generation range in level of detail from those that focus on the low-level realization of the presentation of the story world (e.g., (Lino et al 2010; Luckin et al 2015)) to high-level reasoning about the structure of the discourse, its content, and organization (e.g., (Jhala and Young 2010)). In the context of IS, we adopt a view of discourse as a conversation between system and user (Young 2004; Cardona-Rivera and Young 2013). To that end, some models make use of ideas first developed by philosophers of language to characterize communication as planned, intentional action (Searle 1969). These approaches build on architectures used in natural language discourse generation – typically in the context of instructional text genres – employing planning systems to reason about the construction of discourse plans. These plans, containing sequences of communicative action rather than story action, are designed to communicate the unfolding story to a user (Jhala and Young 2011). A central challenge for this approach is to combine the idioms of narrative and IS with expressive models detailing the comprehension process of a user. Discourse is typically viewed as addressed in ways that impact a reader or viewer's comprehension. One key to future progress in narrative discourse generation will be the effective translation of what we know about narrative comprehension to specifications for the communicative actions in an IS system and how the selection of those actions can prompt specific inferences about story and interaction (Cardona-Rivera and Young 2014).

Causality

Causality is a fundamental feature of narrative: story worlds feature causal laws under which they operate and without which stories would be reduced to mere sequences of events (Forster 1963; Richardson 1997). As observed by Aylett et al. (2011), these causal rules relate not only to world-based causal structures but can also relate to cause in affective changes in story characters.

It is acknowledged that IS systems should preserve story causality – something that has been recognized as the backbone of a meaningful narrative (Barthes 1966), demonstrated to be an important factor in user experience and the construction of a user's mental model of a story (Trabasso and van den Broek 1985) and a component in the degree to which stories remain understandable to users (Riedl and Young 2006, 2010). Indeed, causality plays an important role in all aspects of IS. Illustrations of this include: serving as an indicator of developmental appropriateness of narrative content (Alonso et al. 2011); playing a part in the rereading of interactive texts (Mitchell 2010); serving as the basis for a methodology for coherent design in role-playing games (Lindley and Eladhari 2002); generating scientific narratives from data models (Reitsma 2010); and mapping to existing cognitive models for the assessment of user story understanding (Christian and Young 2004). As mentioned above, representations of causality are the building blocks for many more complex notions that parallel narrative structures like character intentionality (Riedl and Young 2010), conflict (Ware and Young 2014) and character choice (Bahamón and Young 2013).

However, maintenance of causality is far from straightforward given the possibility that user actions may have a negative impact on a planned and causally coherent narrative (for example, user actions might break causal links between events). As a consequence, an enduring challenge in IS has been to find solutions to the problem of maintaining narrative coherence in an interactive environment throughout the course of the development of a story.

Duality of Character and Plot

A story can be considered from the perspective of its featured characters or from a more abstract perspective at the global level of the plot. This distinction is genre- and media-dependent; for instance, the development of cinema has certainly emphasized the perspective of characters, while in early Aristotelian drama there was little emphasis on characters' agency (Cavazza and Pizzi 2006). This duality has filtered through to the technical solutions that underpin IS systems and can be seen in the split between IS frameworks that are character- or plot-centric:

Character-based IS: There exists a perspective in IS which is strongly centered on characters and hence referred to in the literature as character-centric or character-based IS (Cavazza et al. 2002). Such frameworks are frequently simulation based and are populated by virtual characters that are themselves autonomous agents.

Advantages of a character-centric approach include a better representation of individual characters' narrative attributes (character's psychology, affective state, etc.); the provision of a modular description, which better supports the exploration of the impact of changes in story cast roles in a way which would reflect the role played by character design in story creation (McKee 1997); the production of more emergent narratives; and an easy route to generative scalability, which can be closer to some forms of story creation (Charles and Cavazza 2004). However, these advantages are at the expense of control over a narrative's global properties. Further, one of the major issues arising is the problem of simulating the interactions between these autonomous agents in real time.

Plot-based IS: An alternate approach is to see characters as, at most, only semiautonomous agents that are coordinated by a centralized AI system that takes care of aspects such as narrative control. This is usually recognized as a better proposed solution to the problem of authorial control in order to provide improved narrative coherence, compared to emergent systems that are less constrained in their construction of a user's experience over time. Additionally, a plot-centric approach can help facilitate the synchronization of multiple characters, control of user interaction (Riedl et al. 2003), as well as the explicit control of global narrative properties such as suspense (Cheong and Young 2015) and tension (Zagalo, et al. 2004).

The relationship between characters and the narrative they are part of has been recognized as a major problem in IS: the challenge being to achieve some degree of balance between characters' autonomy and the global structure of the plot.

Authoring Content for Interactive Storytelling Systems

One of the emerging issues of Interactive Storytelling has been the actual process by which an Interactive Narrative can be authored and whether it would bear any relationship to authoring processes of linear media. Although the authoring process would concern each module of a typical IS system as described above, we shall concentrate the discussion on narrative generation, one reason being that visual content production can follow a similar process than with linear narratives and interactive media. A large number of Interactive Storytelling prototypes have been inspired by the plot of a preexisting linear narrative (*Who's Afraid of Virginia Woolf* for Façade (Mateas and Stern 2005) to the *Merchant of Venice* (Porteous et al. 2010a)): this can be in part explained by the fact that they set up a precise challenge for narrative consistency as their baseline narratives are well known and also by Truffault's above argument that spectators of the original stories may feel an urge to alter the course of action, motivating them to experiment with an Interactive Narrative. This form of authoring can be termed the "narrative variant" version as it starts with an existing linear story that can be formalized at the level of its narrative actions, then extending this representation to include potential alternative actions ensuring story continuity despite user intervention. Such story creation takes place without a well-identified author, as it consists mainly of a knowledge elicitation exercise in which narrative actions from the original story are formalized in a computational representation supporting narrative generation. Sometimes, this process is assisted by using critical analysis material that helps the transition between the original plot and the knowledge representation. In one case, we have used draft material and commentary from the author himself (Flaubert) to facilitate the formalization of key narrative actions (Cavazza et al. 2009).

Several researchers have advocated that there should be a specific status for Interactive Narrative authors, although details of how this would support the entire creation process have remained rather elusive, and most proposals tend to sway towards branching narratives. Others have sought to resolve the paradox by advocating a theoretical model inspired from improvisational theatre: this can be explained as the use of narrative patterns containing and making sense of spontaneous actions. Similar models have sought inspiration in role-playing games, with the "game master" potentially embodied in drama management module. However, it should be noted that these approaches may limit the paradigms of IS to those requiring permanent user involvement in the narrative: they are also faced with difficulties when turning their theoretical model into requirements for specific software modules. While the status, or the actual need for, a specific author for Interactive Narratives remains elusive, we can consider the production process as similar to the one currently at play in computer games, in which game designers collaborate with programmers without having to concern themselves with technical aspects. Because Interactive Narratives are knowledge-based systems, the translation of high-level narrative actions and behavior specifications into an implemented system should actually be facilitated. The creation of an Interactive Narrative largely becomes a knowledge engineering problem. Some researchers have expressed concern with the difficulty for authors to master AI formalisms (such as planning formalisms), but this simply reproduces the misconception that authors or designers have to take direct control of the technical implementation, something disproved by the game production pipeline. Still, knowledge elicitation can be an arduous task in itself, at several levels. One is the consistency of the planning domain, which can be difficult to maintain manually when the creation of predicates reflects narrative elements rather than real-world, well-defined entities as in traditional planning applications. This process can be assisted by authoring interfaces that check the consistency of the planning domain and support step-by-step execution of the Plan within a visual debugging environment. However, the main challenge of story creation is the ability to control the actual *narrative arc*, which determines major aesthetic properties of the narrative, within the Interactive Narrative context. When using planning for narrative generation, the narrative arc is actually determined by the *plan trajectory*, which corresponds to the shape of the solution plan. This means that a set of high-level constraints on the evolution of the narrative has to be preserved despite user interventions modifying the environment. The chapter on Planning Technologies for Interactive Storytelling contains detailed examples of authoring tools, one in each of the above categories. It will in particular illustrate visual interfaces for controlling the narrative arc based on the use of landmark planning (Porteous et al. 2010b).

Conclusion

We have proposed a framework for Interactive Storytelling based upon a small number of core issues, which have been identified through reviewing recent progress in the field. Our standpoint is that of interactive narrative technologies, and for that reason we have left aside the content production process as well as a discussion of narrative theories as a foundation to Interactive Storytelling. We appreciate that the latter has played an important role in the development of the field, even though Interactive Storytelling cannot be grounded in computational narratology in the same way as natural language processing has developed from computational linguistics. Readers with an interest in narratology and how it relates to technical aspects are referred to Cavazza and Pizzi (2006).

We have also introduced some traditional topics, such as the duality between character and plot, or between story and discourse, which should help the reader to gain access to the abundant literature in the field, including some seminal papers.

The three chapters in this section each address one core issue. On narrative generation, Porteous reviews the role of planning and how it became the dominant AI technique: she also explains the impact that modern planning technologies are having on narrative generation. Cavazza and Charles review user interaction techniques that can support intervention in an Interactive Narrative: they specifically introduce various interaction paradigms, which help categorize existing as well as future approaches. Jhala specifically addresses the relationship between story and discourse and the techniques by which story presentation can be automated so that Interactive Narratives too can benefit from editing and staging strategies. Koenitz discusses how the contributions of humanities to the study of narratives can inform the IS endeavor.

Finally, although we have adopted a personal standpoint on Interactive Storytelling based on our own experience, we trust that this framework should enable the reader to enter the field from a critical perspective, including approaches not discussed here, and develop her own vision and contribution to this exciting endeavor.

Recommended Reading

J. Alonso, A. Chang, C. Breazeal, Values impacting the design of an adaptive educational storybook, in *Proceedings 4th International Conference on Interactive Digital Storytelling (ICIDS 2011)* (Springer, Berlin, 2011), pp. 350–353

R. Aylett, Emergent narrative, social immersion and "storification," in *Proceedings of the Narrative and Learning Environments Conference (NILE 2000)*, Edinburgh, 2000.

R. Aylett, J. Dias, A. Paiva, An affectively driven planner for synthetic characters, in *Proceedings of the Sixteenth International Conference on Automated Planning and Scheduling (ICAPS 2006)*, 2006, Ambleside, UK, pp. 2–10.

R. Aylett, S. Louchart, A. Weallans, Research in interactive drama environments, role-play and story-telling, in *Proceedings Fourth International Conference on Interactive Digital Storytelling (ICIDS 2011)*, (Springer, 2011), Vancouver, British Columbia, Canada, pp. 1–12

J. Arnav, R.M. Young, Intelligent Machinima Generation for Visual Storytelling, Gonzales-Calero, Pedro Antonio; Gomez-Martin, Marco Antonio (Ed.): Artificial Intelligence for Computer Games, pp. 151-170, Springer, New York, 2011.

J.C. Bahamón, R.M. Young, CB-POCL: a choice-based algorithm for character personality in planning-based narrative generation, in *Proceedings of the 2013 Workshop on Computational Models of Narrative* (CMN2013) *a Satellite Workshop of the 35th Annual Meeting of the Cognitive Science Society* (CogSci 2013), Hamburg, 2013, pp. 4–23.

R. Barthes, Introduction a l'analyse Structurale des Récits. *Communications*, 8(1), 1966. (In French), pp.1-27.

R.E. Cardona-Rivera, R.M. Young, A cognitivist theory of affordances for games, in *Proceedings of the Digital Games Research Conference: DeFragging Game Studies (DiGRA2013)*, Atlanta, 2013

R.E Cardona-Rivera, R.M. Young, Games as conversation, in *The working notes of the AIIDE Workshop on Games and Natural Language Processing (GAMENLP 2014)*, Raleigh, 2014

M. Cavazza, D. Pizzi, Narratology for interactive storytelling: a critical introduction, in *Technologies for Interactive Digital Storytelling and Entertainment (TIDSE 2006)* (Springer, Berlin/Heidelberg, 2006), pp. 72–83

M. Cavazza, R. Aylett, K. Dautenhahn, C. Fencott, F. Charles, *Interactive Storytelling in Virtual Environments: Building the "Holodeck" 6th International Conference on Virtual Systems and Multimedia (VSMM'2000)*, Gifu, 2000

M. Cavazza, F. Charles, S.J. Mead, Characters in search of an author: AI-based virtual storytelling, in *Virtual Storytelling Using Virtual Reality Technologies for Storytelling* (Springer, Berlin/Heidelberg, 2001), pp. 145–154

M. Cavazza, F. Charles, S.J. Mead, Character-based interactive storytelling. IEEE Intell. Syst. **17** (4), 17–24 (2002)

M. Cavazza, O. Martin, F. Charles, S.J. Mead, X. Marichal, Interacting with virtual agents in mixed reality interactive storytelling, in *Intelligent Virtual Agents (IVA 2003)* (Springer, Berlin/Heidelberg, 2003), pp. 231–235

M. Cavazza, J.L. Lugrin, D. Pizzi, F. Charles, Madame bovary on the holodeck: immersive interactive storytelling, in *Proceedings of the 15th international conference on Multimedia*, (ACM, 2007), Augsburg, Germany, pp. 651–660

M. Cavazza, D. Pizzi, F. Charles, T. Vogt, E. André, Emotional input for character-based interactive storytelling, in *Proceedings of the 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2009)*, Budapest, 2009, pp. 313–320

F. Charles, M. Cavazza, Exploring the scalability of character-based storytelling, in *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2004)*, New York, 2004, pp. 872–879

S. Chatman, *Story and Discourse: Narrative Structure in Fiction and Film* (Cornell University Press, Ithaca, 1990)

Y. Cheong, R.M. Young, Suspenser: a story generation system for suspense. IEEE Trans. Comput. Intell. AI Games **7**(1), 39–52 (2015)

D. Christian, R.M Young, Comparing cognitive and computational models of narrative structure, in *Proceedings 19th National Conference on Artificial Intelligence (AAAI 2004)*, (AAAI Press, 2004), Menlo Park, CA, pp. 385–390

S. Dow, M. Mehta, B. MacIntyre, M. Mateas, AR façade: an augmented reality interacitve drama, in *Proceedings of the 2007 ACM Symposium on Virtual Reality Software and Technology*, (ACM, 2007), San Jose, CA, pp. 215–216

E.M. Forster, *Aspects of the Novel* (Penguin, Harmondsworth, 1963)
J. Harris, R.M. Young, Proactive mediation in plan-based narrative environments. IEEE Trans. Comput. Intell. AI Games **1**(3), 233–244 (2009)
A. Jhala, R.M. Young, Cinematic visual discourse: representation, generation, and evaluation. IEEE Trans. Comput. Intell. AI Games **2**(2), 69–81 (2010)
C.A. Lindley, M. Eladhari, Causal normalisation: a methodology for coherent story logic design in computer role-playing games, in *Proceedings 3rd International Conference Computers and Games (CG 2002)*, (Springer, 2002), Vancouver, British Columbia, Canada, pp.292–307
C. Lino, M. Christie, F. Lamarche, G. Schofield, P. Olivier, A real-time cinematography system for interactive 3D environments, in *ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, Madrid, 2010
S. Luckin, L. Reed, M. Walker, Generating sentence planning variations for story relling, in *16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, 2015, Prague, Czech Republic
J.L Lugrin, M. Cavazza, D. Pizzi, T. Vogt, E. André, Exploring the usability of immersive interactive storytelling, in *Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology*, (ACM, 2010), Hong Kong, China, pp. 103–110.
M. Mateas, A. Stern, Structuring sontent in the façade interactive drama architecture, in *Artificial Intelligence and Interactive Digital Entertainment (AIIDE)*, Los Angeles, 2005
J. McCoy, M. Treanor, B. Samuel, A.A. Reed, M. Mateas, N. Wardrip-Fruin, Prom week: designing past the game/story dilemma, in *Proceedings of the Conference on the Foundations of Digital Games*, 2013, Chania, Crete, Greece
R. McKee, *Story: Substance, Structure, Style, and the Principles of Screenwriting* (Regan Books, New York, 1997)
J. H. Murray, Hamlet on the holodeck: The future of narrative in cyberspace. New York: Free Press, 1997
A. Mitchell, Motivations for rereading in interactive stories: a preliminary investigation, in *Proceedings of the 3rd Joint Conference on Interactive Digital Storytelling (ICIDS 2010)*, (Springer, 2010), Edinburgh, UK, pp. 232–235
A. Piacenza, F. Guerrini, N. Adami, R. Leonardi, J. Porteous, J. Teutenberg, M. Cavazza, Generating story variants with constrained video recombination, in *Proceedings of the 19th ACM international conference on Multimedia*, 2011, Scottsdale, AZ, USA, pp. 223–232.
J. Porteous, M. Cavazza, F. Charles, Narrative generation through characters' point of view, in *Proceedings 9th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS 2010)*, (IFAAMAS, 2010a), Toronto, Canada, pp. 1297–1304.
J. Porteous, M. Cavazza, F. Charles, Applying planning to interactive storytelling: narrative control using state constraints. ACM Trans. Intell. Syst. Technol. **1**(2), 1–21 (2010b)
J. Porteous, J. Teutenberg, D. Pizzi, M. Cavazza, Visual programming of plan dynamics using constraints and landmarks, in *Proceedings of the 21st International Conference on Automated Planning and Scheduling (ICAPS 2011)*, Freiburg, 2011, pp. 186–193
R. Činčera, J. Roháč, V. Svitáček, Kinoautomat: One Man and His House, Czechoslovakia, (1967)
F. Reitsma, Geoscience explanations: identifying what is needed for generating scientific narratives from data models. Environ. Model. Softw. **25**(1), 93–99 (2010)
B. Richardson, *Unlikely Stories: Causality and the Nature of Modern Narrative* (University of Delaware Press, Newark/London, 1997)
M.O. Riedl, R.M. Young, From linear story generation to branching story graphs. IEEE Comput. Graph. Appl. **26**(3), 23–31 (2006), Melbourne, Australia
M.O. Riedl, R.M. Young, Narrative planning: balancing plot and character. J. Artif. Intell. Res. **39** (1), 217–268 (2010)
M.O. Riedl, C.J. Saretto, R.M. Young, Managing interaction between users and agents in a multi-agent storytelling environment, in *Proceedings 2nd International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2003)*, Melbourne, Australia 2003, pp. 741–748

J. Robertson, R.M Young, Finding schroedinger's gun, in *The International Conference on AI and Interactive Digital Entertainment*, Raleigh, 2014, pp. 153–159.

J. Robertson, R.M. Young, Interactive narrative intervention alibis through domain revision, in *Working Notes of the 8th International Intelligent Narrative Technologies Workshop, Part of the AAAI Conference on AI and Interactive Digital Entertainment*, Santa Cruz, 2015

M.L. Ryan, *Narrative as Virtual Reality: Immersion and Interactivity in Literature and Electronic Media* (Johns Hopkins University Press, Baltimore, 2001)

J.R. Searle, *Speech acts: An Essay in the Philosophy of Language*, vol. 626 (Cambridge university press, London, 1969)

W. Swartout, R. Hill, J. Gratch, W.L. Johnson, C. Kyriakakis, C. LaBore, R. Lindheim, S. Marsella, D. Miraglia, B. Moore, J. Morie, J. Rickel, M. Thiebaux, L. Tuch, R. Whitney, J. Douglas, Toward the holodeck: integrating graphics, sound, character and story, in *Proceedings of the 5th International Conference on Autonomous Agents (Agents-01)*, (ACM Press, New York, 2001)

T. Trabasso, P.W. van den Broek, Causal thinking and the representation of narrative events. J. Mem. Lang. **24**, 612–630 (1985)

F. Truffaut, and G. Scott Helen. Hitchcock. Simon and Schuster, 1985.

M. Ursu, J. Cook, V. Zsombori, R. Zimmer, I. Kegel, D. Williams, M. Thomas, J. Wyver, H. Mayer, Conceiving shapeshifting TV: a computational language for truly-interactive TV, in *Interactive TV: A Shared Experience*, ed. by P. Cesar, K. Chorianopoulos, J. Konstantinos (Springer, Berlin, 2007), pp. 96–106

S.G. Ware, R.M. Young, Glaive: a state-space narrative planner supporting intentionality and conflict, in *Proceedings of the 10th Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE2014)*, Raleigh, 2014, pp. 80–86

R.M. Young, Creating interactive narrative structures: the potential for AI approaches, in *AAAI Spring Symposium in Artificial Intelligence and Interactive Entertainment* (AAAI Press, Menlo Park, 2000)

R.M. Young, The Co-Operative Contract in Interactive Entertainment, Dautenhahm, Kerstin; Bond, Alan; Canamero, Lola; Edmonds, Bruce (Ed.): Socially Intelligent Agents, pp. 229-234, Kluwer Academic Press, 2004

R.M. Young, Story and discourse: a bipartite model of narrative generation in virtual worlds. Int. Stud. **8**(2), 177–208 (2007), (32)

R.M Young, Notes on the use of plan structures in the creation of interactive plot, in *AAAI Fall Symposium on Narrative Intelligence*, (AAAI Press, 1999), Los Angeles, CA

N. Zagalo, A. Barlker, V. Branco, Story reaction structures to emotion detection, in *Proceedings of the 1st ACM Workshop on Story Representation, Mechanism and Context*, (ACM, 2004), New York, NY, pp. 33–38

Planning Technologies for Interactive Storytelling

18

Julie Porteous

Contents

J. Porteous (✉)
School of Computing, Teesside University, Middlesbrough, UK
e-mail: j.porteous@tees.ac.uk

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_71

Abstract

Since AI planning was first proposed for the task of narrative generation in interactive storytelling (IS), it has emerged as the dominant approach in this field. This chapter traces the use of planning technologies in this area, considers the core issues involved in the application of planning technologies in IS, and identifies some of the remaining challenges.

Keywords

Crowdsourcing • GADIN system • Hierarchical task networks (HTNs) • Interactive storytelling (IS) • Automated domain authoring • Boundary problem • Causality • Character-based approaches • Flexibility • Generative power • Goal of • Intent-based narrative planning • Manual domain authoring • Minimal representational assumptions • Narrative control • Narrative planning goals • Narrative plans • Plan failure • Role of • Plan quality criteria • Plan-based narrative generation • Plot-based approaches • Real-time response • Story and discourse • Story progression • Visual programming • Planning domain model • SAYANYTHING system • SCHEHERAZADE system

Introduction

The goal of interactive storytelling (IS) research is the development of multimedia systems in which users can interact and influence, in real time, the evolution of a narrative as it is presented to them.

AI planning was first proposed for the task of narrative generation in IS by Young (1999) because it provided a natural fit for representation with narratives as plans, ensured causality which is important for the generation of meaningful and comprehensible narratives, and provided considerable flexibility and potential generative power. Hence, it is not surprising that during the intervening time, the plan-based approach has been enthusiastically adopted (e.g., the systems reported by Cavazza et al. (2002a), Aylett et al. (2006), Riedl and Young (2010), and Porteous et al. (2013)).

In plan-based IS systems, a plan generation engine is embedded within a planning and execution loop where the planner generates an initial narrative which is “executed” via presentation to the user (e.g., using 2D or 3D graphics (Mateas and Stern 2005; Porteous et al. 2013), filmic content (Piacenza et al. 2011), or text (Orkin and Roy 2012)) and where user interaction can change the state of the planning world and hence require replanning or repair.

Given the real-time nature of IS systems, planner performance is clearly an important issue; however, our intuition is that this is less of an issue than in classical planning benchmark domains (as illustration, contrast the two million nodes explored in the narrative example given in Riedl and Young (2010), with the astronomically large search spaces solved by heuristic search planners – 10^{100} is not unheard of (Hoffmann 2011)) and that other aspects, such as narrative control and authoring, are more central.

These core issues of planning and narrative are reflected in the content of this chapter: we start in the next section with an overview of planning and IS systems to date, discuss issues to do with narrative planning models (i.e., domain and problems) and identify a minimal set of representational assumptions which underpin all applications of planning to IS (section "Plan-based Representation of Narrative Worlds"), and for the remainder of the paper, the focus is two enduring challenges for planning in IS, namely, narrative control (section "Control in Plan-based Interactive Storytelling") and authoring (section "Authoring of Interactive Narrative Planning Models"), which are discussed and state-of-the-art solutions reviewed.

Plan-Based Narrative Generation

Young (1999) first proposed the use of AI planning for the generation of narrative in IS since it offers a rich representation of the causal structure of output narratives along with the story diversity that results from its generative power. Since then plan-based approaches have been enthusiastically adopted.

Early work in IS tended to feature older planning approaches, such as partial-order planning (Riedl and Stern 2006; Young 2000) and hierarchical task networks (HTNs) (Aylett et al. 2006; Cavazza et al. 2002b; Hoang et al. 2005; Karlsson et al. 2006). Indeed, these approaches still feature in some current systems. For example, the LOGTELL system features a partial-order planner for initial plot generation (Da Silva et al. 2010); the intent handling in the IPOCL system is embedded in a partial-order approach (Bae et al. 2011) and HTN's feature in the work of Kelly et al. (2007) and Paul et al. (2011).

The continued use of partial-order planning in IS can to some extent be attributed to the flexibility that it offers. For example, it was shown to provide leverage with respect to aspects such as character intent and believability (Riedl and Stern 2006; Riedl 2004). However, this must be weighed against the poor performance of partial-order planners (especially when compared to state-of-the-art state-space planners). As an illustration of this, experiments on narrative generation with the IPOCL planner report a time of approximately 12.3 hours to generate an example narrative plan (Riedl and Young 2010). It is not clear that such an approach could scale to a realistically sized system or be able to perform within the constraints of a real-time interactive system.

The enduring popularity of HTNs can be explained to some extent by the fact that IS has been viewed as a case of knowledge-intensive planning, whereby a vast amount of prior knowledge needed to be engineered into the planning domain, and this justified the use of specific formalisms such as HTN planning. However, control knowledge can be difficult to specify and maintain precisely because it is embedded (Cavazza et al. 2002a).

More promise is offered in the recent trend in IS to the use of forward state-space planning approaches in the spirit of HSP (Bonet and Geffner 2001) and FF (Hoffmann and Nebel 2001) for use in IS: for example, Barros and Musse (2007) used Metric-FF Hoffmann (2003), Pizzi et al. (2007b) used a variant of HSP in combination with the use of RTA* (Korf 1990), and Porteous et al. (2011a) featured temporal planning with CRIKEY (Coles et al. 2009). The explanation for the adoption of these approaches stems mainly from their spectacular performance gains (e.g., as demonstrated in the biannual ICAPS International Planning Competition), along with convergence on an expressive representation language, PDDL (Fox and Long 2003).

Independently of the particular planning algorithms adopted, there has been a tendency to specialize the planner on one particular aspect of IS. In the main, planning has been applied to the task of story generation, for example (Porteous et al. 2011a; Young 2001). However, planning has also been applied to the task of control, in IDA (Magerko 2007b) and to the planning of discourse aspects (Jhala and Young 2005; Young 2007).

Plan-Based Representation of Narrative Worlds

This section discusses some key issues in plan-based representation of narrative worlds and identifies a minimal set of key representational assumptions that underpin all applications of planning to IS.

Duality of Character and Plot

A narrative can be considered from the perspective of its featured characters or from a more abstract perspective at the global level of the plot. In IS, the tendency has been to distinguish between character- and plot-based approaches, and for planning this has different consequences for representation as discussed below. However, a plan-based approach is neutral with respect to this duality since both stances are compatible with the approach at the representational level which can be seen as a spectrum: at one end the focus is on defining autonomous character behavior for planning from which story will emerge, and at the other end are approaches that use plot requirements to specify the necessary plan-based character behavior.

Plot-Based Approaches

With a plot-based approach, the representation of the narrative world is based on a model of the baseline plot itself. As a consequence, narrative actions are modeled from a plot perspective, and their "execution" can involve one or multiple characters. This is a centralized approach that offers a mechanism for narrative control and hence provides a better solution to the problem of authorial control in order for improved narrative coherence. However, the control provided by a plot-based approach comes at a price: the potential for individual actions to become overcomplicated, along with a reduction in the generative power that more distributed approaches might provide.

Character-Based Approaches

The alternative perspective is one which is strongly centered on characters and hence referred to as character -based (Cavazza et al. 2002b). The key advantage of this approach stems from the promise of greater flexibility and potential for generative scalability when narratives emerge from the interactions between different narrative characters and their roles. However, these advantages must be traded off against the loss of narrative control inherent in a character-based approach which can make it difficult to ensure that generated narratives display desired criteria, such as aesthetic properties.

Narrative Plans

For a large number of story genres, ranging from classic folk tales to modern day stories of heists and bank robberies, the structure of the plot itself can be quite naturally modeled using a plan. In such cases, the goals of the narrative and the actions that characters can take to achieve them are isomorphic to the goals and actions of the plan. In IS, instances of this type of strong connection between the narrative structure and representation have played an important role: from the "bank robbery" scenario discussed by Riedl et al. (2003) through to current systems such as the "riddle" genre of Barros and Musse (2008).

However, it is also possible to make weaker representational assumptions, without implying such a strong connection between the narrative structure and the plan-based representation. For example, in the Madame Bovary-inspired system of Pizzi et al. (2007a), a planning representation was used to support the central character's role rather than problem-solving activity based on primitive intentions and desires, and then planning was used to enforce causality in sequences of narrative actions intertwined with user interactions.

IS Problems from a Planning Perspective

Plan Quality Criteria

The question as to what constitutes criteria for "good" narratives in IS is still open. Current state-of-the-art systems require input from human authors such as assigning interesting authorial goals (Riedl 2009), the design of appropriate initial and goal states (Riedl and Young 2005), or describing desired tension arcs (Porteous et al. 2011b). Minimum length criteria can also play a part, where this implies a tight narrative without superfluous actions. In addition to a quality measure, narrative plans must be produced with hard constraints on their plausibility: every action seen to be performed by a character must appear to have some motivation or intent (Riedl and Young 2010).

IS domains differ markedly from the sorts of benchmarks that have featured in planning research, such as the benchmark domains used in the ICAPS International

Planning Competitions. Indeed, there are some key aspects of these IS domains that go against the accepted planning wisdom: in IS domains optimality (e.g., in terms of makespan) is not essential, rather it is the "shape" of the plan trajectory that is of central importance; goals are not necessarily "real" goals in the classical sense; and in certain situations, the system can actively require plan failure and backtracking. Consider the criteria used to assess plan quality, which sets IS domains apart. Historically, optimality in terms of plan length was the criterion for assessing the quality of plans (Chapman 1987) with more recent extensions to enable the specification of plan metrics (Hoffmann 2003) and adequacy criteria such as preferences and constraints (Baier et al. 2007; Chen et al. 2006). In IS, quality criteria differ since replanning is frequently required in response to user interaction, interesting narrative events may be missed if the emphasis is on generating optimal solutions, and there is frequently a need to generate suboptimal trajectories in order to display narrative concepts such as suspense. Indeed, there is a growing consensus that for these types of domains, quality criteria should be more concerned with the dynamics of the plan and the shape of its trajectory, in other words, the intermediate states that will be traversed when it is executed.

An example of tool support for the authoring of narrative plan quality criteria via a visual programming interface is discussed in section "Authoring Tools: Visual Programming of Trajectory Properties."

Narrative Planning Goals

For the application of planning to IS, the nature of goals differs from classical domains since goals need not equate to the end of the narrative (i.e., serve as the narrative driver) and goals can evolve over time. For instance, the GADIN system of Barber and Kudenko (2007) featured dynamic selection and reselection of story goals in response to narrative dilemmas; the continual multiagent narrative planning approach of Brenner (2010) allowed for temporary sub-goals, which could be violated in the final state of the narrative plan and included a notion of dynamic sub-goal activation; and the brain-computer interfaced system reported by Gilroy et al. (2013) enabled the narrative goal to evolve over time depending on users' affective responses.

Role of Plan Failure

The role of plan failure in IS is very different to that in typical classical planning domains: indeed, in certain situations, an IS system might encourage failure in order to achieve some dramatic effect. In general, planned narratives can fail as a result of user actions. Young (2007) distinguished between actions which don't interfere with the planned sequence and exceptional user actions which require some form of intervention to avoid failure. The model for recovery, proposed by Riedl and Stern (2006) and Riedl et al. (2003), featured tiered replanning with mixed off-line analysis of potential sources of inconsistency and on-line repair that attempts to restore causal coherence, and if all else fails, the system attempts to recover by replanning.

Synergistically, forward state-space search results in benefits for fast replanning since the complete state of the world is always known, thus facilitating replanning directly from the point of failure. This in part provided motivation for the decomposition approach to narrative generation introduced in Porteous et al. (2010b), where narratives are generated via forward search through a series of smaller sub-narratives which are produced incrementally, thus reducing wasted effort should replanning be required (see section "Narrative Control Via Trajectory Properties").

Minimal Representational Assumptions

Based on the previous discussions, a minimal set of representational assumptions can be identified for the development of narrative domain models for planning in IS. These consist of:

Narrative Predicates: a set of predicates which can be used to describe states of the narrative world. These can correspond to object properties, location or properties of characters, or characters' mental states, including emotional ones.

Narrative Actions: which can be used as planning operators for the generation of narrative plans. These specify the ways in which the state of the narrative world can be changed and are expressed as pre- and post-condition operators.

Narrative Goal: a goal condition, expressed as a set of narrative predicates and which represents the default end state for a scene or for the narrative at large (this can be a default goal to be used in the absence of user interactions which might necessitate it being changed).

As an illustration, some examples of these components for a narrative planning domain based on the Aladdin folk tale are shown in Fig. 1. In the absence of any standardized benchmarking narrative planning domains, this is used as illustration due to the familiarity of the tale, and its previous use in narrative research (e.g., Riedl and Young 2010; Teutenberg and Porteous 2013).

Control in Plan-Based Interactive Storytelling

Intent-Based Narrative Planning

An important factor in narrative understanding is that virtual characters must be believable – the audience must suspend belief and the actions of virtual characters must not threaten this (Bates 1994). In other words the audience must perceive them as intentional agents (Dennett 1989). Multiagent simulation-based approaches to narrative generation neatly solve the problem of endowing agents with intentionality: by treating each character as an autonomous agent with its own beliefs and intents, believable character interactions can emerge from a simulation (Aylett

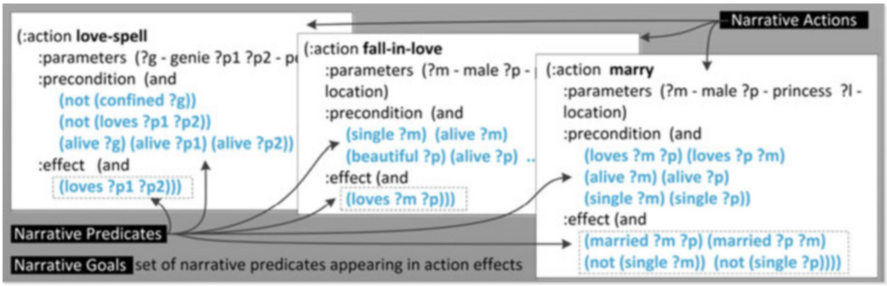


Fig. 1 Illustrative operators from our Aladdin domain represented using PDDL. Operators are specified in terms of preconditions (conditions that must hold for an operator to be visualized in the 3D environment) and post-conditions (effects of the operator). Operators feature typed parameters, such as ?g, ?p1, etc., which are bound to constants at run-time

et al. 2006). The drawback for such systems is in the fulfilling of global narrative goals and author preferences since a character-only simulation relies on emergence and there is no guarantee of even an approximation of criteria other than believability.

In contrast, plan-based narrative generation approaches can struggle with producing an appearance of intentionality. In particular, plot-based narrative generation by a single classical planner that has available to it all character actions and plot events is able to compose sequences of character actions that fulfill global goals that make no sense from any individual character's perspective. To make planning suitable for IS, intent-based planning has been proposed with the IPOCL planner (Riedl and Young 2010). It was originally devised as an extension to a causal link planner that searches the space of partial plans, working backward from the narrative goals. The appearance of intent was ensured by including, in parallel, the planning of frames of commitment that span contiguous subsequences of actions within a plan. The intent for a frame is defined by a character and a target fact that they intend to achieve. The backward chaining nature of the search ensures that characters' actions are causally linked to, and therefore appear relevant to, the final action that closes their frame of commitment. This leads to the generation of narratives where characters' actions have intent, as though planned and performed by autonomous agents, by explicitly representing intent in the narrative and requiring that each action is assigned to some intent. However, intentionality comes at a price, and a notable drawback of the IPOCL planner is the high complexity of the search and long run-time, making it unsuitable for real-time IS use. This is further compounded by the fact that IPOCL has a nonstandard representation language which prevents experimentation with more state-of-the-art planning approaches whose performance might be faster.

Interestingly, Haslum (2012) has recently introduced a compilation for IPOCL narrative domains to classical planning. This meant that the intent-based planning problems could now be tackled by a wide range of classical planners, and for IS it was hoped that this would enable real-time narrative generation. However, the experiments we reported in Teutenberg and Porteous (2013) suggested that narrative generation with the compiled domain models was still too slow to realistically be used in an IS system. Further, the coupling of planning for characters' goals with planning for narrative goals prohibits the generation of certain narratives, most notably those narratives in which one or more characters become unable to complete their intent.

The problem was to develop an approach that would be capable of performing an extensive search of the space of narratives, in order to be able to generate these types of interesting narratives but which would also be capable of performing within the time constraints of a real-time IS system. The solution, which we articulated in Teutenberg and Porteous (2013), was to delegate the task of reasoning about intentions to narrative agents who are responsible for the checking of "narrative action relevance" with regard to their individual intentions. Then, a single narrative planner was used to generate narratives making use of agents' relevant actions. These ideas are fully implemented in a planner, called IMPRACTical (Intentional Multi-agent Planning with Relevant ACTions).

Narrative Control via Trajectory Properties

An enduring challenge in IS is how to control the "shape" of narratives that are automatically generated, and in our work, we have tackled this via an approach that exploits a meta-level abstraction with respect to both time and causality (Porteous et al. 2010). This meta-level is represented using *constraints* – key narrative situations for a domain – that are used as intermediate goals which guide operator selection so that the corresponding trajectory exhibits desired properties (e.g., in narrative these can be associated with pace, suspense, and so on). Please note that while we use the nomenclature of Porteous et al. (2010), we observe the similarity to author goals (Riedl 2009) and also the similarity to the notion of landmarks (Hoffmann et al. 2004). However, unlike landmarks, constraints don't always have to be made true in order to solve the goal, only to conform to desired plan dynamics.

Our planner is a state-space forward search heuristic planner which uses a partially ordered set of constraints to decompose the process of narrative generation into a sequence of subproblems, where each subproblem has a constraint as its goal and the planner generates a narrative for each decomposed subproblem in turn. If the planner can't generate a narrative for a subproblem, then it simply continues with the next constraint, thus ensuring planner continuation. Once all subproblems have been tackled, a final narrative can be assembled by composition of the narratives for each subproblem. However, when our planner is integrated within an IS system, a complete plan need not be output in the traditional sense. Instead operators are sent one at a time for 3D visualization to a user.

In our implemented systems (e.g., Merchant of Venice (Porteous et al. 2010) and NetworkING (Porteous et al. 2013)), this narrative planner operates within a control loop that handles constraint and problem instance selection at run-time. Constraints for a particular narrative world are represented using PDDL3.0 and form a partially ordered set of predicates. For a particular planning instance, a subset of the constraints are selected (e.g., based on user preferences and interaction history). The planner then uses this to drive narrative generation. Initial problem instances are created for each session on the basis of factors such as user preference and enforced variation between user sessions.

Authoring of Interactive Narrative Planning Models

IS systems are dynamic environments within which virtual agents act under the control of system-generated storylines but where real-time interference with the ongoing story can impact on subsequent narrative unfolding. AI planning is well suited for the task of generating stories in such systems; however, the creation of the domain models for these planning-based story generators raises an important practical domain modeling problem: how to author them to ensure they contain sufficient narrative content in order to produce alternative narratives to the baseline plot. Such alternatives can be part of either narrative generation (i.e., the production of multiple story variants) or interactive narrative in its strict sense (i.e., the system responding to dynamic changes made to the story world).

To illustrate this, consider the examples generated using our model of the Aladdin folk tale as shown in Figs. 2 and 3. An initial narrative generated with a domain modeled around a baseline plot might proceed as shown down the left-hand side of Fig. 2 with characters falling in love, casting love spells, and marrying. However, for the system to produce multiple story variants, additional actions are needed for storylines that deviate from the baseline (right-hand side of Fig 3). In addition, to be able to respond to dynamic changes to the story world and to continue the story through to the intended ending, alternative actions are required to generate an alternative course of action from that point onward (Fig. 3).

Manual Domain Authoring

To date the authoring of such interactive narrative domain models has been largely handled manually, a common strategy being to build up the model via systematic consideration of alternatives around a baseline plot. We termed this approach "narrative variant IS" and have used it in our work, for example, with The Merchant of Venice (Porteous et al. 2010).

This method explicitly attempts to create interactive variants of preexisting stories, often classics, for which scholarly analyses exist that facilitate formalization. This strategy of modifying existing narratives can be reduced to a form of remediation (Bolter and Grusin 1999), where the development of a narrative domain starts by modeling a default story as it is traditionally described, i.e., the baseline plot. This default story is represented as a plan, whose goal state corresponds to the default story ending and each main narrative action is represented as a planning operator. There is no assumption that the story itself is equated to a plan for achieving a specific goal. Even so, describing the goal state at different levels of detail ensures

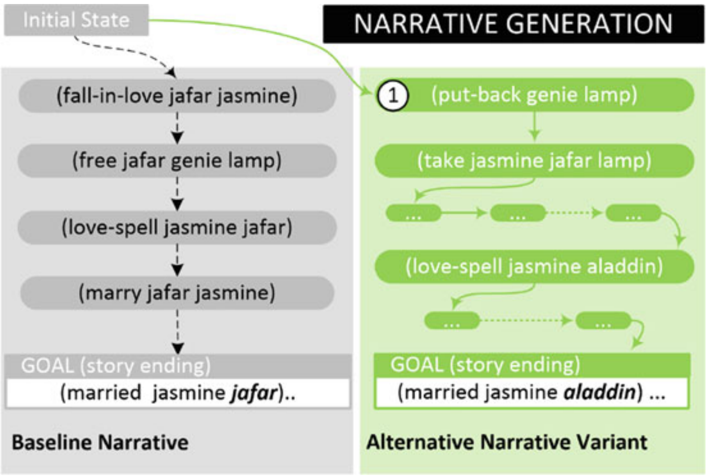


Fig. 2 Example of narrative generation. The figure shows the process of narrative generation with an initial baseline narrative (down the left-hand side) generated using a domain modeled around a baseline story. Also shown in the figure is the generation of an alternative narrative variant (down the right-hand side) which demonstrates the need for additional narrative actions (here, this includes actions such as putting the genie back in the lamp, labeled ①, for a very different story ending with princess Jasmine married to Aladdin)

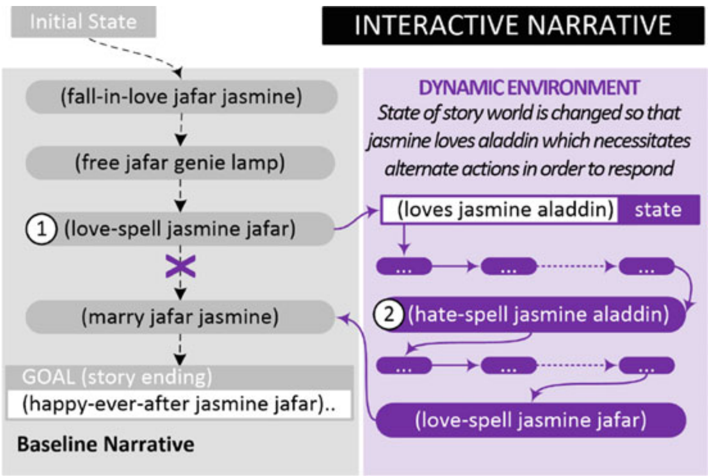


Fig. 3 Example of interactive narrative. The figure illustrates the need for alternate actions in an interactive narrative in order to respond to dynamic changes made to the story world: here, dynamic changes mean that the love spell (labeled ① which should result in princess Jasmine falling in love with Jafar) goes wrong (with the result that Jasmine loves Aladdin), and consequently alternate actions are required to respond appropriately from that point onward (such as casting a spell so that Jasmine hates Aladdin ②)

that various endings are still possible for the same baseline plot. In this context, there exist two determinants of variability around the baseline narrative:

- Run-time modification of the planning domain through user interaction (e.g., removal of facts will invalidate preconditions which would otherwise be satisfied, forcing the story along different avenues)
- Off-line modification of the domain model with addition of new narrative actions that are different but compatible in spirit with the baseline plot

Without such actions, removal of certain facts through interaction may simply result in a dead end, meaning that interaction naturally requires a richer set of actions. Furthermore, even for a given world state, the availability of several compatible actions creates potential for variability triggered by initial conditions or a number of mechanisms for dynamic action selection, such as arbitrarily breaking ties between similar heuristic values in the planning process. The compatibility of operator variants stays within the same semantic context, but this does not mean that the individual operators are directly interchangeable, rather that the generation process will have the possibility to generate varied sequences of operators. Representing narrative actions in the form of planning operators makes explicit their agent-based structure, identifying actors as well as key resources for action – this constitutes a good starting point for the identification of domain predicates which form part of alternative, possibly competing, actions as well as potential targets for interaction.

An indication of the number of different possible narrative variants can be obtained by estimating the size of the narrative space. This depends on the expected narrative path length, n, and the average branching factor, b, (i.e., the number of applicable narrative actions) and can be expressed as b^n. In turn, the branching factor depends on the number of characters in the narrative, c, and the average number of actions that each character has available to them, a, giving, $b = c * a$. As an example, the domain for our Merchant of Venice system featured three central characters with average narrative length of 23 actions and an average of 900 instantiated actions for subplots selected from the baseline play (Porteous et al. 2010). Novel variants around the baseline could be generated in a number of ways, for instance, the operator set contained actions which enabled the plot to unfold in different ways to the baseline (e.g., Shylock could choose to show mercy in the final trial and in so doing avoid the original tragic story ending) and also actions that could be recombined in novel ways, resulting in new semantics.

Automated Domain Authoring

Planning domain model creation is challenging, time-consuming, and an obstacle to the further fielded application of planning technology (Zhuo and Kambhampati 2013). Hence, it is no surprise that automation of the process is a topic of current interest both in the AI planning and narrative research communities. With respect to

AI planning, recent work has been aimed at learning planner action models from correctly observed plan traces (e.g., Amir and Chang 2008; Cresswell and Gregory 2011; Zhuo et al. 2010). However, this work has limited application for narrative domains which do not share the same consistency and alignment with real-world domains as do more traditional planning domains such as “logistics” or “rovers.”

Similarly, in narrative research there has been increasing work aimed at automated creation of content. A popular approach has been the gathering of story elements via crowdsourcing, an approach which can yield abundant content. For example, with the SCHEHERAZADE system, Li et al. (2013) employ this approach to acquire typical story elements which can be assembled as plot graphs and used in a process of story generation. With SCENARIOGEN (Sina et al. 2014), crowdsourcing was used to gather a database of scenarios of everyday activities and likely replacements for use within a serious game context. Orkin and Roy (2012) used a crowdsourcing approach for the hand annotation of logs of the *Restaurant Game* for subsequent use in automating character interactions with human participants in a dialogue-based narrative setting.

An alternative to crowdsourcing aims to obtain narrative content through mining of weblogs and story corpora: the SAYANYTHING system of Swanson and Gordon (2012) selects narrative content on the fly from a corpora of weblogs in response to user text-based interaction, while the approach of McIntyre and Lapata (2009) attempts to generate narratives using knowledge mined from story corpora for a particular genre.

In our work we were motivated to explore automation of our narrative variant approach and have developed an approach to the automated generation of “missing” actions in the domain model that represent opposites of the baseline ones. We developed an automated approach to the identification and creation of such actions, one which can be seen as automating a manual strategy of considering opposite actions and predicates, thus enabling the automated creation of alternate narrative content that departs from a linear baseline narrative (Porteous et al. 2015). Our implemented prototype system, named ANTON, takes as input a planning domain model, identifies missing contrary elements, and uses this to generate new content which is output in an extended version of the domain model. Since it is important that the ANTON-generated content is human readable (e.g., to trace stories during development) and given that contrary relationships form the basis of our approach to action generation, automatic labeling of new content in ANTON is based on antonyms, sourced from a range of linguistic resources such as WordNet and Merriam-Webster.

Authoring Tools: Visual Programming of Trajectory Properties

Earlier we reflected on the nature of quality criteria for narrative plans and observed that these are more concerned with the dynamics of plan trajectories – the intermediate states that will be traversed when it is executed – than the more traditional classical planning domains. For these trajectories, different application-specific interpretations apply: for example, in new media they relate to narrative features such as suspense, pace, and so on.

Based on these observations, in our work we addressed the problem of how to provide authoring support for the specification of the quality criteria for narrative domains. Since controlling trajectory by direct manipulation of the domain model is a considerable challenge and because trajectory suits a visual representation, our solution was to use a visual programming approach. Building on the meta-level of representation of our narrative planning approach (discussed in section "Narrative Control Via Trajectory Properties"), we developed a visual representation of this in the form of a narrative arc for the authoring and exploration of plan dynamics (as reported in Porteous et al. (2011b)).

An overview of our developed system architecture is shown in Fig. 4. Users interact at the meta-level via a narrative arc window (a) and can also explore generated narratives via two visualization windows: an animation window (b) and a timeline window (b). The system also features some hierarchically organized lower-level components. They include the constraints (d) and other PDDL constituents of the domain model (e). Invocation of the planner (f) is driven by the user and enables them to explore the narrative possibilities of different sets of plan dynamics.

Conclusions

Despite the wide-scale adoption of planning, other techniques have been proposed for narrative in IS such as branching narrative structures (Hill et al. 2001; Spierling et al. 2006), behavior trees (Champandard 2008; LLansó et al. 2009), Bayesian networks (Arinbjarnar and Kudenko 2010; Sparacino 2003), case-based reasoning (Gervás et al. 2005; Swanson and Gordon 2010), assumption-based truth maintenance systems (Sgouros et al. 1996), and Monte-Carlo tree search (Kartal et al. 2014).

Nevertheless, on the basis of the following key aspects, our conclusion is that narrative generation for IS is best supported by AI planning:

Causality

The maintenance of local and global story causality is an important dimension of narrativity, as argued by Ryan (2006), which has been observed to be an important factor in user experience (Trabasso and van den Broek 1985). Hence, causality was part of the original rationale for proposing a planning-based approach to narrative in IS (Young 2000) – the ability of the technology to provide appropriate computational narrative structures, cognitive models, and support for shared user-system control of narrative development. The mechanism used in planning algorithms to enforce causal consistency is through the achievement of required preconditions.

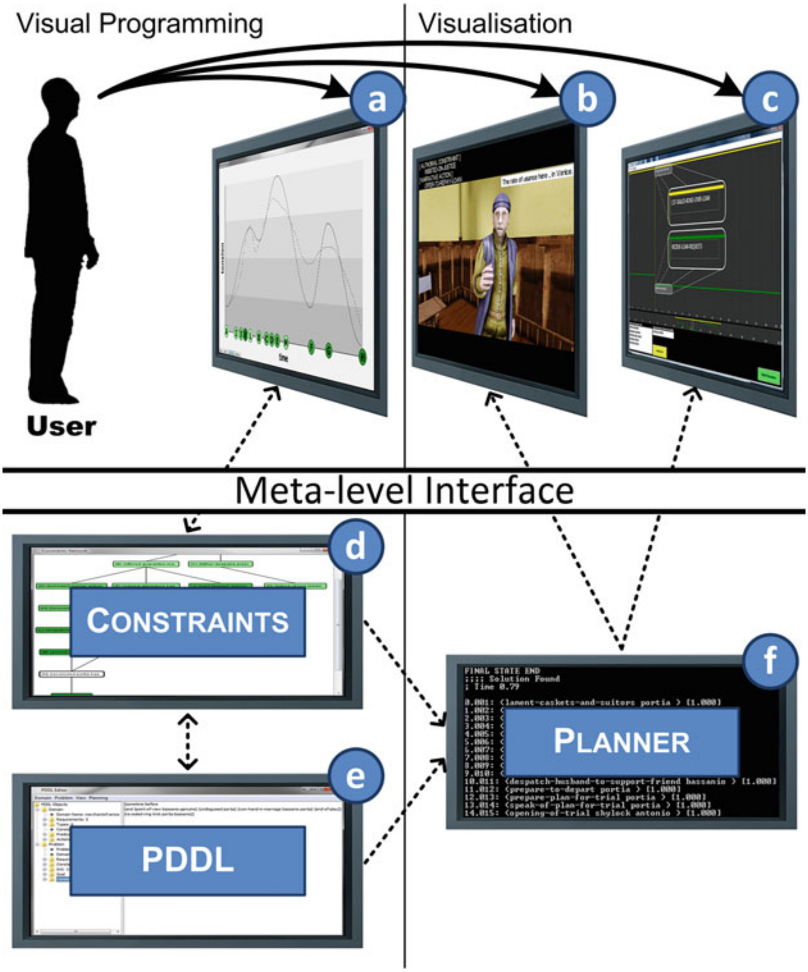


Fig. 4 Visual programming of plan dynamics: user interacts via narrative arc window, drawing desired arcs, adjusting tension levels (**a**); global properties of generated narratives assessed via visualization windows (**b**) and (**c**); lower-level components accessed via interface (**d**) and (**e**); user invocation of planner using specified plan dynamics (**f**)

Generative Power

Regarding shared control of an interactive system, as highlighted by Young (1999), another strength of planning is its generative power – as a consequence of which, it is able to replan in a timely fashion when user actions necessitate changes to the planned narrative. The generative power of a plan-based approach also provides a powerful mechanism to scale up to realistically sized systems (Charles and Cavazza 2004).

Story Progression

In IS, there is a requirement for story progression in the absence of user interaction. This is in contrast to the situation with games that feature narrative content, such as HITMAN (IO Interactive 2000) and HEAVY RAIN (Cage 2010), where the driver for progression comes solely from user interaction; in other words with no user interaction, there would be no progression. A key strength of planning, unlike other technologies, is that it provides a driver for story generation which allows a story to progress regardless of user interaction.

Real-Time Response

Some researchers have proposed the use of reactive planning languages such as ABL (Mateas and Stern 2002) for content generation "on the fly" and have suggested that classical planners lack direct support for this during user interaction via different interaction modalities (Magerko 2007a). However, as already argued, this is not the case for state-of-the-art forward state-space planners, such as HSP (Bonet and Geffner 2001) and FF (Hoffmann and Nebel 2001) which build a plan forward: an approach that actively supports narrative generation in the presence of user interaction. Interestingly, reactive planning approaches have also been advocated in support of interaction based on dialogue, as in FAÇADE (Mateas and Stern 2002) which is largely based on dialogue, even to the point of subordinating plot evolution to dialogic actions. However, the same arguments in favor of planning apply here: the flexible, generative nature of planning makes it possible to respond in real time to user dialogue without sacrificing global narrative properties (Cavazza et al. 2009).

Flexibility

An alternative to plan-based narrative generation that has featured in some IS systems is the use of branching graph narrative structures (Hill et al. 2001; Spierling et al. 2006; Tully and Turner 2004). However, these structures are inflexible and suffer from combinatorial explosion in the number of conditional branches as the amount of possible user interaction increases. They also quickly encounter difficulties propagating the consequences of user interaction. Planning, in contrast, offers additional spatial and temporal flexibility. In addition, if an IS system is to allow seamless user interaction, then the overhead of authoring the narrative is potentially far lower with a planning approach than with a branching narrative approach, which requires the enumeration of all conditional branches as it is authored.

Story and Discourse

In narratology, a number of layered models of narrative have been proposed: a two-layer model where narrative has a content layer, "story," and expression layer, "discourse" (Chatman 1978), or a three-level model which distinguishes between all of the story events, the "fabula," and those story events presented to the audience, the "sjuzhet" (Rimmon-Kenan 2002). Importantly, a plan-based approach provides a consistent technique to manage these different levels of story (Riedl and Young 2010) and discourse (Jhala and Young 2010).

Boundary Problem

Finally, we observe that planning provides a means to support user interaction and in so doing provides a solution to the "boundary problem" which arises when all narrative content is authored and user interaction could potentially go beyond this Magerko (2007b). Planning directly sidesteps this problem, as interactions are dealt with at the level of planning, allowing for the generation and regeneration of narrative content based directly on the consequences of user interaction.

Recommended Reading

E. Amir, A. Chang, Learning partially observable deterministic action models. J. Artif. Intell. Res. **33**(1), 349–402 (2008)

M. Arinbjarnar, D. Kudenko, Bayesian networks: real-time applicable decision mechanisms for intelligent agents in interactive drama, in *Proceedings of IEEE Conference on Computational Intelligence and Games (CIG)*, Copenhagen, 2010

R. Aylett, J. Dias, A. Paiva, An affectively driven planner for synthetic characters, in *Proceedings of 16th International Conference on Automated Planning and Scheduling (ICAPS)*, Cumbria, 2006.

B.-C. Bae, Y.-G. Cheong, R.M. Young, Automated story generation with multiple internal focalization, in *Proceedings of IEEE Conference on Computational Intelligence and Games (CIG)*, Seoul, 2011, pp. 211–218

J.A. Baier, F. Bacchus, S.A. McIlraith, A heuristic search approach to planning with temporally extended preferences, in *Proceedings of 20th International Joint Conference on Artificial Intelligence (IJCAI)*, Hyderabad, 2007

H. Barber, D. Kudenko, Dynamic generation of dilemma-based interactive narratives, in *Proceedings of 3rd Artificial Intelligence and Interactive Digital Entertainment Conference (AIIDE)*, Stanford, California, 2007

L. Barros, S. Musse, Improving narrative consistency in planning-based interactive storytelling, in *Proceedings of 3rd Artificial Intelligence and Interactive Digital Entertainment Conference (AIIDE)*, Stanford, California, 2007

L.M. Barros, S.R. Musse, Towards consistency in interactive storytelling: tension arcs and dead-ends. Comput. Entertain. **6**, 1–17 (2008)

J. Bates, The role of emotion in believable agents. Commun. ACM **37**, 122–125 (1994)

J.D. Bolter, R. Grusin, *Remediation: Understanding New Media* (MIT Press, Cambridge, MA, 1999)

B. Bonet, H. Geffner, Planning as heuristic search. Artif. Intell. **129**, 5–33 (2001)

M. Brenner, Creating dynamic story plots with continual multiagent planning, in *Proceedings of 24th National Conference on Artificial Intelligence (AAAI)*, Atlanta, Georgia, 2010

D. Cage, *Heavy Rain,* Atari, Quantic Dream (2010)

M. Cavazza, F. Charles, S. Mead, Emergent situations in interactive storytelling, in *Proceedings of ACM Symposium on Applied Computing (SAC)*, Madrid, 2002a

M. Cavazza, F. Charles, S.J. Mead, Character-based interactive storytelling. IEEE Intell. Syst. **17** (4), 17–24 (2002b)

M. Cavazza, D. Pizzi, F. Charles, T. Vogt, E. André, in *Proceedings of 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, Budapest, 2009

A. Champandard, *Behaviour Trees for Next-Gen Game AI* (2008). https://aigamedev.com/insider/presentations/behavior-trees/

D. Chapman, Planning for conjunctive goals. Artif. Intell. **32**, 333–377 (1987)

F. Charles, M. Cavazza, Exploring the scalability of character-based storytelling, in *Proceedings of 3rd International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, New York, 2004

S. Chatman, *Story and Discourse: Narrative Structure in Fiction and Film* (Cornell University Press, Ithaca, 1978)

Y. Chen, B. Wah, C.W. Hsu, Temporal planning using subgoal partitioning and resolution in SGPlan. J. Artif. Intell. Res. **26**, 323–369 (2006)

A.I. Coles, M. Fox, K. Halsey, D. Long, A. Smith, Managing concurrency in temporal planning using planner-scheduler interaction. Artif. Intell. **173**(1), 1–44 (2009)

S. Cresswell, P. Gregory, Generalised domain model acquisition from action traces, in *Proceedings of 21st International Conference on Automated Planning and Scheduling (ICAPS)*, Freiburg, 2011

F.A.G. Da Silva, A.E.M. Ciarlini, S.W.M. Siqueira, Nondeterministic planning for generating interactive plots, in *Proceedings of 12th Ibero-American Conference on Advances in Artificial Intelligence*, Bahía Blanca, 2010

D. Dennett, *The Intentional Stance* (MIT Press, Cambridge, MA, 1989)

M. Fox, D. Long, PDDL2.1: an extension to PDDL for expressing temporal domains. J. Artif. Intell. Res. **20**, 61–124 (2003)

P. Gervás, B. Díaz-Agudo, F. Peinado, R. Hervás, Story plot generation based on CBR. Knowl. Based Syst. **18**(4–5), 235–242 (2005)

S. Gilroy, J. Porteous, F. Charles, M. Cavazza, E. Soreq, G. Raz, L. Ikar, A. Or-Borichov, U. Ben-Arie, I. Klovatch, T. Hendler, A brain-computer interface to a plan-based narrative, in *Proceedings of 23rd International Joint Conference on Artificial Intelligence (IJCAI)*, Beijing, 2013

P. Haslum, Narrative planning: compilations to classical planning. J. AI Res. **44**, 383–395 (2012)

R. Hill, J. Gratch, W. Johnson, E. Kyriakakis, C. Labore, R. Lindheim, Toward the holodeck: integrating graphics, sound, character and story, in *Proceedings of 5th International Conference on Autonomous Agents*, Montreal, QC, 2001

H. Hoang, S. Lee-Urban, H. Munoz-Avila, Hierarchical plan representations for encoding strategic game AI, in *Proceedings of 1st Artificial Intelligence and Interactive Digital Entertainment Conference (AIIDE)*, California, 2005

J. Hoffmann, The metric-FF planning system: translating "ignoring delete lists" to numeric state variables. J. Artif. Intell. Res. **20**, 291–341 (2003)

J. Hoffmann, Everything you always wanted to know about planning (but were afraid to ask), in *Proceedings of 34th Annual German Conference on Artificial Intelligence (KI)*, Berlin, 2011

J. Hoffmann, B. Nebel, The FF planning system: fast plan generation through heuristic search. J. AI Res. **14**, 253–302 (2001)

J. Hoffmann, J. Porteous, L. Sebastia, Ordered landmarks in planning. J. Artif. Intell. Res. **22**, 215–278 (2004)

ICAPS, *International Planning Competitions*. http://ipc.icaps-conference.org/

IO Interactive, *HITMAN*, Eidos Interactive. Square Enix (2000)

A. Jhala, R.M. Young, Cinematic camera control using discourse planning techniques, in *Proceedings of 20th National Conference on Artificial Intelligence (AAAI)*, Pittsburgh, Pennsylvania, 2005

A. Jhala, R.M. Young, Cinematic visual discourse: representation, generation, and evaluation. IEEE Trans. Comput. Intell. AI Games **2**, 69–81 (2010)

B. Karlsson, A. Ciarlini, B. Feijó, A. Furtado, Applying a plan-recognition/plan-generation paradigm to interactive storytelling, in *Proceedings of ICAPS Workshop on AI Planning for Computer Games and Synthetic Characters*, Cumbria, 2006

B. Kartal, J. Koenig, S.J. Guy, User-driven narrative variation in large story domains using monte carlo tree search, in *Proceedings of 13th Int. Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, 2014

J.P. Kelly, A. Botea, S. Koenig, Planning with hierarchical task networks in video games, in *Proceedings of ICAPS-07 Workshop on Planning in Games*, Rhode Island, 2007

R.E. Korf, Real-time heuristic search. Artif. Intell. **42**(2–3), 189–211 (1990)

B. Li, S. Lee-Urban, G. Johnston, M. Riedl, Story generation with crowdsourced plot graphs, in *Proceedings of 27th National Conference on Artificial Intelligence (AAAI)*, Bellevue, Washington, 2013

D. LLansó, M. Gómez-Martín, P. González-Calero, Self-validated behaviour trees through reflective components, in *Proceedings of 5th Artificial Intelligence and Interactive Digital Entertainment Conference (AIIDE)*, California, 2009

B. Magerko, A comparative analysis of story representations for interactive narrative systems, in *Proceedings of 3rd Artificial Intelligence and Interactive Digital Entertainment Conference (AIIDE)*, California, 2007a

B. Magerko, Evaluating preemptive story direction in the interactive drama architecture. J. Game Dev. **2**(3), 25–52 (2007b)

M. Mateas, A. Stern, A behavior language for story-based believable agents. IEEE Intell. Syst. **17** (4), 39–47 (2002)

M. Mateas, A. Stern, Structuring content in the façade interactive drama architecture, in *Proceedings of 1st Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE)*, California, 2005

N. McIntyre, M. Lapata, Learning to tell tales: a data-driven approach to story generation, in *Proceedings of 47th Meeting of the Association for Computational Linguistics (ACL)*, Singapore, 2009

Merriam-Webster, *Dictionary and Thesaurus Online*, http://www.dictionaryapi.com/

J. Orkin, D.K. Roy, Understanding speech in interactive narratives with crowdsourced data, in *Proceedings of 8th Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE)*, California, 2012

R. Paul, D.K. Charles, M. McNeill, D.M.G. McSherry, Adaptive storytelling and story repair in a dynamic environment, in *Proceedings of. 4th International Conference on Interactive Digital Storytelling (ICIDS)*, Vancouver, 2011

A. Piacenza, F. Guerrini, N. Adami, R. Leonardi, J. Teutenberg, J. Porteous, M. Cavazza, Changing video arrangement for constructing alternative stories, in *Proceedings of 19th ACM International Conference on Multimedia*, Arizona, 2011

D. Pizzi, M. Cavazza, J.-L. Lugrin, Extending character-based storytelling with awareness and feelings, in *Proceedings of 6th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, Hawaii, 2007a

D. Pizzi, F. Charles, J.-L. Lugrin, M. Cavazza, Interactive storytelling with literary feelings, in *Proceedings of 2nd International Conference on Affective Computing and Intelligent Interaction (ACII)*, Lisbon, 2007b

J. Porteous, M. Cavazza, F. Charles, Narrative generation through characters' point of view, in *Proceedings of 9th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, Toronto, 2010a

J. Porteous, M. Cavazza, F. Charles, Applying planning to interactive storytelling: narrative control using state constraints. ACM Trans. Intell. Syst. Technol. (ACM TIST) **1**(2), 1–21 (2010b)

J. Porteous, J. Teutenberg, F. Charles, M. Cavazza, Controlling narrative time in interactive storytelling, in *Proceedings of 10th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, Taipei, 2011a

J. Porteous, J. Teutenberg, D. Pizzi, M. Cavazza, Visual programming of plan dynamics using constraints and landmarks, in *Proceedings of 21st International Conference on Automated Planning and Scheduling (ICAPS)*, Freiburg, 2011b

J. Porteous, F. Charles, M. Cavazza, NetworkING: using character relationships for interactive narrative generation, in *Proceedings of 12th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, Saint Paul, MN, 2013

J. Porteous, A. Lindsay, J. Read, M. Truran, M. Cavazza, Automated extension of narrative planning domains with antonymic operators, in *Proc. of 14th Int. Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, 2015

M. Riedl, Narrative Generation: Balancing Plot and Character, PhD thesis, Department of Computer Science, North Carolina State University, 2004

M. Riedl, Incorporating authorial intent into generative narrative systems, in *Proceedings of AAAI Spring Symposium on Intelligent Narrative Technologies*, California, 2009

M. Riedl, A. Stern, Believable agents and intelligent story adaptation for interactive storytelling, in *Proceedings of 3rd International Conference on Technologies for Interactive Digital Entertainment (TIDSE)*, Darmstadt, 2006

M. Riedl, R.M. Young, Open-world planning for story generation, in *Proceedings of 19th International Joint Conference on Artificial Intelligence (IJCAI)*, Edinburgh, 2005

M.O. Riedl, R.M. Young, Narrative planning: balancing plot and character. J. Artif. Intell. Res. **39**, 217–267 (2010)

M. Riedl, C. Saretto, R.M. Young, Managing interaction between users and agents in a multi-agent storytelling environment, in *Proceedings of 2nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, Melbourne, 2003

S. Rimmon-Kenan, *Narrative Fiction: Contemporary Poetics*, 2nd edn. (Methuen Routledge, New York, 2002)

M.L. Ryan, *Avatars of Story* (University of Minnesota Press, Minnesota, 2006)

N.M. Sgouros, G.K. Papakonstantinou, P. Tsanakas, A framework for plot control in interactive story systems, in *Proceedings of 13th National Conference on AI (AAAI)*, Portland, Oregon, 1996

S. Sina, A. Rosenfeld, S. Kraus, Generating content for scenario-based serious games using CrowdSourcing, in *Proceedings of 28th National Conference on Artificial Intelligence (AAAI)*, Québec, 2014

F. Sparacino, Sto(Ry)Chastics: a bayesian network architecture for user modeling and computational storytelling for interactive spaces, in *Proceedings of 5th International Conference on Ubiquitous Computing (UbiComp)*, Washington, 2003

Ulrike Spierling, Sebastian A. Weiß, Wolfgang Müller, Towards accessible authoring tools for interactive storytelling, in *Proceedings of 3rd International Conference on Technologies for Interactive Digital Storytelling and Entertainment (TIDSE)*, Darmstadt, 2006

R. Swanson, A.S. Gordon, A data-driven case-based reasoning approach to interactive storytelling, in *Proceedings of 3rd Joint International Conference on Interactive digital storytelling (ICIDS)*, Edinburgh, 2010

R. Swanson, A.S. Gordon, Say anything: using textual case-based reasoning to enable open-domain interactive storytelling. ACM Trans. Interact. Intell. Syst. **2**(3), 1–35 (2012)

J. Teutenberg, J. Porteous, Efficient intent-based narrative generation using multiple planning agents, in *Proceedings of 12th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, Saint Paul, MN, 2013

T. Trabasso, P.W. van den Broek, Causal thinking and the representation of narrative events. J. Mem. Lang. **24**, 612–630 (1985)

G. Tully, S. Turner, Integrated decision points for interactive movies, in *Proceedings of 2nd International Conference on Technologies for Interactive Digital Storytelling and Entertainment (TIDSE)*, Darmstadt, 2004

WordNet, *A lexical database for English*, https://wordnet.princeton.edu/wordnet/

R.M. Young, Notes on the use of plan structures in the creation of interactive plot, in *AAAI Fall Symposium on Narrative Intelligence*, Cape Cod, North Falmouth, Massachusetts, USA, 1999

R.M. Young, Creating interactive narrative structures: the potential for AI approaches, in *AAAI Spring Symposium on Artificial Intelligence and Entertainment*, Stanford, CA, USA, 2000

R.M. Young, An overview of the mimesis architecture: integrating intelligent narrative control into an existing gaming environment, in *AAAI Spring Symposium on Artificial Intelligence and Interactive Entertainment*, Stanford University, California, USA, 2001

R.M. Young, Story and discourse: a bipartite model of narrative generation in virtual worlds. Interact. Stud. Soc. Behav. Commun. Biol. Artif. Syst. **8**, 177–208 (2007)

H.H. Zhuo, S. Kambhampati, Action-model acquisition from noisy plan traces, in *Proceedings of 23rd International Joint Conference on Artificial Intelligence (IJCAI)*, Beijing, 2013

H.H. Zhuo, Q. Yang, D.H. Hu, L. Li, Learning complex action models with quantifiers and logical implications. Artif. Intell. **174**(18), 1540–1569 (2010)

User Interaction for Interactive Storytelling 19

Marc Cavazza and Fred Charles

Contents

Abstract

User interaction is a central component of Interactive Storytelling, yet it has often been neglected, as precedence is often granted to the pursuit of narrative generation techniques. This chapter presents several paradigms of interaction using examples drawn from fully implemented Interactive Storytelling systems and proposes an empirical classification of modes of interaction based on the nature of

M. Cavazza (✉)
School of Electronics and Digital Arts, University of Kent, Canterbury, UK
e-mail: M.O.Cavazza@kent.ac.uk

F. Charles
School of Computing, Teesside University, Middlesbrough, UK
e-mail: f.charles@tees.ac.uk

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_57

user involvement. First is provided a context for the need for multimodal interaction with the story world and the influence of the user's intervention onto the dynamics of story generation. Then, the requirements of affective interaction within the context of Interactive Storytelling are covered and illustrated through traditional multimodal interaction, as well as physiological interfaces. Finally, the potential of Brain-Computer Interfaces (BCI) to support Interactive Storytelling is discussed, in particular through the unification of user experience, user input, and affective filmic theories.

Keywords
Multimodal interfaces • Affective interfaces • Physiological computing • Brain-Computer Interfaces

Paradigms and Technology

Interactive Storytelling (IS) assumes, almost by definition, the existence of interaction mechanisms that will support user intervention in reaction to observed narrative progression, with the ultimate objective of altering the course of action. It can be observed that a significant fraction of IS research has described systems and experiments that are closer to narrative generation than true IS, in the sense that little user interaction actually took place in real time during the narrative itself. Narrative generation in these cases still supports story variability through the initial parameterization of the system and the role of the initial state of the narrative environment. Still, it is fair to propose that proper IS requires mechanisms for user interaction that are tailored to the visual presentation and staging of the virtual narrative. Interaction can be studied from the perspective of the narrative (what prompts it and what changes it brings upon it), as much as from the perspective of the interaction modalities themselves. From a narrative perspective, interaction can target every element that can affect the unfolding of the story, but key to the IS experience is the fact that this process should be an integral part of the user experience. In other words, the medium itself should be the user interface, thus precluding the use of menus or any meta-story level which turns the user into a director. The scope and target of interaction are thus determined by the medium itself, which can be divided into characters and objects.

Interaction is irrevocably linked both to the mechanisms of narrative generation and to the visual interface. User influence is integrated by modifying the information available to the narrative generator, which is a representation of the world state at large (from the concrete instantiation of objects to the abstract beliefs, emotions, or motivations of characters). This is why there should be some level of intelligence in the way in which user interactions are interpreted in terms of the real-time changes they impose on the story world: This intelligence is common to the various modalities which can support interaction and should provide a unified approach to the interaction component of IS.

Physical Interaction

Physical interaction is made possible by the nature of the interface through which the narrative is presented, which is constituted by a virtual world. This mode of interaction has become common place since it was popularized in 3D games and has developed widely accepted formalisms for the identification of affordances, various types of object manipulation and their integration in the gameplay and overall user experience. Physical interaction becomes intelligent when a direct connection exists between the virtual world and a more abstract representation (planning domain, world state) that defines the behavior of the system. This relies on a number of mechanisms, a fundamental one being the real-time recognition of interaction events. These mechanisms have evolved from low-level graphical events based on collision detection, to offer some elementary conceptualization of change semantics in the environment. This feature has played an important role in the adoption of game engines as development environment for IS, since most game engines – and in particular the Unreal® series of game engines, from Unreal Tournament® to the latest Unreal® Development Kit (UDK) – incorporate some form of event detection and control manipulated using an often proprietary, scripting language. The existence of event-based systems in game engines has played a fundamental role in the integration of AI technologies as they support a discretized representation of actions (see Fig. 1). Our first Interactive Narrative (IN) (Cavazza et al. 2001) used physical interaction as its main mode of user intervention. In this interactive micro-sitcom, inspired by characters in the TV series *Friends*, physical interaction with objects was a means of influencing the course of events. One typical example was for the user, using the native interaction mechanisms of the game engine (navigation and direct interaction with the story world objects), to "steal" onstage objects perceived to be of narrative importance. The absence of these objects would lead to failure of some intended actions and would force the narrative into different directions. This constituted an early form of IS that emphasized the ability to interact in what was otherwise seen as a self-automated 3D animated story.

There is little direct physical interaction between the user and characters in IS, unlike in computer games, even if it need not be a kind of fighting. This is due to a combination of factors. Firstly, the type of embodiment in a 3D environment often makes physical interaction difficult without imposing constraints on navigation and controls which reintroduce a game-like experience. Secondly, it can be reasonably said that the most compatible physical interactions are those which preserve the user's visual perspective on the characters and the environment and are compatible with his dual role of spectator and actor. For instance, nonverbal behavior through body language can constitute a modality for interaction, taking different forms depending on the diegetic nature of the interaction. In our virtual Madame Bovary system, turning your back on the central character, Emma Bovary, while she is talking is interpreted as a lack of interest and a negative emotional input.

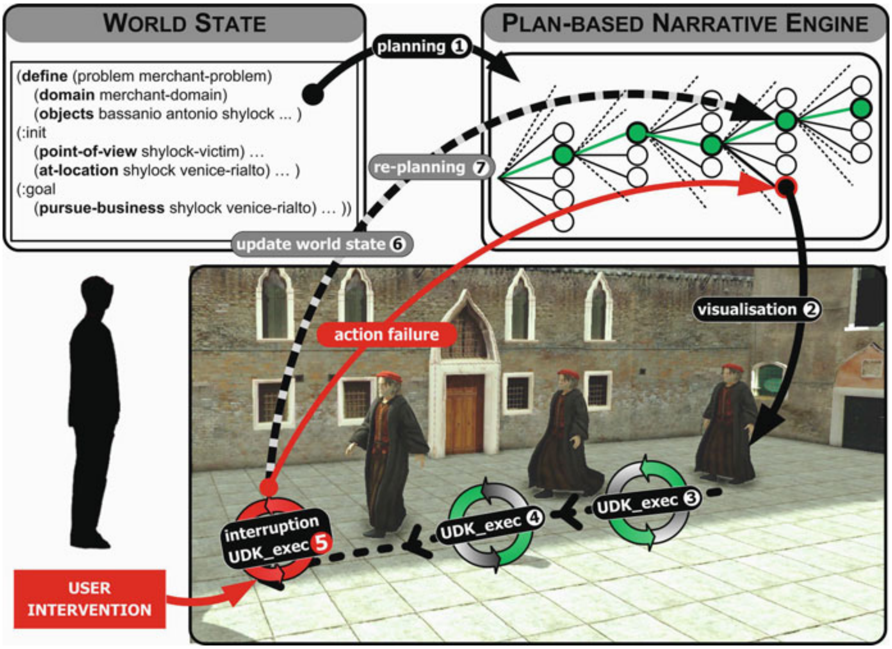


Fig. 1 A *Merchant of Venice* example showing user interruption of action staging: narrative domain information is input to the narrative engine (*1*); narratives are generated and planned actions are sent for visualization in the 3D world (*2*); actions are visualized under the control of UDK (*3–4*); user intervention causes an interruption of UDK and triggers notification of action failure to the narrative engine and updating of the current world state (*5*); once the world state is updated (*6*), replanning is activated (*7*)

On the other hand, interaction with objects in a virtual world has significant potential since objects themselves can have a strong narrative meaning and act as interaction affordances. Barthes (1966) used the notion of *dispatchers* to characterize these objects of core narrative importance, although he made reference to a discourse level of presentation in line with the notorious Chekov's gun. In the context of INs, there is less scope to attract attention to an object at discourse level as a form of foreshadowing (despite attempts at implementing this feature (Bae et al. 2008)), and narrative objects are most often seen as contextual affordances taking into account the course of action. For instance, objects in the environment that could play the role of a gift can be a target for interaction, whether the user's action would be to steal the gift (Cavazza et al. 2002) or offer it (Pizzi and Cavazza 2007).

Two modes of physical interaction can be described in this case: object interaction and object-mediated character interaction. Direct object interaction is a major modality of non-diegetic interaction, whether the user operates in god mode or ghost mode. It often involves the removal of an action resource from the environment in an effort to impair potential narrative actions being performed by some character. It should be observed that while the authors implemented this feature as early as 2001, as one potential mechanism for interaction, recent user experiments in immersive INs have confirmed its popularity with experimental subjects (Lugrin et al. 2010).

Object-mediated character interaction generally takes place when interaction is diegetic, and a typical case of this would be the handing of objects on stage to other characters while the user is themselves impersonating one of the characters. An example of this can be found in our VR Emma system (Lugrin et al. 2010), which features the ability to offer flowers as a gift but also, in a way that demonstrates the symmetrical role of object-mediated interaction, by accepting or not accepting a gift from a character.

Multimodal Interaction

Although some possible form of physical interaction with characters was described, language-based interaction appears to be a more natural form of intervention, even across the various interaction paradigms listed above. When impersonating a character, the user would interact through dialogue, within the technical limitations of state-of-the-art language technologies. This is a case where interaction should be diegetic, which constrains the contents of user utterances, if one assumes a compliant user. (The issue of cooperation is expected to differ greatly according to the interaction paradigms: users acting from within the story are expected to abide by genre conventions because of the diegetic constraint. There is obviously no such expectation when the user intervenes in god mode or ghost mode.) For instance, user-character dialogue in the MRE system (Rickel et al. 2002; Swartout et al. 2006) becomes part of the narrative simulation and is facilitated by the user having to follow the communication style required by the task.

In a mixed-reality narrative, inspired by the James Bond motion picture series, the authors have implemented a different form of multimodal dialogue (Cavazza et al. 2004), based on a combination of expressive gestures and user utterances, where the multimodal emphasis is not so much on reference to the environment as on the nonverbal behavior of the user. These experiments took the diegetic form of interaction to an extreme, where the user themselves becomes part of the visual medium, something that can only be achieved in a mixed-reality installation.

Linguistic input can also support other interaction paradigms, in particular those in which the user is influencing the story from a spectator's perspective. This is one of the historical interfaces for IS, in which the user shouts advice to the story characters. The authors have implemented this form of input in our first *Friends* IN (Cavazza et al. 2002) using a simplified spoken grammar with elements of keyword spotting for increased robustness. In this case, user interaction is clearly non-diegetic, as user utterances are not meant to constitute any part of the story itself, simply the medium through which user influence takes place.

Affective Interaction

The Rationale for Affective Interaction with Interactive Narratives

To a large extent, the entertainment experience of narratives is mediated by their ability to elicit emotions. INs which provide an advanced level of visualization naturally share this property with traditional media, although interaction provides an additional mechanism through the feeling/awareness of agency it introduces. With the recent developments in affective interfaces that directly capture a user's emotional state, it became possible to envision interactive storytelling systems designed around users' affective responses. The main criterion for categorizing such affective interactive storytelling systems is whether emotions are captured spontaneously or are part of some sort of acting by the user, as in Cavazza et al. (2003) and Lugrin et al. (2010).

The definition of emotional models is a major challenge for these systems, because of the discrepancy between the basic emotions captured by interface technology and those associated with narration, which tend to be more complex. In this section, physiological computing techniques (Fairclough 2009) have offered new directions to address this problem; there is first a need to review contemporary theories of affective responses to narratives, which have been developed in the context of traditional media like film.

Affective Filmic Models

There has been substantial research in the mechanisms by which traditional media, in particular film, elicit emotional responses in their viewers. It is worth noting that the two main competing affective filmic theories reproduce the character/plot dichotomy identified as a defining concept in interactive narrative. Plot-based approaches attribute affective responses to story progression, the shape of the narrative arc, or discourse-level phenomena. For instance, Cheong and Young (2008) have described a suspense model based on story progression. The affective filmic theory of Smith (2003) posits that many filmic elements act as emotional cues including discourse elements and in particular filmic idioms (e.g., camera positioning and shot structure) and editing. Smith describes affective filmic response as a cyclical process in which emotional cues contribute to building up a specific mood, which in turn renders the viewer more susceptible to further cues. The emotional categories he describes are both traditional ones such as fear and more complex mood descriptions such as depression. One of the elements of Smith's theory is that emotions can exist without an object or a goal. Conversely, the character-based model assumes that user's emotions are dictated by the relationship they establish with the story characters. The most elaborate character-based theory has been developed by Tan (1996) and explicitly references empathy. This constitutes a higher-level category that does not fully determine lower-level responses, whether dimensional or categorical.

Emotional Speech Recognition

Spoken interaction provides the most natural form of intervention in a narrative; it is compatible with most IS paradigms and is essential to those in which user interaction should be diegetic. However, there exist many technical limitations when considering the state-of-the-art in speech recognition and natural language understanding. Besides, when user utterances have to be consistent with the story genre, and the genre becomes more literary, they inevitably grow in linguistic complexity, posing additional challenges to automatic processing. In the search for an interaction mechanism that would allow unrestricted linguistic expression, the authors have investigated the use of emotional speech recognition (Vogt et al. 2008). In this context, the authors have shown that emotional speech recognition allowed users to respond to the character Emma in our Madame Bovary system. The limitation imposed by the small set of emotional categories recognized could be offset by a contextual use of these categories, leveraging on the emotional categories used in the planning domain itself.

Interactive Narrative and Passive Physiological Sensing

The main interaction paradigm for IN is based on agency: whether as an actor or a spectator, the user influences the unfolding of the story through her intervention. However, monitoring affective states makes it possible to devise new paradigms, such as narrative adaptation, where story progression and pacing respond in real time to the user's affective response. This requires an integrated approach, in which narrative generation and user input refer to the same emotional model. The authors have developed a prototype inspired from the filmic theory of Smith, and based upon a medical drama as a baseline narrative, whose visual content supports complete 3D narratives (visualized through the UDK® engine) of up to 7 min in duration.

The narratives generated by the system and their 3D visualizations which are shown to users, utilize the notion of Smithian cues (Smith 2003), where the atmosphere of actions is used to prime user responses to important narrative events. The presentation of the narratives to viewers features dramatic visualizations of narrative events such as patients being in critical medical situations or arguments between characters, with different staging styles corresponding to the intended level of affective content.

The authors have adapted Smith's original mood-cue approach to make emotional description more specific and amenable to physiological sensing. Two types of affective responses have been considered: one determined by the immediate content of scenes and the other by story progression. Physiological input uses both galvanic skin resistance (GSR), as a measure of arousal, and facial EMG (fEMG), as a measure of valence. The tension and suspense of the core narrative are determined by the intensity of measured arousal of the user through GSR, while the system manages the overall pacing of the narrative by inserting a variety of reorderable and recombinable scenes illustrating plot background and character relationships. EMG signals are used to measure positive and negative response to categories of actions, which are selectively preferred or avoided in future narrative adaptation. This mapping also takes advantage of the different response times of GSR and EMG. The generation of cues over time and the processing of user's emotional responses reflect the cyclical model of Smith, without however recurring to an explicit representation of the user's mood. A detailed description of the system's results and evaluations can be found in Gilroy et al. (2012). It shows that users respond both to emotional cues and the overall story pacing, confirming the initial hypothesis and the system's ability to process instantaneous responses as well as built-up affective states. There is a certain paradox in this form of "passive interaction," which removes user's agency from Interactive Narrative; however, this approach may precisely constitute a promising alternative when considering interaction over prolonged periods of time or continuous affective input. In the next section, the authors consider affective interaction under the opposite character-based paradigm, based in particular on targeted affective responses, such as empathy and anger.

Brain-Computer Interfaces: Empathic and Anger-Based Interaction

According to Tan (1996), the bond established by spectators with fictional characters is a major mechanism for eliciting affective responses. Tan explicitly mentions the concept of empathy as central to emotional response, and there is further evidence from neuroscience studies of empathic responses to films (Raz et al. 2012). The possibility of using empathy to support a unified approach to Interactive Narrative was explored, one in which narrative generation would generate situations prone to induce empathy in the user, and at the same time would be able to detect the expression of empathy as an affective input modality. The authors used the same medical drama as a background narrative, with several adaptations in the emotional categorization of narrative actions. The main driver is the situation of the feature character (a junior female doctor) towards which the user should feel empathy, hence the tagging of narrative actions as a function of their impact (positive or negative) on the character's situation. Consequently, there is no need for emotional cues to be included in the narrative or for various editing techniques to support them. From a narrative generation perspective, the default story is one in which her situation constantly deteriorates as the baseline story unfolds. User intervention is conceived of as a kind of support from the user, which when successful reverts the deterioration of the character's situation. The authors use a Brain-Computer Interface (BCI) to allow the user to express empathy. This is based on neuroscience findings showing that prefrontal cortex (PFC) asymmetry is a marker of the affective dimension of approach (Davidson 1992) and a proxy for empathy (Light et al. 2009). Furthermore, PFC asymmetry is amenable to control via neurofeedback making it a good candidate for use as an input technique. The difficulties inherent to BCI led us to focus user interaction at a specific stage of the narrative, which is determined dynamically: as the situation of the main character deteriorates, the user is prompted for support, which triggers a short (30s) neurofeedback session (see Fig. 2). During this session, the user expresses empathy

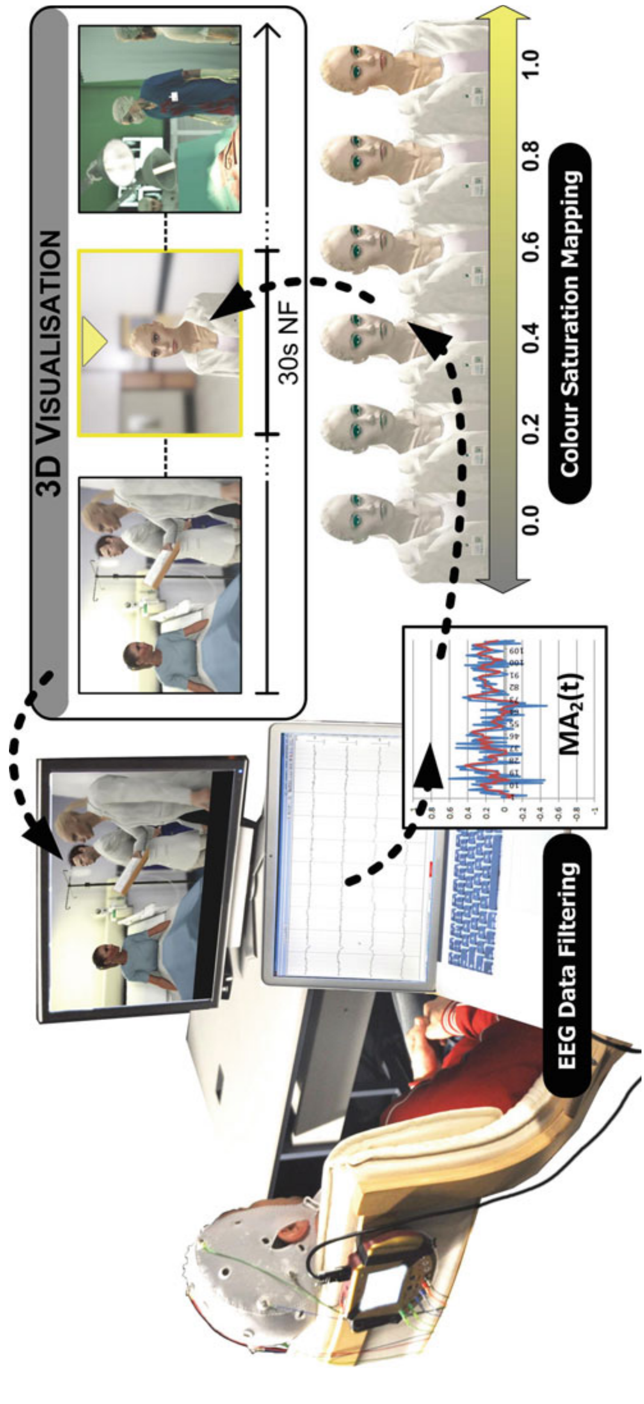


Fig. 2 Setup of our Interactive Narrative BCI prototype: (1) the user watches the narrative generated in real time; (2) during 30 s of neurofeedback epoch, $MA_2(t)$ is mapped to the color saturation of the character in need of support

towards the character, which translates into increased left PFC asymmetry. The visual feedback channel, through which neurofeedback is implemented, consists of the appearance of the character itself: a distressed character fades to gray, regaining its colorful appearance. User intervention is considered successful when the users maintain their level of asymmetry clearly above threshold over the neurofeedback epoch, and this measure of success is used as a parameter for narrative generation for the remainder of the story. In other words, if users manage to express support, the situation of the character in the story will actually improve. The authors have implemented several versions of this system, initially using EEG input in the alpha band to measure PFC asymmetry and support neurofeedback (Gilroy et al. 2013), which resulted in a 73 % success rate for the interaction mechanism (Cavazza et al. 2014), despite minimal training for the users. Recently, the authors have turned to fNIRS as an input mechanism, which is well suited to capturing PFC activity and benefits from a more stable, albeit slightly delayed signal.

Perspectives

While there is no definite answer on which filmic affective theory better accounts for observed phenomena or is more appropriate to the Interactive Narrative context, the above experiments can be analyzed from a system integration perspective. The use of empathy as a unifying concept provides a level of integration of all aspects from narrative generation to user input mechanisms which is difficult to match, as it unifies theory and practice. On the other hand, the strict emphasis on characters may not be appropriate to all genres and may limit the ability to take sophisticated editing into account. The latter aspect is, however, one well supported by Smith's theory and in agreement with the growing interest in editing as a main influencer of affective and cognitive mechanisms alike (Smith et al. 2012). However, owing to the real-time nature of narrative generation, it remains unclear whether Interactive Storytelling will be able to reach a level of editing sophistication comparable to those from which affective filmic theories have been developed.

The Nature of Influence: Local Versus Global

Not enough research has been dedicated to user motivation for interaction during an IN. In those paradigms where user involvement is continuous, it can be assumed that motivation rests within the participation itself, although, as evidenced with Façade (Mateas and Stern 2002), an interest in the story conclusion can also be maintained. This is a case where narrative influence appropriates some aspects of game-related tasks. The situation is more complex in those paradigms in which the user has more freedom in their decision to interfere with the story. The main balance to be maintained is between the immediate impulse to interfere with an ongoing action and a more strategic influence dictated by a global understanding/appreciation of the plot, where a user attempts to modify the overall course of events. Our experience

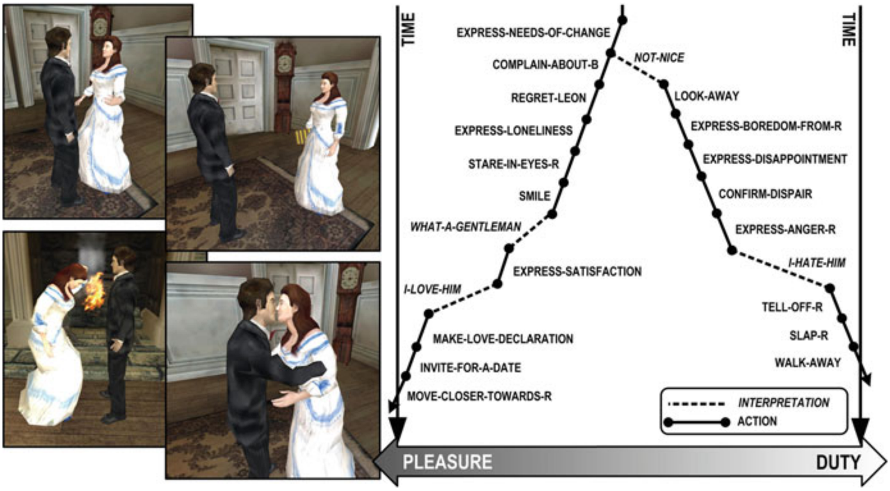


Fig. 3 The role of the pleasure-duty spectrum in generating narrative in our Bovary system: the sequence of operators leading towards pleasure (left-hand segment) corresponds to the baseline plot of the novel, while the sequence leading towards duty (right-hand segment) represents a plausible variant, consistent with the genre. In this example, the presence of interpretation operators with the same polarity as other operators in the segment (*dashed line*) leads more quickly to the narrative ending

favors a genre-related interpretation. In narrative genres where situational aspects play a major role, such as comedy, the user may be prompted to interfere with ongoing actions for immediate effect, more so than they would in more dramatic genres, where interventions may be increasingly strategic and aimed at changing the overall course of action, possibly even altering the ending.

Influencing the global progression of a narrative also requires models onto which this progression can be mapped. Drama, in particular when inspired by classical stories, tends to revolve around high-level dimensions, values, and morals, a finding which was independently emphasized by Damiano and Lombardo (2009). This is also a genre, or a family of genres, in which intervention may swing the course of action towards certain dimensions or values. The strong polarity of a novel such as Madame Bovary, where the fate of the central character Emma Bovary can be seen as evolving along the duty/pleasure dimension (Mettinger 1994) makes it possible to describe possible endings along that main dimension (Pizzi et al. 2007). More precisely, the authors have posited that this dimension could be applied not only to the story as a whole but also somehow recursively to individual sections or chapters.

Figure 3 illustrates two distinct evolutions of generated narratives which are composed of two types of operators: actions and interpretations (a form of evaluation made by the character over its current situation). While the narrative leading to pleasure (left) corresponds to the actual baseline of the novel, the one on the right represents a plausible alternative where Emma returns to her family duty. The interpretation operators, which are often triggered by user interactions, either change the current polarity of the duty/pleasure dimension (if in opposition with current direction) or lead more quickly to the narrative ending (if they have the same polarity).

Conclusions

It is one of the paradoxes of IS research that many published systems incorporate very limited interactivity in the traditional sense, i.e., user intervention in the narrative. This could be explained by a widespread assumption that narrative generation remains the cornerstone of interactive narratives and that existing user interaction techniques are ready to be integrated once progress in narrative generation is deemed sufficient to support complete interactive narratives. The authors have challenged this dominant assumption on several grounds. Firstly, interaction paradigms are fully part of the design of an interactive narrative, as they condition the user experience. They are also dependent on the performance of component interaction technologies, albeit in complex fashion: while some basic technologies, such as spoken dialogue, may not yet be ready to support unconstrained user interaction, these constraints may be incorporated into the design of interaction paradigms. This suggests that it may be counterproductive to consider narrative generation in isolation.

The authors have illustrated multiple interaction paradigms that can be classified along two main dimensions: one is the level of involvement of the user in the story, which determines the frequency of interventions, and the other is her relation to the visual narrative itself. The richness of possibilities is often underestimated, probably because the use of game engines as main visual environment has imported their default interaction mode; however, immersive media whether VR or AR provide alternative contexts for the relationship between multimodal interaction and narrative visualization.

Finally, it appears that affective interaction and interactive narratives could develop a symbiotic relationship. Interactive narratives constitute a privileged test bed for affective interfaces in that they can implement various sorts of affective feedback loops in which users' emotional responses can be studied. Similarly, affective input covers a wide spectrum of user interaction paradigms, from passive spectator (Gilroy et al. 2012) to user acting (Lugrin et al. 2010), in plot-based or character-based approaches (Gilroy et al. 2012, 2013).

Recommended Reading

G. Aranyi, F. Charles, M. Cavazza. Anger-based BCI using fNIRS neurofeedback. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology (UIST '15)*. ACM, New York, NY, USA, pp. 511–521 (2015). http://dx.doi.org/10.1145/2807442.2807447

B.-C. Bae, R.M. Young, A use of flashback and foreshadowing for surprise arousal in narrative using a plan-based approach, in *Proceedings 1st Joint International Conference on Interactive Digital Storytelling* (*ICIDS*), Springer-Verlag, Berlin, Heidelberg, pp. 156–167 (2008)

R. Barthes. Introduction Á L'analyse Structurale Des Récits. In: L'analyse structurale du récit, Communications, Paris, Seuil, 8, 7–33 (1966) (in French)

M. Cavazza, F. Charles, S.J. Mead, Characters in search of an author: AI-based virtual storytelling, in *Proceedings 1st International Conference on Virtual Storytelling* (*ICVS*), Springer-Verlag, Berlin, Heidelberg, pp. 145–154 (2001)

M. Cavazza, F. Charles, S. Mead. Interacting with virtual characters in interactive storytelling, in *Proceedings 1st International Joint Conference on Autonomous Agents and MultiAgent Systems* (*AAMAS*), IFAAMAS, Bologna, Italy, pp. 318–325 (2002)

M. Cavazza, O. Martin, F. Charles, S.J. Mead, X. Marichal, Interacting with virtual agents in mixed reality interactive storytelling, in *Proceedings of the 4th International Workshop, IVA 2003*, Kloster Irsee, Germany, 15–17 Sept, pp. 231–235 (2003)

M. Cavazza, F. Charles, S. Mead, Multi-modal acting in mixed reality interactive storytelling. IEEE Multimedia **11**(3), 30–39 (2004)

M. Cavazza, G. Aranyi, F. Charles, J. Porteous, S. Gilroy, I. Klovatch, G. Jackont, E. Soreq, N.J. Keynan, A. Cohen, G. Raz, T. Hendler, Towards empathic neurofeedback for interactive storytelling, in *Proceedings of 2014 Workshop on Computational Models of Narrative* (*CMN 2014*), 31 July –2 Aug, Quebec City (2014)

Y.G. Cheong, R.M. Young, Narrative generation for suspense: modeling and evaluation, in *Interactive storytelling* (Springer, Berlin/Heidelberg, 2008), pp. 144–155

R. Damiano, V. Lombardo, A unified approach for reconciling characters and story in the realm of agency, in *Proceedings 1st International Conference on Agents and Artificial Intelligence* (*ICAART*), Porto, Portugal, pp. 430–437 (2009)

R.J. Davidson, Anterior cerebral asymmetry and the nature of emotion. Brain Cogn. **20**(1), 125–151 (1992)

S.H. Fairclough, Fundamentals of physiological computing. Interact. Comput. **21**(1), 133–145 (2009)

S.W. Gilroy, J. Porteous, F. Charles, M. Cavazza, Exploring passive user interaction for adaptive narratives, in *Proceedings of the 2012 ACM International Conference on Intelligent User Interfaces* (*IUI 2012*) (ACM, New York, 2012), pp. 119–128

S.W. Gilroy, J. Porteous, F. Charles, M. Cavazza, E. Soreq, G. Raz, L. Ikar, A. Or-Borichov, U. Ben-Arie, I. Klovatch, T. Hendler, A brain-computer interface to a plan-based narrative, in *Proceedings of the Twenty-Third international joint conference on Artificial Intelligence* (*IJCAI '13*), ed. by F. Rossi. (AAAI Press, Palo Alto, 1997–2005, 2013)

J. Gratch, Why you should buy an emotional planner, in *Proceedings 3rd International Conference on Autonomous Agents. Workshop on Emotion-Based Agent Architectures* (*EBAA*), ACM, New York (1999)

S.N. Light, J.A. Coan, C. Zahn-Waxler, C. Frye, H.H. Goldsmith, R.J. Davidson, Empathy is associated with dynamic change in prefrontal brain electrical activity during positive emotion in children. Child Dev. **80**(4), 1210–1231 (2009)

J.-L. Lugrin, M. Cavazza, D. Pizzi, T. Vogt, E. André, Exploring the usability of immersive interactive storytelling, in *Proceedings 7th ACM Symposium on Virtual Reality Software and Technology* (*VRST*), ACM, New York, pp. 103–110 (2010)

M. Mateas, A. Stern, A behavior language for story-based believable agents. IEEE Intell. Syst. **17** (4), 39–47 (2002)

A. Mettinger, *Aspects of Semantic Opposition in English* (Oxford University Press, Oxford, 1994)

D. Pizzi, M. Cavazza, Affective storytelling based on characters' feelings, in *Proceedings Intelligent Narrative Technologies: Papers from the AAAI Fall Symposium*, AAAI Press, Palo Alto (2007)

D. Pizzi, F. Charles, J.-L. Lugrin, M. Cavazza, Interactive storytelling with literary feelings, in *Proceedings 2nd International Conference on Affective Computing and Intelligent Interaction* (*ACII*), Springer-Verlag, Berlin, Heidelberg, pp. 630–641 (2007)

G. Raz, Y. Winetraub, Y. Jacob, S. Kinreich, A. Maron-Katz, G. Shaham, I. Podlipsky, G. Gilam, E. Soreq, T. Hendler, Portraying emotions at their unfolding: a multilayered approach for probing dynamics of neural networks. Neuroimage **60**(2), 1448–1461 (2012)

J. Rickel, S. Marsella, J. Gratch, R. Hill, D.R. Traum, W.R. Swartout, Toward a new generation of virtual humans for interactive experiences. IEEE Intell. Syst. **17**(4), 32–38 (2002)

G.M. Smith, *Film Structure and the Emotion System* (Cambridge University Press, Cambridge, 2003)

T.J. Smith, D.T. Levin, J. Cutting, A window on reality: perceiving edited moving images. Curr. Dir. Psychol. Sci. **21**, 107–113 (2012)

S.K. Sutton, R.J. Davidson, Prefrontal brain asymmetry: a biological substrate of the behavioral approach and inhibition systems. Psychol. Sci. **8**(3), 204–210 (1997)

W. Swartout, J. Gratch, J.R. Hill, E. Hovy, S. Marsella, J. Rickel et al., Toward virtual humans. AI Magazine **27**(2), 96 (2006)

E.S. Tan, *Emotion and the structure of narrative film: Film as an emotion machine* (L. Erlbaum Associates, Mahwah, 1996)

T. Vogt, E. André, N. Bee. EmoVoice – a framework for online recognition of emotions from voice, in *Proceedings 4th IEEE tutorial and research workshop on Perception and Interactive Technologies for Speech-Based Systems: Perception in Multimodal Dialogue Systems*, Springer-Verlag, Berlin, Heidelberg, pp. 188–199 (2008)

N. Zagalo, A. Torres, V. Branco. Passive interactivity, an answer to interactive emotion, in *Entertainment Computing* (*ICEC*), Springer-Verlag, Berlin, Heidelberg, pp. 43–52 (2006)

Part V

Networking in Games

Commodity Video Game Technology in Teletherapy

20

Gary Ushaw, Richard Davison, and Graham Morgan

Contents

Abstract

Motion detection devices are readily available as part of domestic gaming system. Many rehabilitative regimes require a patient to repeat specific movements in order to exercise the affected limbs. Incorporating these movements into games is of interest as the exercises can take place in the domestic environment without

G. Ushaw (✉) • R. Davison • G. Morgan
School of Computing Science, Newcastle University, Newcastle, UK
e-mail: gary.ushaw@newcastle.ac.uk; graham.morgan@ncl.ac.uk

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_26

continual medial oversight, and the use of a gaming context encourages engagement with the exercise regime, leading to increased dosage. This chapter discusses the motion capture devices available commercially, presenting a benchmark for comparison, and provides a case study of a game for rehabilitation of the upper limb after stroke. The work is resented in a manner which is intended to be applicable to the development of any game rehabilitation.

Keywords
Commodity video game technology • Motion capture devices • Stroke patients, rehabilitative game for • Microsoft Kinect and Kinect • Nintendo Wii MotionPlus • Sixense Entertainment • Sony PlayStation Move • Teletherapy, commodity video game technology

Introduction

The diversity and widespread adoption of domestic gaming systems has led to much interest in utilizing action video games in the field of teletherapy. The availability of motion detection devices for these systems has further increased interest in utilizing commodity video game technology in rehabilitation of impaired dexterity.

Rehabilitative gaming has a number of benefits. Firstly, the game should provide a more involving context, so that the patient enjoys the rehabilitative process and is more likely to be engaged in it. Secondly, the exercises can easily be carried out at home as gaming systems which include motion capture devices are relatively commonplace and affordable. Finally, data on the patient's progress and recovery can be collected as the software monitors the patient's actions. The suitability of domestic gaming systems to meet the first two requirements of engagement and accessibility seems intuitively feasible. However, the choice of device, and analysis of the input to meet the fidelity requirements of data collection for medical investigation and monitoring, requires discussion at the outset of any rehabilitative gaming project.

Many motion capture devices are available for domestic use, including Microsoft Kinect, Nintendo Wiimote Plus, Sony PS Move, and Sixense Hydra. In this chapter a benchmark for domestic motion capture devices is described and used to measure and compare the fidelity of the data and the reliability of the systems. The results should serve as a recommendation for future projects utilizing domestic gaming devices for rehabilitation.

Further to the choice of gaming system most suited to a teletherapy project, many choices must be made during the design and development of a rehabilitative game to ensure that the patient is sufficiently engaged to allow for meaningful monitoring of progress and rehabilitation. The second part of this chapter discusses one such rehabilitative game, describing the choices made in its implementation based on tenets and knowledge from both the wider games industry and medical practitioners. The game was designed by a team of professional game developers and medical experts. The discussion is intended to provide a framework of recommendations to be considered when applying serious gaming to the field of teletherapy.

Motion Capture Devices

Motion capture devices have become a commonplace aspect of commercial gaming systems. Such devices are provided, either as a standard feature of the system or as an affordable add-on, by all three of the major gaming console manufacturers (Nintendo, Sony, and Microsoft). There are also motion gaming devices readily available for use with personal computers (such as the Sixense Hydra). This profusion of motion tracking systems has led to their successful employment in many rehabilitative studies. In this section, the functionality of each of the widely available devices is described in turn. The compatibility of the devices with PC technology is also addressed, as teletherapy projects are more likely to be developed for personal computer than console (due to the accessibility of the hardware and SDKs).

Nintendo Wii MotionPlus

Nintendo was the first to market with a motion sensing device for a domestic gaming console. The technology proved disruptive to the game industry's accepted norms, achieving huge success for the Wii console with family audiences. The Wii system consists of a sensor bar (to be placed on or under the television set) and one or more handheld remote (or Wiimote). There are two aspects contributing to the system's motion sensing ability. Firstly, the sensor bar has an infrared LED cluster at each side, and the Wiimote motion controllers contain an infrared camera for tracking their position relative to the sensor bar. Secondly, the controllers each contain three-axis accelerometers and three-axis gyroscopes which are used for dead-reckoning calculation of position, orientation, and velocity. An official SDK for utilizing Wiimote technology with a PC has not been released. However, bespoke software can be written which intercepts and interprets the data stream from the devices via a Bluetooth wireless connection.

Sony PlayStation Move

The PlayStation Move from Sony consists of one or more handheld controllers (or "wand") and a RGB camera to be mounted on the television or monitor. At the top of each wand is an illuminated sphere. The camera tracks this sphere in three dimensions; the distance from the camera is measured from the size of the sphere in the image. Each wand also contains a three-axis accelerometer, thee-axis gyroscope, and a geomagnetic sensor. Consequently the system can track both the position and orientation of the wand. The geomagnetic sensor is used to calibrate the measurements against the earth's magnetic field, thereby addressing cumulative errors from the other sensors. The combination of camera tracking and dead reckoning is intended to provide motion tracking of the wand whether it is visible to the camera or not. The system knows what color the wand sphere is illuminated with, so the image processing software is searching for a circular shape of a known color (a much less computationally intensive activity than searching an image for body parts or non-illuminated controllers). This means that the latency is much less than in other systems where more complex calculations are required. A potential drawback of applying the PlayStation Move to domestic rehabilitation is that the official SDK (the “Move.Me”) requires the use of a PlayStation 3 as a server, in addition to the PC that is running the application. This could lead to an unwieldy setup for home use.

Microsoft Kinect and Kinect 2

Microsoft’s Kinect motion sensing system directly monitors the user’s body, so no handheld devices are required. The system consists of a sensor bar which contains a RGB camera, an infrared laser pattern projector, and an infrared camera. Three-dimensional points in space for each bone of a humanoid skeleton are triangulated using the stereo pair of the infrared projector and camera. Three images are made available by the system: color, infrared image, and an interpolated inverse depth image. Positional information for each bone in the detected skeletal model is made available by the Kinect SDK. The orientation of the bones can then be calculated from the joint hierarchy. Depending on the lighting conditions, and the complexity of the image, the system delivers motion tracking information for up to 20 skeletal joints per user at up to 30 frames per second (Smisek et al. 2013).

A second generation of Kinect device was released for Microsoft’s Xbox One in 2013 (and for Windows8 in 2014). The updated system incorporates a wider field of vision camera, capable of calculating the position and orientation of 25 joints per user, as well as their heart rate and facial expression. The color map resolution has been increased, and the depth map has been augmented by time-of-flight technology contained in the new camera hardware affording a unique depth value per pixel. The combination of this technology and resolution should result in more accurate tracking as compared to the original Kinect device. Both Kinect and Kinect 2 are supported via official SDK for the Windows environment, making it an easily accessibly device for development of PC-based rehabilitative applications.

Sixense

Sixense Entertainment has introduced a motion tracking device for use with PC applications, which uses magnetic field motion tracking to provide continuous position and orientation information. This is in contrast to the camera and infrared-based systems developed by the console manufacturers. The use of electromagnetic fields is a well-established technology for reliably measuring three-dimensional space (Hansen 1987). Devices which are reliant on cameras are prone to interruption due to line-of-sight issues; such problems do not occur with magnetic tracking. The Sixense comprises of a base unit, which is connected to the PC via USB, and wireless controllers which the user holds while performing the movements. The base unit contains three orthogonally orientated magnetic coils which emit an electromagnetic field, providing the reference for the position and orientation of the controllers. Similarly each controller contains three smaller orthogonal coils whose position and orientation are measured relative to the emitter's coils. Consequently each controller broadcasts three-dimensional position and orientation data to the base unit. The use of three controllers, one in each hand, and one tucked into the belt or pocket at waist level, provides further data on the patient's base position and motion. The device and the SDK are designed to be utilized with PC hardware.

Benchmarking the Devices

At the outset of a project in rehabilitative gaming, a vital decision must be made. Identifying the most appropriate motion capture device to utilize is of paramount importance. Each available device has performance advantages and disadvantages. This section describes a benchmark for assessing these devices in terms of the range and fidelity of captured movement data. Results are presented from carrying out the benchmarking test, leading to recommendations on which devices are most suitable to the particular demands of different rehabilitative projects.

The Experiment

Three easily defined and easily repeated movements are used as the benchmarks of this study. The first benchmark movement is to move the hand in a vertical circle, at full arm's length, in front of the subject (i.e., a circle in the coronal plane). The second is to hold the arm at full length in front of the subject and turn 360° on the spot, so that the hand moves in a horizontal circle at shoulder height around the subject's position (i.e., a circle in the transverse plane). The third involves the subject swinging the arm in a vertical circle, keeping the arm at full length (i.e., a circle in the sagittal plane). A number of different subjects were used in the testing of each device. The devices must be able to cope consistently with patients of various shapes and sizes if they are to be used successfully in rehabilitation.

Markers were placed at 1 m intervals up to 4 m, with the sensing device placed at a height of 120 cm. Users stood with their heels inline to each of these markers in turn. At each distance, the user stretched their arm fully forward and rotated it fully 360° in each axis – for the coronal and sagittal planes, this took the form of solely shoulder rotation, while the transverse plane necessitated the user to rotate their entire body using their legs and feet while keeping their back straight. For each plane, the rotation was performed three times, with the second rotation used to calculate results, in order to reduce side effects from the user speeding up, slowing down, or otherwise moving the controllers differently as the trial started and came to an end.

Metrics

Each device calculates 3D position and orientation in some way, so it is desirable to measure the quality of both. For a measurement of positional accuracy, only the relevant two axes for each plane from each trial run are considered, creating a sequence of 2D coordinates that represent the vertices of a 2D polygon. Each trial's data is then centered on the 2D origin, according to the centroid of the polygon, calculated from all *n* positions p in the sequence *s* as:

$$\text{Centroid(s)} = \frac{p_0 + p_1 + \cdots + p_n}{n}$$

Once the captured data has been transformed in this way, two metrics are determined by which to rate the quality of the data: circularity and orientation drift. The circularity of a data set *s* is defined using the following common shape factor (where *a* is the area of the shape and *p* its perimeter):

$$\text{Circularity(s)} = \frac{4\pi a}{p^2}$$

As the coordinates are two-dimensional, the area and perimeter of the polygon can be calculated using the Shoelace formula:

$$\text{Area(s)} = \frac{1}{2}\left|\sum_{i=1}^{n-1} p_i x p_{i+1} y + p_n x + p_1 y - \sum_{i=1}^{n-1} p_i y p_{i+1} x + p_n y + p_1 x\right|$$

$$\text{Perimeter(s)} = \sum_{i=1}^{n-1} \left\|\overrightarrow{p_{i+1} - p_i}\right\| + \left\|\overrightarrow{p_n - p_1}\right\|$$

The accuracy of orientation is calculated in a similar way. By transforming the relevant forward basis vector for the input device by an orientation from the recorded sequence, a vector pointing in the direction of the device is created. By projecting this vector onto the plane being tested, a 2D position is formed. For directions that lie exactly on the plane, the point lies at a distance of 1.0. Deviations in the direction away from the plane result in projected points closer to the origin. The projected direction vectors over the course of one movement form a 2D polygon that are tested using the circularity measurement described previously. This method does not take into account deviations in rotation *around* the axis, but still serves as a useful metric of orientation quality.

As the circularity calculation provides values in the range [0,1], the circularity of position and orientation can be used to calculate a measure of data quality as a simple arithmetic mean. Finally, a base of expected noise is measured, by recording a 30 s sample of the device at rest. The standard deviation in position on all three axes, and cosine of angle difference from the forward basis vector, is then calculated.

A sequence of 3D positions p and orientations q are recorded, with timestamps of each sample. Positions are converted from the internal metric of the device to meters, while orientations are stored as quaternions. Orientations are transformed such that the local-space "forward" reference frame of the device points along the user's outstretched arm – this is necessary as the Wii MotionPlus uses a coordinate system where the z-axis is "up," rather than "forward."

The Kinects are entirely camera based. This results in a lack of information when the camera cannot see the user's hand. The PS Move uses camera tracking for positional data, so it too suffers from a dropout in positional information when the Move "wand" is occluded. Although orientation can still be tracked from the sensors contained within the Move device, for the purposes of this experiment they were ignored when the position was not determinable, to preserve the 1:1 mapping of positions to orientations. This is expected to be most evident when recording movements in the transverse plane, as the user's body obscures the camera's view of the hand / wand, with some additional occlusion expected with the sagittal plane recordings. While any loss of data is undesirable, the ability for the devices to redetect and track the desired movement after obscuration is a useful metric when considering the efficacy of motion devices.

Steadiness

Steadiness measurements were made at a distance of 2 m. Table 1 shows the sum of the changes between the samples in a 30 s recording of the device at rest and the change between the first and last samples. Immediately obvious is the poor result in movement from the Wii Remote Plus – the sum distance is several orders of magnitude larger than any other, due to accumulated error in double integration of position from acceleration. Both the MoveMe and Sixense have large sum distances recorded, but, due to the low change in final position, it can be inferred that this was due to a high frequency jitter, rather than a large drift over time. The Wii Remote Plus reports an acceleration value without any processing for gravity; this was accounted for by taking a short sample of the accelerometer vector at rest, and subtracting this value, transformed by the orientation of the device. Evidently this was not sufficient to reliably remove gravity from the accelerometer reading, and further processing would be required to create a stable result.

Table 1 Device steadiness comparison

Device	Sum distance	Sum angle	Distance traveled	Angle traveled
Kinect 1	0.0842	0,693	0.00369	0.00961
Kinect 2	0.954	149.878	0.000908	0.719
PS Move	2101.93	8.835	8.944	0.00169
Sixense	227.936	1.219	4.271	0.00825
MotionPlus	4281.13	2.254	4272.98	0.202

Table 2 Sample counts for each device

Device	Coronal	Transverse	Sagittal
Kinect 1	70.125	117.0	56.625
Kinect 2	36.381	75.864	34.9
PS MoveMe	140.583	220.231	144.333
Sixense	137.667	188.267	142.933

The Kinect 2 shows a large accumulated sum angle, despite a low distance traveled. As the orientation is determined from the skeleton (calculated by the Kinect software), this is a curious result, as it suggests that the "parent" joint of the skeletal hierarchy has had a large drift or jitter in its position. The low distances for the Kinect devices are at odds with the other tested devices, suggesting additional filtering of position. The low changes in angle over time compared to position for all devices is a good indicator of the relative difficulties in determining these values.

Data Sample Rate

The devices communicate with the host PC in a variety of ways: Kinect via direct USB connection, Sixense wirelessly to a USB connected base unit, and PS Move via Bluetooth to a PlayStation 3 running the MoveMe server, which then contacts the host computer via TCP/IP. Due to this, and the differing ways in which position and orientation are derived, it is useful to compare update rates as reported from the interface software. The mean sample counts for each device across all recorded trial runs are collated in Table 2.

The Kinect devices update at a significantly lower rate than the MoveMe and Sixense devices, which both report at a similar rate. The effects of this are discussed later.

Data Circularity

The circularity metric calculations for each device are shown in Fig. 1. The results from each device are discussed in more detail in this section (Table 3).

Wii MotionPlus

The position of the Wii MotionPlus device can be determined in two ways: either via double integration of its accelerometer data over time or by processing the position of the IR light sources from the tracker bar, as seen by the Wii's onboard camera (broadly, the closer together the points, the further away the device is from the tracker bar). Unfortunately, the experiment as designed is unsuitable for position determination via the tracker bar, as the bar will never be in the line of sight when rotating in the coronal plane and only briefly when rotating in the transverse or sagittal planes. Accurate derivation of position from accelerometer data is notoriously difficult, as any noise or bias in the accelerometer data will quickly accumulate into large errors in position. As gravity will be detected by the

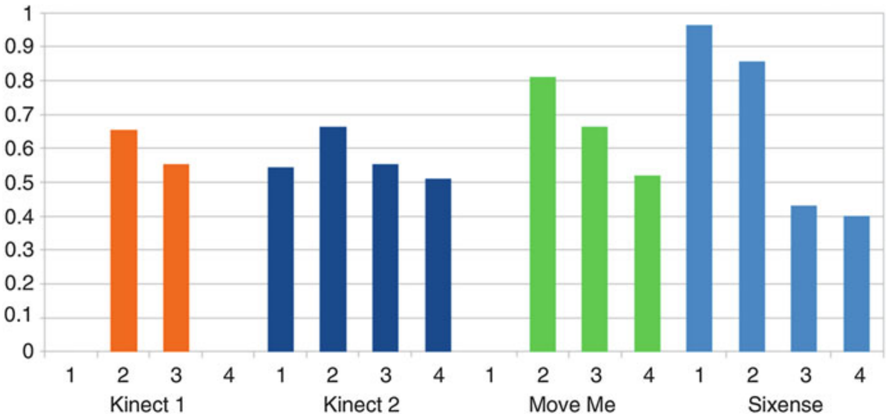


Fig. 1 Device circularity comparison at distances of 1 m, 2 m, 3 m, and 4 m

Table 3 Device circularity comparison. Distances are measured in meters (dashes indicate distances for which no data could be recorded for that device)

Device	Distance (m)	Orientation	Position	Quality
Kinect 1	1	–	–	–
	2	0.7048	0.6035	0.6542
	3	0.5124	0.5958	0.5541
	4	–	–	–
Kinect 2	1	0.5294	0.5283	0.5289
	2	0.6558	0.6849	0.6703
	3	0.5392	0.5555	0.5474
	4	0.5181	0.4888	0.5034
PS Move	1	–	–	–
	2	0.9752	0.6458	0.8105
	3	0.9409	0.3863	0.6636
	4	0.7634	0.2417	0.5025
Sixense	1	0.9805	0.9487	0.9646
	2	0.8749	0.8354	0.8552
	3	0.4476	0.4153	0.4315
	4	0.4967	0.3017	0.3992

accelerometers, it must be accounted for, and any inaccuracy in this calculation will again result in large errors in position.

Despite the poor positional output, a test trial of rotating in the transverse plane shows excellent stability, as shown in a plot in Fig. 2, of the x and z components of the transformed forward vectors of each frame. As positional data could not be reliably calculated, no further tests were performed using the Wii MotionPlus device.

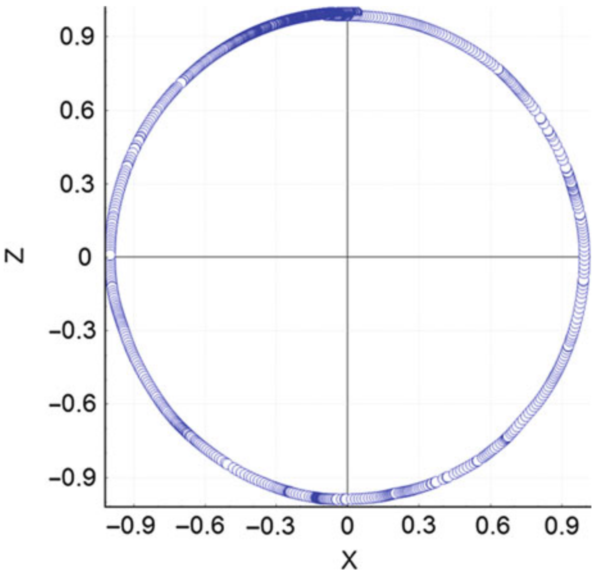


Fig. 2 Wii orientation accuracy

Kinect and Kinect 2

Trials of the Kinect devices reveal that both have practically identical performance at 2 m and 3 m; however, the Kinect 2 benefits from updated optics and processing that allow it to track users at 1 m and 4 m. Figure 3 illustrates this, with blue bars representing Kinect V2, and orange bars Kinect V1.

The Kinect scores at 2 m are noticeably lower than those of the other tested devices (0.65 for Kinect 1, 0.52 for Kinect 2, compared to 0.81 for PS MoveMe and 0.85 for Sixense). Splitting up the quality metric on a per-axis basis in Table 4 clarifies where the weaknesses in the Kinect sensor lie.

Both Kinect devices display poor mean performance in the transverse plane (Kinect 1 quality 0.204, stdev 0.23; Kinect 2 quality 0.154, stdev 0.143). This is unsurprising as rotations in this plane lead to the tracked hand becoming obscured. However, comparison with the MoveMe in the transverse plane (mean quality 0.68, Stdev 0.288), a similarly camera-based solution, reveals particular weakness in this plane.

The method whereby each device calculates its data must be considered. Both Kinects calculate an entire hierarchical skeleton of joint data, with the position of the wrist bone relying on the rest of the arm being correctly tracked, thus requiring more of the users arm to be visible than with the MoveMe, where the bright, uniquely colored ball of the PS Move is found via simple computer vision techniques.

The quality scores of both Kinect devices are limited by the low sample rate of the devices. Kinect 2 in the transverse plane at 2 m is shown in Fig. 4. When compared to Fig. 5, the lower sample density becomes apparent (as shown in Table 5). The data also shows a phenomenon unique to the Kinect devices: as the

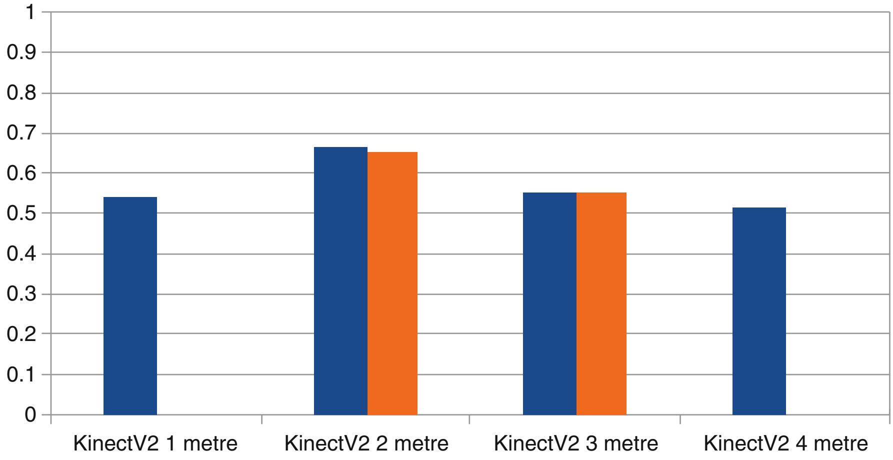
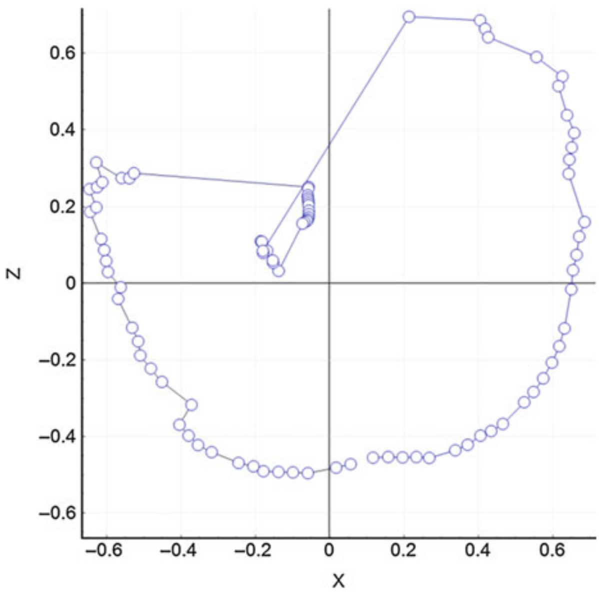


Fig. 3 Kinect quality distance comparison

Table 4 Kinect quality per axis at 2 m

	Coronal	Transverse	Sagittal
Kinect 1	0.839	0.204	0.807
Kinect 2	0.834	0.154	0.784

Fig. 4 Plot of Kinect 2 at 2 m in transverse plane



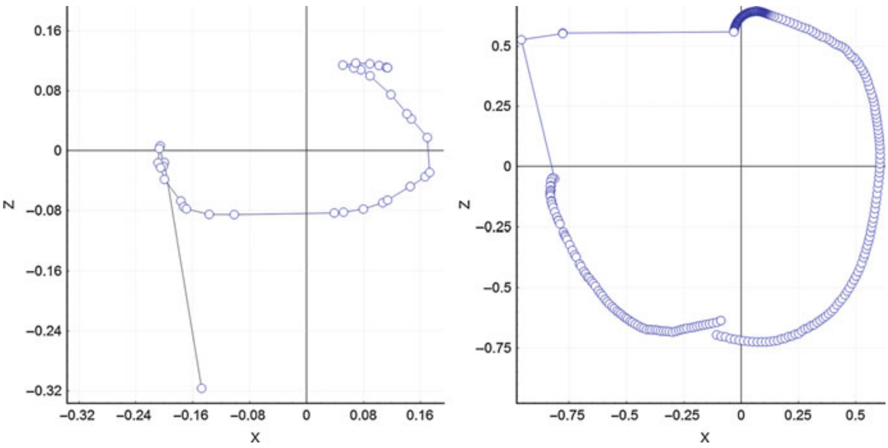


Fig. 5 Plot of PS Move at 4 m and 2 m in transverse plane

Table 5 Kinect sample count per axis

	Axis	1 m	2 m	3 m	4 m
Kinect 1	Coronal	–	80.80	52.33	–
	Transverse	–	133.40	89.67	–
	Sagittal	–	62.40	47.00	–
Kinect 2	Coronal	34.80	35.56	40.17	29.00
	Transverse	69.83	79.00	77.00	78.00
	Sagittal	36.00	31.57	34.50	38.50

user rotates such that their hand becomes obscured, the Kinect devices attempt to seek out the hand elsewhere in the image, frequently picking up the left hand or other part of the body, thus the cluster of data points toward the center of the plot. This is in contrast to the PS MoveMe, which simply stops tracking until the ball is detected again.

PlayStation Move

During the PS MoveMe tests, the 1 m point was not successfully recorded, as the field of view of the camera was insufficient. Recorded orientation information was excellent, with orientation circularity of ~0.95 in all 3 planes at 2 m and 3 m, dropping to 0.76 at 4 m. Unlike the other devices, orientation is derived from sensors on board the handheld device, rather than a camera image or magnetic field, and so is unaffected by distance. However, a reduction in image position tracking leads to fewer orientation samples, impacting the circularity. Position circularity at 2 m is good (circularity 0.646, stdev 0.178), but displays a rapid drop-off at distance.

Tracking of the PS MoveMe starts to degrade at 3 m. Degradation of z-axis tracking is most prominent, while movement in the coronal plane remains accurate. Figure 5 shows an example of the PS MoveMe in the transverse plane.

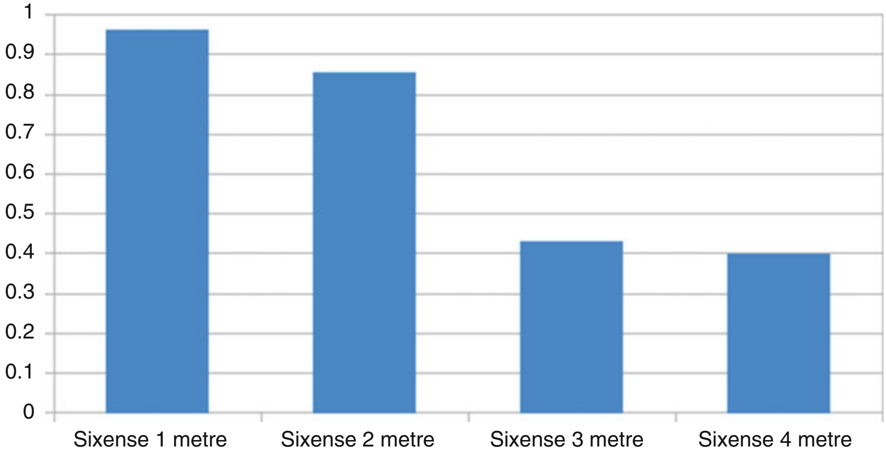


Fig. 6 Sixense quality distance comparison

Immediately obvious is the large gap in the top left quarter of the graph, with the beginning of the gap denoting the point at which the move was obscured, and the end the point at which tracking was regained, with an additional tracking fault. Data is also more ovular, with an approximately 2:1 ratio between minimum and maximum point lengths in the plot axis (position circularity, 0.45). This can be compared to the right-hand plot, in which a transverse plane recording at 2 m shows an approximate 1:1 ratio (position circularity, 0.807). The position detection limit of the MoveMe software appears to be approximately 4 m. It was noted that further movement in the z-axis caused a cutoff effect, where no further movement in the axis is seen.

Sixense

The Sixense is unique among the controllers as it can track position without the use of a camera. Its performance degrades quickly with distance, however. This can be seen in a graph of its combined score metric vs. distance, in which a combined score of 0.96 drops by 50 % at 3 m (Fig. 6). This performance degradation is the worst over distance, and beyond 2 m the Sixense displays the worst quality metric of all devices. At a distance of 1 m (or even less, as the device is not constrained by a camera’s field of vision), however, the device is the best performer, generating almost perfect circularity for both orientation and position.

The results collated in Table 5 reveal a weakness in the sagittal plane. Table 6 shows the standard deviation of collected samples for each axis and distance. The quality of data extracted from the Sixense at 1 m has particularly low deviation in quality in the transverse plane. At 4 m, the sagittal plane results in a consistently poor-quality metric. The extent to which the Sixense degrades is best seen visually – Fig. 7 shows recorded data in the sagittal plane at 1 m and 4 m, with

Table 6 Sixense circularity axis comparison

Axis	1 m	2 m	3 m	4 m
Coronal	0.9592	0.9521	0.4361	0.6648
Transverse	0.9847	0.9512	0.7023	0.3588
Sagittal	0.9497	0.6622	0.1560	0.1740

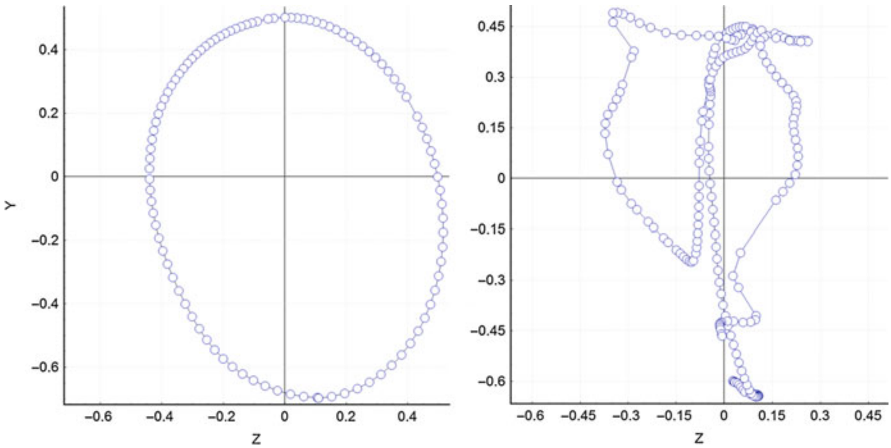


Fig. 7 Plot of Sixense at 1 m and 4 m in sagittal plane

Table 7 Sixense quality deviation comparison per axis

Axis	1 m	2 m	3 m	4 m
Coronal	0.0257	0.0225	0.3634	0.3739
Transverse	0.0064	0.0243	0.1657	0.2119
Sagittal	0.0198	0.1837	0.2114	0.1075

a breakdown in accuracy, making detection of a circular gesture impossible at the greater distance (Table 7).

Recommendations

Domestic motion capture devices have great potential for rehabilitative gaming. A benchmark for measuring the suitability of such devices has been presented. That benchmark has been applied to the commonly available systems and each has been shown to have strengths and weaknesses. Those relative strengths are now discussed and recommendations made as to how best to utilize the devices to the needs of a specific rehabilitative project.

Of note is the excellent quality of data obtained from the Sixense device when used at distances close to the base unit, showing high accuracy in both position and orientation across all three planes. This suggests the Sixense would make an excellent device to use in situations where a high degree of fidelity in movement detection is required, such as detection of specific limb movements and poses. The device begins to heavily degrade at distance, so careful considerations must be made as to the environment it is used in. Applications deployed via a laptop or tablet would be ideal conditions for the Sixense device.

The primary drawback of the PS MoveMe system is its reliance on expensive external hardware and a network connection, but in cases where greater working distances are required, the device produces quality positional and orientation information at 2 m, with degradation in performance beyond that. At 2 m, all trial participants could be fully detected with limbs outstretched, making the PlayStation MoveMe a good choice for living room interaction on a larger screen. Further work will be required to determine whether tighter control of the environment, in regard to any patches of light within the PS MoveMe's field of vision, would result in improved performance at range.

Although the Kinects score less well, notably in the transverse plane, it should be pointed out that these devices provide data on the entire shape of the body. When used at a distance of 2 m, with no hand occlusion, these devices produce quality data. At 2 m, both Kinect devices are very similar, both displaying a weakness in recalculating body tracking sufficiently to determine hand position during rotations in the transverse plane. As with the MoveMe, further testing is required to determine the sensitivity of the Kinect to differing lighting conditions; however, several trials with the Kinect 2 showed the device tracking the left hand as the user turned away from the camera, suggesting that processing of the camera image for facial detection to infer orientation could have a positive effect on reducing false readings.

It is hoped that these results will serve as a recommendation for future projects utilizing domestic gaming devices for rehabilitation.

Design and Development of a Gesture-Based Video Game for Rehabilitation of Upper Limb After Stroke

The development of a rehabilitative game for stroke patients is now described at length, focusing on the design choices made. The game was designed by a team of professional game developers and medical experts and is intended to address the rehabilitation of arm movement in stroke victims.

A series of coordinated bimanual movements have been identified as providing optimum and appropriate exercise for recovering stroke victims. These movements are used as the basis of the video game. The Sixense three-dimensional motion detection device is utilized to track the patient's upper limbs. The motion is used as the input for a professionally developed video game, which was designed specifically to utilize the prescribed exercises, and provides feedback on the patient's progress and recovery.

This section focuses on the implementation details which ensure that the game encourages the correct rehabilitative movements and that it is sufficiently engaging to hold the patient's attention promoting compliance with the exercise regime. At each step of design and development, tenets from professional game development were incorporated into the overall medical goals of the project. Each of those tenets is discussed within the context of the game and presented as advice for the development of serious games for health more generally. Conclusions and recommendations are drawn which can be utilized in the development of further teletherapy projects.

Motivation for the Project

Stroke is a major global health problem that is predicted to increase as the population ages (Murray et al. 2010). Many stroke survivors suffer from hemiparesis, which is a partial paralysis of one side of the body caused by damage to the brain. Improvements to upper limb function can be achieved with intense, repetitive, and challenging exercise of the affected limb (Langhorne et al. 2009). A series of coordinated bimanual movements have been identified which together form the functional basis for activities of daily living (Kimmerle et al. 2003). Stroke victims suffering from impaired upper limb mobility are encouraged to perform these exercises on a daily basis to aid rehabilitation. However, these exercises are, by their repetitive nature, time-consuming for the therapist and boring for the patient, which can result in less exercise being carried out than is recommended.

The use of video games to promote rehabilitative exercise has been suggested in numerous works, in order to keep the patient's attention and provide a more interactive experience (Davison et al. 2014; Rego et al. 2010; Flores et al. 2008; Saposnik and Levin 2011). Such an approach has a number of benefits: the game should provide a more involving context, so that the patient enjoys the rehabilitative process and is more likely to be engaged in it; the exercises can easily be carried out at home as gaming systems which include motion capture devices are relatively commonplace and affordable; and data can be collected on the patient's recovery as the software monitors the patient's actions. Work to date on implementing such an approach has taken two distinct tacks – either existing commercial games have been used as the basis of the study (Joo et al. 2010; Lange et al. 2009) or bespoke games have been developed as part of the research project (Burke et al. 2009a, b; Morrow et al. 2006; Ushaw et al. 2013). Utilizing an existing commercial video game in this context has limitations as the game itself has not been developed with rehabilitation or monitoring in mind; the movements which are practiced are those chosen by the game designers to most suit their gameplay, so are not necessarily the movements the medical professionals would choose to promote rehabilitation. Furthermore, extracting data from a commercial game is impossible on a console or without access to the source code and tool-chain on a PC. However, commercial games tend to be engrossing and immersive; factors which encourage the patient to persevere with the exercise. Conversely, the games developed specifically for the purpose of rehabilitation are designed to react to the prescribed exercise movements and can collate data pertaining to the patient's progress (Graziadio et al. 2014). A game developed for rehabilitation offers the opportunity to include adaptive gameplay, whereby the game detects when the patient is performing less well and dynamically changes parameters of the game to suit the patient's abilities (Pirovano et al. 2012; Cameirão et al. 2008; Rossol et al. 2011; Borghese et al. 2012). The downside of developing a bespoke application is that the quality of the game itself tends to be significantly lower than a commercial title for obvious budgetary reasons, so it may not be as enjoyable and involving as a commercially developed game.

The project involved the development of a video game designed for rehabilitation of stroke-related upper limb disability, which is of sufficiently high caliber to be comparable to commercial games. The game uses three-dimensional motion sensing technology to detect the movement of the patient's upper limbs and presents the player with an ongoing series of challenges based on practicing the coordinated bimanual movements described in Kimmerle et al. (2003). The presentation of the game (graphics, animation, audio) is of a high quality, having been developed by a commercial games studio. A video game design ethos widely used in the games industry, which suggests that frequent in-game rewards encourages a player to remain engaged and continue playing (Koster 2013), is employed. This approach, while more costly, results in a game which requires the patient to practice the prescribed exercises and presents an interactive experience which encourages the patient to engage regularly in their home.

The overall aim of the project is to develop a validated system which provides health professionals with the capability of remote monitoring of home-based upper limb rehabilitation through interaction with a bespoke action video game. A widespread measure of upper limb rehabilitation is the Chedoke Arm and Hand Activity Inventory (CAHAI) score (Barreca et al. 2004), which is calculated from a therapist's visual assessment of the patient carrying out a set of exercises. If this score can be predicted automatically from the motion capture data used by a game, as it is being played in the patient's home, then considerably more frequent monitoring can occur without the therapist having to be present.

System Design

The starting point for the design of the game was the range of moves which the patient must practice, so the first topic to consider is a discussion of why the moves were selected, and the devices and software which were used to detect them. The architecture of the system is also presented, focusing on the way in which the three-dimensional space is monitored, the progression of the exercise regime as the patient's dexterity improves, and the direct feedback that is provided to the patient while carrying out the exercises. The work involved in ensuring that the game feels as polished as a commercial release is also addressed, as this leads to an increased level of engagement by the patient.

The Exercises

It has been established that upper limb rehabilitation after stroke is promoted by frequent and repetitive exercise of the affected limb (Langhorne et al. 2009).

However, it is also established that conventional rehabilitation programs carried out domestically, with no supervision from therapists, suffer from low compliance by patients (Touillet et al. 2010). Anecdotally the patient is more likely to lose interest as the only direct feedback may be an increasing list of figures in a spreadsheet. Video games have been mooted as a source of greater compliance (Rego et al. 2010), as feedback is more interesting (e.g., higher scores and progress through challenges), so the patient remains engaged with the process of rehabilitation at an abstracted level.

In Kimmerle et al. (2003), a series of coordinated bimanual movements of the upper limbs are identified as those which form the functional basis for daily living. The motions which were selected for the patient to perform while playing the action video game comprise 100 separate upper limb patterns based on combinations of these movements. These identified movements are used to control the action in ten distinct mini-games, each with easy, medium, and hard difficulty settings (based on the complexity of the movement required of the patient). The method whereby the progression of the games, and their rehabilitative movements, is fed back to the player is discussed later. The rehabilitative actions which were identified involve both movement of the upper limb through the three-dimensional space and rotation of the wrist, elbow, and shoulder joints around their axes of freedom (e.g., supination and pronation of the wrist, as well as flexion and extension). Consequently, a motion tracking device was required which could reliably recognize these types of movement.

The Motion Tracking Device

As discussed earlier in this chapter, the choice of motion tracking device for a teletherapy project is vital. Key requirements when designing this game were a motion tracking device which would not only enable the tracking of the three-dimensional movement and rotation of limbs but also include a SDK which allows complete control over the tracking algorithms and access to the motion data which is recorded. This requirement of accessible data quickly ruled out the use of a home console, so the development platform was decided to be the PC. A comparison of the various affordable motion capture devices available for use with PC led to the decision that the Sixense was optimal for the project's purposes. Briefly stated, neither the Wiimote nor the Kinect offers the three-dimensional rotational input that was required, and the PS Move, while offering the full set of input data, requires an ungainly hardware combination of both a PlayStation 3 console and a PC in order to officially use the SDK. Further to this, as the application is intended to be used at relatively close quarter, via a laptop, the accuracy of motion detection at around 1 m was vital.

The Sixense motion control system from Sixense Entertainment Inc. uses magnetic motion tracking to provide continuous position and orientation information. The use of electromagnetic fields is a well-established technology for reliably measuring three-dimensional space (Hansen 1987). Three controllers were used, one in each hand and one tucked into the belt or pocket at waist level to provide further data on the patient's base position and motion. The base unit is located at shoulder height in front of the patient (but not blocking the game screen), so that it is in the center of the full range of vertical upper arm freedom of movement. The distance between the patient and the base unit is also important, as magnetic field measurement is known to decay in strength and to distort as the distance between the source and sensor increases (Zachmann 1997). The optimal distance was determined to be around 60 cm (i.e., approximately an arm's length); if the sensors are detected to have gone beyond 80 cm from the detector, an on-screen message instructs the patient to move closer to the screen and base unit.

The Game Design

The game which was designed consists of a series of scenes involving circus performers carrying out their acts in a big top tent; the patient is instructed to perform specific move combinations, and, if they succeed, the circus performer successfully completes the act. The more successful the acts that are performed, the more the player can progress through the game onto increasingly more complex moves.

Each circus act sequence is self-contained, and this modular nature offers some significant advantages. Importantly the game is structured in a way that the software knows which action the motion analysis algorithms are searching for at any time. A more generalized game wherein the patient may choose to perform any of a number of actions at a particular time would require considerably more complex motion analysis algorithms. The algorithms would need to test any player movement against the full range of possible exercises, which would take up a significant amount of the computing resources available on a standard laptop. Another key advantage arises as the order of the sequences can easily be changed, in response to play testing by the initial test groups of patients. This means that movements which are found to be more difficult can be pushed further back in the progression, and easier movements can be brought forward. The standalone nature of each circus act also greatly aided the testing and bug-fixing phase of the project. As the player's actions can only influence one sequence at any time, there is limited opportunity for more complex system-wide issues to occur. Furthermore each sequence can attain a level of polish and professionalism independently of other sequences, leading to a more complete feeling to the game, which encourages the patient to engage for longer periods of time.

Asymmetrical Motion Capture

A key point to incorporate when developing a system for motion capture of the upper limbs of stroke victims is that the dexterity of the paretic limb is generally considerably less than that of the non-paretic limb. For example, if an exercise requires the patient to raise both arms as high above their head as possible, the patient is likely to be able to raise the unaffected limb to a significantly greater height than the limb requiring rehabilitation. Consequently the motion capture needs to be treated as asymmetric.

This asymmetry is achieved through the use of two distinct bounding volumes in the three-dimensional space, one related to each arm. All movements are then interpreted in terms of percentages of those bounding volumes.

At the start of the first session with the game, the patient must go through a calibration process which defines the extent of each bounding box. The patient is instructed to reach as far as possible along each axis in turn; the process is repeated three times and the average reach taken as the extent of the bounding box along each axis. The bounding box dimensions for each limb are saved as part of the patient's profile for subsequent sessions. The patient also has the opportunity to recalibrate at the start of a subsequent session; this allows the therapist to recommend that the recalibration takes place as the patient has improved the reach of their paretic limb.

Each of the 100 rehabilitative moves is defined in terms of the percentile position within the bounding box of each Sixense controller, as well as the orientation of the controller. For example, a move may entail the movement of each limb between around 20 % toward the lower vertical limit of the bounding box and 60 % toward the upper limit. For an able-bodied person, these values will represent approximately equal distances for both arms, but for a recovering stroke victim, the distance for the paretic limb is likely to be significantly less. Consequently when tracking the motion of the sensors in three-dimensional space, it is the movement as a percentage of the bounding box dimensions which is used as a comparison. The move is declared successful if the measurements match the percentage-based parameters of the prescribed exercise. In that way, a patient with severely limited dexterity in one arm can still progress through the game and practice the exercises. Further note that this reduction in reach is not communicated by the in-game avatar which instructs the patient on which exercises to carry out. The avatar's reach remains constant, which encourages the patient to feel successful and to continue with the exercises.

In-Game Help and Feedback

Visual aids are provided to instruct the player on the exercise which must be practiced during each action sequence. These aids are presented as part of the game's graphical user interface (GUI) in real time. An avatar is displayed to the right of the screen, which is animated to show the movement that the patient is expected to perform in a sequence. The avatar repeats the motion throughout the game sequence. The avatar's hands are prominently colored, whereas the rest of the avatar is plain white; this focuses the patient's attention on the movement of the hands which must be replicated during the exercise.

There are also two colored rings displayed on the avatar in the GUI. These show the current position of the patient's hands – i.e., they represent the position information from the two magnetic sensors in the hands, relative to the sensor on the belt. The guide rings are color coded to match the avatar hand colors (green for left, blue for right – note that the avatar is a mirror image of the patient, as the patient faces the screen on which the avatar is displayed). It is important to note that the distance of the rings from the center of the avatar is proportional to the full extent of the patient's bounding box. If the player's paretic limb is extended to its full ability, then the guide ring will be at the edge of the circular region in the GUI, in keeping with the approach we take of treating distances as percentages of the bounding boxes for each limb. This has the additional benefit of providing positive feedback to the patient, showing that their current full reach is acceptable for progress in the game.

It can be desirable for rehabilitative games to self-adapt to the patient's ability (Cameirão et al. 2008; Pirovano et al. 2012). This entails monitoring the player's attempts at specific rehabilitative motions and, if the patient is repeatedly failing, to modify the parameters used in the detection routines to be more forgiving. For example, Pirovano et al. (2012) use AI algorithms to decide whether to allow more leeway in how much the feet are moving if a patient is struggling to progress. As the patients' ability is to be monitored remotely, a consistent set of parameters was required, so adaptive gameplay was not an option. The in-game help and feedback described in this section partly address this, coupled with the progression of difficulty in the sequences of moves from easier through to difficult.

Further to the information provided by the avatar, we provide an option for the patient to view a video of a therapist carrying out the required rehabilitative movement. The patient may choose to watch this video at any time, and the game itself will pause while the video is played back on screen. If the patient has failed to meet the requirements of an action over a number of timed attempts, the option to view the explanatory video becomes more prominent in the GUI, encouraging the patient to study how to achieve the move without overt suggestion of requiring help or failure.

Maintaining Player Compliance

While there is much literature pertaining to the use of video gaming in therapeutic applications, the works tend to focus on attempting either to identify relevant uses of existing video games (Joo et al. 2010; Lange et al. 2009) or have included a comparatively basic video game designed specifically for rehabilitation but developed by a research lab on a budget which is tiny compared to that of commercial games (Burke et al. 2009a). While both approaches have yielded very interesting and promising results, the decision was taken early in this project to develop a video game of commercial quality which was designed specifically with rehabilitation in mind. The overall focus is to produce a game which will aid the long-term recuperation of many patients, so it was felt that an appealing gaming experience was important.

The game involves high-quality animations and music, as well as quirky and engaging characters. These serve the purpose of giving the rehabilitation sessions a sheen of fun and entertainment. The game could probably have been built solely around the basically colored two-dimensional avatar and the two guide rings, but it was felt that this would not be a sufficiently engaging experience for the patient to keep coming back for repeated sessions. This level of polish comes at a price, of course; not only was the game considerably more expensive and time-consuming to develop, but the laptop itself must devote a significant proportion of its computing resources to the graphics and audio, leaving less computational power for the gesture recognition and data handling.

Various elements of mainstream game design have been utilized to encourage the patient to maintain compliance with the exercise regime. These are based on the idea that an enjoyable game presents the player with a continuous set of rewards and attainable challenges (Koster 2013). We include a series of "achievements," comparable to those on Xbox or Steam; these are awarded when the player hits certain thresholds (e.g., 50 h of play or successfully popping 100 balloons) and consist of a GUI element and a musical cue. The player can review achievements gained and those yet to attain from the menu screens. We also employ a high score table, for each challenge, so the player can see progress and aim for improvement. The scores are based on the number of moves successfully completed during a mini-game (i.e., there is no gradation of success of any particular move). These elements combine to encourage the patient to continue with the exercises and to sustain a level of abstraction between playing the game and carrying out repetitive exercise.

Implementation and Evaluation

The game was developed as a system to provide remote monitoring by health professionals of home-based upper limb rehabilitation. A part of the purpose of this study was to derive clinically relevant measures of upper limb function from analysis of the movements made by patients while playing the game. Some requirements of that work have direct bearing on the implementation so are discussed in this section to provide guidance on developing a game with medical professionals.

Assessment Build

The game is required to gather data on the patient's movement for analysis. This is achieved through an assessment build of the game. The assessment build consists of a sequence of 12 to 40 exercises which the patient plays through one after the other, with no retries of failed exercises or choices about difficulty or ordering. The raw data provided by the Sixense motion control system is recorded and packaged for analysis; i.e., no real-time analysis or filtering of the data takes place. In effect, the patient uses the main build of the game as the rehabilitative aspect of the process, practicing the various exercises, and the assessment build is then used to feedback progress to the medical practitioners.

The first task for the patient during an assessment is the calibration process. The patient is instructed to carry out a series of motions which record the extent of the reach of each limb and where the midpoints are in all three dimensions. It is important to start each assessment with this calibration as the patient's reach may have improved since the previous assessment. The patient is then instructed to attempt each of the 12 to 40 assessment moves in turn. The raw data from each of the sensors (position and orientation) is stored throughout the exercises. No filters are applied, and the data is not interpreted in terms of percentages of the bounding boxes, so that any assumptions made in processing the input stream when attempting to recognize moves do not color the analysis. If the player fails to achieve an action within the time limit, the data is still recorded and the game moves on to the next exercise.

When the assessment is complete, the full data is packaged and labeled with a unique ID for the patient and the date. It can then either be collected by a therapist or uploaded automatically to the cloud, for analysis. The movements which comprise the assessment build have been chosen as those which give the therapists the most information about the progress of the patient's rehabilitation. They are not necessarily the moves which most contribute toward that rehabilitation. The assessment moves are distributed throughout the main version of the game, along with all the other rehabilitative moves.

Ordering the Moves

The rehabilitative actions which the patient is asked to perform are split into three categories: easy, medium, and difficult. Furthermore, within these categories there is a progression as more complex moves are introduced. The initial ordering was defined according to the therapists, but extensive play testing from patients revealed that some of the moves were incorrectly rated in terms of the challenge they presented. This required the order of the moves to be changed in later builds of the game. The system architecture has been developed to be both modular and data driven. This meant that reordering the moves simply involved changing the ordering in a data file: this is a common approach in well-constructed video game architectures. The exact content of the assessment build is also governed by a data file, so again it can be changed with ease as requirements evolve over the period of study.

The overall structure of the game is governed by a series of levels which are gradually unlocked as the patient progresses through the rehabilitative exercises. Initially, for example, only the easiest challenges are available; completing an easy challenge causes a more difficult challenge to become available. A level selection screen is also included. Again this is a fairly standard implementation within a video game; however, it has additional benefits in terms of rehabilitation. Presenting the player with only a few exercises to begin with encourages focus on those exercises which are initially most therapeutic and achievable, without overwhelming the new patient with too many options. Also, presenting the exercises as a series of levels which must be attained through success in earlier challenges gives the patient a series of goals to aim for. In gaming, it is important to provide a player with a constant stream of rewards, in order to maintain interest, and a classic example is unlocking extra levels as the game progresses (Koster 2013). This encourages the patient to not only continue with the exercises that have already been practiced but to strive toward more complex exercises.

Identifying the Paretic Limb

The exercises are designed to exercise the paretic limb; in many instances this entails different movements for the non-paretic limb. However, the affected arm could be either the left or right, so the software must take this into account. The GUI must show the avatar moving each limb appropriately, and the motion recognition algorithms must track the correct limb for each part of the exercises.

At the start of the calibration process, the patient is instructed to hold controller 1 in the good hand, to hold controller 2 in the hand with limited dexterity, and to clip controller 3 to their belt. The Sixense system uses a set of LEDs on each controller to show how they are numbered (based on the order in which they were detected at system start). The software then detects whether the controller in the paretic limb is to the left or to the right of the controller in the belt and in the non-paretic limb. Once this is established, the detection algorithms track the relevant hand for each half of the exercises – i.e., we consider the two limbs, and their associated bounding boxes, as the paretic and non-paretic limbs, rather than the left and right limbs. Also, if necessary, the graphics in the GUI for the avatar are flipped around the vertical axis. As the GUI consists of two-dimensional textures in an orthographic projection, this is a simple matter of reflecting the texture coordinates around the vertical center of the texture maps.

Patient Profile Data

The game saves profile data for each patient that plays it. This data serves two functions related to gameplay dynamics and ease-of-use and to patient monitoring.

The first type of data is that related to the player's parameters and progress. So that the patient does not have to go through the calibration process every time the game is played, the parameters of the bounding boxes are saved to the profile and reloaded on subsequent exercise sessions (the patient may use the application several times a day and is more likely to do so if the exercises can be accessed directly with minimal setup time). The profile also includes an indication of whether the paretic limb is the left arm or the right arm. Additionally game progress data is saved, which includes a list of whether each level has been unlocked in the progression, values pertaining to the achievements, and preferences on audio volume, brightness of the screen, etc.

A further set of profile data is generated for each session which contains details of the patient's progress and engagement. This data consists of information such as how long, and how often, each exercise was practiced, the success rates of the exercise attempts, and the scores attained. This data is collected by the therapist for further consideration and is clearly marked with unique identifiers for the patient and date. Further to this, the raw data from the Sixense input device is saved to a set of data files for further analysis by researchers on the project if required.

Conclusions Drawn from the Project

The development of a video game for rehabilitation of upper limb dexterity in stroke victims has been described in detail. The section has focused on the aspects of the technology and the design which contribute toward the game's rehabilitative nature, both in terms of providing the medically relevant exercise and feedback and presenting an entertaining and encouraging medium for that exercise. The game utilizes readily available motion capture technology and runs on a standard laptop, so that it can be used as often as the patient likes within the home.

The game was designed to incorporate the specific exercises that have been identified as contributing to upper limb rehabilitation, and it was developed to a commercial standard. This combination is intended to provide a sufficiently entertaining experience for the target patients so that they persevere with the course of exercises. The motion capture algorithms were designed specifically to monitor the asymmetrical upper limb movement exhibited by stroke survivors. The instructions given by the game, and the feedback on patient progress, are couched in familiar video game graphical user interface elements and are also presented in a way which does not penalize, or belittle, patients with more severe dexterity issues. This is achieved through the use of a separate bounding box for each limb, and measuring all movement as a percentage of the extents of the bounding box, rather than as absolute values. Familiar game design elements are also employed to create a progression of difficulty, and reward, as the patient's dexterity improves over the course of time.

The game embeds its rehabilitative exercises within a charming and friendly context of circus performances, utilizing state-of-the-art computer graphics, animation, and audio. Anecdotal reportage from the patient test group shows that the game is successfully encouraging more exercise and therefore higher compliance with the rehabilitative program.

Recommendations for Commodity Game Technology in Teletherapy

The chapter has addressed the use of commodity video game technology in teletherapy. The issues have been discussed in a practical manner relating them to the development of a game for rehabilitation of stroke and to a benchmark for motion capture input devices. A number of recommendations can be made from these studies, which can be applied to future projects applying serious gaming to teletherapy.

The choice of input device is of paramount importance in designing a serious game for rehabilitation. The benchmark described in this chapter has been applied to each of the currently commercially available devices, and each has been shown to have its own advantages.

- The PlayStation Move system demonstrates highest fidelity at range of two to four meters, making it an ideal option for a game played on the television in a family room environment. The disadvantage of this system is that it requires both a PlayStation console and a PC to operate in tandem.
- The Kinect motion capture system performs reasonably well at around 2 to 3 m, and the Kinect 2 extends this range to between 1 and 4 m. While the fidelity is not as that measured for PS Move, the Kinect systems have many additional features (face recognition, heartbeat monitoring, full skeletal tracking) which may be of interest in a specific teletherapy project.

- The Sixense motion sense controllers have the highest fidelity at closer distances (up to two meters), making them the ideal choice for a laptop-based application where the patient is closer to the screen. Further to this, as the Sixense uses magnetic fields, there are no issues with line of sight to the controllers (which mars the functionality of the other systems considered).

A rehabilitative game should be developed with both the medical requirements in place and the need for the game to be sufficiently involving that the patient makes maximum use of it. In other words, if the user is sufficiently entertained to regard the process as a game rather than a rehabilitative program, then the chances of continued engagement are increased.

- Feedback to the player should imply success. In the project described in this chapter, a patient with stroke does not necessarily have the full reach of the paretic limb. This is not reflected in the avatar or the instructions – if the patient is reaching as far as possible, then the avatar should show complete success.
- A modular game design has multiple advantages:
  - It is easier to change the order of levels or challenges, depending on difficulty testing or individual patient requirements.
  - It is a less complex task to bring each level or challenge up to a high-quality level, and to fix bugs, without affecting other sections of the game.
  - Assessment of each challenge can lead to removing the poorly performing ones, or identifying which need more work, without affecting the overall game.
- As much in-game help and feedback as possible is of high importance. If a patient is struggling to meet a challenge, or has forgotten what to do, the instructions should be readily available, on screen, and simple to follow. Incorporating videos of the actions being carried out can be a useful way of implementing this.
- The game itself should be entertaining and involving; animations and sound effects should be amusing. Engaging professional game developers is a route toward this.
- Commercial games keep player interest with rewards. This can be replicated in a teletherapy project. Methods of implementing this include high score tables, achievements, and daily challenges.
- Minimal setup time is also important. A patient is more likely to engage if the game is "plug and play." In particular the game should remember the patient's details so that progress can continue from one session to another without a lengthy configuration process.

Finally, it should be noted that games in teletherapy can be utilized to gather a vast amount of raw data on the patient's state and progress. Real-time gaming allows this data to be uploaded to the cloud or otherwise collected for further analysis.

It is hoped that this chapter will serve as a guideline to the development of gaming projects for teletherapy. The recommendations made, while based on actual projects, are intended to be applicable generally. Many design choices must be made when developing a game that will provide some form of rehabilitation or monitoring. These choices must ensure that the game will both provide the medical intention and engage the potential users to the point where they regard themselves as players rather than patients during the process.

Recommended Reading

S. Barreca, C.K. Gowland, P. Stratford, M. Huijbregts, J. Griffiths, W. Torresin, M. Dunkley, P. Miller, L. Masters, Development of the Chedoke Arm and Hand Activity Inventory: theoretical constructs, item generation, and selection. Top. Stroke Rehabil. **11**(4), 31–42 (2004)

N.A. Borghese, M. Pirovano, R. Mainetti, P. L. Lanzi, An integrated low-cost system for at-home rehabilitation, in *18th International Conference on Virtual Systems and Multimedia* (IEEE, Coventry, 2012), pp. 553–556

J.W. Burke, M.D.J. McNeill, D.K. Charles, P.J. Morrow, J.H. Crosbie, S.M. McDonough, Optimising engagement for stroke rehabilitation using serious games. Vis. Comput. **25**(12), 1085–1099 (2009a)

J.W. Burke, M.D.J. McNeill, D.K. Charles, P.J. Morrow, J.H. Crosbie, S.M. McDonough, Serious games for upper limb rehabilitation following stroke, in *Conference in Games and Virtual Worlds for Serious Applications* (IEEE, 2009b), pp. 103–110.

M.S. Cameirão, S. Bermudez i Badia, E.D. Oller, P. F. Verschure, Using a multi-task adaptive VR system for upper limb rehabilitation in the acute phase of stroke, in *Virtual Rehabilitation* (IEEE, Vancouver, 2008), pp. 2–7

R. Davison, S. Graziadio, K. Shalabi, G. Ushaw, G. Morgan, J. Eyre, Early response markers from video games for rehabilitation strategies. ACM SIGAPP Appl. Comput. Rev. **3**(14), 36–43 (2014)

E. Flores, G. Tobon, E. Cavallaro, F. I. Cavallaro, J. C. Perry, T. Keller, Improving patient motivation in game development for motor deficit rehabilitation, in *Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology* (ACM, Yokohama, 2008) pp. 381–384

S. Graziadio, R. Davison, K. Shalabi, K.M.A. Sahota, G. Ushaw, G. Morgan, J.A. Eyre, Bespoke video games to provide early response markers to identify the optimal strategies for maximizing rehabilitation, in *Proceedings of the 29th Annual ACM Symposium on Applied Computing* (ACM, Geungju, 2014), pp. 20–24

P.K. Hansen, Method and apparatus for position and orientation measurement using a magnetic field and retransmission. USA Patent 4,642,786, 10 Feb 1987

L.Y. Joo, T.S. Yin, D. Xu, E. Thia, P.F. Chia, C.W.K. Kuah, K.K. He, A feasibility study using interactive commercial off-the-shelf computer gaming in upper limb rehabilitation in patients after stroke. J. Rehabil. Med. **42**(5), 437–441 (2010)

M. Kimmerle, L. Mainwaring, M. Borenstein, The functional repertoire of the hand and its application to assessment. Am. J. Occup. Ther. **57**(5), 489–498 (2003)

R. Koster, *Theory of Fun for Game Design* (O'Reilly Media, North Sebastopol, 2013)

B. Lange, S. Flynn, A. Rizzo, Initial usability assessment of off-the-shelf video game consoles for clinical game-based motor rehabilitation. Phys. Ther. Rev. **14**(5), 355–363 (2009)

P. Langhorne, F. Coupar, A. Pollock, Motor recovery after stroke: a systematic review. Lancet Neurol. **8**(8), 741–754 (2009)

K. Morrow, C. Docan, G. Burdea, A. Merians, Low-cost virtual rehabilitation of the hand for patients post-stroke, in *International Workshop on Virtual Rehabilitation* (IEEE, New York, 2006), pp. 6–10

C. J. Murray, T. Vos, R. Lozano, M. Naghavi, A.D. Flaxman, C. Michaud, ... L. Bridgett, Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study. Lancet. **380** (9859), 2197–2223 (2010)

M. Pirovano, R. Mainetti, G. Baud-Bovy, P.L. Lanzi, N.A. Borghese, Self-adaptive games for rehabilitation at home, in *IEEE Conference on Computational Intelligence and Games* (IEEE, Granada, 2012), pp. 179–186

P. Rego, P.M. Moreira, L.P. Reis, Serious games for rehabilitation: a survey and a classification towards a taxonomy, in *5th Iberian Conference on Information Systems and Technologies* (IEEE, Santiago, 2010), pp. 1–6

N. Rossol, I. Cheng, W.F. Bischof, A. Basu, A framework for adaptive training and games in virtual reality rehabilitation environments, in *Proceedings of the 10th International Conference on Virtual Reality Continuum and Its Applications in Industry* (ACM, Santiago, 2011), pp. 343–346

G. Saposnik, M. Levin, Virtual reality in stroke rehabilitation a meta-analysis and implications for clinicians. Stroke **42**(5), 1380–1386 (2011)

J. Smisek, M. Jancosek, T. Pajdla, 3D with Kinect, in *Consumer Depth Cameras for Computer Vision* (Springer, London, 2013), pp. 3–25

A. Touillet, H. Guesdon, G. Bosser, J.M. Beis, J. Paysant, Assessment of compliance with prescribed activity by hemiplegic stroke patients after an exercise programme and physical activity education. Ann. Phys. Rehabil. Med. **53**(4), 250–265 (2010)

G. Ushaw, E. Ziogas, J. Eyre, G. Morgan, An efficient application of gesture recognition from a 2D camera for rehabilitation of patients with impaired dexterity, in *HEALTHINF* (INSTICC, Rome, 2013), pp. 315–318

G. Zachmann, Distortion correction of magnetic fields for position tracking, in *Proceedings of Computer Graphics International* (IEEE, Washington, DC, 1997), pp. 213–220

Network Support for Mobile Gaming

21

Armir Bujari, Marco Furini, and Claudio E. Palazzi

Contents

Abstract

Mobile gaming has become very popular thanks to its entertainment nature and to the widespread popularity of high-end mobile devices. The game scenario is very challenging as the support of mobile games is not as easy as one may think. In particular, the traffic generated by mobile games has specific network requirements that need to be satisfied; otherwise, the playability of the game might be annoying instead of pleasant. With this vision, we overview the network support solutions for mobile games. In particular, we characterize the mobile game network requirements and present studies that propose a suitable network support for these games.

A. Bujari (✉) • C.E. Palazzi
Department of Mathematics, University of Padua, Padua, Italy
e-mail: abujari@math.unipd.it; cpalazzi@math.unipd.it

M. Furini
University of Modena and Reggio Emilia, Modena, Italy
e-mail: marco.furini@unimore.it

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_24

Finally, we overview an interesting case study where mobile games can be played in an environment composed of fixed and mobile networks.

Keywords
Interactive games • Wireless network • Access point • Broadcast message • Energy consumption • MAC layer retransmissions • MANETs • Network changes • SAP

Introduction

The mobile gaming market continues to expand quickly and is equaling the revenues of the console game market. It follows that mobile games have become one of the most important digital platforms for gamers and publishers alike: as of 2014, mobile games account for over one third of monthly spending among digital gamers in the United States alone, and it reached the amount of US$21 billion revenue (Newzoo 2014).

Different actors are contributing to the success of this market: giants of the game industry, like Electronic Arts and Ubisoft, are transforming console game into mobile game; novel companies, like King and Zynga, are releasing more and more titles for the mobile gaming market; and cell phone industries are producing devices with more and more appealing multimedia features. As a result, the mobile gaming scenario is filled with a wide range of games: from very simple graphic games to cutting-edge 3D graphics and from single- to multiplayer games (Furini 2008; Palazzi and Maggiorini 2011).

The most common mobile gaming scenario is composed of players (from a few to hundreds of thousands) who use their own device (e.g., a cell phone, a game console) to play a game, of a game server (or of a series of game servers) that is in charge of running the game properly, and of a network that connects devices and server(s). To play the game, each player must interact with the server (and/or with the other players), and the server has to process each player's input, has to compute the new game state, and has to communicate this state to each player. Indeed, the game normally proceeds through states, and to ensure that all players will perceive the same game evolution, the state of the game should be the same in the whole scenario. Unfortunately, the support of mobile games is not easy as one may think, and the variation of the network latency may cause players to experience a game state inconsistent with the one of the game server or with the one of other players. Needless to say, this is a critical issue.

Indeed, mobile games have peculiar characteristics (interactivity, consistency, fairness, scalability, and continuity) that require special services from the network (e.g., very low end-to-end delays). For instance, to provide interactivity, the time elapsed from the game event generation at a certain node and its processing time at every other node participating in the same game session must be kept under a certain interactivity threshold; consistency requires all the players to view the same game state; fairness requires the network to not affect the chances of winning; and continuity requires the network to avoid interrupted game sessions usually caused by disconnections, handoffs, or any other wireless-/mobility-related issue. If the network does not support these peculiar characteristics, the performance of the mobile games might result very poor.

In essence, the traffic produced by mobile games has timing constraints (i.e., it is time sensitive), and the network is required to meet these constraints in order to well support these games. Unfortunately, in a mobile scenario, the service provided by the network might be not sufficient to meet the mobile game requirements. For instance, in an infrastructure-based scenario, the Access Point used to reach the Internet could be employed for heterogeneous flows and applications, both elastic and real time, thus introducing delays that would jeopardize the performance of interactive games. Similarly, if a node moves out of the AP transmission range, it loses all the ongoing sessions. The problems are exacerbated in infrastructure-less networks (e.g., ad hoc, mobile, and vehicular) that are prone to disconnection, energy consumption, etc. For these reasons, the game scenario has been studied by researchers and practitioners in the attempt to provide an effective and efficient supporting scenario (Furini 2007; Griwodz 2002).

In this chapter we focus on network support for mobile games, as depending on the networking performance, the user may experience a pleasant mobile game or an annoying one that will discourage him/her from playing again. In particular, we present what are the mobile game network requirements and what are the current studies that aim at providing a suitable network support for these games. Finally, we describe a particular architecture that seems to be a promising solution to support games in a mobile scenario.

The remainder of the chapter is organized as follows. Section "Mobile Game Network Requirements" describes the typical network requirements of a mobile game, and section "Network Support for Mobile Games" presents the services provided by current wireless networks. Section "Network Architectures" presents the characteristics of the most used architectures able to support mobile games, and section "Smart Architecture: A Case Study" presents a promising holistic solution for the support of mobile games in the wireless scenario. Finally, section "Conclusions" concludes this chapter.

Mobile Game Network Requirements

The mobile game scenario is composed of games that require special services from the network, as the delay, the jitter, and the transmission rate directly affect the quality perceived by players. Before detailing the network requirements of online games, it is worth describing a practical example of problems that might arise when playing online games. Figure 1 shows the frame evolution of an Armagetron game session (Armagetron 2014), where user Cla (blue cycle and wall) plays against user Eu (green cycle and wall). The game's rules are simple: each player controls a light cycle that leaves a wall behind it wherever the cycle goes; the cycle cannot stop and can turn only at 90° angles. The goal of the game is that of having all other players

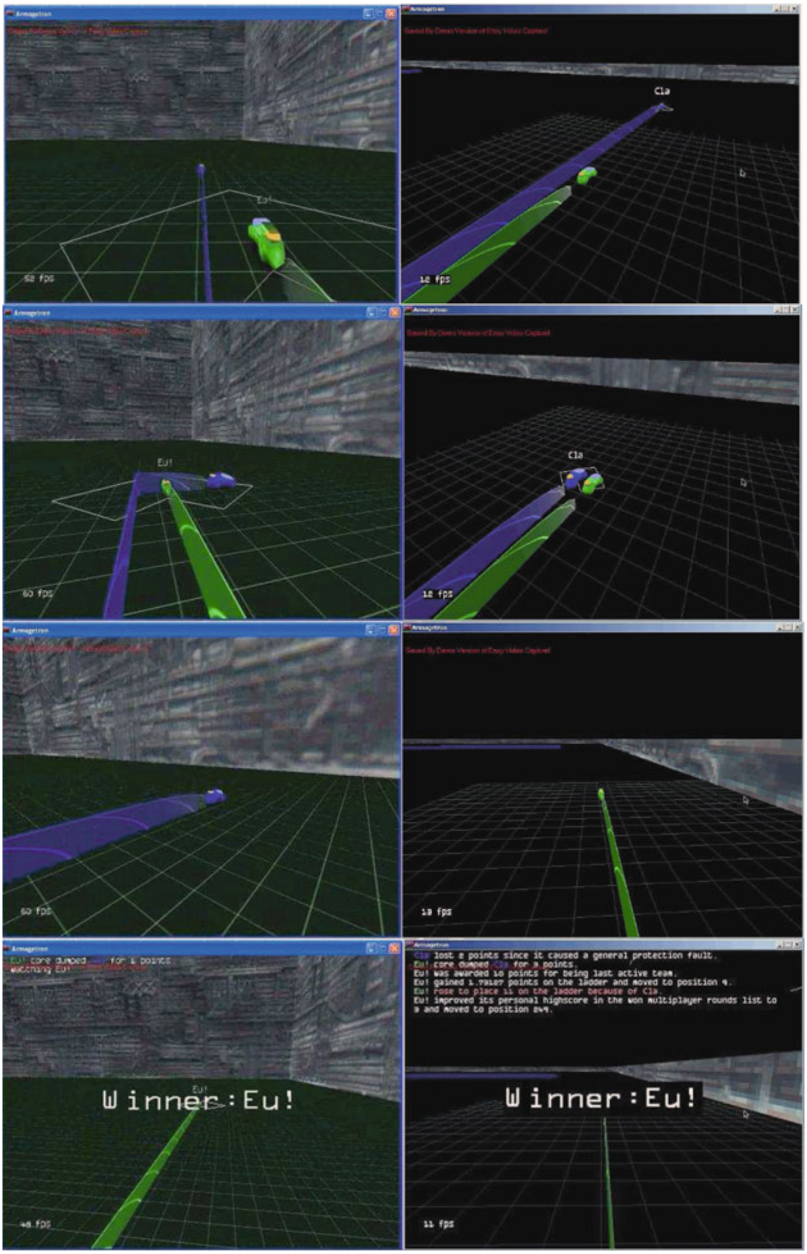


Fig. 1 Frame sequence of an Armagetron game session: *Cla*'s view (*left column*, blue player) versus *Eu*'s view (*right column*, green player); evident lag differences generate inconsistencies and unfairness

Fig. 2 Frame sequence of a game session as seen by a client: other participants are positioned in the game arena with a lag that depends on the distance between clients and server

crashing into some wall while avoiding hitting others' own wall. Speed helps players in trapping other players, but the only way to speed up a light cycle is to drive very close to a wall.

In the left column of Fig. 1, there are subsequent frames as seen by the player Cla, who is connected to a certain game server with 180 ms of round-trip time (RTT). Instead, frames on the right column correspond to the game as seen by the player Eu, who is connected to the same server with only 20 ms of RTT. Frames on the same row relate to the same instant of the game action; it is hence very easy to notice the inconsistency between the two game state views by simply comparing the position of one light cycle with respect to the other. In particular, during the second and third row of frames, player Cla believes he has just won (in the left frame of the second row, Cla sees Eu hitting his wall), whereas after a couple of seconds (last frame row), he realizes that the server has declared Eu as the winner. This is a clear sign of inconsistency and unfairness. Player Cla would surely refrain from renewing his subscription to the game. Even if the game were free, player Cla would probably stop playing to avoid the frustration of sure, yet undeserved, failing.

The reason for this inconsistency and unfairness is in the different RTT experienced by the two game connections and embodies an example of how the network may affect the playability of a mobile game. Indeed, the real game evolution is the server's view, whereas Cla and Eu visualize their own information by combining local player's movements with (differently delayed) game server's updates. As a result, Eu sees the game action evolving in advance with respect to Cla, as the former is much closer to the game server than the latter. To clarify this aspect, Fig. 2 presents another practical example on how network delays might affect consistency. The picture presents both the position of a certain participant as seen by another playing client (the full figure avatar) and the position as seen by the server (the human-shaped light boxes). It can be noted that transmission lag creates a difference in the avatars' positions as perceived by the server and by each of the various clients.

Inconsistency and unfairness are only two among the various problems that a network may cause to a mobile game. Indeed, in general, online games require the network to support five main features that are intrinsically correlated: interactivity, consistency, fairness, scalability, and continuity.

Interactivity: also known as responsiveness, it refers to the delay between the generation of a game event in a node and the time at which other nodes become aware of that event. In order to ensure an enjoyable playability to the final user, the time elapsed from the game event generation at a certain node and its processing time at every other node participating in the same game session must be kept under a certain interactivity threshold (Armitage 2003; Panten and Wolf 2002). Unfortunately, variable network congestion conditions may suddenly slow down the game fluency on the screen. Moreover, players in the same virtual arena could be so numerous that some game servers may experience impulsive computational load and lose interactivity. These problems are obviously amplified when plunged into a wireless, mobile scenario. In fact, wireless characteristics and node mobility generate a high variability in experienced delays and network traffic and also generate a high churn rate that worsens the aforementioned problems (Nandan et al. 2005a).

Consistency: it refers to the simultaneous uniformity of the game state view in all nodes belonging to the system. Also in this case, the network response might affect the consistency and playability of the game. The easiest way to guarantee absolute consistency would be that of making the game proceed through discrete locksteps. Having a single move allowed for each player and synchronizing all agents before moving toward the next round surely grants absolute consistency but, on the other hand, impairs the system interactivity. A trade-off issue between consistency and interactivity needs thus to be solved in order to develop a proficient game platform.

Fairness: every player should have the same chances of winning the game session, regardless of different network conditions. In this context, relative delays have to be considered as important as absolute ones. Simultaneous game evolution with identical speed should be guaranteed as much as possible to all participants. To this aim, it has been demonstrated how increasing the interactivity degree of the game platform may lead also to improved fairness (Ferretti et al. 2005). However, it is known that the wireless environment generates peculiar delays and related unfairness problems. These problems came from the nature of the shared wireless medium and from location dependency. In fact, if we view a node and its interfering nodes to form a "neighborhood," the aggregate of local queues at these nodes represents the distributed queue for this neighborhood. This aggregate queue is not FIFO served, as flows sharing the queue have different, dynamically changing priorities determined by the topology and traffic patterns. Thus, they get different feedback in terms of packet loss rate and packet delay when congestion occurs (Xu et al. 2005).

Scalability: it regards the capability of the system in providing efficient support to a large community of players. Indeed, it is of primary interest for game companies to have revenues generated by a very high number of customers. Besides, humans are social beings that enjoy the presence of others and the competition against real adversaries. Yet, especially in the case of fast-paced games, when the interactivity threshold cannot be met, scalability is sometimes sacrificed by denying the access to some users depending on their experienced delays (Rakion 2014). Therefore, by sustaining interactivity, one can also provide a higher scalability degree in terms of both the number and the geographic dispersion of players allowed to participate to the same virtual arena. Discussing scalability, it is particularly interesting to notice that wireless connectivity provides means to widen the set of potential players as it is no longer required to be wired to the Internet to engage online games. Think of the possibility to establish outdoor online gaming sessions through 3G, WiMAX, vehicular IEEE 802.11p, ad hoc networks, mesh networks, etc. Clearly, this brings novel issues into the scenario, such as disconnections and energy consumption (Marfia and Roccetti 2010; Roccetti et al. 2010).

Continuity: it is concerned with having game sessions not interrupted by disconnections, handoffs, or any other wireless-/mobility-related issue. Indeed, players would be very frustrated by having their game session continuously interrupted and restarted (maybe after a while). This problem may happen also when trying to exploit the new wireless capabilities of popular smartphones to create proximity-based gaming sessions. Indeed, we can imagine players forming an ad hoc network to engage in an outdoor multiplayer game based on the connectivity means of their smartphones. Yet, players' movements and different energy consumption among devices may create detached cluster of nodes (Kokkinos et al. 2011; Yu et al. 2003). To this aim, proposed solutions regard games featured with short game sessions, having quick/smart handoff mechanisms, and considering route and server migration (Chen et al. 2006; Palazzi et al. 2010a; Zhu et al. 2011).

Network Support for Mobile Games

To ensure a good playability in the mobile game scenario, the network should provide a service able to guarantee the five features described in the previous section. To this aim, many researchers have focused their studies on the issues encountered in a wireless environment. In the following, we analyze problems and solutions that have emerged while supporting mobile games through infrastructure-based and infrastructure-less wireless networks.

Infrastructure-Based Wireless Network

In this type of network, there is an Access Point (AP) that has access to the Internet, and all the other mobile nodes connect to the AP to have Internet access. In this scenario, the AP offers to all engaged mobile nodes the same functionalities. An example of this network can be found in the home scenario (but also in public areas), where a Wi-Fi network can be used to establish an interactive gaming session or to connect to remote players or game servers. There are two main drawbacks in this scenario: (i) if a mobile node moves out of the transmission range of an AP, it loses its connectivity to the Internet and therefore the mobile game stops working, and (ii) the AP could be employed to manage the traffic produced by both elastic and real-time applications, thus introducing delays that would compromise the playability of interactive games (Palazzi et al. 2007).

To mitigate these problems, researchers proposed to extend the range of a wireless infrastructure with the employment of mesh networks, which represent a way to create wireless chains among APs. In this scenario, if the moving node enters in an area covered by a new AP, it might be able to connect to the Internet again. However, it has to deal with handoffs that probably terminate its ongoing sessions. This problem can be solved through a smooth handoff that seamlessly transfers the connectivity from the old to the new AP before disconnection. A well-known protocol to perform this task is represented by Mobile IP that transfers packets from the old AP to a new one through routing triangulation (Perkins 2002). Therefore, this network topology can hypothetically support infrastructure-based wireless mobile games where players are located and moving in an area wider than one single AP's transmission range, but the problems due to the management of heterogeneous traffic flows remain.

To mitigate these problems, many research papers focused on IEEE 802.11 and presented analysis, problems, and solutions. Unfortunately, the vast majority of them focused on a throughput/losses point of view (Bianchi 2000; Bottiliengo et al. 2004; Heusse et al. 2003), and therefore they cannot be applied to the mobile game scenario. In fact, the performance of mobile games (and of real-time applications in general) depends on the measured per-packet delay and jitter and not on the achieved throughput/losses (Beigbeder et al. 2004; Conti et al. 2002). For this reason, different studies focused onMAC layer retransmissions, showing that MAC retransmissions can be wasteful and potentially harmful for time-sensitive applications as they introduce delays (Palazzi et al. 2005b); however, other studies showed that without retransmissions implemented at the link layer, the loss rate becomes unacceptable for any application (Nam et al. 2003; Xylomenos and Polyzos 1999). Similarly, another important issue is that persistent TCP-based flows (e.g., downloads) are responsible for performance deterioration of concurrent UDP-based flows (e.g., interactive games). Indeed, the continuous search for more bandwidth performed by TCP's congestion control algorithm creates queues at buffers, thus augmenting the per-packet delay of any flow sharing the same channel. With the aim of ensuring low per-packet delays, while preserving downloading throughput, Palazzi et al. (2006b) proposed a solution based on a smart AP (SAP). The idea is to monitor the entire traffic passing through and, for each TCP flow, to compute the maximum transmission rate so as to not exceed the channel capacity and accumulate packets in queue. Only standard features and protocols are used to facilitate deployment: the maximum transmission rate for each flow is enforced by modifying on the fly the advertised window in TCP ACK packets. We observe that an insightful analysis of this issue with respect to online games is still missing, as well as efficient solutions aimed at reducing queuing delay over wireless links (Hole and Tobagi 2004; Palazzi et al. 2005b).

Infrastructure-Less Wireless Network

In an infrastructure-less network (aka ad hoc networks) wireless nodes can communicate with any other node in its transmission range without the need of any AP. Transmissions can happen both directly between two nodes that are close to each other and through multi-hop when a node needs to send a message to another node that is out of its range. Ad hoc networks have received the attention of researchers and practitioners thanks to their high deployment flexibility, low cost, and robustness, which make them perfectly suitable for a whole plethora of scenarios where the infrastructure is missing, e.g., away from towns, in areas hit by a major disaster or just not covered by APs, in military battlefields, etc. In this scenario, Bononi et al. (2009) showed that ad hoc networks can be used to support both real-time and non-real-time applications provided that each node manages the outgoing traffic with specific transmission scheduling and with (at least) two different queues: one dedicated to the traffic generated by elastic applications and another one dedicated to the traffic coupled with timing constraints.

A particular type of ad hoc network is the one where nodes move (Corson and Macken 1999). This scenario is known as mobile ad hoc networks (MANETs) and represents a typical outdoor gaming scenario, where players' devices dynamically connect to each other creating an ad hoc network to support their multiplayer games. This scenario is very challenging for two main reasons: disconnections and energy consumption. MANETs are prone to disconnections. Due to mobility, one or more nodes may get out of range of the original ad hoc network, becoming unable to reach the game server node and continue playing. Even worse, the server node itself may get out of range for the rest of the network, thus interrupting the game session for everybody.

Energy consumption might be a major concern when considering outdoor games based on the connectivity of small devices because of the limited amount of energy stored in their batteries. This limited energy can be quickly consumed by game-related computation, visualization, and communication. Clearly, the worst energy consumption is experienced by the node that is also the server of the game session: that node will experience a much faster decline of its energy reserve as it will have to receive all game events from other players, compute game state updates, and transmit these updates back to the other players. Many solutions have been proposed to ensure data transmissions over MANETs while aiming at low energy consumption (Yu et al. 2003). These solutions are generally based on having nodes alternating sleep/awake modes, and therefore they cannot be exploited for interactive mobile games due to the timing constraints of the produced traffic. Other solutions exist that aim at saving nodes' energy through smart routing in a MANET (Kokkinos et al. 2011; Yu et al. 2003). Yet, they do not consider any delivery delay options, thus not necessarily ensuring interactivity. Moreover, the scenario with one of the nodes also embodying the game server is not considered. In this scenario, one single node (i.e., the game server) receives and generates most of the traffic. The application itself and the network architecture impose unbalanced energy consumption. Thereby, the server (and player) node will run out of energy much before the other player nodes, yet interrupting the gaming session for every node. To address this issue, Palazzi et al. (2010a) proposed to utilize three kinds of nodes in the game network: active servers, backup servers, and players. In essence, the traffic produced by mobile games has timing constraints (i.e., it is time sensitive), and the network is required to meet these constraints in order to well support these games. In this solution, all servers collect game events sent by other players and continuously update their game state view. However, only the active server forwards the current game state to all the players (including the backup servers); this limits the energy consumption of backup servers until they are called to become active.

Another particular type of ad hoc network is the one where nodes move at high speed. In this case, the network is named vehicular ad hoc networks (VANETs) and represents a scenario where mobile games will probably play a successful role (Palazzi et al. 2010b). Similar to the MANET scenario, in a VANET, players' devices dynamically connect to each other creating an ad hoc network to support the game. In particular, transmissions can follow either a pull or a push model (Nandan et al. 2005b). With the former, vehicles explicitly ask to receive certain data from other vehicles, whereas with the latter, messages are proactively broadcast to every node in a certain area of interest. Both models can involve multi-hop transmissions. Since the shared nature of the wireless channel, communications involving several nodes in the same VANET are clearly more efficient if performed through broadcasting (i.e., push model). Having many connected players moving fast in their cars certainly represents one of the most challenging and interesting context for infrastructure-less mobile games. Fundamental issues for mobile games such as the quick propagation of game events among players are exacerbated in this context and require special attention (Beigbeder et al. 2004; Biswas et al. 2006); similarly, interference and sudden congestion may be exacerbated by the very high mobility of vehicles causing the management of these networks to be tougher than the one of MANET (Casteigts et al. 2011). Moreover, this scenario is very challenging for two main reasons: management of network changes and the broadcast of a message.

Network changes. The management of network changes in a fast mobility scenario represents a tough issue both for ensuring connectivity and for the high variability in concurrent network traffic (Zhu et al. 2011). Indeed, since the topology of a vehicular network (vehicles around a certain player and their concurrent data traffic) may continuously and quickly change, the traffic may experience sudden delays that will likely jeopardize the interactivity of the gaming session (Palazzi et al. 2010b).

Broadcast of a message. The broadcast of a message also represents a tough issue (Yang et al. 2004), as its delivery is not fast as it should be. Experts report that main reasons behind a slow broadcast delivery are due to a nonoptimal number of hops experienced by a message to cover all the involved cars and, more in general, to an excessive number of vehicles that try to simultaneously forward the message (Biswas et al 2006; Fasolo et al. 2005; Korkmax et al. 2004; Yang et al. 2004). To tackle this problem a theoretically optimal broadcast algorithm has been proposed (Zanella et al. 2004). However, there are practical difficulties in the implementation of this algorithm as it would require a complete and continuously updated knowledge of the network topology. For instance, with N cars, the algorithm employs as many as O(N log N) control messages (Wan et al. 2002). It goes without saying that this is not a scalable solution. A backoff mechanism that reduces the frequency of message retransmissions when congestion is causing collisions is proposed in Yang et al. (2004). In Biswas et al. (2006), instead, as soon as a car receives a broadcast message from a following vehicle along a strip, it refrains from forwarding it as the reception of this message is a clear confirmation that subsequent cars have already received it. Unfortunately, both these two schemes do not consider a very important factor in determining the final propagation delay of a message: the number of hops a broadcasted message traverses before covering its whole area of interest. The solution presented in Bononi and Di Felice (2006) utilizes a distributed proactive clustering scheme to dynamically create an efficient virtual backbone infrastructure in the vehicular network. The backbone formation process takes into consideration both the current distance among candidate backbone vehicles and the estimated lifetime of the wireless connection among neighbor backbone members. The attempt is that of improving the robustness and lifetime of connections among backbone members even in a highly mobile scenario as a vehicular network. To minimize the number of hops, Korkmax et al. (2004) proposed to individuate the farthest car within the source's backward transmission range, which has to forward the message. To this aim, jamming signals are emitted by each car with a duration that is directly proportional to the distance between the considered car and the message's source. The car with the longest jamming signal is clearly the farthest car from the source. Even if this guarantees a minimum number of hops to cover the whole area of interest, the time wasted to determine the next forwarder through jamming signals could make this scheme not suitable for a mobile game scenario. Fasolo et al. (2005) proposed to assign different contention windows to each car to have different waiting times before propagating the broadcast message. Nodes set their respective contention windows with an inverse proportion of the distance from the sender, thus needing less forwarders (and transmissions) to cover a certain area. Yet, this scheme assumes that there is a unique and constant transmission range for all the cars in every moment; this is obviously not realistic in a VANET because of its high and fast mobility. Other solutions have hence been devised to solve this shortcoming. In particular, similar automatic transmission range estimators are proposed in Palazzi et al. (2007) and Roccetti et al. (2010) to assess the actual transmission range for every car in the platoon. More in detail, the former exploits this information to have the farthest vehicle (i.e., the farthest relay) in the sender's transmission range becoming the next forwarder of the message. Instead, the latter uses this information to assign the forwarding task to the farthest spanning relay. In both cases, the computed transmission range estimation is used to support a multi-hop broadcasting scheme for message exchange able to dynamically adapt to the different (transmission range) conditions a vehicular network may encounter.

Network Architectures

To support mobile games, many researchers focused on the design of network architectures to ensure a good playability. According to their characteristics, these proposals might be grouped into three different categories (centralized client–server, fully distributed, and mirrored game server). In the following, we present the main advantages and disadvantages of these categories when applied to a mobile game scenario.

Centralized client–server. It is an architecture composed of a single authoritative point (the server) which is responsible to run the main logic of the game, execute players' commands, enforce consistency, send back to the client the new game state update, etc. Clients have only to receive the new game state, render it on the screen, and forward player's commands. The server can be both a single computer and cluster of computers in order to increase the performance of the system (Butterfly 2014). The centralized client–server architecture represents the simplest solution for authentication procedures, security issues, and consistency maintenance (Gautier and Diot 1998; Quake 2014; Ultima 2014). This architecture, assuming to have N simultaneous players, generates a number of messages in the order of O(N), but the presence of a single authoritative point represents a unique bottleneck that limits its efficiency and scalability.

Fully distributed. It is an architecture that well represents the peer-to-peer paradigm: all the involved nodes share the same intelligence and are equally responsible for running the whole logic of the system; each client has to autonomously update the game state view based on its player's commands and on game actions received from other players. The main advantage in employing a fully distributed architecture is that of spreading the traffic load among many nodes, thus generating a more scalable and failure-resilient system (Gautier and Diot 1998; Safaei et al. 2005). However, this approach requires terminals endowed with higher computational capabilities, and identical copies of the current game state need to be stored at each node. Therefore, it is necessary to introduce some complex coordination mechanism among peers; in fact, this scheme has to be distributed over the set of involved nodes and has to be able to guarantee the coherence of all game state views. The exchanged messages could hence rise to the order of $O(N^2)$, where N is the number of simultaneous players. Finally, authentication, cheating, and general consensus among all the peers are harder to be addressed than when a centralized architecture is employed. However, the fully distributed architecture is generally preferred for infrastructure-less networks such as MANETs and VANETs because of its ability to deal with high mobility of nodes that continuously changes the topology of the network (Palazzi et al. 2010a).

Mirrored game server. It is an architecture that represents a hybrid solution able to efficiently embrace all the positive aspects of both centralized client–server and fully distributed architectures (Palazzi et al. 2006a). When employing this architecture, the game state servers (GSSs) are interconnected in a peer-to-peer fashion over the Internet and contain replicas of the same game state view. Players communicate with their closest GSS through the client–server paradigm. Each GSS gathers all the game events of its engaged players, updates the game state, and regularly forwards it to all its players and GSS peers. There are three main advantages in employing a mirrored game server architecture: (i) the absence of a single point of failure, (ii) the networking complexity that is maintained at the server side, and (iii) the possibility to easily implement authentication procedures. Even if synchronization is still required to ensure the global consistency of the game state held by the various servers, this requirement is made easier with respect to fully distributed architectures thanks to the lower number of involved nodes. Assuming to have N simultaneous players and M GSSs, for example, the generated game messages amount to O(N+M), which is again O(N) unless considering the unlikely case of having more servers than players. The presence of multiple high-performance GSSs helps in distributing the traffic over the system and reduces the processing burden at each node (Safaei et al. 2005). Moreover, having each player connected to a close GSS reduces the impact of the player-dependent access technology (e.g., dial-up, cable, DSL) on the total delay experienced (Jehaes 2003). Indeed, the communication among players results mainly deployed over links physically connecting GSSs, which can exploit the fastest available technology (e.g., optical fibers) to reduce latency. As a result, through this architecture, it becomes simpler to adopt efficient solutions for the trade-off among the five main features of a mobile game (i.e., interactivity, consistency, fairness, scalability, and continuity). For instance, Palazzi et al. (2005a) suggested that during a game session some events can lose their significance as time passes, and therefore discarding superseded events for processing fresher ones may be of great help for delay-affected GSSs, achieving high interactivity degree without compromising consistency. Furthermore, for very fast-paced games, little inconsistencies are not highly deleterious for players' fun. In these cases, even some non-superseded game event could be dropped when dropping all superseded ones is not yet sufficient to maintain an adequate level of responsiveness.

Given the available architectural solutions, we can hence infer that when considering a (moderately) mobile network of players supported by infrastructure and by Access Points, the hybrid solution of mirrored servers probably represents the best choice. Conversely, in case of no available infrastructure or very high mobility of players (e.g., cars' passengers), then a fully distributed solution may be the only feasible option.

Smart Architecture: A Case Study

In this section we review the Smart Architecture, an interesting case study where mobile games can be played in an environment composed of fixed and mobile networks (see Fig. 3). The Smart Architecture embraces the positive aspects of the

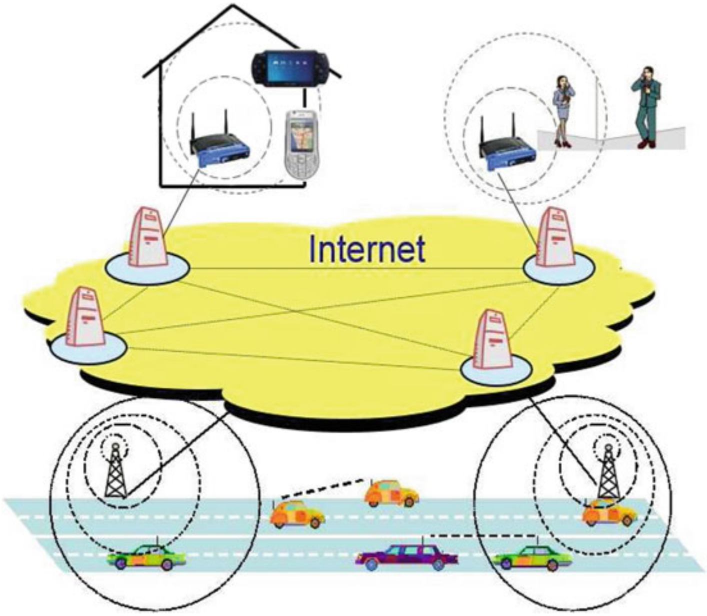


Fig. 3 Hybrid architecture for distributed game entertainment in heterogeneous scenarios with mobile users (even car passengers)

centralized client–server and of the fully distributed architectures, and therefore it can be considered a hybrid between the two architectures (Cronin et al. 2004). Experiments showed the benefits of using the hybrid architecture when supporting mobile games. In the following, we review the two main components of the architecture: (i) a smart synchronization mechanism for mirrored game servers and (ii) a smart AP able to avoid last-mile queuing delays that would jeopardize interactivity just one step before delivering.

Smart Synchronization Mechanism Among GSS

To speed up the synchronization among GSSs, the Smart Architecture employs a smart synchronization mechanism (SSM) that exploits the semantics of the game to discard few game packets in order to preempt interactivity loss when intense network traffic or excessive computational load is slowing down some GSS (Palazzi et al. 2006a; Safaei et al. 2005). The idea behind the SSM takes inspiration from the active queue management approaches employed in RED (Floyd and Jacobson 1993) and RIO (Clark and Fang 1998) and utilizes a uniformly distributed dropping function. Yet, the parameter taken under control by GSSs is the time elapsed from the generation of the game event, which is named Game Time Delivery (GTD). In fact, upon each packet arrival, each GSS determines the GTD of the arrived event, namely, sample_GTD, and feeds a low-pass filter to compute

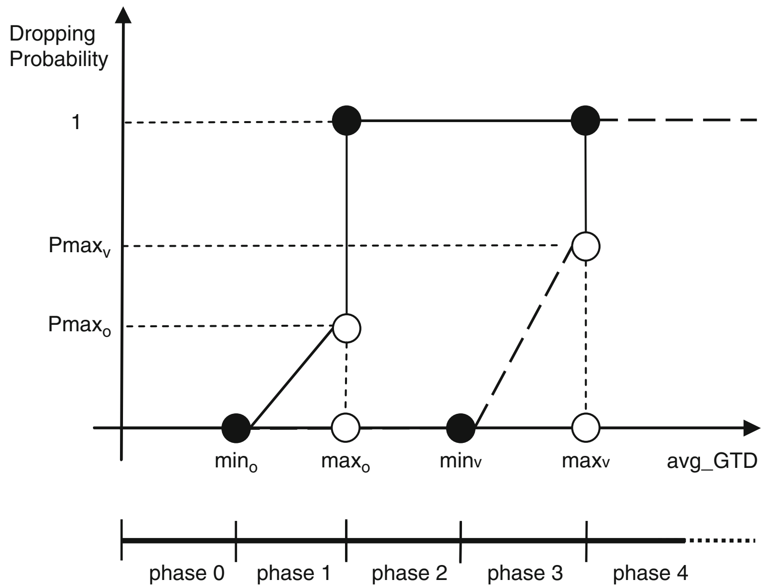


Fig. 4 Discarding probability functions

the updated average GTD, namely, avg_GTD. When avg_GTD exceeds a certain threshold, the GSS drops superseded events with a certain probability p, without processing them. If avg_GTD exceeds a subsequent limit, p is set equal to 1, and all superseded events waiting for being processed are discarded.

Indeed, during a game session some events can lose their significance as time passes, i.e., new actions may make the previous ones irrelevant. For example, where there is a rapid succession of movements performed by a single agent in a virtual world, the event representing the last destination supersedes the previous ones.

To ensure an adequate playability degree even to fast and furious class of games, a further dropping probability function is provided in order to discard even non-superseded game events when dropping all the superseded ones is not yet sufficient to maintain an adequate level of responsiveness.

Figure 4 depicts the two discarding functions of the SSM. Three parameters (and three phases) characterize each of the twin algorithms: min_o, max_o, and $Pmax_o$, for superseded events, and min_v, max_v, and $Pmax_v$ for non-superseded ones.

A Smart AP

Even if SSM coupled with a mirrored game server architecture is proficient in maintaining a high degree of responsiveness among GSSs, still problems may arise at the edges of the considered topology, where users in their homes or along a street may be engaged in an online game through an AP (see Fig. 3). Concurrent traffic may generate queues that build up at the AP, thus delaying the game event delivery and wasting all the interactivity patrimony created by the SSM. In particular, as previously mentioned, some applications may be particularly harmful toward online gaming traffic as they may increase queuing delays to such an extent that interactivity may be completely compromised (Palazzi et al 2005b).

The smart AP aims at achieving the best performance for both elastic and real-time applications by appropriately limiting the advertised window for TCP flows (Palazzi et al 2006b). An optimal trade-off between throughput and low delays could be achieved by maintaining the sending rate of TCP flows high enough to efficiently utilize all available bandwidth but, at the same time, limited in its growth so as to not utilize buffers. As a result, the throughput is maximized by the absence of packet loss, while the delay is limited by the absence of queuing.

In essence, the maximum sending rate for each TCP flows at time t, namely, TCPubrate(t), can be represented by:

$$TCPubrate\ (t) = \frac{(C - UDPtraffic\ (t))}{\#TCPflows\ (t)} \tag{1}$$

where *UDPtraffic (t)* is the amount of bandwidth occupied by UDP-based traffic at time t, *#TCPflows (t)* is the concurrent number of TCP flows, and C is the total capacity of the bottleneck link. This upper bound can be enforced to all TCP flows sharing the same wireless link by having the corresponding AP appropriately modifying the advertised window of passing-through TCP flows.

To compute (1), a comprehensive knowledge of the flows transiting through the bottleneck (i.e., the last hop links) is needed: the total capacity of the channel, the aggregate amount of current UDP traffic, and the number of TCP flows currently active on the wireless link. This information is indeed possessed by the AP as all flows have to pass through it.

A mathematical model evaluated the Smart Architecture. In particular, the model considered a lognormal distribution for the DTD game traffic (Park and Willinger 2002), and the simulated scenario was the one of Fig. 3: seven interconnected GSSs with a network latency between the two farthest GSSs of 90 ms, with GSS that transmits to the other GSS game updates related to their engaged clients every 30 ms.

Results were gathered collecting the total latency experienced by game events reaching one of the clients connected through an IEEE 802.11g AP to one of the GSS. The same AP was also in charge of handling traffic coming from other applications run on different devices that were simultaneously sharing that wireless link: a UDP-based video stream, a UDP-based live video chat, and a TCP-based downloading session. The video stream and video chat applications were simulated by injecting in NS-2 real traces corresponding to high-quality MPEG4 Star Wars IV and VBR H.263 Lecture Room-Cam, respectively, as available in Movie (2014).

In the following charts, REG represents the case where a regular synchronization scheme is adopted by GSSs, whereas SMA represents the case employing the Smart Architecture. Finally, GIT (Game Interactivity Threshold) represents the maximum

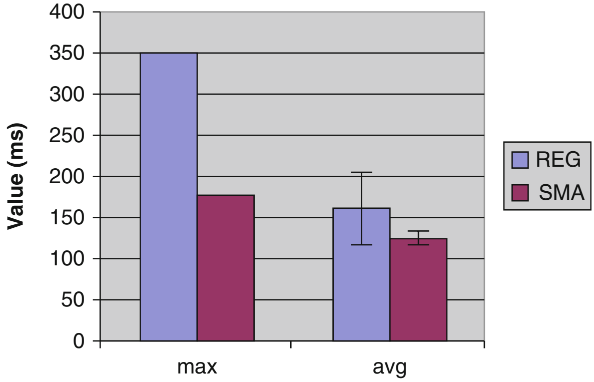


Fig. 5 SSM evaluation: statistical values for the delivery delay of game events (packets) from their generation to the GSS engaging the considered player

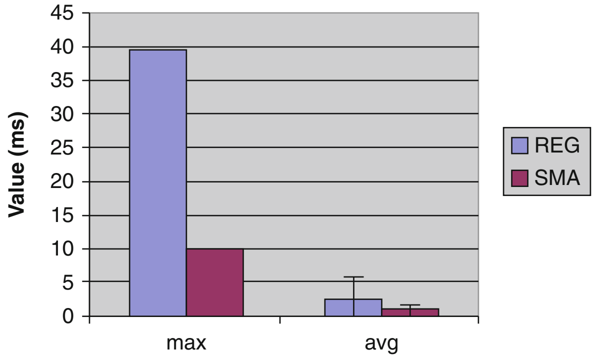


Fig. 6 Smart AP's evaluation: statistical values for the delivery delay of game events (packets) from the AP to the considered player

delay that a game event can experience from its generation to its delivery to the gaming device to preserve interactivity. As a reference, the threshold of 150 ms was considered based on related scientific literature for interactive online games (Pantel and Wolf 2002).

Focusing on the interactivity benefits provided by the first component of the Smart Architecture, i.e., the SSM coupled with a mirrored game server architecture, Fig. 5 shows the maximum and the average with the standard deviation bars of the delivery time that game events experience to reach the GSS that supports the considered player. The chart reports statistical values about the time elapsed from the generation of a game event and its delivery to the GSS that will then forward it to the considered player. Clearly, SMA outperforms REG, which demonstrates the effectiveness of FS in quickly synchronizing GSSs.

Figure 6 shows the impact of the wireless last hop on the game interactivity. In particular, the chart reports statistical values related to time elapsed from the moment when the AP receives the game update from its GSS and the moment when the considered player actually receives it on his/her mobile device. Basically, Fig. 6 highlights the effectiveness of the smart AP versus a traditional one.

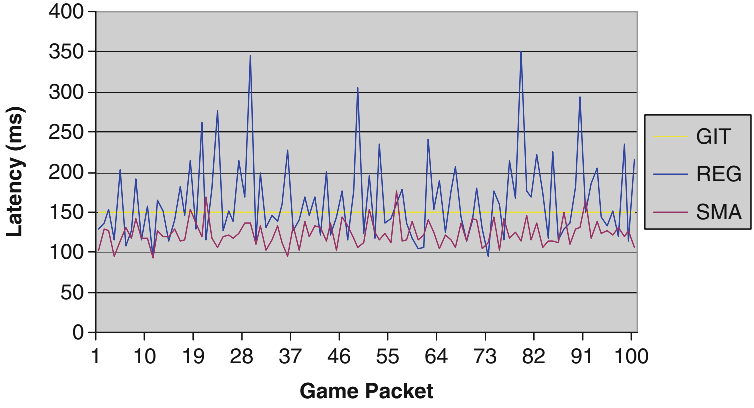


Fig. 7 Smart Architecture's comprehensive evaluation: instantaneous delivery latency of game events (packets) over the whole game platform

Indeed, Fig. 7 confirms that combining the two components of SMA will produce a sum of the positive benefits seen in Fig. 5 and Fig. 4. In particular, Fig. 7 reports the delivery time of 100 subsequent game events through the whole gaming platform. The regular configuration of the game platform (REG) is compared with the configuration including both FS and smart APs (SME). The outcome clearly demonstrates how SMA outperforms REG. Moreover, SMA is able to keep the game event delivery time almost always under the interactivity threshold (GIT).

Conclusions

The mobile gaming scenario is becoming very popular thanks to the widespread availability of mobile devices equipped with powerful networking and multimedia features. In this chapter we highlighted that mobile games produce time-sensitive traffic (i.e., traffic coupled with timing constraints) and require the network to meet these constraints in order to well support their peculiar characteristics: interactivity, consistency, fairness, scalability, and continuity.

Unfortunately, the support of mobile games is not as easy as one may think, as the service provided by the network might be not sufficient to meet the mobile game requirements. For this reason, many researchers focused their studies on this challenging scenario with the goal of finding solutions to support the peculiar characteristics of the mobile games. Throughout this chapter, we overview the network support solutions for mobile games and we focused on an interesting solution specifically designed to support the playability of games in the mobile scenario.

References

Armagetron. A Tron clone in 3D. http://armagetron.sourceforge.net/. Accessed 31 Oct 2014

G. Armitage, An experimental estimation of latency sensitivity in multiplayer Quake 3, in *Proc. of IEEE International Conference on Networks (ICON)* (Sydney, 2003), pp. 137–141

T. Beigbeder, R. Coughlan, C. Lusher, J. Plunkett, E. Agu, M. Claypool. The effects of loss and latency on user performance in unreal tournament 2003, in *Proc. of ACM Network and System Support for Games Workshop* (NetGames 2004) (Portland, 2004)

G. Bianchi, Performance analysis of the IEEE 802.11 distributed coordination function. IEEE J. Select Area Commun. **18**(3), 535–547 (2000)

S. Biswas, R. Tatchikou, F. Dion, Vehicle-to-vehicle wireless communication protocols for enhancing highway traffic safety. IEEE Commun. Mag. **44**(1), 74–82 (2006)

L. Bononi, M. Di Felice, A cross-layered MAC and clustering scheme for efficient broadcast in VANETs, in *Proc. of 2-th IEEE International Workshop on Mobile Vehicular Networks (MOVENET 2006)* (Pisa, 2006)

L. Bononi, L. Donatiello, M. Furini, Real-time traffic in ad-hoc sensor networks, in *Proc. of the IEEE International Conference on Communications* (Dresden, 2009)

M. Bottigliengo, C. Casetti, C. Chiasserini, M. Meo, Short-term fairness for TCP flows in 802.11b WLANs, in *Proc. of IEEE INFOCOM 2004* (Hong Kong, 2004)

Butterfly Grid Solution for Online Games. http://www.butterfly.net, Accessed 31 Oct 2014

A. Casteigts, A. Nayak, I. Stojmenovic, Communication protocols for vehicular ad hoc networks. Wireless Commun. Mobile Comput. John Wiley & Sons, Ltd. **11**(5), 567–582 (2011)

L.-J. Chen, T. Sun, G. Yang, M. Gerla, USHA: a simple and practical seamless vertical handoff solution, in *Proc. of IEEE Consumer Communications and Networking Conference (CCNC 2006)* (Las Vegas, 2006)

D.D. Clark, W. Fang, Explicit allocation of best-effort packet delivery service. IEEE/ACM Trans. Netw. **6**(4), 362–373 (1998)

M. Conti, L. Donatiello, M. Furini, Design and analysis of RT-ring: a protocol for supporting real-time communications. IEEE Trans. Ind. Electron. **49**(6), 1214–1226 (2002)

S. Corson, J. Macker, Mobile ad hoc networking (MANET): routing protocol performance issues and evaluation considerations (IETF RFC 2501, 1999)

E. Cronin, A.R. Kurc, B. Filstrup, S. Jamin, An efficient synchronization mechanism for mirrored game architectures. Multimed. Tools Appl. Springer **23**(1), 7–30 (2004)

E. Fasolo, R. Furiato, A. Zanella, Smart broadcast algorithm for inter-vehicular communication, in *Proc. of Wireless Personal Multimedia Communication (WPMC'05)* (IWS 2005, Aalborg, 2005)

S. Ferretti, C.E. Palazzi, M. Roccetti, G. Pau, M. Gerla, FILA, A holistic approach to massive online gaming: algorithm comparison and performance analysis, in *Proc. of ACM GDTW 2005*, (Liverpool, 2005), pp. 68–76

S. Floyd, V. Jacobson, Random early detection gateways for congestion avoidance. IEEE/ACM Trans. Netw. **1**(4), 397–413 (1993)

M. Furini, An architecture to easily produce adventure and movie games for the mobile scenario. ACM Comput. Entertain. **6**(2), 19:1–19:16 (2008)

M. Furini, Mobile games: what to expect in the near future, in *Proc. of EuroSis GAMEON Conference on Simulation and AI in Computer Games* (Bologna, 2007)

L. Gautier, C. Diot, Design and evaluation of MiMaze, a multi-player game on the Internet. in *Proc. of IEEE International Conference on Multimedia Computing and Systems (ICMCS'98)* (Austin, 1998)

C. Griwodz, State replication for multiplayer games, in *Proc. of ACM NetGames 2002* (Braunschweig, 2002)

M. Heusse, F. Rousseau, G. Berger-Sabbatel, A. Duda, Performance anomaly of 802.11b. in *Proc. of IEEE INFOCOM 2003* (San Francisco, 2003)

D.P. Hole, F.A. Tobagi, Capacity of an IEEE 802.11b wireless LAN supporting VoIP, in *Proc. IEEE International Conference on Communications (ICC 2004)* (Paris, 2003)

T. Jehaes, D. De Vleeschauwer, T. Coppens, B. Van Doorselaer, E. Deckers, W. Naudts, K. Spruyt, R. Smets, Access network delay in networked games, in *Proc. of ACM NetGames 2003* (Redwood City, 2003)

P. Kokkinos, C. Papageorgiou, E. Varvarigos, Multi-cost routing for energy and capacity constrained wireless mesh networks. Wireless Commun. Mobile Comput. John Wiley & Sons, Ltd. **11**(3), 424–438 (2011)

G. Korkmax, E. Ekici, F. Ozguner, U. Ozguner, Urban multi-hop broadcast protocol for inter-vehicle communication systems, in *Proc. of the 1st ACM Workshop on Vehicular Ad-hoc Networks (VANET 2004)* (Philadelphia, 2004)

G. Marfia, M. Roccetti, Dealing with wireless links in the era of bandwidth demanding wireless home entertainment, in *Proc. 6th IEEE International Workshop on Networking Issues in Multimedia Entertainment (NIME′10) – 2010 I.E. International Conference on Multimedia and Expo* (ICME′10, Singapore, 2010)

Movie Trace Files. http://www-tkn.ee.tu-berlin.de/research/trace/ltvt.html. Accessed 31 Oct 2014

C.H. Nam, S.C. Liew, C.P. Fu, An experimental study of ARQ protocol in 802.11b wireless LAN, in *Proc. of IEEE Vehicular Technology Conference (VTC 2003)* (Orlando, 2003)

A. Nandan, S. Das, G. Pau, M.Y. Sanadidi, M. Gerla, Cooperative downloading in vehicular ad hoc wireless networks, in *Proc. of IEEE/IFIP International Conference on Wireless On demand Network Systems and Services* (WONS, St. Moritz, 2005a)

A. Nandan, S. Das, B. Zhou, G. Pau, M. Gerla, Ad-torrent: digital billboards for vehicular networks, in *Proc. of IEEE/ACM International Workshop on Vehicle-to-Vehicle Communications*, vol. 1, (V2VCOM, San Diego, 2005b), pp. 286–294

Newzoo: Market Research for the Games Industry Online. http://www.newzoo.com. Accessed 31 Oct 2014

C.E. Palazzi, S. Ferretti, S. Cacciaguerra, M. Roccetti, A RIO-like technique for interactivity loss avoidance in fast-paced multiplayer online games. ACM Comput. Entertain. **3**(2) (2005a)

C.E. Palazzi, G. Pau, M. Roccetti, M. Gerla, In-home online entertainment: analyzing the impact of the wireless MAC-transport protocols interference, in *Proc. of IEEE WIRELESSCOM 2005* (Maui, 2005b)

C.E. Palazzi, S. Ferretti, S. Cacciaguerra, M. Roccetti, Interactivity-loss avoidance in event delivery synchronization for mirrored game architectures. IEEE Trans. Mult. **8**(4), 847–879 (2006a)

C.E. Palazzi, S. Ferretti, M. Roccetti, G. Pau, M. Gerla, What's in that magic box? The home entertainment center's special protocol potion, revealed. IEEE Trans. Consum. Electron. **52**(4), 1280–1288 (2006b)

C.E. Palazzi, M. Roccetti, S. Ferretti, G. Pau, M. Gerla, Online games on wheels: fast game event delivery in vehicular ad-hoc networks, in *Proc. of 3rd IEEE V2VCOM 2007, IEEE Intelligent Vehicles Symposium 2007* (Istanbul, 2007)

C.E. Palazzi, A. Kaiser, N. Achir, K. Boussetta, A preliminary evaluation of backup servers for longer gaming sessions in MANETs, in *Proc. of the ACM International Workshop on DIstributed SImulation & Online gaming (DISIO 2010) – ICST SIMUTools 2010* (Torremolinos, 2010a)

C.E. Palazzi, M. Roccetti, S. Ferretti, An inter-vehicular communication architecture for safety and entertainment. IEEE Trans. Intell. Transp. Syst. **11**(1), 90–99 (2010b)

C.E. Palazzi, D. Maggiorini, From playgrounds to Smartphones: mobile evolution of a kids game, in *Proc. of the 8th IEEE Communications and Networking Conference (CCNC 2011)* (Las Vegas, 2011)

L. Pantel, C.L. Wolf, On the impact of delay on real-time multiplayer games, in *Proc. of the ACM 12th International Workshop on Network and Operating Systems Support for Digital Audio and Video* (Miami, 2002)

K. Park, W. Willinger, Self-Similar Network Traffic and Performance Evaluation. John Wiley & Sons, Inc., New York 2002

C. Perkins, IP Mobility Support for IPv4. IETF RFC 3344 (2002)

Quake Forge Project. http://www.quakeforge.org/. Accessed 31 Oct 2014

Rakion. http://rakion.softnyx.net/. Accessed 31 Oct 2014

M. Roccetti, G. Marfia, A. Amoroso, An optimal 1D vehicular accident warning algorithm for realistic scenarios, in *Proc. IEEE Symposium on Computers and Communications* (ISCC′10) (Riccione, 2010)

F. Safaei, P. Boustead, C.D. Nguyen, J. Brun, M. Dowlatshahi, Latency driven distribution: infrastructure needs of participatory entertainment applications. IEEE Commun. Mag. **43**, 106–112 (2005) (Special Issue on "Entertainment Everywhere: System and Networking Issues in Emerging Network-Centric Entertainments Systems", Part I)

Ultima Online. http://www.uo.com. Accessed 31 Oct 2014

P.J. Wan, K. Alzoubi, O. Frieder, Distributed construction of connected dominating set in wireless ad hoc networks, in *Proc. of IEEE INFOCOM 2002* (New York, 2002)

K. Xu, M. Gerla, L. Qi, Y. Shu, TCP unfairness in ad hoc wireless networks and a neighborhood RED solution. Wireless Netw. ACM. **11**(4), 383–399 (2005)

G. Xylomenos, G.C. Polyzos, TCP and UDP performance over a wireless LAN, in *Proc. of IEEE INFOCOM '99* (New York, 1999)

X. Yang, J. Liu, F. Zhao, N. Vaidya, A vehicle-to-vehicle communication protocol for cooperative collision warning, in *Proc. of ACM 1st Annual International Conf. on Mobile and Ubiquitous Systems: Networking and Services (MobiQuitous '04)* (Boston, 2004)

C. Yu, B. Lee, H.Y. Youn, Energy efficient routing protocols for mobile ad hoc networks. Wireless Commun. Mobile Comput. John Wiley & Sons, Inc. **3**(8), 959–973 (2003)

A. Zanella, G. Pierobon, S. Merlin, On the limiting performance of broadcast algorithms over unidimensional ad-hoc radio networks, in *Proc. of IEEE WPMC04* (Abano Terme, 2004)

K. Zhu, D. Niyato, P. Wang, E. Hossain, D.I. Kim, Mobility and handoff management in vehicular networks: a survey. Wireless Commun. Mobile Comput. John Wiley & Sons, Inc. **11**(4), 459–476 (2011)

Opportunistic Networking for Games and Entertainment 22

Dario Maggiorini, Laura Anna Ripamonti, and Christian Quadri

Contents

Abstract

These days we are witnessing a tremendous increase in the availability of personal communication devices that are able to provide ubiquitous connection. As a result, we can also see an increasing demand for mobile services where content is personalized based on user location and context. Mobile multiplayer gaming is today provided as a service and, of course, is not an exception.

D. Maggiorini (✉) • L.A. Ripamonti • C. Quadri
Department of Computer Science, University of Milan, Milan, Italy
e-mail: dario@di.unimi.it; ripamonti@di.unimi.it; christian.quadri@unimi.it

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_22

Unfortunately, differently from other legacy services, location- and context-based gaming strictly requires near-field communication to interact with nearby players in order to create teams and arenas. Since currently adopted technologies suffer from scalability (Bluetooth) or energy (WiFi) constraints, Opportunistic Networks (ONs) have already been addressed as a viable solution to involve potential players which are located in the surrounding area. Nevertheless, it is not yet clear how player experience will be affected by the increased delay and probabilistic message forwarding introduced by an ON. In this chapter the aforementioned phenomenon will be addressed in order to achieve a better understanding of the problem, and, thanks to a case study, guidelines for game designers will be provided to actually deliver a compelling and intriguing experience through an opportunistic game.

Keywords
Delay Tolerant Network (DTN) • Energy consumption • Mixed reality games • Mobile Ad hoc NETwork (MANET) • Network behavior and game design • Opportunistic games • Scalability issues • Opportunistic networks (ONs) • Public Transportation System (PTS) • Quality of Experience (QoE) • Multiplayer games

Introduction

In these last years, the growing availability of personal mobile devices with ubiquitous communication capabilities has boosted the demand for mobile services. These mobile services are nowadays typically offered based on location and context, opening uncountable possibilities for innovative application scenarios. Gaming applications and services are, of course, no exception to this trend, also thanks to an immediate applicability (collaborative and competitive multiplayer is a standard feature for modern games) leading to successful commercial products that have been developed over the years (such as Nintendogs by Nintendo and Metal gear solid: portable ops. by Konami). When provisioning services to mobile personal devices, the usual approach is to connect the device to a cellular network and provide customized information based on the user's location or feedback. Proximity detection between users must be performed by a centralized server working on information supplied by the network operator or by the user herself. As a matter of fact, the 3G/LTE network infrastructure sometimes struggles to keep pace with location-based applications, due to its limited scalability (Qureshi et al. 2006) and pull-based service model. With respect to cellular networks, even if we forget about scalability issues, a number of additional limitations can be observed impacting the deployment of large-scale gaming services over 3G or LTE. Bandwidth is usually not a problem, since many games are designed (to minimize costs) to use network resources sparingly. Nevertheless, the problem of *goodput* has been addressed in the past (e.g., by Pentikousis et al. 2005) especially in relation to connection time drawbacks (Curcio and Lundan 2002; Lakaniemi et al. 2001; Xiaoying et al. 2012); in particular, practical experiments seem to indicate ad hoc WiFi between devices as a more attractive platform than 3G (Gass and Diot 2010; Pralhad et al. 2010). Another problem is generated by connection availability (*coverage*), which may not be evenly distributed on the map even in western countries, as can be observed in Fig. 1. In the figure, a map of the city of Milan (Italy) is proposed with an outline of the network coverage of 3G (left) and LTE (right). As we can see, the coverage is spotted at best as soon as a user moves out of city center. LTE, on the other hand, is not currently granted in every location inside the city center. Finally, we do have issues related to *energy consumption*: a game must share a mobile phone with many other useful (and more important for the user) applications, such as an agenda or a messaging subsystem: power conservation is paramount when developing any mobile application. The problem of energy usage in cellular networks has already been extensively addressed in the past (e.g., by Le and Jukka 2010; Perrucci et al. 2009) and offloading to WiFi seems to be a promising solution for power conservation (Ristanovic et al. 2011; Sharma et al. 2009). With the upcoming massive deployment of LTE technology, many of the aforementioned problems have been addressed and, in some cases, partially resolved; nevertheless, there is still a huge number of devices in use (both smartphones and 3G-enabled portable consoles) which are not supporting LTE. As a consequence, developers who see the market benefits to develop cross-platform applications do not feel comfortable about requiring real-time network interaction over a cellular network. A practical example of this trend can be found in the PlaystationVita: the flagship mobile console from Sony (2011). The Vita is 3G enabled, but 3G real-time gaming is unsupported by most of the games and does not meet the favor of many users. If we go back to the more general issue of providing a gaming service, we also have to say that many cell-phone providers are offering the same mobile services from within their core network, making it difficult to share information between users of competing providers (Dedrick et al. 2003). For all the above reasons, many service and data providers are looking today for innovative solutions, offering a scalable, carrier-independent, location-based, and cheaper-than-3G data distribution.

In the specific case of entertainment applications, game designers usually address the above problem by adopting a near-field short-range communication paradigm, where all mobile gaming devices connect to each other as a Mobile Ad hoc NETwork (MANET). While this is a reasonable solution on the small scale, bringing it to a larger area or connecting more players calls for a different kind of peer-to-peer communication paradigm. One viable solution to involve a huge number of players on a wide area is the adoption of a delay-tolerant network (DTN) (Fall 2003) where unplanned and unpredictable contact opportunities between players are exploited to perform data exchange and forwarding. Following this kind of approach, it is possible to create an Opportunistic Network (ON) (Boldrini et al. 2008; Cai and Eun 2009; Hui et al. 2005, 2008; Sandulescu

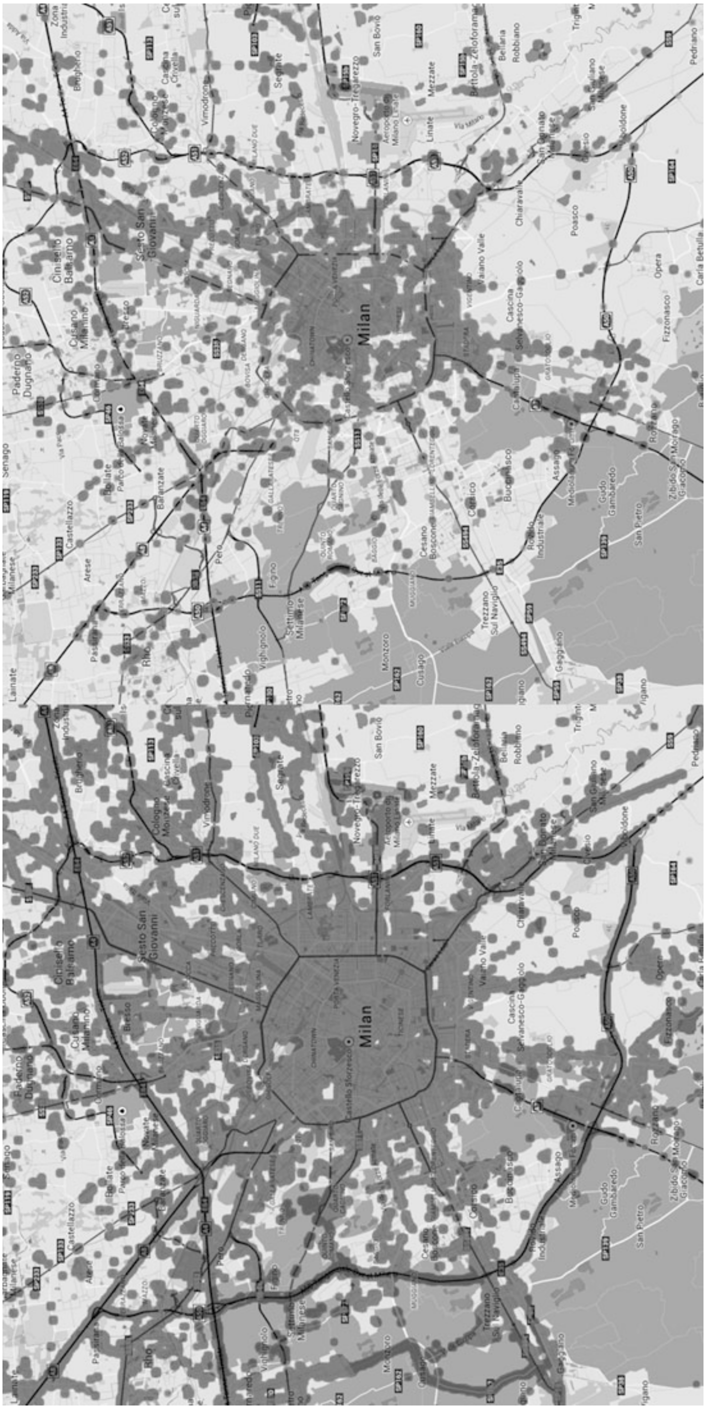


Fig. 1 Cellular network coverage in the area of Milan (Italy) as of November 2014; 3G (*left*) and LTE (*right*) are reported (Source: opensignal.com)

and Nadjm-Tehrani 2008) between mobile consoles. Games designed for an ON can take advantage of contacts – both intentional and unplanned – between players as part of the gameplay: this will make a game inherently location/context dependent and will naturally scale up to a virtually unlimited number of players. On the other hand, interaction can be fast and real time only during encounters (meaning, while the encounter lasts), and unpredictable delays as well as data loss may take place any moment. If the game is designed to require some kind of group communication, the direct connection between two devices must be extended by means of some routing strategy. Of course, lacking a supporting infrastructure, the only option with ONs is to exploit encounters to also relay data to nodes that are not currently in the radio range. To perform this operation in the most efficient way possible, modern opportunistic routing strategies try to exploit node movement patterns. In the specific case of mobile games, this requires to deal with human mobility patterns. Human mobility models have already been extensively studied in the past (Brockmann et al. 2008; Gonzales et al. 2008), and results have been extended to ONs in order to predict traffic generation (Ekman et al. 2008; Lee et al. 2009; Rhee et al. 2008, 2011) or to optimize data forwarding strategies (Chaintreau et al. 2006; Hui et al. 2005). Nevertheless, routing performance in ONs is still bounded by players' density, and offered end-to-end delay is in the order of minutes if not hours. As bottom line, the availability of neighboring nodes and the distribution of inter- and intracontact times pose severe constraints to the gameplay and, therefore, heavily affect the way the game is designed in order to produce a compelling user experience.

Given all the considerations above, it should be clear by now that contact-based games make some sense (and revenues) where the density of the users is such to grant a favorable environment. For this reason, current studies and products are usually targeting an urban environment, if not even a campus or building. In the following, in order to keep the discussion at a reasonable level of complexity, an urban environment will be considered. From a game designer's point of view, the main problem in this context is about forecasting the feasibility of a given gameplay when applied to a real – and complex – urban environment.

The goal of this chapter is to provide a clear understanding of the technological problem about implementing a game on top of an ON as well as to provide practical directions to game designers about constraints impacting on gameplay and game mechanics when designing a game requiring contact among users on at least an urban-wide scale. Moreover, some considerations will be drawn about planning and deployment strategies.

The rest of this chapter is organized as follows. Firstly we discuss related works inherent to the design of games based on random encounters between players. Then, we will be delving into the major technical challenges related to the design of opportunistic games. After introducing the problem of evaluating the quality of experience in opportunistic games we will propose a case study to better understand guidelines we will be providing to game designers in the following section. Finally, we summarize this chapter and outline prospective evolutions on the topic.

Related Work

Unfortunately, and to the best of our knowledge, current literature on the topic of designing games over ONs is very limited. The same problem we are discussing here has been addressed in very general terms in Maggiorini et al. (2012a), where game deployment feasibility was studied without considering constraints imposed by the environment. In Maggiorini et al. (2012b) the main concern is network scalability, but limited discussion is devoted to design issues while in Maggiorini et al. (2013) there is an effort to correlate player experience with network performance. In all cases, only a subset of the problem has been addressed, without a complete study about design and deployment. A first organic, and more comprehensive, analysis seems to be Maggiorini et al. (2014), where a more exhaustive simulated environment is proposed together with an urban-wide alternate reality game designed to run on the opportunistic network offered by an urban public transportation system. In particular, in this last work we can find a number of considerations about how to correlate network performance with game design decision. Nevertheless, this last paper is still limited to the very specific case – albeit significant – of an urban ON deployed on top of a Public Transportation System (PTS) and offers no suggestions about how to correlate game decisions with the city topology.

In the following two subsections, we are going to provide separate discussions for ON as a technological platform and multiplayer games as a challenging application for wireless networks.

Opportunistic Networks in Urban Environment

Disruption/delay-tolerant networks (DTNs) are network infrastructures where an end-to-end path between the source-destination couple may not exist all the time, due to an intermittent availability of participating nodes. Despite the fact that DTNs have been introduced only in 2003 (Fall 2003), a considerable effort has been devoted by the scientific community on defining their architecture and services (Jain et al. 2004; Lindgren et al. 2003; Nelson et al. 2009; Yuan et al. 2009). In their basic form, DTNs define buffering and routing strategies to minimize the delay introduced by the discontinuous connection availability.

ONs, on the other hand, are a special case of DTNs where, due to node mobility, and end-to-end path may not exist at all for a long period or could change while a packet is forwarded toward its destination. Taking benefit from node mobility, unplanned contact opportunities are exploited to implement routing/forwarding (Boldrini et al. 2008; Cai and Eun 2009; Hui et al. 2005, 2008; Sandulescu and Nadjm-Tehrani 2008). As a matter of fact, there is currently a somewhat limited interest in ONs covering small areas while many scientists are actively working on ONs covering very large areas. This is due to the fact that in the former case the network diameter may be comparable to the nodes' transmission range, leading back to the basic DTN model managed by buffers. Instead, in the latter case, a node waiting for content may be active, but unreachable, for a very long time, and a full-fledged opportunistic routing solution is actually required.

The first contributions on a large scale (De Oliveira et al. 2008; Demmer and Fall 2007; Pentland et al. 2004) focused on rural environments in developing regions where a number of villages are spread over a large territory. The common goal of all the referenced projects is to provide network access for elastic non-real-time applications, so that the local population may enjoy basic Internet services (e.g., e-mail and non-real-time web browsing). In these cases, the set of neighbors for every node is usually small yet does not change frequently over a span of time; failure to deliver a packet is generally the result of a missed encounter and not due to unpredictable node mobility.

Campus networks (see, e.g., Balasubramanian et al. 2007; Banerjee et al. 2010; Burgess et al. 2006; Zhang et al. 2007) are designed to serve students and faculties who commute between colleges or from/to nearby towns. A relatively small number of nodes, when compared to a full-fledged urban environment, usually characterizes these kinds of services. The main contribution in this direction is represented by Burgess et al. (2006), where five colleges are linked with nearby towns and to one another over an area of 150 square miles.

Scaling up in terms of number of nodes, we find urban environments where nodes are densely deployed as the result of people commuting inside the city. ONs in urban environments (such as Ahmed and Kanhere 2007; Jetcheva et al. 2003; Sede et al. 2008) are usually characterized by many, short, contact opportunities. In Jetcheva et al. (2003), authors propose a commercial application called Ad Hoc City. Ad Hoc City employs a multitier wireless ad hoc network architecture to provide elastic Internet access. Local Internet access is managed by access points (APs) covering the geographic area. Messages from mobile devices are carried to an AP and back using an ad hoc backbone that exploits buses. Authors verified the validity of the proposed approach against real movement traces by King County Metro bus system in Seattle, WA. Later on, using the same real data, the authors (Ahmed and Kanhere 2007) proposed a cluster-based routing algorithm for intracity message delivery. Finally, the contribution from Sede et al. (2008) uses data from the public transportation system of Shanghai to test the performance of a forwarding mechanism adopting a probabilistic routing strategy, where probabilities are related to intracontact times.

As we can observe above, the majority of contributions dealing with ONs in urban environments are strongly intertwined with vehicular networks. Moreover, many of them exploit the Public Transportation System (PTS) already present in the city because the node mobility is quasideterministic (Giaccone et al. 2009): buses move along predetermined paths, following (most of the time) a known schedule; this kind of behavior is making the resulting ON more manageable.

Multiplayer Games in Urban Environments

Multiplayer games are play experiences where players meet and interact with each other. This interaction may imply *competition* in order to demonstrate skill superiority or *collaboration* to achieve a shared goal through a coordinated interaction with a virtual environment. When limited to small user groups and spaces, multiplayer games can be considered as an alternative form of standard game genres, where game mechanics require some sort of coordination between participants. Nevertheless, on urban scales, they can assume a completely different meaning and focus on exploiting existing technologies, relationships, and places as platforms for gameplay.

In both collaborative and competitive approaches, the communication pattern is not changing in any sensible way, and game design decisions related to network constraints are not going to be influenced either. Nevertheless, the collaborative variety proved to be more interesting in terms of player involvement and may also have a useful side effect in fostering social interaction and promoting usage for public services. In particular, collaborative multiplayer games played on a urban-wide scale may take the form of collaborative mixed reality games, which are "goal directed, structured play experiences that are not fully contained by virtual or physical worlds" (Hansen et al. 2012). Moreover, we can notice that collaborative mixed reality games can assume multiple forms, whose game–real world interfaces can range from "augmented reality" to "real environment" (Milgram et al. 1994). As a result, they will include paradigms like context-aware games, alternate reality games (ARGs), social network–based games, and augmented reality–based games. Moreover, due to the fact that they "play" with the boundaries of traditional gamespaces, mixed reality games are able to blend game mechanics with our everyday life (Flintham et al. 2003; Montola 2011).

Among the most notable examples of mixed reality games, we can list the following examples that, besides being collaborative, are also pervasive:

- Entertainment applications that exploit sensor data – collected from cameras, accelerometers, gyroscopes, and GPS – as part of the gameplay (e.g., Biblion and iSpy for iOS devices and Facebook, Foursquare, Ingress)
- Games whose narrative is built by players using diverse media elements and which usually require players to collaborate in order to solve puzzles/challenges (e.g., Evoke, I love bees, The lost experience, Cruel 2 be kind, Pac Manhattan)
- Games that exploit players' social networks as resources for the gameplay (e.g., Farmville, Spent, and Oregon Trail for Facebook, DropIn for LinkedIn)
- Games that overlay information on some depiction of reality (e.g., Magic: Eye of Judgment for PS3, Face Riders for Nintendo 3DS, Parallel Kingdom for iPhone)

Finally, it is important to underline that not only a shared definition of collaborative mixed reality games is yet to come, but the genre is also missing the benefits of a coherent body of focused investigation and practice about the design criticalities that should be properly addressed to create compelling and fun games (Hansen et al. 2012). As a matter of fact, a game designer confronting with this genre should have a clear understanding of the effects deriving from the interaction among game mechanics, technologies, and social engagement strategies at work in the game in order to convey novel game experiences. This means, in our specific case, that the game designer should be aware of all the crossed effects and problems deriving from the design of a collaborative game mechanic, supported by a – partially – mediated social interaction among players taking place through a wireless infrastructure and issues typical of Human-Computer Interface disciplines (HCI), such as the design of effective touchscreen interfaces (Coulton et al. 2006).

For the above reasons, focusing on collaborative games is advisable in order to cover a broader spectrum of applications and to allow a game designer to be more creative in building an enticing gaming experience.

Technical Challenges in Designing Opportunistic Games

In this section we will be discussing challenges and constraints imposed by current technology and urban environments, in order to provide strategic insights to game designers.

Target Platform Constraints

Selecting a mobile platform implies a selection between the adoption of smartphones or portable game consoles. Game consoles are, from a technical point of view, specifically designed for the task at hand. Unfortunately, due to their nature of proprietary devices, the general want of cross-platform software does not make them very attractive for a large-scale deployment. Moreover, smartphones, as multifunction interoperating devices, have proven to be much more widespread and can be found in large numbers whenever we are cruising the city or commuting from house to campus. As an added value, software development kits (SDKs) for mobile phones are much cheaper when compared to the ones for portable game consoles, and distribution platforms are also more accessible, especially for independent developers.

For the above motivations, it should be advisable to consider primarily a smartphone as target platform, unless, of course, the game is intended as an advertisement for the platform itself or is taking advantage of dedicated hardware. From the remainder of this chapter, where not explicitly stated otherwise, the reader can assume to use programmable smartphones as deployment platforms.

Communication Technology Constraints

From a communication standpoint, there are currently a number of technical limitations to a contact-based game on a smartphone. Firstly, purely peer-to-peer communication is possible thanks to a number of technologies, but none of them seems to offer a feasible service for gaming. On many mobile phones, WiFi in ad hoc mode is unavailable to programmers due to tethering limitation imposed by mobile network operators. WiFi P2P (also called WiFi Direct) can alleviate the lack of ad hoc mode for standard WiFi also thanks to the fact that only one of the devices is required to comply with this new standard. WiFi P2P is already available on Intel “Centrino 2” systems and the latest generation of Android smartphones. Network capacity may vary depending on the underlying transmission technology with data rates from 6 Mbps (802.11a) to 150 Mbps (802.11n).

Bluetooth is more accessible than WiFi for ad hoc networking but requires by design direct interaction from the user in order to start communication. This behavior prevents advanced forms of seamless *push* communication with occasional unpaired devices and requires a setup phase, which the player might be unwilling to face at every session. Moreover, Bluetooth imposes the presence of a *master* node (which is also a single point of failure) and a maximum of eight *slave* nodes, with major scalability impact. As a matter of fact, Bluetooth is a viable solution for gaming only when players meet on purpose and not when the aim is to build an occasional group based on a common shared location. Bluetooth is already available on almost every smartphone with a data rate ranging from 1 (version 1.2) to 24 Mbps (version 4.0).

NFC, as the last contender, is not a transport service up to the task: actually, when two devices *bump* together, two short messages are exchanged (in the same way as RFID does), and the communication is over unless those two devices go out of range and close back again. As a matter of fact, the exchange of a single message is not enough to implement any interaction protocol. Of course, the NFC handshake can be used to drive the first phase of a Bluetooth setup, but player explicit acknowledgment is still required unless the latest version of Android O.S. is in use (4.4, at the time of writing). Unfortunately, this last case currently covers only a fraction of the installed base: 30 % (Google 2014).

If we compare the technologies described above in the context of networked games, we have to take into account connection setup time, network capacity, connection model, and energy consumption. Connection setup time sees WiFi and NFC as winners: a few hundreds of milliseconds (depending on announce rates and security handshake such as WPA2 for WiFi) against human solicited intervention followed by a 6 s handshake is no match. If the game must run seamlessly on contact or proximity WiFi and NFC are viable options; otherwise, for games requiring an agreement between players or an explicit acknowledgment before playing, Bluetooth can also be considered. Network capacity is not an issue with any technology but for NFC, which allows exchanging only one message. As a matter of fact, networked games are always designed to minimize the generated traffic in order to help users without a flat-rate contract. Network usage for highly interactive DSL-based games ranges from 80 to 200 Kbps; portable games are supposed to use much less than that. The connection model sees WiFi as a favorite again. WiFi provides a one-to-many connection infrastructure with a broadcast data distribution. Bluetooth also provides a one-to-many (one-to-eight, actually) connection, but data distribution is not broadcast, and the master node must provide some sort of application-level routing with the associated energy consumption. NFC is limited to a one-to-one connection model, which is useful only for a two-player game. Energy consumption evaluation may be difficult since it depends on both network access and CPU usage: CPU may become an issue in case of optional cryptography (such as WEP vs WPA2) and when application-level routing is required (like in Bluetooth). If we focus our evaluation to energy spent for network access all technologies are limited to 100 mW due to FCC regulation; in particular, Bluetooth for indoor use (class 2, with a maximum range of 10 m) uses 2.5 mW. Of course, the actual energy usage will depend on how long the transmission system will be in use. For longer gaming sessions Bluetooth seems to be the best choice while WiFi should be preferred for shorter burst-like sessions; NFC, of course, reports the lowest energy consumption due to its one-to-one and single-message communication.

Scalability Issues

Scalability is also a key factor for gaming applications. Differently from other kinds of applications, here we are mostly concerned for downward scalability, rather than upward. The challenge here should be about keeping the game entertaining when few players are present, rather than having the service running under the onslaught of a huge number of users.

From a game design perspective this means that, depending on the game timescale, the player might need offline content to be able to play alone at times. As a result, game mechanics must provide synchronization mechanisms between users and some way to balance power between hard-core players and casual ones.

Quality of Experience in Opportunistic Games

Quality of Experience (QoE) is a measure of a player's experiences with the game and is a more holistic evaluation when compared with the more narrowly focused user experience, which is focused on the game interface and implementation. In other words, we try to understand the level of satisfaction the player gets from the gaming activity in its completeness. The easiest way to perform an evaluation is to identify some kind of performance index for actions that may be linked to a positive experience. Among all the possible indexes, a wise choice is to use the number of possible actions a user can perform over time intervals (i.e., the contact opportunities) and the level of advancement a player is able to cover in a given time span. The first index is measuring the *level of involvement* a player perceives from the game; we can assume that a user likes to feel part of the system and in having many opportunities to interact with other players. The second index is measuring the *level of rewarding* a player gets from playing the game; we can assume that a user likes to attain a steady and continuous progress inside the game structure. It is worth to be noted that, while the level of involvement may depend solely on movement patterns and players' density, the level of rewarding is actually in control of the game designer, which have to put to good use even a limited number of actions/contacts. Every game should then be designed in a way that, taking into account the average number of contacts (constrained by the environment) and the operations which can be performed during each contact, will lead the player to a significant – and rewarding – progress in the game.

It is important to outline that the selection of the target users has a strategic importance here. Target user base should not be identified just by a common driving interest or need: in order to effectively share the same space multiple times they should also show a comparable (or, at least, synchronized) movement pattern. Under this perspective, designing opportunistic games to be played while commuting or when sharing a common timetable (such as in a school or factory) may actually benefit a lot to the user experience.

Understanding the Problem: A Case Study Through Simulation

In order to better understand the constraints posed to game designers by a real opportunistic network we are going to address a case study. Since an actual urban-wide ON has not been implemented yet, this case study needs to be simulated using a reasonable realistic setup. Following the guidelines we mentioned before and in order to better compare to existing literature, the simulation will represent an ON exploiting a PTS as supporting infrastructure. Moreover, to be able to evaluate the level of rewarding for players, a game mechanic to be implemented using the ON must also be defined. To this purpose, a mobile game inspired from a real – and standalone – one is used as meaningful for the broader category of collaborative mixed reality games. In this game, interaction is loosely real time, and data is exchanged exploiting short extemporaneous contacts: that way we will respect the constraints imposed by the technology and by the interaction patterns among candidate players.

A detailed discussion about the simulation and the underlying system can be found in Maggiorini et al. (2014).

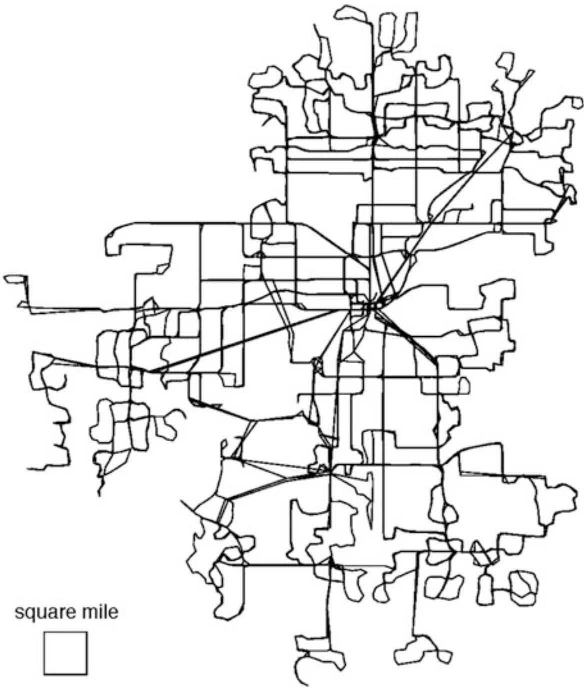
Simulation Setup

The simulation environment described here takes into account the technological constraints we already outlined before.

In order to be as realistic as possible, simulations are using the actual topology of three large cities: Milan (Italy), Edmonton (Alberta, Canada), and Chicago (Illinois, U.S.A.), and the timetables of their respective PTS. Milan (Italy) is a medium-size town (typical for many European cities) with a complex transportation system extending above and below ground. The ground transportation system inside the city spans 69 lines with a very high bus density. As it can be observed from Fig. 2, the overall city structure is clearly not Manhattan like; crosses between bus lines may occur at any time, and there is no constant space between contacts. The second city we consider is Edmonton, whose PTS is depicted in Fig. 3. It is larger than Milan and has more bus lines (a total of 94) but with a lower density; the city map is

Fig. 2 Public transportation lines in Milan

Fig. 3 Public transportation lines in Edmonton



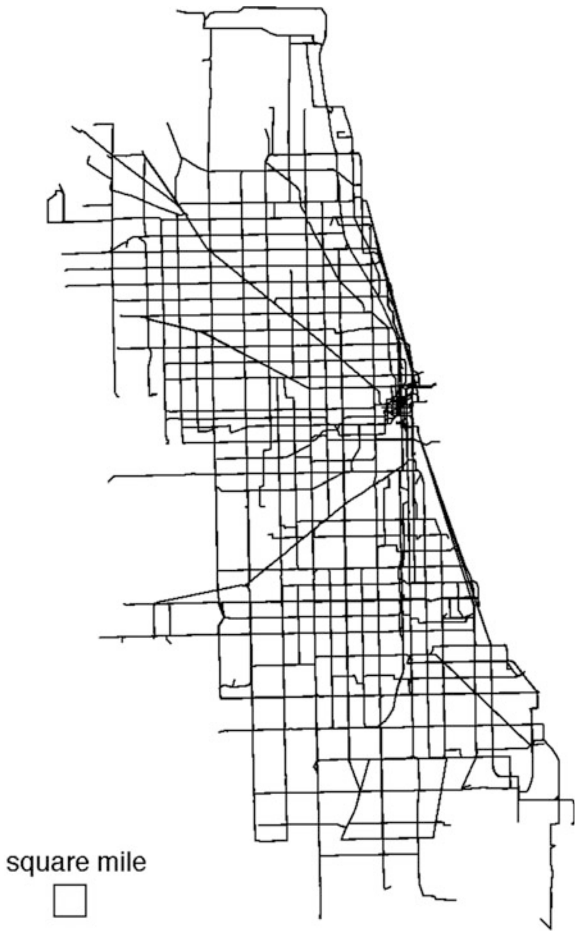


Fig. 4 Public transportation lines in Chicago

Table 1 Features of PTS layouts for the three cities

	Milan	Edmonton	Chicago
City size in square miles	49	134	380
Number of lines	69	94	142
Line length (mean ± st. dev.) in miles	9.9 ± 3.6	12.9 ± 6.7	15.3 ± 7.0
Line density	13.85	9.05	5.73

mostly regular, with some winding paths in the suburbs. The last city is Chicago, which is useful to draw considerations on scalability due to its immense size. The PTS for this last city can be seen in Fig. 4. Chicago has many bus lines (142), but with a very low density, and a Manhattan-like plan. Table 1 presents a summary of the three cities and their PTS layouts. In the table, bus line density is calculated as the mean number of miles covered by buses per square mile.

Players will be commuting inside the city using the public transportation system. In Milan, housing in downtown is quite uncommon: the usual condition for commuters is to live in the suburbs or just outside the city and to spend the day working/studying in or close to the city center. Home locations are randomly selected using a polar coordinate system from the city center using a Gaussian distribution for the radius. In Edmonton and Chicago, the concept of suburbs and city center differs from Milan: there is no clear "outside area" from a geometrical center, especially for Chicago, which is along the coastline. In these two cities, home locations are centered on each user's intended destination and follow a uniform distribution; as a matter of fact, if commuting takes place via PTS in such a large environment with low line densities, it is quite unrealistic for a user to cross the entire city to go to work; remember we are not considering subways here.

A player will spawn at her home location and move by foot using a constant speed to the closest bus stop. From there, the bus stop which is closest to the destination will be reached using one or more bus lines selected with a minimum hop algorithm (i.e., minimizing the number of transits between lines). From the last bus stop to the final destination the player will walk again. After stopping for a random interval of time at the destination, the player will leave and take the trip to go back home. On the way back, the same approach for path selection as before will be applied; nevertheless, the backward path might be different due to asymmetric bus lines. Departure times in both directions are selected using a Gaussian distribution.

Players are classified into two groups offering different behaviors: students and workers. Students leave home in the morning to move toward one of the main universities in the city and will move back home in early afternoon. University campuses in each city are selected based on popularity and number of subscribers. In Milan we have Università Statale di Milano, Università di Milano-Bicocca, Politecnico di Milano, Università Bocconi, and Università Cattolica del Sacro Cuore; in Edmonton we have MacEwan University (City Centre Campus), University of Alberta (in downtown), and Athabasca University; finally, in Chicago, we have East–West University, Roosevelt University, Chicago University, and DePaul University. Workers, on the other hand, will simply leave home in the morning to go to the office and come back in late afternoon.

In Milan, home locations are generated using a Gaussian distribution with $\mu =$ 6 km and $\sigma = 2$ km from the geometrical city center. These values are coming from the actual topology of the city. In Edmonton and Chicago a distance from the destination of 5 and 10 km has been used with a uniform distribution. These last two values are coming from the practical observation of house sharing announcements relative to the considered universities at the beginning of the fall quarter in 2013.

Student destinations are selected randomly picking a university from a list while office location for workers is following a Gaussian distribution with $\mu = 1.5$ km and $\sigma = 1$ km in Milan and uniformly in the areas identified on the map as "downtown" and/or "business district" for Edmonton and Chicago.

Departure and return times are also following a Gaussian distribution. Departure times in the morning have parameters $\mu = 7$ A.M. and $\sigma = 30$ min for everyone.

Students will come back from school with $\mu = 1$ P.M. and $\sigma = 15$ min while workers will vacate the office with $\mu = 5$ P.M. and $\sigma = 60$ min.

When dealing with transmission links, all devices are using IEEE 802.11b with an available bandwidth of 11 Mbps and a radio range of 10 m. Link capacity is accounted using a token bucket model. Communication takes urban canyons into account: only line-of-sight contacts are allowed. To provide connection, access points acting as hubs between users are deployed on buses; this way, there is no need to create ad hoc networks (WiFi or Bluetooth) between phones. A user simply registers while boarding, and the access point will provide a local area network for device discovery and data exchange. When multiple connection points are available at the same time, each device will just connect to one of them following its local rules (e.g., the signal strength).

Mobile Alchemy

Mobile Alchemy is an opportunistic collaborative game inspired by a game for mobile devices, Alchemy, which is available for Android and iOS platforms. In Alchemy, the user starts with four basic elements (fire, water, air, and earth) and combines them together in order to get other – more complex – ones. The total number of elements is 380 and the goal of the game to discover all possible combinations.

In Mobile Alchemy, each user starts from the same four basic elements and combines them following the same set of rules of Alchemy. Differently from the standalone version, a limited number of elements can be stored in the device, and combinations cannot take place on a single device, but proximity with another user is required. When another device is in range, the player can browse its elements and try to make a new combination. If a new combination can be created, the local element will be destroyed and replaced by the result while the remote element is left untouched.

Mobile Alchemy is designed to be played while commuting on buses, with the purpose of increasing the interest of commuters in adopting public transportation systems instead of private cars (hence with the side effect of contributing in the reduction of airborne pollution and traffic). Buses will serve both as meeting points for users and as data mules. Every coach will carry around users' elements and an unlimited reserve of the four basic elements to avoid players to get stuck in the game.

To simulate the gaming activity, all players have a local storage of four slots and will start in the morning with a set of the four basic elements. While commuting, each player will keep scanning elements carried by in-range players and will make a combination whenever possible, giving priority to those leading to yet unknown elements. This behavior will simulate the greedy exploration attitude typical of many casual players for puzzle game. Combinations can take place on trip or at a bus stop when waiting for transit. While on trip, each player on the bus is able to combine her elements with the elements of all other players on the same bus (or the four basic elements carried by the bus, as a second option). While waiting for transit, a player can try to combine her elements with those *transported* by each bus stopping by, even if she is not going to catch it.

In order to avoid creating all compounds in a single run, players on trip can perform only one combination for each bus stop; on the other hand, players waiting for transit will be allowed one combination for each bus stopping by.

Whenever a player obtains an element that does not allow further combinations, she can discard it and replace it with one of the basic elements carried by the bus.

Network Behavior and Game Design

In this section we are going to discuss the correlation between network behavior in the proposed case study and game design constraints. The behavior observed in the simulation can suggest a number of best practices for game designers.

Involvement

As already mentioned, the level of involvement of each player is not connected to game mechanics but depends on movement patterns and density of players. For this reason, the following guidelines can be applied to any opportunistic multiplayer game where nodes follow a movement model similar to a PTS.

Performance indexes of contact opportunities for minimum, average, and maximum population are briefly summarized in Table 2. The numbers in the table are quite positive with respect to the scalability considerations we made earlier in this chapter. In particular, the number of contact opportunities scales up exponentially with the number of players. This is positive, because it means that game designers are not strictly required to put an upper limit to teams: an interaction group can be as big as it needs to be. Moreover, even with the lowest population values, a minimum level of interaction opportunities seems to be guaranteed. A designer can always rely on a minimum number of interactions to exploit for game progression; mechanics to let a user playing alone are required only if the lower bound is too low, depending on the game. The hugest number of contacts is always reported in Edmonton, which despite its size and line density sits in the middle of the scale. This feature is, most probably, depending on topology: buses in Edmonton meet at large hubs providing the last stop for a good number of different lines; users waiting for transit are definitely more likely to have a lot of contact opportunities. Game

Table 2 Total number of contacts between players; population is mixed: 50 % students and 50 % workers

	Number of players		
	1000	10,000	20,000
Milan	21,606	2,049,669	8,259,124
Edmonton	193,599	19,637,955	78,165,799
Chicago	76,642	7,693,059	30,889,160

Table 3 Average contact time between players; population is mixed: 50 % students and 50 % workers

	Number of players		
	1000	10,000	20,000
Milan	214.67	222.06	227.00
Edmonton	498.66	512.44	510.66
Chicago	1035.00	986.36	1009.70

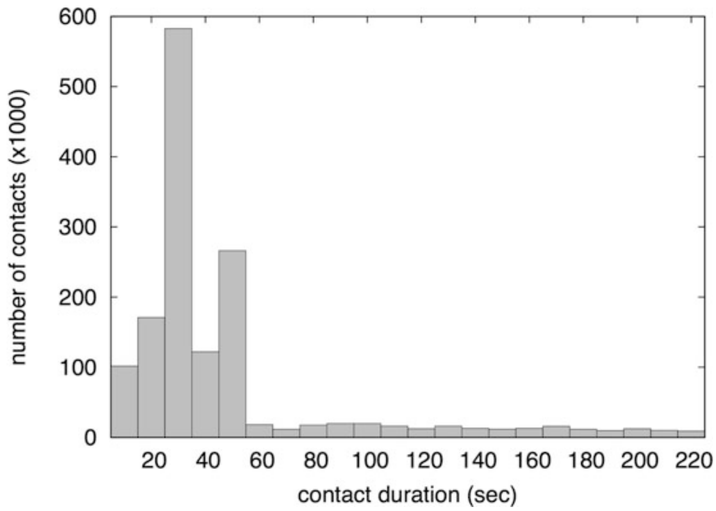


Fig. 5 Distribution of contact duration in the PTS of Milan with a mixed population of 10,000 players

designers do not need to create games to be used at hubs or requiring hubs to play; if a gathering point is present, an opportunistic game will inherently benefit from its presence.

Other design guidelines can be devised analyzing the average contact time between users (Table 3). Average contact duration seems to be, in all cases, unrelated to the population size, opening up to support some degree of real-time interaction in contact-based games: players may be able to actually do something moderately complex and not to just exchange a bundle. Even in the worst case, in Milan, 4 min are more than enough for a fruitful interaction involving a sequence of actions between players. Unfortunately, the actual distributions of contact times report a different situation: the majority of contacts last for less than 60 s (see Fig. 5). More precisely, only 37 % of contacts last for more than 60 s. This is somehow a limitation for the gaming activity but still a reasonable perspective time for a fruitful interaction. A different situation can be found in Chicago (see Fig. 6), where 54 % of contacts last longer than 10 min, making it – technically – the best case. This is most probably due to longer routes with few stops shared by a large number of users, especially on the coastline. Anyway, by observing Fig. 6 (where the bucket size is set to 1 min for sake of clarity), we can also see that 41 % of

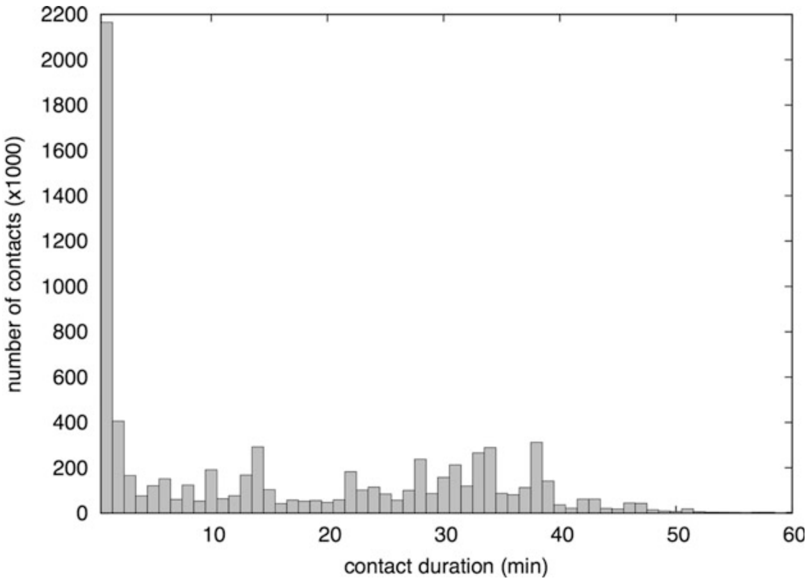


Fig. 6 Distribution of contact duration in the PTS of Chicago with a mixed population of 10,000 players

contacts last for 1 min or less. After the first minute, the following buckets report a number of contacts that is comparable to Milan, but values drop to negligible values just after 50 min. From a game designer perspective, real-time interaction is still feasible, but its scalability is not following the same rules as the number of contact opportunities. Moreover, the actual topology and the movement patterns seem to be minor elements in the picture. Game designers are encouraged to insert some degree of real-time interaction in their games, but to be on the safe side, this interaction should be bounded to 30–60 s.

By considering a population with a homogeneous behavior it is possible to come up with even more hints for game designers. Since workers and students do not follow the same movement patterns, it might be that people converging to a limited number of interest points have a higher chance to use the same bus and see each other on the way. In the simulation, considering a population of students only is actually improving all performance indexes, and by a closer look at the traces, it seems that students' movement pattern is a favorable one: the PTS attitude to *cluster* students inside buses is improving the situation in a tangible way. Unfortunately, and despite this improvement, the observed behavior regarding the escalation of the number of contacts and the average contact time is still very similar, including the contact time distribution. From a game designer perspective, we can say that players exposing similar movement patterns may benefit from a better experience without requiring different game mechanics. More importantly, there is no real need to tune game design to the movement pattern that, in this specific case, is influenced by the PTS organization.

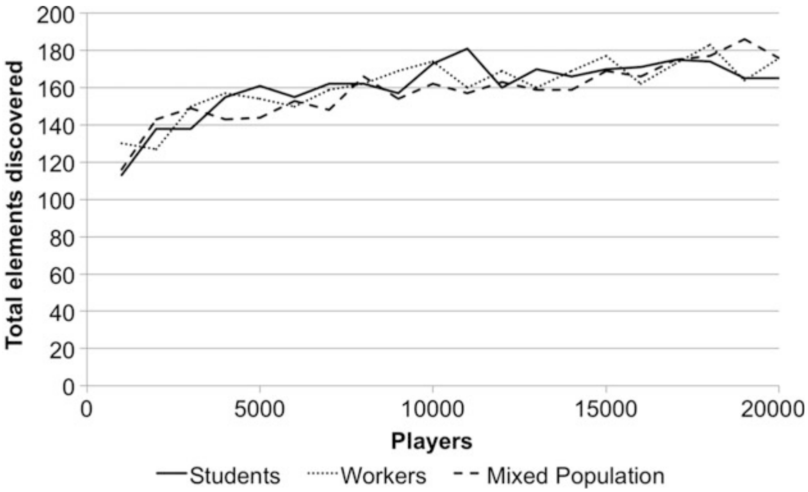


Fig. 7 Total discovered elements in Chicago

Rewarding

While discussing rewarding, game mechanics play a role; guidelines proposed in this subsection should be applicable to any multiplayer collaborative games with formal elements similar to those present in Mobile Alchemy.

The total number of distinct elements discovered by all the players can be a good measure of how far in the game the global population of players is allowed to go in one single day; i.e., it can indicate the general level of consensus the game is achieving. From simulations, in all cities this index seems to be upper bounded to around 180 elements, very far away from the complete list of 380. A summary for the city of Chicago is reported in Fig. 7; other cities and user populations expose very similar behaviors. Moreover, even by increasing the population 20 times, there is a moderate improvement in the game: the number of discovered elements will be just 60 % more. This is a clear indication that this game performance index is constrained by the environment itself (i.e., playing while commuting) or by the game mechanics rather than by the number of users or the city topology. Moreover, in all cases, there is no significant difference between the three populations; it seems that users' habits are not very significant in designing this kind of games. Game designers working on this kind of games should try to tune game content in order to allow users to explore the entire game environment within a reasonable time span.

Let us now talk about the level of personal rewarding. Game designers should also take care of how much every player gets involved in the game. Assuming that a player takes pleasure in completing the game as much as possible, her level of rewarding can be measured by means of the average number of elements discovered during a day. A summary for Chicago is reported in Fig. 8; once again, the behavior

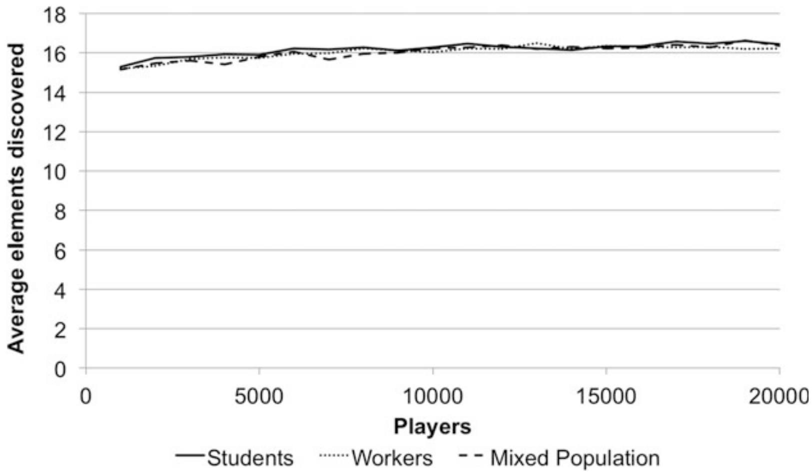


Fig. 8 Average number of elements discovered by single users in Chicago

observed in other cities is extremely similar and provide the same upper bound. As a matter of fact, there is actually a benefit for each user from being able to play in a bigger group, but this benefit is extremely small. As we can see, despite the fact that the population has been increased 20 times, there is a negligible improvement in the personal discoveries at the end of the day. This is underlining again that, for this kind of games, the level of rewarding is not dependent on the number of users and the city topology. Once again, and in all cases, it is not possible to perceive a significant difference between the three populations: one more time, user habits do not seem to influence how much each player will like this kind of games.

The previous suggestion to game designers still stands: they should try to reward the players in game even for a limited number of discoveries, without assuming the number of discoveries will scale up with the population size. In particular, linking personal rewarding with an increment of the player base (the classic "bring your friend to advance" approach) may not always lead to the desired result, depending on the underlying community.

From a game design standpoint, the situation we just described poses severe restrictions to game mechanics. Players get usually rewarded when completing sequences of puzzles or on reaching the end of a sequence of chained quests. Collaborative contact-based games should then be designed with limited requirements – in terms of time to use and actions to perform – to get in-game achievements. In the case of Mobile Alchemy (and games following similar rules), players should be rewarded for intermediate activities and time spent in the process, rather than only for reaching a final element. On the other side, the good news is that, up to now, customization of game mechanics in order to adapt to a particular PTS and/or city plan does not seem to be a requirement. Moreover, contents offered should not really need to be proportional to the number of players or to their density.

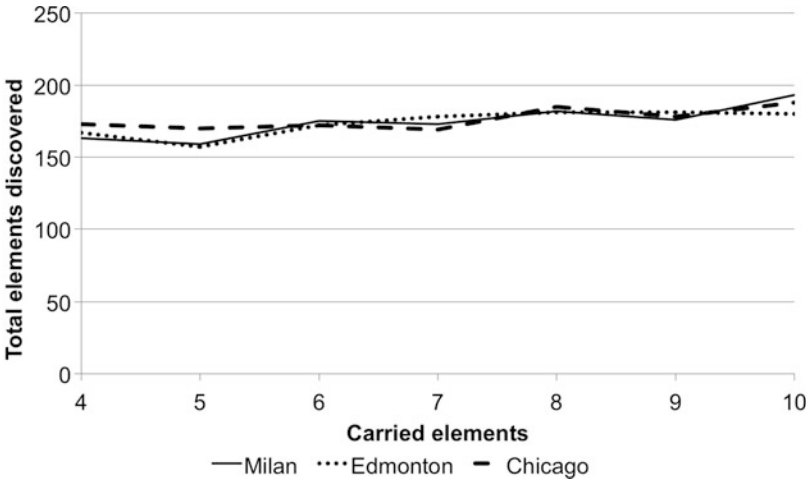


Fig. 9 Total discovered elements while increasing transport slots

One possible explanation for the behaviors we observed so far is that users are not actually competing for resources: all the players progress in the game at the same rate, regardless of their density in the city and the way the PTS is organized; whenever a new element is discovered it spreads quickly, without giving any real advantage to anyone. It looks like players keep using the system to build the same elements repeatedly. This latter aspect, too, is something that should be handled carefully by a game designer, since a good game usually should provide *unequal* results to players.

A good question now is how to overcome the indexes' upper bound (the lack of variety for elements produced, in this specific case) and increase the level of rewarding in the game. As already hinted before, game designers can work on two aspects: *i*) change the game mechanics and (*ii*) get around the structural limitations imposed by PTS movement pattern.

With respect to game mechanics – and without changing the game – mobile devices can increase or decrease local resources; in this specific case this means to change the number of transport slots available for elements. These resources are in direct relationship with options (choices) the player can face while playing. At the beginning of the day, each player will fill her slots with a random sequence of basic elements making sure that each basic element will be present al least once. Figures 9 and 10 report the evolution observed with a population of 10,000 students when the transport slots are gradually increased from 4 to 10.

Increasing the transport slots is actually boosting the total number of (distinct) discovered elements. Albeit limited, this gain is greater than the one obtained from pushing another 10,000 players into the system. Moreover, it can also be observed that the average player may now get very close to double the number of elements discovered during the day. On this last point, large cities seem to be favorite, with Edmonton leading by a short head.

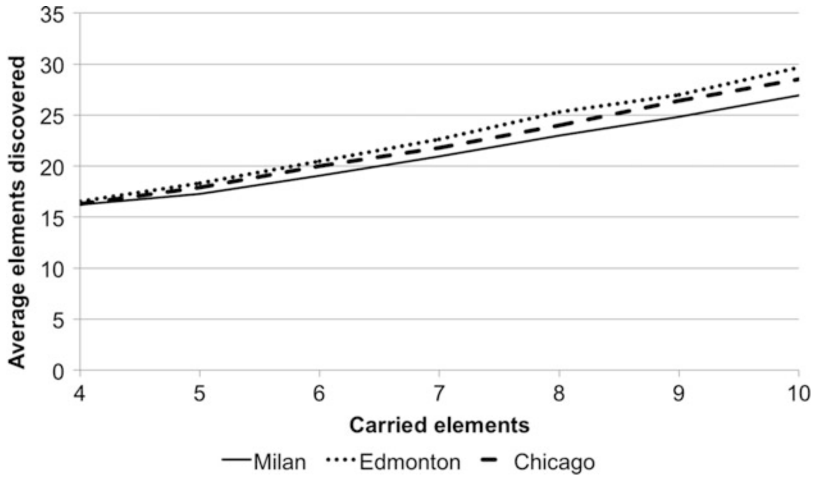


Fig. 10 Average discovered elements while increasing transport slots

While it is clear that giving more options to the users allows them to better exploit the game, best practices in traditional game design discourage giving a player too many concurrent options and suggest five of them as optimal upper bound. Game designers should try to increase the number of options for each player while carefully avoiding to make the game too difficult to manage, otherwise the player may get frustrated and quit the service.

With respect to the second option mentioned above, limitations coming from PTS movement patterns, these may come from players struggling to meet a partner to perform a combination. This can be due to a number of reasons, varying from the constrained movement pattern imposed by the PTS to the fact that even an extremely high number of players will be a negligible fraction of the commuting population in the city.

As reported by simulation results, the average number of successful data exchanges between players is quite limited (see Fig. 11) and follows, in all cases, a profile very similar to the one depicted in Fig. 8. This means that, if another player is in range, it is very likely we will discover a new combination thanks to the complex elements this player is carrying around. On the other hand, the share of combinations a player is required to perform alone (i.e., using the basic element transported by the bus) is higher with small populations and, in our case study, remarkable in Milan (see Fig. 12), where – probably due to its high line density – the presence of more travel options is reducing the probability to meet other players during off-peak hours. With a larger population, we observe a reduction in the interactions with buses up to a lower bound around 5 %; this value is shared among all cities.

Following Figs. 11 and 12, game designers should try to keep to a minimum the number of interactions between players that are strictly required in order to progress in the game. When designing this kind of games on an urban scale, intensive

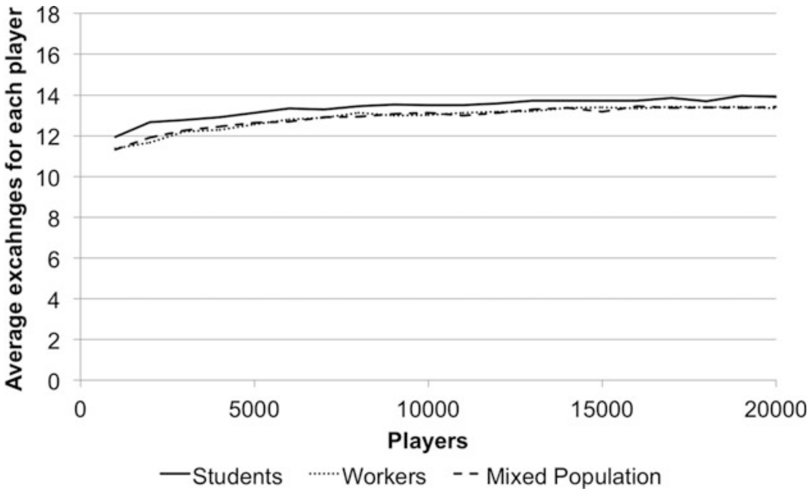


Fig. 11 Average number of combinations performed with other players in Milan

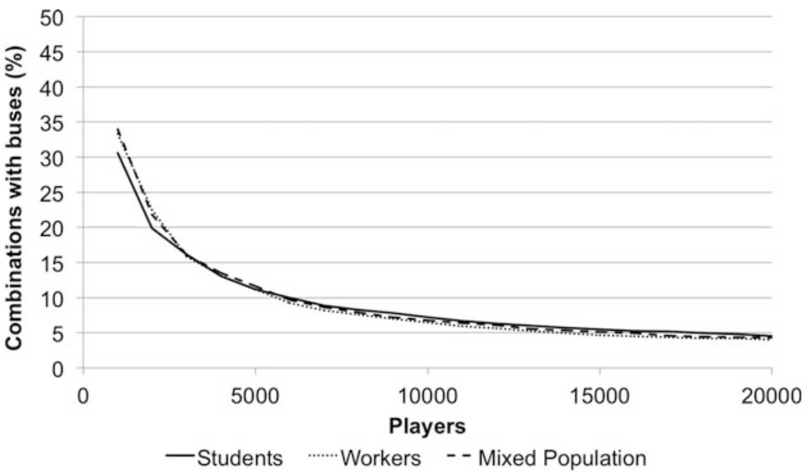


Fig. 12 Percentage of combinations users are performing alone, using elements carried by buses in Milan

interaction between players should not be a fundamental part of the core of the game mechanics, but rather, it can make a good addition to extend the single-player experience.

One last hint for game designers is coming out from observing at what time each player performs the last combination. This information helps us to understand the actual engagement of users in the game. The cumulative distribution functions of the time of the last successful combination in the case of 10,000 players for the city of Milan are reported in Fig. 13. The other two cities provide a very similar behavior.

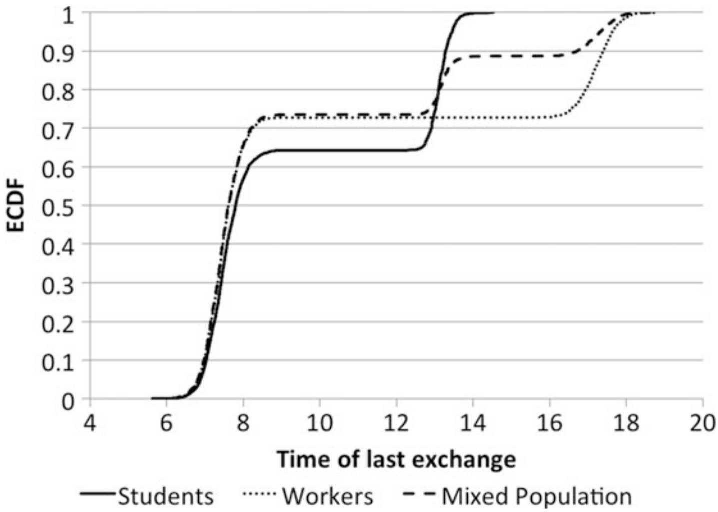


Fig. 13 ECDF of the time of last successful combination for 10,000 users in the city of Milan

As reported by the simulations, 70 % of players stop playing in the morning upon arrival at school or office. This index is a bit lower (65 %) if we consider only the students. Nevertheless, a vast majority of users can enjoy Mobile Alchemy only in the morning, and for a limited time. The lack of new combinations in the afternoon is due to the fact that the majority of users are carrying the same – complex – elements and the system reaches a sort of deadlock. The reduction of the user population due to different times to go back home does not seem to be an issue. The only way to evolve from this steady state could be to distribute new – and easy to combine – elements during lunchtime.

From a game designer perspective, it is unrealistic to think of this kind of games as a closed system: a continuous stream of content from the outside is required as a stimulus to keep the game rolling. If this stimulus is not provided within hours, game evolution quickly reaches a steady state and withers.

Conclusion

In this chapter we addressed the complex problem of designing a game to be run on top of an opportunistic network. In particular, after discussing general issues and understanding technical challenges, a meaningful case study has been proposed in order to see a practical application for a number of practical game design guidelines. The proposed case study focuses on the simulation of an opportunistic network deployed on three different public transportation systems using parameters coming from the actual urban settings of Milan (Italy), Edmonton (Alberta, Canada), and Chicago (Illinois, U.S.A). The game design guidelines described here aim to improve the Quality of Experience perceived by players intended as both level of involvement and level of rewarding. Moreover, we also outlined a number of features from a quantitative point of view. First, there is a degree of independence between player activity and user density. Second, interaction between users may be limited but scales up with the number of available options in the game. Finally, the game cannot be considered as a closed system but must constantly receive content from the outside in order to keep entertaining the players.

For the future, a number of issues still need to be addressed by the scientific community. First of all, we are lacking a taxonomy for opportunistic games. This taxonomy will help us to generalize best practices and apply them to game genres rather than single case studies. Moreover, even limiting the discussion to urban level, the resulting system is very complex and cannot be evaluated completely without an actual test bed. Last but not least, researchers should be looking for a way to integrate in the system other urban elements, such as taxis and pedestrians. The inclusion of additional urban elements will help make the system truly pervasive and may foster attention to local services or cultural heritage, implementing some sense of *urban gamification*.

Recommended Reading

S. Ahmed, S. Kanhere, Cluster-based forwarding in delay tolerant public transport networks, in *Proceedings of the 32nd IEEE Conference on Local Computer Networks* (2007), pp. 625–634

A. Balasubramanian, B. Levine, A. Venkataramani, DTN routing as a resource allocation problem, in *Proceedings of the 2007 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications* (2007), pp. 373–384

N. Banerjee, M.D. Corner, B.N. Levine, Design and field experimentation of an energy-efficient architecture for DTN throwboxes. IEEE/ACM Trans. Networking **18**(2), 554–567 (2010)

C. Boldrini, M. Conti, A. Passarella, Modelling data dissemination in opportunistic networks, in *Proceedings of the Third ACM Workshop on Challenged Networks* (2008), pp. 89–96

D. Brockmann, L. Hufnagel, T. Geisel, The scaling laws of human travel. Nature **439**, 462–465 (2008)

J. Burgess, B. Gallagher, D. Jensen, B.N. Levine, Maxprop: routing for vehicle-based disruption-tolerant networks, in *Proceedings of the 25th IEEE International Conference on Computer Communications* (2006)

H. Cai, D.Y. Eun, Aging rules: what does the past tell about the future in mobile ad-hoc networks? in: *Proceedings of the Tenth ACM International Symposium on Mobile Ad hoc Networking and Computing* (2009), pp. 115–124

A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, J. Scott, Impact of human mobility on the design of opportunistic forwarding algorithms, in *Proceedings of the IEEE International Conference on Computer Communications* (2006), pp. 1–13

P. Coulton, O. Rashid, W. Bamford, Experiencing "Touch" in mobile mixed reality games, in *Proceedings of the 4th Annual International Conference in Computer Game Design and Technology* (2006)

I.D.D. Curcio, M. Lundan, SIP call setup delay in 3G networks, in *Proceedings of the Seventh International Symposium on Computers and Communications* (2002), pp. 835–840

C.T. De Oliveira, R.B. Braga, D.M. Taveira, N.C. Fern, O.C.M.B. Duarte, A predicted-contact routing scheme for brazilian rural networks, in *Proceedings of the Electrical Engineering Program, COPPE/UFRJ* (2008)

J. Dedrick, V. Gurbaxani, K. Kraemer, Information technology and economic performance: a critical review of the empirical evidence. ACM Comput. Surv. **35**(1), 1–28 (2003)

M. Demmer, K. Fall, Dtlsr: delay tolerant routing for developing regions, in *Proceedings of the 2007 Workshop on Networked Systems for Developing Regions* (2007), pp. 1–6.

F. Ekman, A. Keränen, J. Karvo, J. Ott, Working day movement model, in *Proceedings of the 1st ACM SIGMOBILE Workshop on Mobility Models* (2008), pp. 33–40

K. Fall, A delay-tolerant network architecture for challenged internets, in *Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications* (2003), pp. 27–34

M. Flintham, S. Benford, R. Anastasi, T. Hemmings, A. Crabtree, C. Greenhalgh, N. Tandavanitj, Where on-line meets on the streets: experiences with mobile mixed reality games, in *Proceedings of the of the SIGCHI Conference on Human Factors in Computing Systems* (2003), pp. 569–576

R. Gass, C. Diot, An Experimental Performance Comparison of 3G and Wi-Fi, in *Proceedings of the 11th International Conference Passive and Active Measurement, Lecture Notes in Computer Science Volume 6032* (2010), pp. 71–80

P. Giaccone, D. Hay, G. Neglia, L. Rocha, Routing in quasi-deterministic intermittently connected networks, in *Bioinspired Models of Network, Information, and Computing Systems*, vol. 39 (2009), pp. 126–129

M.C. Gonzalez, C.A. Hidalgo, A.L. Barabasi, Understanding individual human mobility patterns. Nature **453**, 779–782 (2008)

Google. Android dashboard: platform versions (2014), http://developer.android.com/about/dashboards/. Accessed 11 Nov 2014

D.L. Hansen, Z.O. Toups, L.E. Nacke, A. Salter, W. Lutters, E. Bonsignore, Mixed reality games, in *Proceedings of the Conference on Computer-Supported Cooperative Work and Social Computing* (2012)

P. Hui, A. Chaintreau, J. Scott, R. Gass, J. Crowcroft, C. Diot, Pocket switched networks and the consequences of human mobility in conference environments, in *Proceedings of the ACM SIGCOMM First Workshop on Delay Tolerant Networking and Related Topics* (2005)

P. Hui, J. Crowcroft, E. Yoneki, Bubble rap: social-based forwarding in delay tolerant networks, in MobiHoc, in *Proceedings of the 9th ACM International Symposium on Mobile Ad hoc Networking and Computing* (2008), pp. 241–250

S. Jain, K. Fall, R. Patra, Routing in a delay tolerant network, in *Proceedings of the 2004 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications* (2004), pp. 145–158

J. Jetcheva, Y.C. Hu, S. PalChaudhuri, A. Saha, D. Johnson, Design and evaluation of a metropolitan area multitier wireless ad hoc network architecture, in *Proceedings of the Mobile Computing Systems and Applications* (2003), pp. 32–43

A. Lakaniemi, J. Rosti, V.I. Raisanen, Subjective VoIP speech quality evaluation based on network measurements, in *Proceedings of the IEEE International Conference on Communications*, vol. 3 (2001), pp. 748–752

W. Le, M. Jukka, Energy consumption analysis of WLAN, 2G and 3G interfaces, in *Proceedings of the of IEEE/ACM Int'l Conference on Green Computing and Communications* (2010), pp. 300–307

K. Lee, S. Hong, S. Kim, I. Rhee, S. Chong, SLAW: a mobility model for human walks, in *Proceedings of the IEEE International Conference on Computer Communications* (2009), pp. 855–863

A. Lindgren, A. Doria, O. Schelén, Probabilistic routing in intermittently connected networks. Mobile Comput. Commun. Rev. **7**(3), 19–20 (2003)

D. Maggiorini, C. Quadri, L.A. Ripamonti, On the Feasibility of opportunistic collaborative mixed reality games in a real urban scenario, in *Proceedings of the International Conference on Computer Communications and Networks* (2012a), pp. 1–5

D. Maggiorini, C. Quadri, L.A. Ripamonti, Scaling online collaborative games to urban level, in *Proceedings of the IFIP Wireless Days Conference* (2012b), pp. 1–6

D. Maggiorini, L.A. Ripamonti, A. Bujari, Palazzi C E (2013) Evaluating design constraints for proximity-based games in a real urban topology, in *Proceedings of the IEEE International Conference on Multimedia and Expo*, pp. 1–6

D. Maggiorini, C. Quadri, L.A. Ripamonti, Opportunistic mobile games using public transportation systems: a deployability study. Multimedia Syst. J. **20**(5), 545–562 (2014)

P. Milgram, H. Takemura, A. Utsumi, F. Kishino, Augmented reality: a class of displays on the reality-virtuality continuum, in *SPIE Telemanipulator and Telepresence Technologies* (1994), pp. 282–292

M. Montola, A ludological view on the pervasive mixed-reality game research paradigm. Pers. Ubiquit. Comput. **15**(1), 3–12 (2011)

S. Nelson, M. Bakht, R. Kravets, Encounter-based routing in DTNs, in *Proceedings of the IEEE International Conference on Computer Communications* (2009), pp. 846–854

K. Pentikousis, M. Palola, M. Jurvansuu, P. Perala, Active goodput measurements from a public 3G/UMTS network. IEEE Commun. Lett. **9**(9), 802–804 (2005)

A. Pentland, R. Fletcher, A. Hasson, Daknet: rethinking connectivity in developing nations. Computer **37**(1), 78–83 (2004)

G.P. Perrucci, F.H.P. Fitzek, G. Sasso, W. Kellerer, J. Widmer, On the impact of 2G and 3G network usage for mobile phones' battery life, in *Proceedings of the European Wireless Conference* (2009), pp. 255–259

D. Pralhad, H. Xiaoxiao, D.R. Samir, Performance comparison of 3G and metro-scale WiFi for vehicular network access, in *Proceedings of the 10th ACM Conference on Internet Measurement* (2010), pp. 301–307

A. Qureshi, J. Carlisle, J. Guttag, Tavarua: video streaming with WWAN striping, in ACM Multimedia 2006 (2006)

I. Rhee, M. Shin, S. Hong, K. Lee, S. Chong, On the Levy walk nature of human mobility, in *Proceedings of the IEEE International Conference on Computer Communications* (2008), pp. 924–932

I. Rhee, M. Shin, S. Hong, K. Lee, S.J. Kim, S. Chong, On the levy-walk nature of human mobility. IEEE/ACM Trans. Networking **9**(13), 630–643 (2011)

N. Ristanovic, J.Y. Le Boudec, A. Chaintreau, V. Erramilli, Energy efficient offloading of 3G networks, in *Proceedings of the 8th IEEE International Conference on Mobile Adhoc and Sensor Systems* (2011), pp. 202–211

G. Sandulescu, S. Nadjm-Tehrani, Opportunistic DTN routing with window-aware adaptive replication, in *Proceedings of the 4th Asian Conference on Internet Engineering* (2008), pp. 103–112.

M. Sede, L. Xu, L. Da, W. Min-You, L. Minglu, S. Wei, Routing in large-scale buses ad hoc networks, in *Proceedings of the Wireless Communications and Networking Conference* (2008), pp. 2711–2716

A. Sharma, V. Navda, R. Ramjee, V.N. Padmanabhan, E.M. Belding, Cool-Tether: energy efficient on-the-fly wifi hot-spots using mobile phones, in *Proceedings of the 5th International Conference on Emerging Networking Experiments and Technologies* (2009), pp. 109–120

Sony. Playstation®Vita (2011), http://www.playstation.com/psvita/. Accessed 11 Nov 2014

Z. Xiaoying, L. Feng, Y. Jishen, Analysis of user-perceived web quality of service delay, in *3rd IEEE International Conference on Network Infrastructure and Digital Content* (2012), pp. 319–324

Q. Yuan, I. Cardei, J. Wu, Predict and relay: an efficient routing in disruption-tolerant networks, in *Proceedings of the 10th ACM International Symposium on Mobile Ad hoc Networking and Computing* (2009), pp. 95–104

X. Zhang, J. Kurose, B.N. Levine, D. Towsley, H Zhang, Study of a bus-based disruption-tolerant network: mobility modeling and impact on routing, in *Proceedings of the 13th Annual ACM International Conference on Mobile Computing and Networking* (2007), pp. 195–206

QoE and Latency Issues in Networked Games

23

Jose Saldana and Mirko Suznjevic

Contents

J. Saldana (✉)
Department of Electrical Engineering and Communications EINA, Aragon Institute of Engineering Research (I3A), University of Zaragoza, Zaragoza, Spain
e-mail: jsaldana@unizar.es

M. Suznjevic
Department of Telecommunications, Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, Croatia
e-mail: mirko.suznjevic@fer.hr

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*, DOI 10.1007/978-981-4560-50-4_23

Abstract

The rise of the Internet opened new possibilities for computer games, allowing real-time interaction between players in different parts of the world. Online games permit a number of people to compete in a shared virtual world. However, the synchronization and the maintaining of a coherent game state to be shared by the applications of all the players is not a trivial problem: different sources of latency appear and may cause inconsistencies between the game states observed by each of the players. Different genres of online games present specific latency requirements, depending on the game dynamics, its characteristics, and the level of interaction between the players. This chapter discusses the different mechanisms that companies use in order to overcome the problem of network latency when providing online games: the use of low-bandwidth traffic flows, the different protocols used at transport level, the architectures employed, the distribution of the hardware resources, the mechanisms for hiding the effect of the network to the players, etc. In addition, the different techniques used for estimating the user's Quality of Experience from network parameters are surveyed. Although latency is the most important parameter, other ones such as packet loss, delay variation (jitter), or bandwidth are also considered. Different QoE-enhancing mechanisms, such as client-side prediction or server delay compensation, are summarized. Other scalability-related techniques are also explained.

Keywords

Online games • Quality of Experience • Latency • Network Impairments • Real-time interaction

Introduction

The popularity of computer games is growing, and nowadays the gaming industry is a solid and well-established one. Although some people still think that computer games are just toys for kids, this is no longer true: according to the Entertainment Software Association (ESA) *Top Ten Industry Facts 2013* report, the average age of a game player in the USA is 30, and he/she has been playing games for 13 years.

There are many reasons and motives why people play digital games. Research has confirmed the impact of several factors such as challenge, freedom to act in a virtual word, opportunities to socialize with other people, etc. (Deci and Ryan 1985). Regarding the challenging aspects, computer games try to find the equilibrium between two extremes: boredom, which is caused by trivial and non-challenging tasks, and anxiety, which presents itself when the task is perceived as too hard or impossible. In such equilibrium, users can be fully immersed in the game and enter the "state of flow" (Chen 2007). Developers use different methods for making the game engaging, like the creation of immersive, complex, and important stories (e.g., saving the world), high scores, records of completing certain activities in the game (achievements), etc. Some other components have a social dimension such as enabling multiplayer interactions with other players, competitiveness tools (e.g., ranking lists), etc. These social components are certainly essential: if you "save the world" but no one notices it in the real world, the reward is perceived as significantly lower than in the case you can share your accomplishments with someone else. Similarly, showing off the new powers and weapons of your virtual avatar to other players can be considered as one of the motivational aspects of the game.

As a consequence, more and more people no longer want to play *against the machine*, but they want to fight with and against real people (and beat them), because it is more challenging and rewarding. And if you defeat in a virtual world someone you know in the real world, it is even better! This effect can be seen even on mobile devices, where the majority of the games produced are simple ones intended for one player and off-line play, but the games producing most revenue are multiplayer ones.

In the first years of computer games, player-versus-player competition was only possible in a very limited way: many games and consoles included two (or more) controllers, thus allowing real-time competition between players. It was certainly very limited: the players had to be in the same place and their number was reduced (two or four). Nevertheless, the experience was better than that of turn-based games, where only one player could use the controller at a certain moment.

However, the rise of the Internet opened new possibilities, since it allowed the interaction between computers in different parts of the world. In principle, the Internet was not designed as a real-time network, and it was not able to guarantee an upper bound for packet delivery delay. In fact, the first widely deployed services (e.g., e-mail, file transfer, virtual terminal connection) tolerated some amount of delay. However, in the last decades, as more and more households were connected to the Internet (with, e.g., 14 or 33 kbps modems), the Internet has been utilized for supporting real-time services as Voice over IP (VoIP) or video conferencing, where delay is critical. In this context, multiplayer online games became popular worldwide for the first time during the 1990s.

Different limits for the delay of real-time services have been found: for example, in 1996, the International Telecommunications Union (ITU) reported in their Recommendation G.114, *One-way transmission time* that a one-way delay of 150 ms could be considered as the limit for a comfortable voice conversation. As the delay grows, the probability of the two call participants interrupting each other becomes higher.

The terms "latency" and "delay" are normally used interchangeably. In the hardware business, in general, "latency" basically means "inherent delay" (i.e., the delay caused by the underlying technology). By contrast, "delays" are due to other holdups (e.g., packet processing, queuing due to network congestion, or retransmission of data due to packet loss). In networking terms, such delays can be incurred on every router on the network path being taken by the data. In this chapter we will use the term "latency" meaning all the time required by the information for traveling from one application layer to another.

When a game developer is designing a non-online title for a console, everything is under their control: the hardware, the software developing tools, and the methodology are well known. So their duty ends when the game box is sold to the player.

However, when a game includes an online mode, the company is also in charge of the network support, and they have to maintain an infrastructure including a number of servers and network resources in order to let the game work. So selling the game box is just a new beginning. In addition, a new problem appears; although the software and the hardware where it runs are well known, there is something between the different devices that is not under the company's control – the network.

Companies employ different techniques in order to provide games in a scalable, reliable, and profitable way, with the main objective of providing a good Quality of Experience. This concept not only includes Quality of Service, which is directly related to the parameters that can be measured from the network. The concept of Quality of Experience also includes the expectations, feelings, perceptions, cognition, and satisfaction of the user (Rehman-Laghari et al. 2011). This means that two players may feel very different when their game is affected by the same latency: the most experienced player will feel the problem as more severe than the novel one (Suznjevic et al. 2013).

This chapter surveys the most extended practices used by game providers when supporting networked games and the methods used to estimate subjective quality of the game users. There is a special difficulty with this kind of service, since gamers show a very demanding profile: they always want the best in terms of graphics, speed, frame rate, sound, etc. As far as network is concerned, gamers have also proven to be very difficult customers to deal with. In Chambers et al. (2005), different game traces from a number of servers were studied and some conclusions were devised: players are impatient when connecting to a server, they present short attention spans, they are not loyal to a server (when they can select between different servers of the same game), and they reveal when they lose interest, so game providers can detect waning interest and react to it. One consequence of this is that game workloads are only predictable over short-term intervals. Finally, predicting the success of an online game has revealed as a very difficult task, so the provision of the resources is a complicated problem, which is stressed at launch time of a new title.

The Impact of Network on the Shared Virtual World

As previously said, when the whole game is located on a local device, be it a PC, a console, or something else, a player can interact with the virtual world of a game in "real time," meaning that the time which passes since the player makes an action until the virtual world presents a response to that action is very low (lower than the limits of human perception, usually just a couple of milliseconds).

However, when we introduce the notion of a shared multiplayer virtual world, we also introduce additional latency in receiving the response, caused by the network. We will illustrate this using a simple case of a duck-hunting game using a client–server architecture with three clients (Fig. 1). We will consider that each of the three clients experiences 50, 100, and 150 ms of one-way network delay from their computers to the server, respectively. Let us assume that the second client shots the target at a certain moment. The time it takes the action to travel to the

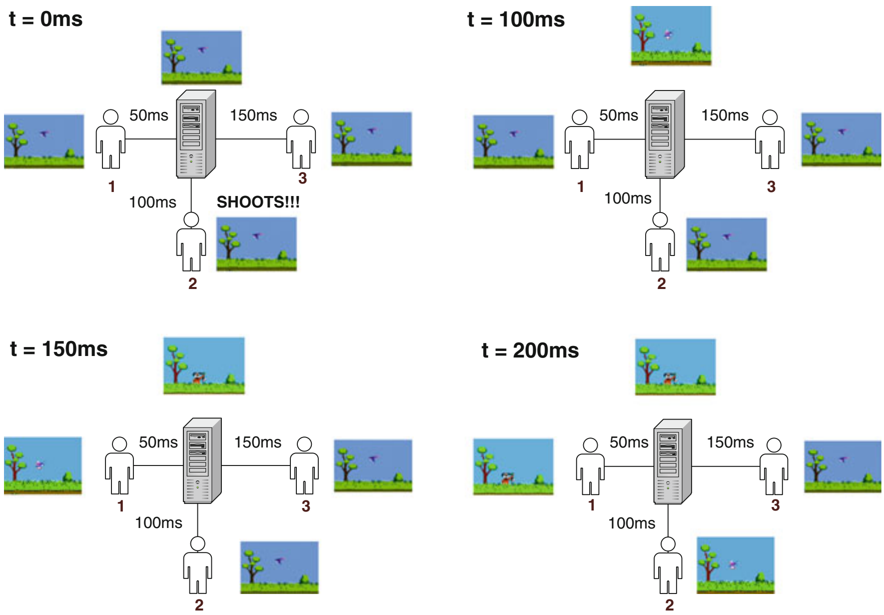


Fig. 1 Scheme of a game with three players experiencing different amounts of latency

server and back is 200 ms. Thus, at 100 ms mark, only the server will know that the duck has been shot; at 150 ms, client 1 will know that the duck has been shot, and at the server side the dog has already retrieved the duck. At 200 ms, client 2 will see he/she has hit the duck, while client 3 still sees the duck alive.

There are two basic approaches for handling the synchronization of the state of the distributed virtual world. The first on is “lockstep” deterministic simulation which ensures that all clients have the same state through delaying the execution of commands for a time period which is needed for all the participants in the virtual world to get the information regarding that command. Such approach is used, for example, in *StarCraft 2*, where all the actions of the users are delayed for the default value of 200 ms. This approach is viable in games which lack very high and strict latency requirements, such as strategy games. In our example this would increase the response time from action to reaction of client 1 for 200 ms, which is higher than the latency imposed by the network itself. The second approach is to execute the updates as soon as they arrive to the client, which is employed in very fast-paced games like First-Person Shooters. In our example this approach would increase the delay in which reaction is received by client 2 by the value of network latency, but this approach also creates three different versions of the state of the virtual world, which are separated in time on the different client machines. These discrepancies between client-side states of the virtual world can cause unexpected behaviors, anomalies, and reduction of the Quality of Experience of the end users. As we will see, game developers employ various mechanisms to deal with these anomalies.

As it can be deduced from the example, the worst enemy of gamers, from a network point of view, has a name, *latency*, and this is stressed for the most interactive game genres, where movements are fast, and some tens of milliseconds do matter. This problem is shared with other real-time services, and different means for reducing it are being proposed. In fact, the Internet Engineering Task Force (IETF), in charge of developing the Internet protocols, is looking for proposals and organizing meetings with the aim of reducing the latency of the Internet (Ford 2014).

Online Game Classification

Online Game Genres

Games can be grouped into different genres, based on their mechanics and concepts. While game genres are quite general terms with a bit blurred border, certain tendencies regarding the network traffic of each genre can be observed. The NPD Group, in its *Software Category Definitions* report (2008), established 13 different game categories. However, not all these categories include the possibility of playing online. In the present chapter, we will only focus on those categories where online playing is a must, or those where networked versions of the games are usual. According to this, the game genres considered are:

– First-Person Shooters (FPS): is a game genre which yielded the first multiplayer online game (*Maze War*) and was one of the first genres which became popular worldwide with *Wolfenstein 3D*. In a FPS a virtual world (often very limited in size) is shared by some tens of players grouped into teams, each of them controlling an avatar, with the aim of accomplishing a mission or killing all the enemies. The view of the user is first person, meaning that the player sees the hands and the weapon of his/her avatar and the part of the scenario in front of him. The main characteristic of this genre is the high interactivity it requires: movements and shots are really fast, and the aim of the player matters. For example, average Time To Kill (TTK) or average time since the weapon is fired until the opponent is slain, in a recent FPS (*Call of Duty: Ghosts*, released in 2013) is just 161 ms for the most popular assault rifles. Studies have shown that these games need very low latency values for providing a satisfactory QoE.
– Massively Multiplayer Online Role-Playing Games (MMORPG): In these games the virtual world is usually very big, and the number of real users sharing it is also huge (typically several thousands). In addition, a number of NPCs (Non-Player Characters) controlled by AI (Artificial Intelligence) are present in the virtual world. Each user controls a persistent avatar which can learn new abilities and earn weapons and equipment. A wide range of activities can be performed in the game: different quests assigned by NPCs allow the player to improve the skills of his/her avatar; trading with other characters is also allowed, using virtual money; group missions are included, where players have to cooperate in order to accomplish a difficult challenge; and finally, some areas in a

virtual world permit a number of players to fight against another group. Although the interactivity of these games is not as critical as in FPSs (e.g., the aim is not as important, since you first click on the objective and then select the spell or weapon to use against it), the speed and the delay also matter and can decide the result of a fight.

– Real-Time Strategy (RTS): A number of players (typically up to 10) share a virtual world, and each of them controls a civilization including army units, buildings, and structures. The typical actions of the player include gathering different resources, making new buildings, technological development, or creating armies. The control of the virtual characters is indirect, meaning that the player only selects the characters and assigns tasks to them (e.g., gathering food, attacking an enemy town, etc.). The term "real time" means that they are not turn-based, but all the players act at the same time, each one from a computer.
– Multiplayer Online Battle Arena (MOBA): In these games, considered as a subgenre of RTSs, two teams fight in order to conquer the battlefield. Each player controls a character and has to cooperate with other real players and also with other ones controlled by the computer. The control is more "direct," taking into account that only a virtual avatar is directly managed by the player.
– Sports: This genre groups very different games, including car racing, soccer, etc.

Delay Sensitivity of Different Genres

As remarked in Feng et al. (2005), gamers are very sensitive to delay, but the maximum tolerable delay varies with the game genre. Thus, some studies have taken a simple approach, just based on latency, and ignoring other network impairments as jitter (i.e., variation of the network latency) or packet loss. This can be enough in a first approach, although some more advanced models for estimating subjective quality also consider other network impairments, as we will see.

In Dick et al. (2005), a survey with a number of FPS games was carried out. As a result, it was reported that a one-way delay of 80 ms could be acceptable for most of the users (they rated it as "unimpaired"). In Henderson and Bhatti (2003), it was shown that delay has an impact on the decision of a player to join a game server, but the influence on the decision of leaving the server is lower.

Regarding MMORPGs, the study carried out in Ries et al. (2008) showed that the quality level rated by the players dropped from "excellent" to "good" for one-way delays greater than 120 ms. In Chen et al. (2006a), it was reported that the duration of game sessions declined for values of the one-way delay about 150 or 200 ms. Some studies have also been carried out for RTSs: the reported delays range from 200 ms (Cajada 2012) to 500 ms (Claypool and Claypool 2006) of one-way delay.

A final remark has to be added here: the game genre is not the only parameter to be taken into account when estimating the maximum delay that a player may tolerate. Some studies have shown that the experience of the player increases his/her delay sensitivity, i.e., an experienced player gets annoyed more easily than a novel one (Suznjevic et al. 2013). In addition, the behavior and the skill level of the other players in the party also have an influence on the subjective quality. All in all, the figures presented have some degree of uncertainty and may not only depend on network latency.

Architectures of Online Gaming Systems

The synchronization of a game is not a trivial problem: all the clients share a virtual world and it must be consistent, taking into account that the network latency experienced by each player may have a different value, and it may also vary during the party.

Two basic models have been in use for online games: client–server and peer to peer (p2p). In the client–server model, there is one central entity which holds a "true" copy of the virtual world state (usually a server), and all the other entities are under its authority. As a consequence of network latency, desynchronization may occur, and it may happen that the game state calculated by a client is corrected by the server. In peer-to-peer games, this central authority does not exist, so clients have to synchronize between them. In the history of online games, both models have been used, but in the last decade a clear dominance of the client–server architecture can be observed (Feng et al. 2005). There are some exceptions where a peer-to-peer scheme is used, as, e.g., *StarCraft 1* (1998), where up to eight people could play together (Claypool et al. 2003; Lee 2012). Some academic research also considered the peer-to-peer model for an experimental game called *MiMaze* (Gautier and Diot 1998). In addition, the possibility of using a peer-to-peer architecture for a popular MMORPG (*World of Warcraft*) was considered in Miller and Crowcroft (2010), but the results showed that it was not a good idea, since it would occasionally saturate residential connections and would increase latency with respect to client–server solutions. In the remainder of the text, we will focus on the client–server model.

Client–Server Model

There are several reasons for companies to employ client–server architectures. First of all, the fact of having a single "authority" makes synchronization easier. The server is the authority and the clients must obey. However, cheating in online games is also possible: for example, if the time stamp of an action is modified before a packet is sent to the network, a shot can be considered as previous with respect to the enemy's one, thus giving an advantage to the cheater. In that sense, a server establishing a single time source can avoid the problem. Another clear advantage appears, i.e., billing is easier if the company controls the server: if the player does not pay (or does not register properly), he/she may not be allowed to play.

Companies employ different policies regarding the control of the server. They can be classified as:

- The server application is distributed with the game (either included in the game itself or as a separate application) and can be installed and controlled by the player (e.g., *Warcraft 3, Counter-Strike 1*). The user is then able to create a dedicated server, or even one of the clients may act as the server. This option allows the users to create LAN parties, where latency can be really low, or to host the games over the Internet, i.e., users may create a high number of servers (with public IP addresses) on which other users can play. If this policy is employed, the game developer does not have full control, but neither has to invest in the hardware and network infrastructure. Over the time, different services for finding available servers have been developed (e.g., www.gametracker.com), and they significantly help the players to find servers which satisfy their demands for network latency, game type, etc. This option was used by the authors of Feng et al. (2005) to obtain detailed statistics of *Counter-Strike 1*, generated in a public server created by them.
- The dedicated server is included in the released application, but it is not controlled by the user. This means that any player may act as the server, but the game developer company performs a sort of "orchestration" between the users, selects the one with the best connection, and assigns his/her client application the role of the server. This process is transparent to the player, who may or may not act as the server and is not aware of that fact. This has a clear advantage for the company: they only have to join a number of users who will play together, but the processing and network charge required for supporting a party is "externalized" and assigned to one of the players. As a result, the scalability is clearly increased, but this has a counterpart: what happens if the user acting as the server leaves the party? Fig. 2 shows a traffic capture of *Call of Duty* for *Playstation 2* obtained in May 2012, where an interval when the game stopped can be appreciated. At $t = 500$ s, the player who is acting as the server (*server #1*) decides to leave the party, so there is a gap until a new player is selected to act as the server. During the gap, the rest of the players watch a *server migration* message in their screens, and they have to wait until the device of another player takes the role of the server (*server #2*). Another problem of this policy is related to the unfairness: the player acting as the server has a clear advantage, because he/she experiences null network latency.
- Server fully controlled by the developer/publisher (e.g., *World of Warcraft*). This is the most common tendency nowadays for certain types of games such as MMORPGs. The servers are placed, controlled, and maintained by the game company which enables complete control, cheating prevention, and advanced business models (e.g., selling virtual items for real currency). There have been some cases in which reverse engineering techniques have been employed in order to decrypt the traffic, or to create a private version of the whole game server (this has happened, e.g., to *World of Warcraft*). This is also known as "server emulation" and has many legal issues (Debeauvais and Nardi 2010).

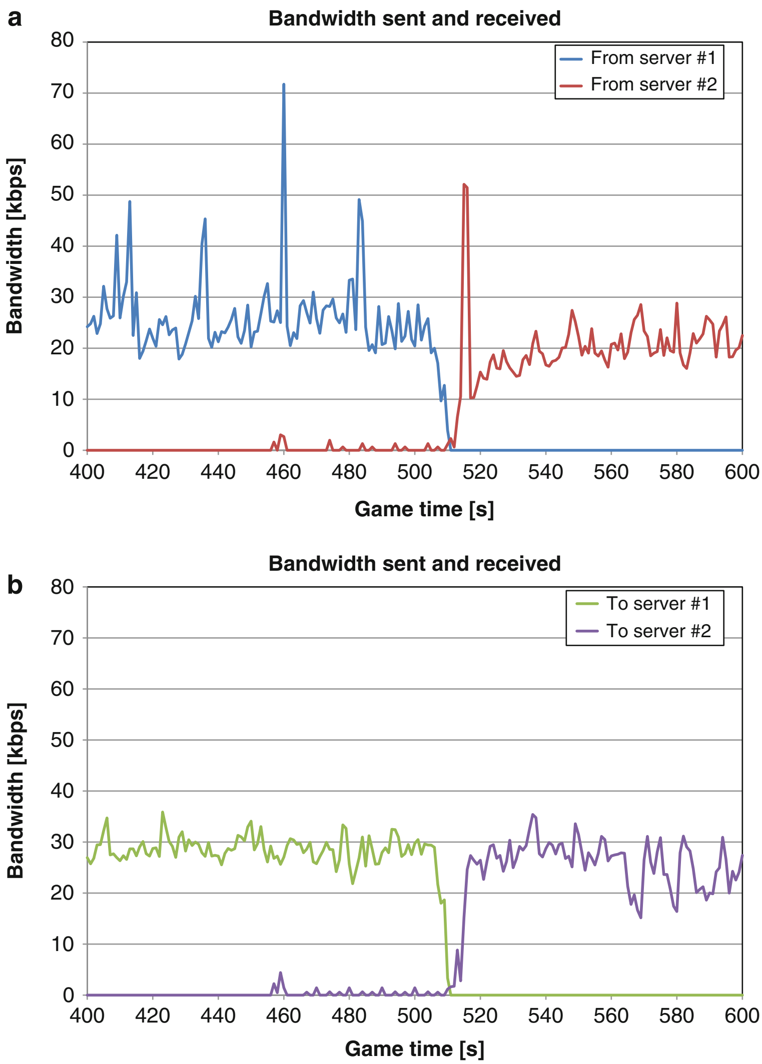


Fig. 2 Traffic capture during a server migration in Call of Duty

However, these versions usually do not work with the last versions and updates of the client and may contain bugs and viruses.

– In some Real-Time Strategy games (e.g., *StarCraft 2*), a peer-to-peer model is employed, but there is still a server. The server acts as a traffic aggregator: it receives the traffic from each client and sends the aggregated input to the other ones. It does not hold the game state (so there is not a central authority), but it forwards the game commands received from each player to the others. *StarCraft 2* is based on lockstep deterministic simulation: the method is based on each client queuing the commands received from every other player and executing them in the future (12 frames, which may be equivalent to 200 ms). The advantage of this communication approach between the clients for RTS is clear: they require very low bandwidth, since they just send the player's commands, instead of sending the updated position of each unit (there may be thousands of them). The presence of the server also reduces the sent bandwidth, since each client only sends its player's commands once, and the server is the entity in charge of forwarding them to the other clients. As a counterpart, an observable input delay appears, since units do not respond immediately to player's commands. In addition, the slowest player (i.e., the one with the highest latency) slows down the game for all of them. The reasons for including a server in *StarCraft 2* are not only related to traffic scaling. The server is also in charge of authentication, storing player's data, matching players with similar skill levels, and performing anti-cheating mechanisms. It is also able to reduce piracy, since the players have to use the server.

General Characteristics of Game Network Traffic

The main functions needed for executing an online game can be divided into: gathering the player's commands, executing the logic of the virtual world based on those commands, rendering the virtual scene, and displaying the information to the user. The information which is transferred between the client and the server is dependent on the location of these functions. In the majority of traditional games, the game logic is executed on the server, while the remaining functions are done on the client. In that case the information traveling from the server to the client comprises updates of the virtual world state which are then rendered and displayed on the client. From the client, only the commands of the player are sent to the server. In a more recent approach, commonly called "cloud gaming," both game logic and virtual world rendering are located on the server, from which a high-quality video stream is sent towards the client, while the player's commands are transferred from the client to the server.

In addition to player commands and inputs, and virtual world state refreshes and video streams, other flows may appear, such as chat or audio for player communication, authentication, accounting, etc. Although some games have inbuilt VoIP systems, many players use stand-alone applications (e.g., *TeamSpeak*, *Ventrilo*, *Skype*) when playing. Finally, game updates and downloads of elements as maps or some textures created by the players may also require additional traffic flows. In traditional online games, the game content (scenarios, characters, textures, etc.) has to be stored in the computer of the player in order to avoid its sending during the game. This often results in high storage requirements in the hard disk of the user. Two examples: The *World of Tanks* v9.2 folder requires 28 GB, whereas *World of Warcraft* v5.4.8 uses 24.8 GB. Game clients and supplementary data have been traditionally distributed over CDs or DVDs, but today it is also common that the game clients are just downloaded from the Internet. These flows in which game clients are obtained are not real time, but just typical Internet file downloads. However, these flows are not the ones directly supporting the game action, so they are out of the scope of this chapter.

Network Traffic Characteristics of Traditional Games

As said in the introduction, game providers control the two extremes of the communication: they manage the server and they develop the game client application. However, they do not have a strict control over the network connecting both elements (usually the Internet). So they have to drastically limit the traffic to be exchanged during the game to what is strictly necessary, looking for a wide penetration, even in countries where bandwidth is scarce. This tendency has not changed over the years: Figs. 3 and 4 show the traffic profile of two First-Person Shooter games (which are considered, traffic wise, the most demanding game genre): 5000 packets of a client–server flow of *Quake 2* (1997) are compared

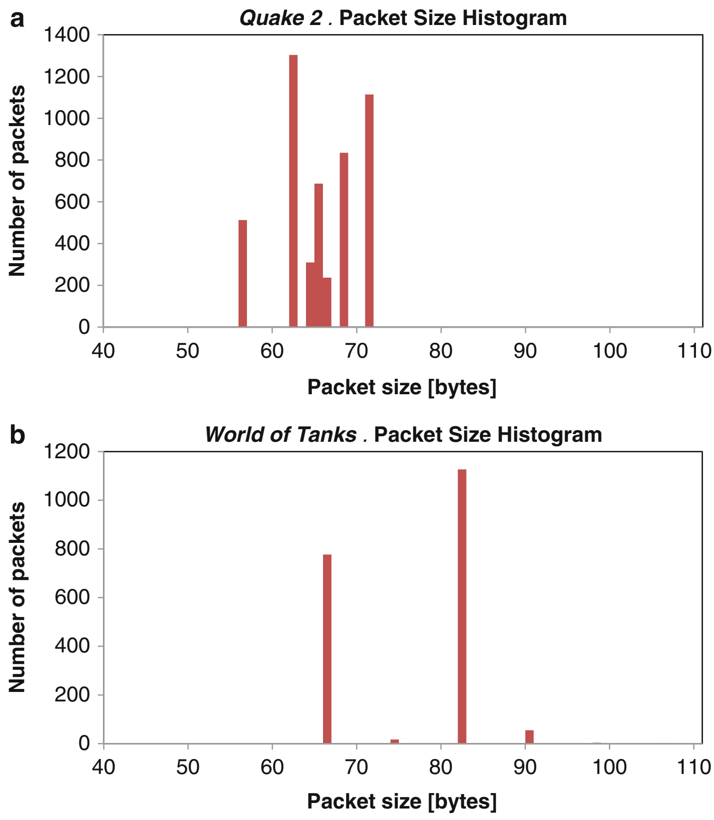


Fig. 3 Packet size histogram of (**a**) Quake 2 (1997) and (**b**) World of Tanks (2014)

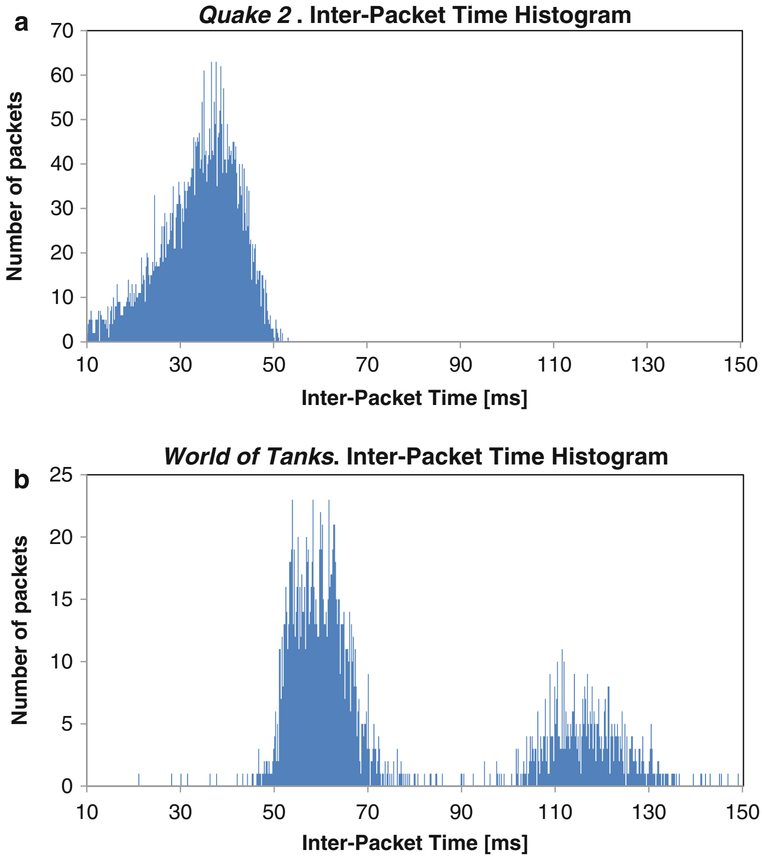


Fig. 4 Inter-packet time histogram of (**a**) Quake 2 (1997) and (**b**) World of Tanks (2014)

with the traffic of *World of Tanks*, using 2000 packets captured in 2014. As it can be seen, high rates (inter-packet time of about 30–50 ms) of small packets (60–80 bytes) are employed in both cases. The bandwidth of a client–server flow is roughly 16 kbps for the former and 8 kbps for the latter. A high number of game updates are needed in order to maintain the illusion of the virtual world which is responsive in the real time (i.e., below the time for which the human perception will notice the delay). In general, for the high Quality of Experience, the networked aspect of the game should be hidden from the player, and the game should run like it is stored on the local computer.

Thus, as illustrated in Fig. 5, the typical structure of an online game consists of a number of client applications and a server. During the game, no information is directly exchanged between the clients. Each client sends a flow to the server including the movements of the gamer. The server calculates the next state of the game and sends it to each player. In this scenario, three "bottlenecks" can be identified, namely:

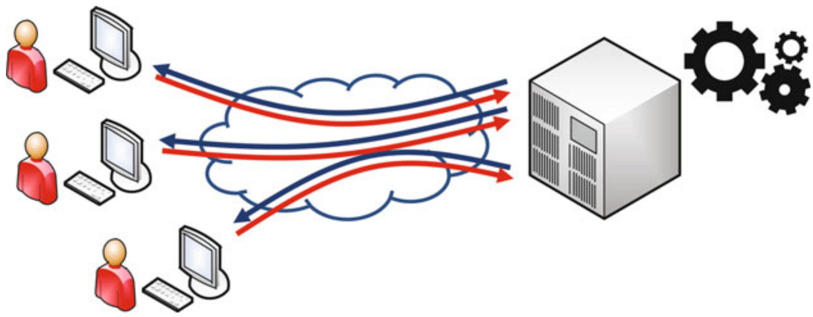


Fig. 5 Client/server infrastructure and traffic flows exchanged during a party

- The uplink resources, which must be enough so as to permit the information of each gamer to arrive in time to the server
- The processing capacity of the server, in charge of calculating the next state of the game
- The downlink capacity, necessary to send the state of the virtual world to each player

Two of these bottlenecks are related to the network, which begins with the access connection of the player, includes a number of Internet links, and ends with the connection of the server. The most stringent link is most likely the access link of the player, assuming that the game provider has correctly dimensioned his/her network.

There is a scalability issue derived from the use of a client/server architecture: while the amount of information received by the server grows linearly (if the server manages a party with N players, it will receive a flow reporting the movements of each one), the amount of information to be generated by the server grows with the square of N. Certainly, the server has to report to each of the N players the actions of the rest ($N-1$) of the players, so the number of flows is $N \bullet (N-1)$. This was the reason behind the limitation of the number of players in First-Person Shooters to some tens: the downlink of residential connections. As reported in Feng et al. (2005), some of these games were *designed to saturate the narrowest last-mile link*. As we will see, different techniques can be employed to overcome this limitation in the genres where the number of players sharing a virtual world must be higher.

Traffic Profile: High Rates of Small Packets

Despite their different genres, networked games share a set of common characteristics. The first one to be highlighted is that online games generate long-lived flows.

Nevertheless, the duration of the session depends on the game genre. In Feng et al. (2005), the traffic of a First-Person Shooter server was analyzed, and the results showed that the typical connection of a player lasted about 30 min in a server. At the same time, 99 % of the sessions lasted less than 2 h. In contrast, an MMORPG was studied in Tarng et al. (2008), and the average session time measured was 2.8 h.

According to their real-time requirements, they tend to generate traffic profiles consisting of high rates of small packets, and this is stressed with the interactivity level of each genre. According to this, First-Person Shooters are usually the ones presenting the highest rates and the smallest payloads.

As said in the introduction, the traffic profile based on high rates of small packets does not result in a high bandwidth usage. Many of these games can be played with a residential connection of some tens of kilobits per second. The "historical reason" (i.e., a 33 kbps modem should suffice to play the game in the 1990s) still persists since, regrettably, this may be the bandwidth amount available in some developing countries nowadays. This low-bandwidth usage is shared by FPSs (Ratti et al. 2010; Feng et al. 2005), MMORPGs (Svoboda et al. 2007; Suznjevic and Matijasevic 2012), MOBAs, and RTSs, which bandwidth usage is even lower: from 4 to 6 kbps, according to Claypool (2005). In the case of MMORPGs, the bandwidth may strongly vary depending on the activity performed by the player. Some studies have divided player actions into different categories (e.g., *trading*, *questing*, *player* vs. *player*), and the difference in terms of bandwidth can be up to five times (Suznjevic et al. 2009). If the player is fighting, his/her application is expected to generate more bandwidth than when he/she is just trading (buying or selling certain virtual goods). This is also related with the number of players involved in the action: the server has to send more information to the client if the number of players surrounding his/her avatar is high.

These usually low bandwidth requirements increase the market penetration of networked games, which are not only targeted for affluent countries. In fact, in developing countries, many people who do not have Internet connection at home still play these games in Internet cafés, very popular in these areas (Furuholt et al. 2008). In Batool and Mahmood (2010) and Gurol and Sevindik (2007), the profile of the users of Internet cafés was studied, reporting that more than 50 % of them do play online games there.

Another characteristic of these traffic flows has to be highlighted: they are very inefficient in terms of ratio of useful (game) information and signaling information (headers). Taking into account that an IPv4/UDP (the Internet Protocol version 4 and the User Datagram Protocol) header requires 28 bytes, if many of the payloads are about 20–60 bytes long, it can be seen that the efficiency is really low. This is the reason why some optimization techniques, based on header compression and multiplexing, have been proposed (Saldana et al. 2013), as an adaptation from VoIP optimization standards (Thompson et al. 2005).

Network Protocol: TCP or UDP?

UDP is more suitable than TCP (Transmission Control Protocol) for real-time services. UDP just sends the data and does not expect any feedback from the destination. In fact, RTP (Real-Time Protocol), the protocol used for sending VoIP streams, is designed for traveling on UDP datagrams. So in a first approach, the most intuitive option for an online game would be to select UDP at the transport level. In contrast, TCP is a closed-loop protocol initially designed for bulk transfers: it tries to get the maximum bandwidth amount while maintaining fairness and avoiding the saturation of the links. For this aim, it integrates a number of mechanisms, which rely on the reception of acknowledgment (ACK) packets from the destination, in order to control the transmission rate and to retransmit lost packets.

UDP prioritizes the continuous arrival of packets against the reliability of TCP. The disadvantage of UDP is that the service has to be designed with some tolerance to packet loss. As it happens in VoIP, where the possibility of losing some samples is assumed, and robust codecs are designed with the capacity to interpolate and regenerate lost information, UDP-based games have to implement some policies for hiding packet loss to the player. We will discuss them later.

However, some online games do use TCP, and this is somewhat related with the game genre. Some clear trends can be observed: FPSs always generate UDP flows. Many MMORPGs use TCP, although some of them use UDP (Chen et al. 2006b). The most popular MOBA (*League of Legends*) uses UDP. RTSs may use TCP or UDP.

When a game uses TCP, the initial assumption behind this protocol, "I have to transmit this information as fast as possible" is no longer true, taking into account that the information to be transmitted does not exist previously, but is generated as the user plays. This makes the behavior of TCP very different from what could be expected. In the literature, the normal use of TCP (e.g., for downloading a file) has been described as *network-limited,* since the network is the element setting the limit, whereas the use of TCP for an online game is called *application-limited* (Wu et al. 2009), taking into account that the application generates a continuous flow requiring a certain amount of bandwidth. The difference is illustrated in Fig. 6, where the evolution of TCP sending window (which controls the transmission rate) is shown. In (a), it can be observed that, during a file download (simulated in NS2) using FTP (File Transfer Protocol), the window follows a *sawtooth* shape: TCP increases its rate until network congestion occurs, and then it folds back, and begins the increase again. In contrast, (b) has been generated with the traffic model for *World of Warcraft* (an MMORPG) presented in Suznjevic et al. (2014), and it can be observed that the size of the sending window is always low. There are even certain moments where TCP has nothing to send.

Additionally, the use of TCP stresses the inefficiency of game flows. In Svoboda et al. (2007), it was reported that these games set to 1 the PUSH bit of the TCP

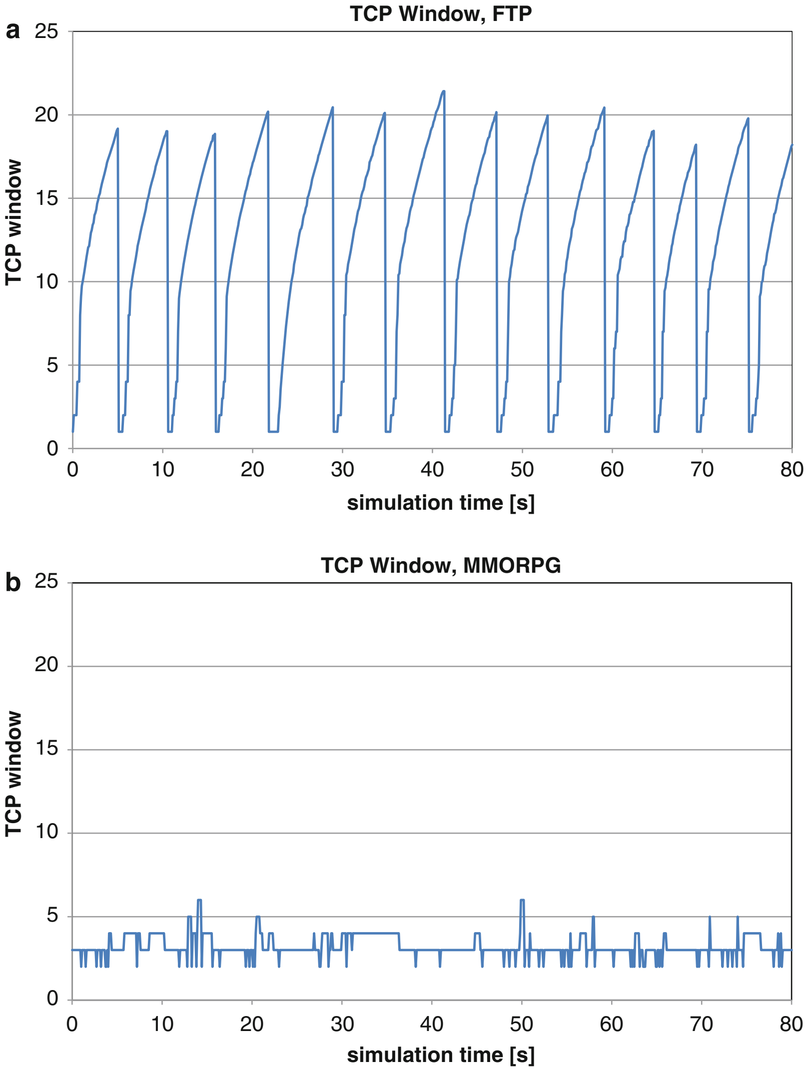


Fig. 6 Sending window of TCP for (**a**) an FTP download; (**b**) a client–server flow of World of Warcraft

header, thus making the protocol send the packet as soon as possible, without waiting for having a full payload. A bidirectional TCP connection is established between the client and the server, and ACKs are piggybacked into normal packets. However, in the traffic traces used in the study, 56 % of the packets were *pure* ACKs, i.e., they had no payload.

Client-Based Versus Cloud Games

The increase of the bandwidth available in households in some countries has led to a new model for supporting computer games, namely, *cloud gaming*. In this model, the game is run on a remote server (both game logic and virtual scene rendering), and the client is no longer a big application, but just a thin one in charge of reproducing a video stream and transmitting the player commands to the server. This has some clear advantages:

- Any game can be played without a previous installation. The model can be "select and play."
- The hardware requirements of the client device are drastically reduced, since it only reproduces a stream generated by the server (Claypool et al. 2012).
- From a business point of view, the subscription model is clearly the most suitable one. It should be noted that the player can not only play multiplayer games but also individual ones, so he/she will even pay for using an individual game online.

As a counterpart, the server infrastructure to be deployed by the game distributor is bigger, since each frame to be sent to a player has to be calculated on the server side. Also, bandwidth required is obviously higher, typically higher than 3 Mbps which is much higher than traditional games. Fig. 7 shows a comparison of bandwidth usage of various genres of traditional games and a game implemented as a cloud game. Cloud gaming can require up to 1000 times more network bandwidth than traditional games!

Estimating Quality of Experience

Quality of Experience (QoE) has been defined by the COST action IC 1003 Qualinet (2013) as the degree of delight or annoyance of the user of an application or service. It results from the fulfillment of his or her expectations with respect to the utility and/or enjoyment of the application or service in the light of the user's personality and current state. QoE has many aspects that comprise it and many factors which influence it. Influence factors can be grouped into *system*, *context*, and *user* factors. Games are a specifically hard topic for QoE research due to the complexity of the playing behavior and the additional factors which have an impact on the QoE (e.g., players' expertise in the game). This work is focused on one subset of system factors – network parameters – but it should be taken into account that many other factors impact the final QoE.

It can be said that VoIP was the first real-time service which became popular on the Internet. In order to accomplish its real-time requirements, it has to send small packets with a high cadence. The round-trip time (the time a packet takes to go to the other side and return) has to be kept low in order to make it possible for two people to maintain a normal conversation without interrupting each other. The one-way delay is also known as *mouth-to-ear* latency, and it not only includes

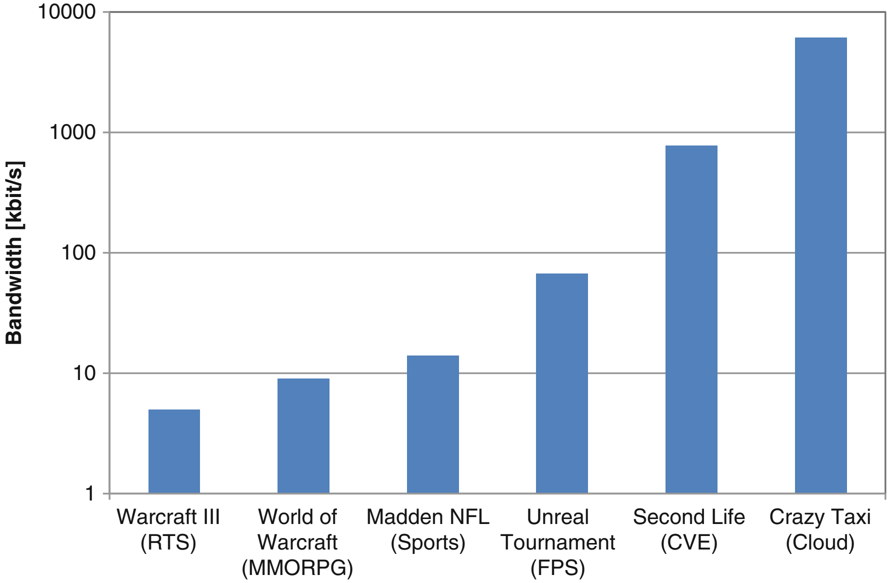


Fig. 7 Bandwidth usage of traditional game genres and cloud game

network delay, but also the latency caused by the codec (e.g., 20 ms for including two samples) and the equipment. The ITU defined a standard way to measure one-way transmission time in their G.114 Recommendation.

Online games' traffic has many similarities with that of VoIP: they also generate small packets with high cadences, so a comparison can be established between mouth-to-ear latency and player-to-server delay. Players usually talk about "lag" or "ping" as the round-trip time. This parameter is of primary importance for them, and in fact, many games (e.g., *Counter-Strike*) include the "ping" in the score table, in addition to the hits achieved by the player.

When defining the quality of a voice call, the ITU not only tried to capture objective parameters as, e.g., delay or loss of information, but it also developed a model able to estimate the subjective quality, taking as an input the objective parameters. The E-model, defined by ITU-T Recommendation G.107, calculates a score, called the "R factor," ranging from 0 to 100, which can be directly translated into a Mean Opinion Score (MOS) from 1 (bad) to 5 (excellent) (Cole and Rosenbluth 2001).

The difference between games and voice is clear: a conversation is always similar, because the problem is to transmit the sounds from the speaker to the hearer in a clear and non-delayed way. However, when considering a game, many factors intervene:

- The interactivity requirements of the game: As we have seen before, each game genre has its own characteristics and delay requirements. For example, in a First-

Person Shooter, a delay of 300 ms may jeopardize the game, whereas this delay can even be not noticeable in a Real-Time Strategy game.

- The way the game is developed: The implementation of different mechanisms (we will discuss them later) may hide the effect of network impairments to the player. Some advanced techniques, as smoothing the transition between the state predicted by the client and the real state received from the server, may significantly reduce the player's nuisance caused by delay.
- The transport protocol employed: If TCP is used, no information will be lost between the client and the server. However, if packet loss occurs in the network, this will require retransmissions and may cause additional latency: it may happen that one TCP packet is lost in the network but subsequent ones arrive correctly. In this case, the packets already arrived will be stopped by TCP till the new copy of the lost packet arrives, since TCP has to deliver packets in order.

The development of a subjective quality model requires a number of real people to participate in tests with the service. There are two basic methodology approaches which can be used: (a) single stimulus, in which for one gaming session players give an Absolute Category Rating (ACR) of QoE for some specific parameter values (e.g., MOS tests) or (b) double stimuli in which the users compare playing sessions with different parameter settings and decide which one is better (e.g., paired comparisons). Also, there are approaches in which bodily reactions are measured to establish QoE values (e.g., EEG, face expression, etc.). Therefore, a single model able to produce a subjective quality score for all online games, using as input the network parameters, cannot be developed. However, some specific models have been developed for certain games (Ries et al. 2008; Wattimena et al. 2006), and some more generic models have been proposed (Ubicom 2005).

Another approach some studies have taken is to perform "objective measurements" (Kaiser et al. 2009): a number of identical "bots," i.e., virtual avatars controlled by Artificial Intelligence, are placed in the same virtual scenario, and a number of parties between them are performed. If the number of parties is high enough, then the score will be the same for all the bots. Then, different network impairments (latency, jitter, packet loss) are added to one of the bots, and another set of tests is performed. The performance degradation of the network-impaired bot can then be statistically characterized. However, in this chapter the focus is on estimation of subjective quality of real players.

Network Impairments Considered

When network operators post their offers, they usually summarize everything into a single parameter: the maximum connection throughput measured in megabits per second. This is a parameter everyone can understand, and it is of primary importance when using the Internet. However, as long as online games are concerned, the maximum throughput may not result the most important parameter: as said before, online games do not require a high throughput (we are not talking about cloud-based games).

These network impairments are often known as "Key Performance Indicators (KPI)." In online games, the three mainly considered are latency, jitter, and packet loss. Between them, it is clear that the most important one is latency due to jitter and packet loss having small values in today's networks.

These impairments are strongly related between them, and they have sometimes been studied together for many online games. For example, the effect of delay and jitter was studied in Dick et al. (2005) for two FPS, a sports and an RTS game. Its effect is very different depending on the considered game and genre.

Latency

Latency can be defined as the time required for transmitting player's information from the application's layer of the client to the application's layer of the server. This would be the one-way delay. If the inverse path (from server to client) is also considered, then we have the round-trip time, informally known as "ping" by computer gamers. Many games (if not all) measure the value of the ping and report it to the user, taking into account its crucial importance when playing an online game. It is usually the only network parameter presented to the user.

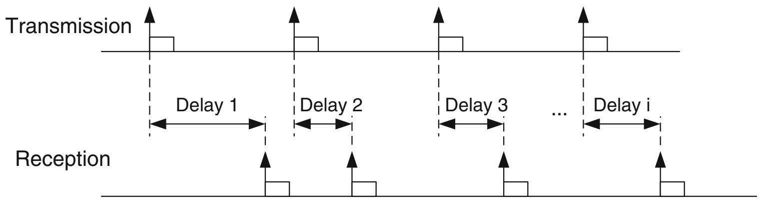
If the latency is high, the virtual world is not responding in real time and the player perceives his/her actions being "late." This severely impacts the QoE of the player and his/her immersion in the virtual world. The situation gets even more complex when we have multiple players with different latencies. These differences in latency can cause inconsistencies and anomalies in the virtual world, the most known being "shooting behind the wall problem." Let us assume that a player with a high delay is just running to hide from an enemy with lower latency, and the enemy shoots him while running across the open space. While the information about that shot reaches the player with high latency, in his/her section of the virtual world, he/she will have some time to, e.g., go round a corner, before the "you have been killed" message from the server arrives. So the perception of this lagged player will be something like "I have been shot around the corner which breaks the laws of physics." We will illustrate this problem in the section about QoE-enhancing mechanisms included in games.

When a networked game is played in a LAN, latency can be kept really low, in the order of 1 ms. In fact, some players organize the so-called LAN parties in order to play together with high bandwidth and also low latency levels. Some of these LAN parties are small, with a group of people going to other's houses and bringing their own computers. In Asia the games are, in general, played in large Internet cafes which also have their own LANs, and the games are often played in those LANs. Other LAN parties can be organized as big events with hundreds (or even thousands) of players in the same place for some days (e.g., a weekend), where public competitions and tournaments may also be held. The possibility of playing with really low latency levels is also a factor attracting people to these events. One of the biggest LAN parties is *DreamHack*, held twice a year in Sweden. *DreamHack* holds the world record for the largest LAN party in the world with over 22,000 players and has also held a record for the fastest Internet link in the world (120 Gbps) up to 2012 when it was beaten by the second largest computer festival in the world: The *Gathering* in Norway.

However, the problem we are considering in this chapter refers to the case where networked games are played through the Internet, with each player running his/her own computer at home. This case covers the vast majority of the gaming activity in the world. In this case, the concept of "latency budget" can be used (Ford 2014), taking this approach: there is a limit in the maximum delay a player is prone to tolerate, and there are different "sources of latency" contributing to it. These sources can be divided into the next categories:

- Generation delay: It refers to the delay between the physical event and the availability of the data. In the case of an online game, it is negligible, since the event to be captured is a key stroke, the movement of the mouse or the game controller, and it can be assumed that they are almost instantaneous.
- Packet transfer delay: It accounts for the propagation of the information between the source and the destination. It first includes the propagation delay, which refers to the part caused by the speed of light: for example, if the distance between the client and the server is 5,000 km, an unavoidable one-way delay of 16 ms appears. This gets increased in optical fiber: since its index of refraction is higher, the speed of light is reduced to about 200,000 km/s. Other contributions to transfer delay are buffering in intermediate routers and the time required for sending the bits at a certain rate (transmission delay). This may be negligible on gigabit links, but not in a residential access (e.g., a DSL uplink of 1 Mbps would require 12 ms for transmitting a packet of 1,500 bytes). All in all, packet transfer delay not only depends on the distance but on the number of intermediate routers, the congestion level of the traversed networks, the bandwidth of the links, and even the size of the packets.
- Processing delay: In online games, it is mainly caused by the response time of the server. This can be small in client games, but it may be significantly higher in the case of cloud games, since the server has to render the scene, code it in a video, and send it to the client application, which has to decode and display it to the user (this subpart can also be called "playout" delay). In Huang et al. (2013), a comparative study including two commercial cloud game platforms was presented, and the processing delay was in the order of 100–300 ms.

Many actions are being deployed in order to reduce or mitigate the contributions to the latency, some of them more effective than others. For example, the problem of "bufferbloat" (Gettys and Nichols 2011), i.e., the excessive delay caused by over-dimensioned router buffers, is being issued in the last years through mechanisms such as Active Queue Management (AQM) (Adams 2013). Geo-location of content sources (game servers in our case) is also a method for drastically reducing latency.



$$IPDV = \frac{\sum \left| delay_i - delay_{i-1} \right|}{n - 1}$$

Fig. 8 Definition of Inter-Packet Delay Variation (*IPDV*)

Variation of Latency

The statistical variation of the delay between different packets of a flow is sometimes known as "jitter." It may be caused by network congestion, which is translated into a different queuing delay for each packet on a flow. Another source of this variation can be that different packets of the same flow may not follow the same network path. If the traffic of an online game shares the bandwidth of a residential connection with, e.g., a TCP file download, this may also cause high delay variations, taking into account that TCP throughput typically follows a sawtooth pattern (see Fig. 6a).

Different definitions of the variation of latency can be found in the literature. In RFC 3393, the IETF defines the Inter-Packet Delay Variation (IPDV) for two packets inside a stream as "the difference between the one-way-delay of the selected packets." In RFC 3350, the method for measuring the jitter between each pair of RTP packets is defined: "The interarrival jitter J is defined to be the mean deviation (smoothed absolute value) of the difference D in packet spacing at the receiver compared to the sender for a pair of packets." In *netem*, a network emulator included in Linux, the jitter is expressed as the standard deviation of the delay, following a normal distribution. In Brun et al. (2006), the end-to-end jitter is formally defined as the "expected absolute value of the sum of inter-packet delay variations introduced by each node along the path between the source and the destination." The definition of Inter-Packet Delay Variation (IPDV) is illustrated in Fig. 8.

In Voice over IP, the jitter is compensated with a playout buffer where packets are buffered and played in a steady manner. Thus, the jitter is translated into an additional delay and packet loss, which happens when a packet arrives too late to be played. In fact, the formula for calculating the R factor of ITU's E-model does not consider jitter as an input. Something similar happens with video streaming or Internet TV, but the buffer can be in the order of seconds in these cases, since they do not present the same interactivity requirements.

Taking into account the tight interactive requirements of online games, a playout buffer is not practical, so jitter does affect them, and it is in fact used as an input variable for some subjective quality models for First-Person Shooters (Wattimena et al. 2006) or MMORPGs (Ries et al. 2008). For example, if the jitter is high enough to produce packet reordering, then it is translated into packet loss, since the game will not take into account, e.g., a position update prior to the last one received. In some Real-Time Strategy games, as already said, commands are queued and executed about 200 ms later due to the algorithm for state synchronization between all participating distributed clients. This could be considered as a sort of "playout buffer."

Packet Loss

Packet loss is another impairment which appears in networks. Different causes can prevent a packet from arriving to its destination in an IP network. Obviously, wireless networks have a higher packet loss rate than wired ones, although they implement different retransmission mechanisms at lower levels. Packet loss in wired networks is mainly caused by buffers: they may drop packets when they become full, but some policies (Active Queue Management Policies as, e.g., Random Early Drop) may discard packets according to some statistical probabilities even if the buffer is not full.

For TCP-based games, packet loss is not directly considered as a network impairment, since this transport protocol will retransmit the lost packets. However, this is translated into an additional "spikes" of delay, since the packets arrived correctly will have to wait until the new copy of the lost packet arrives.

When the game is based on UDP, it is assumed that some packets may get lost. In VoIP robust codecs are used, able to reconstruct the original signal even in the absence of some of the packets. In online games, different techniques are employed so as to hide the effect of packet loss to the end user.

Many of the QoE-enhanced mechanisms are designed for hiding the effect of packet loss to the player. The client and the server may be able to do a prediction if a packet is lost, based on the current movement of the avatar. However, if the forecasted position and the one arrived in the next packet are not the same, then the player may observe a sudden and abrupt movement of the scenario. Although this can be mitigated by smoothing algorithms, it reduces the quality experienced by the player.

Different First-Person Shooter games show very different degrees of robustness against packet loss. In Zander and Armitage (2004), two games were tested with real people, and the results regarding packet loss were very different: while *Halo* stopped working when packet loss was roughly 4 %, the users of *Quake III* did not experience any degradation in the quality even with a packet loss of about 35 %.

Bandwidth

Although bandwidth is not a network impairment by itself, its scarcity may be translated into high amounts of latency, jitter, and packet loss. In addition, background traffic (i.e., bandwidth utilization) is also a parameter to be taken into account: a strongly congested network is not good for playing online games.

Cloud games require high amounts of bandwidth, especially in the downlink, so if the connection is not able to provide it, the quality of the video needs to be reduced (in terms of resolution or frames per second), thus reducing the subjective quality in turn. Other option is to suffer packet loss which severely degrades the quality of the video stream (blocking and blurring effects) and therefore of the game as well.

Developed QoE Models

As remarked in Ubicom (2005), different benchmarks exist for some computer applications:

- Benchmarks for computer performance when running office applications are developed by certain companies (e.g., *PCMark* by *Futuremark*).
- The ITU's E-Model can be used for estimating subjective quality in VoIP.
- Video quality can be estimated by Peak Signal-to-Noise Ratio (PSNR) or with other estimators, called perceptual visual quality metrics (PVQMs), which may include user's perception of video artifacts (Lin and Kuo 2011).
- The main parameter for file-download applications is the throughput (or the time required for downloading the data).

However, as far as online games are concerned, it is very difficult to define a single benchmark able to summarize the quality experienced by the user. The quality of an online game may depend on two groups of system parameters: the hardware where the game runs and the network. In some cases (i.e., game consoles) the hardware is fixed, and games are specifically designed for it, so the only variable parameters (leaving apart the display and the speakers) are those related to the network. As said in Ubicom (2005), a benchmark able to summarize the effect of the network for online games should be able to:

- Capture the relationship between the objective metrics and the user's satisfaction with the game.
- Accurately represent the effect of the network and the other players.
- It should also be reproducible.

These QoE models should consider the network impairments having a real influence on the player's experience with the game. For example, if the game has a very good method for concealing packet loss, this parameter may not be included, or included with a low influence.

These models are obtained using a cohort of volunteers who play the game while network impairments are added. After playing, they fill a survey about the experienced quality, and then regression techniques are employed to devise a mathematical model able to obtain a Mean Opinion Score, using as an input the value of each of the impairments. They usually consist of a polynomial formula, where each of the network impairments is weighted using a coefficient. Different statistical methods can then be used in order to estimate how well the model corresponds to the subjective quality results expressed by the users.

An example is the "impairment factor" proposed in Ubicom (2005):

$$\text{Impairment factor} = (W_L \times L + W_J \times J)\ (1 + E) \tag{1}$$

where W_L and W_J are the weighting factors for latency (L) and jitter (J), respectively, and E is the packet loss rate.

Other models obtain similar formulae for the impairment factor: in Wattimena et al. (2006), only delay and jitter are considered when obtaining the impairment factor, taking into account that the considered game (a First-Person Shooter) has a very good method for concealing packet loss.

$$\text{Impairment factor} = 0.104 \cdot \text{ping} + \text{jitter} \tag{2}$$

In Ries et al. (2008), something similar happens: packet loss is not considered, taking into account that a TCP-based game is being tested, so lost packets are retransmitted. The formula is in this case:

$$\text{MOS} = 5.17 - 0.012 \cdot \text{delay} - 0.018 \cdot \text{jitter} \tag{3}$$

However, the models have many limitations:

- One model is only valid for a single title. Some generic models have been proposed, but the weighting factors have to be tuned according to each game, and new subjective tests are required.
- The models involve a simplification. For example, it has been demonstrated that the influence of a network impairment is modified depending on the activity performed by the player (Suznjevic et al. 2013). In this study, in which real players were involved, the impact of packet loss was stronger for the players fighting with other real ones in *dungeons* than for the ones *questing* (this activity is not normally performed in group).
- Most models do not take context and user parameters into account.

All in all, it can be said that the subjective quality estimation for online games is a much more complex problem than in the case of Voice over IP. Although some similarities can be found, a universally accepted formula for measuring it does not exist. This is sometimes translated into a simplification of the problem, just talking about latency and neglecting the effect of other network impairments as jitter or packet loss.

QoE-Enhancing Mechanisms Employed in Game Support

Delay-Related Methods

Client-Side Prediction

As remarked in Bernier (2001) and Oliveira and Henderson (2003), different methods can be employed to combat delay in online games. The so-called client-side prediction has been largely used in First-Person Shooters. It can be divided into *input prediction* and *dead reckoning*, where input prediction hides the latency for the client-controlled actions, while dead reckoning hides the latency of other participating players.

For input prediction we can consider this sequence: the client generates a player command, which is transmitted to the server, where the next status of the game is calculated and sent to all the clients, and each of them renders the scene for its player. However, during this interval, something has to be shown to the player. The method consists of performing the movement of the client locally, just assuming that the server will accept the command. The drawback of the method is that, if the server's response does not fit with the client prediction, a perceptible shift in the position of the avatar will be appreciated by the player.

Dead reckoning is basically an algorithm for estimating the position of an entity in the virtual world based on its previous position, orientation, speed, acceleration, and other parameters. For dead reckoning, after receipt of the first state protocol data unit (PDU) for an entity (e.g., the character of another player), each client in the virtual world net begins moving the entity by applying the agreed-upon dead reckoning algorithm. Its movement is refreshed by receiving subsequent PDUs. If a spike of delay or packet loss appears for packets carrying PDUs related to some of the entities, each copy of the virtual world will continue to show movement of those entities according to agreed dead reckoning algorithm until the next update has been received. In addition, when an inconsistency between the server status and the one predicted by the client is found, some games are able to make the transition to the new status less abrupt, using smoothing algorithms.

Although client-side prediction may cause inconsistencies, it is able to increase the QoE, since it hides network latency to the user. Thus, it is widely employed in online games, and some titles allow the user to tune its parameters (Bernier 2001).

Server Side Delay Compensation

In addition to client prediction, the server may perform a compensation method in order to correctly merge virtual realities which are out of sync due to network delay. The server stores the history of the recent player positions (e.g., servers running *Valve's Source* engine store players' positions for 1 s), and when it has to calculate the new state, it first estimates the moment when the action was performed in the client version of the virtual world state. In other words, the server "rewinds time" according to the particular client's latency, calculating execution of a particular command (e.g., whether the player's shot has managed to hit the target). The formula is

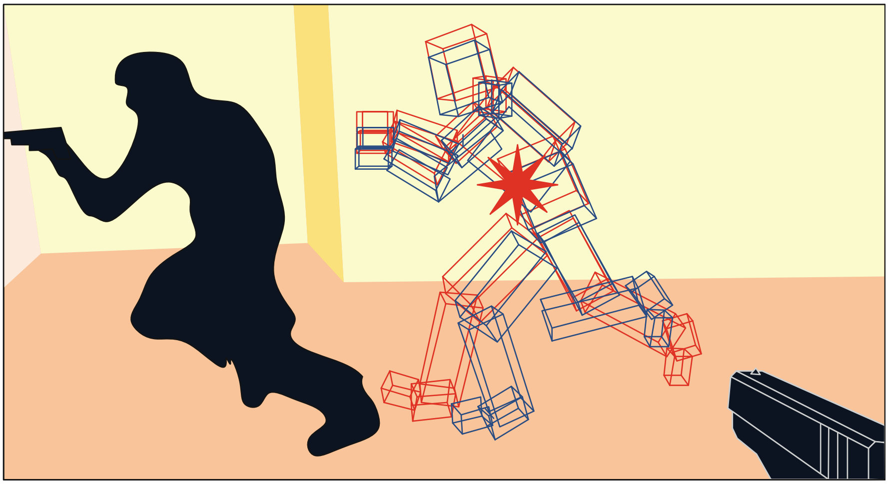


Fig. 9 Server side latency compensation

$$\text{Command Execution Time} = \text{Current Server Time} - \text{Packet Latency} - \text{Client View Interpolation}$$

It uses the execution time of the command, the latency of the player, and the movement prediction the client is using at that moment.

Figure 9 illustrates the case in which the client has 200 ms of latency (RTT). It shows the screenshot taken on the server right after the server confirmed the hit. The red hitbox shows the target position on the client where it was 100 ms + interplay period ago. While the player's command (i.e., shooting at the target) traveled to the server, the target continued to move to the left. After the user command arrived on the server, for the calculation whether or not that was a hit, the server restored the target position (blue hitbox) based on the estimated command execution time. The server traces the shot and confirms the hit.

This sometimes can again result in anomalies as illustrated in Fig. 10 (on behalf of simplicity, client prediction is not considered): Jack's avatar is hidden between two buildings when Wang appears running forward. Jack makes a good shot towards Position 1 just when Wang is there. This shot is transmitted to the server. Simultaneously, Wang's computer transmits its new position to the server. When Jack's shot arrives to the server, it "goes back in time" just the time estimated as Jack's latency, and then it calculates if the shot has hit anyone. Wang was at Position 1 at that moment, so the server concludes that Wang's avatar has been hit, and transmits this information to both players. As a result, Wang may perceive that he has been shot "around the corner."

This anomaly is further illustrated in Fig. 11. In this case player being shot at (Wang) would be receiving a hit, when in his version of the virtual world he is already behind the cover, and it is physically impossible to get hit due to the cover

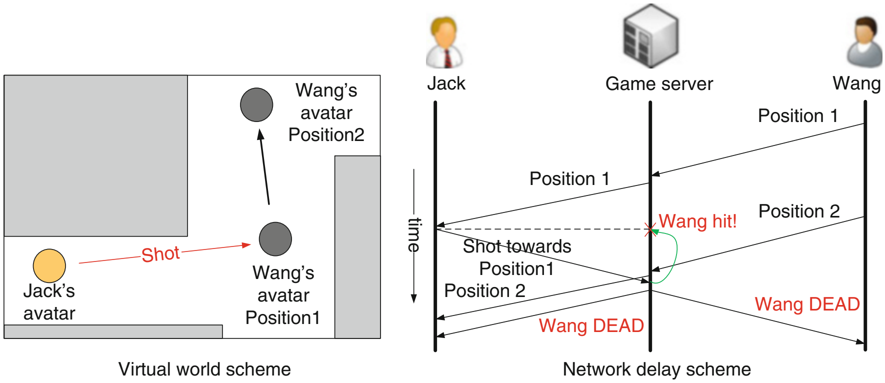


Fig. 10 Scheme of the server delay compensation mechanism

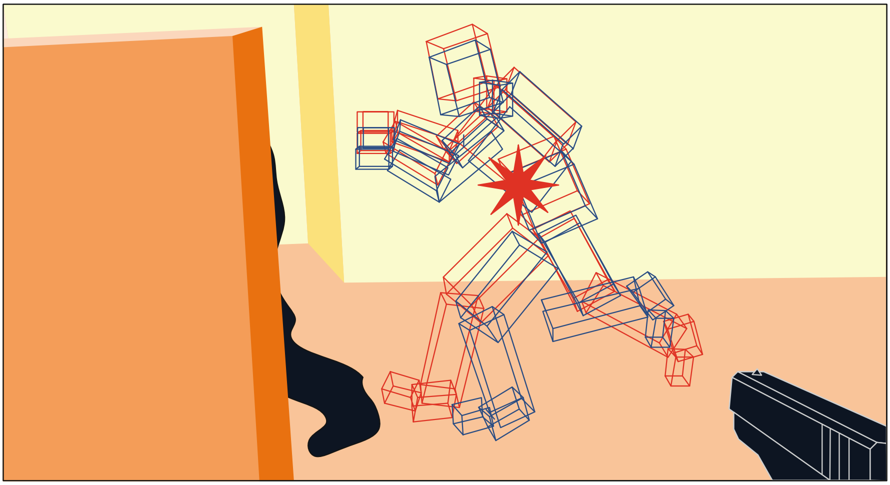


Fig. 11 Anomaly caused by sever side lag compensation

provided by the wall. Wang may feel a bit annoyed with this apparently unfair behavior of the game.

This problem could be solved if the client was able to send a "hit" message instead of a "shoot" one. However, servers cannot allow the clients to report hits, because of security reasons: some clients (or man-in-the-middle proxies) could then inject fake hits in order to increase the game performance of certain malicious players.

Geographical Server Distribution

A natural method for reducing latency is the optimization of server locations. The closer the server is to the player (physically), the transfer part of the latency is lower. Many games report the geographical place where their servers are, and they encourage the gamers to play in servers near their location. Sometimes this can be done automatically: the player enters in a queue, and an algorithm tries to match groups of players with similar skill levels and geographical locations, before starting a party between them. This happens in many First-Person Shooters, where a party can be short and many players do not need to play against other specific people.

In other cases, players are allowed to select the server where they want to play. They may be able to select a server in another zone, but in that case they are aware of the possibility of having a higher latency.

Scalability-Related Methods

In certain game genres (e.g., First-Person Shooters or Real-Time Strategy), every client receives from the server information of the status of every avatar in the virtual world. This is good for consistency, since, e.g., a map including the positions of the rest of the players (or only those in the same team) can be displayed. However, in this case, the amount of information to be sent by the server grows with the square of the number of players. Some studies have modeled the traffic depending on the number of players (Branch et al. 2008; Lang et al. 2005).

This scalability problem sets a limit on the number of players that can share a virtual scenario in these genres, and this problem persists: for example, *Counter-Strike 1* (1999) was able to manage 32 players, and, 21 years later, *World of Tanks*' (2010) normal parties involve 30 people.

However, one of the main attractive points of other genres (MMORPG) is its "massiveness." The fact of sharing a virtual world with a huge number of people is especially interesting for the players as it gives them the feeling of participation in a large virtual world with its economic and social specifics. Players may join a guild, participate in missions with other people, have a list of virtual friends, show the new weapons of their fancy avatars to lots of real people, etc. In fact, many of these games include thousands of people in the same virtual world. Thus, game developers implement different mechanisms in order to deal with this hard scalability problem.

Area of Interest

If all of the players in a massive virtual world received real-time information regarding all of the other players, that would impose impossible requirements both on the network bandwidth and on the processing capabilities of the server. Therefore, an Area of Interest (AOI) can be defined around the avatar, also including its field of view. Then, only the information generated by other avatars in player's AOI is sent to him/her. This allows the game developer to maintain consistency between all the players and at the same time to keep the processing load and the traffic to be sent to the player in reasonable limits. This results in a statistical model of the traffic which depends on the number of players in the server: this effect was studied in Suznjevic et al. (2014), where a model for an MMORPG (*World of Warcraft*) with a different statistical behavior depending on the number of players in a server was presented.

Sharding

A virtual world can be distributed across game servers in two ways: (a) one logical instance of the virtual world is created for all players and is spanned across all the machines of the game server farm (e.g., *EvE Online*); (b) the virtual world is replicated on more than one logical instance, called *shard*. Shards partition the player base across several logical instances of the virtual world (players in different shards cannot interact with each other), and through that the computational load is reduced. It should be noted that in recent times the limits of the shards in the most popular MMORPGs have been weakened. For example, in *World of Warcraft* (*WoW*) it is possible to communicate with players in other shards, or even other games, using Blizzard's overlay communication system *battle.net*. Also, in *WoW* since recently it is possible for certain zones to be shared across different shards, so players from different shards are now able to play together.

Zoning

Scalability on a single shard is achieved through a parallelization technique called *zoning*. This technique partitions the virtual world into geographical areas called *zones*, which can be handled independently by separate machines, as depicted in Fig. 12. In the past, zones had borders (e.g., invisible walls) with transition spots between them, such as portals, because transition between zones required certain time. These events had a bad influence on the players' immersion in the virtual world, which led to the development of the technique called *seamless zoning*, in which zones might be divided by some geographic markers (e.g., mountains), but there are no loading screens when players cross from one zone to another.

Mirroring

Another technique, called *mirroring*, targets parallelization of game sessions with a large density of players located and interacting within each other's AOI. "Hot spots" appear in virtual worlds, and they typically include major cities or zones in which gathering of players is common. To address this problem, mirroring is performed by distributing the load by replicating the same game zone on several servers. In each replicated server, the state for a subset of entities (i.e., active entities) is calculated, while the remaining entity (i.e., shadowed entities) states are calculated in the other participating servers and are synchronized across servers, as shown in Fig. 12. The overhead of synchronizing shadowed entities is much lower than the overhead of computing all entities as active entities (Muller et al. 2006).

Instancing

Instancing is a technique which can be perceived as a simplification of mirroring, or even sharding on a smaller scale. This technique distributes the session load onto

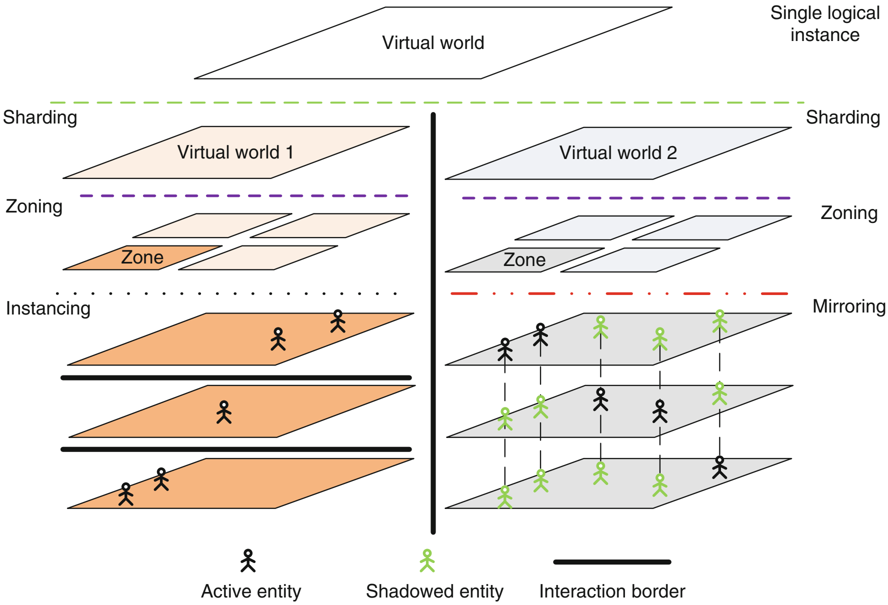


Fig. 12 Scalability techniques for online games

multiple parallel instances for the highly populated zones. The instances are independent of each other, which means that two avatars from different instances will not be able to interact with each other, even if they are located at coordinates within their AOI. This technique is common in many MMORPGs. All scalability techniques, including instancing, are depicted in Fig. 12.

Conclusion

Companies employ different techniques to provide games in a scalable, reliable, and profitable way, with the main objective of achieving a good Quality of Experience level. This concept not only includes Quality of Service, which is directly related to the parameters that can be measured from the network, but also captures the expectations, feelings, perceptions, cognition, and satisfaction of the user. There is a special difficulty with this kind of service, since gamers show a very demanding profile. This chapter has discussed the different mechanisms that companies use in order to overcome the problem of network impairments (latency, packet loss, etc.) when providing online games.

The synchronization and the maintaining of a coherent game state to be shared by the applications of all the players is not a trivial problem: different sources of latency appear and may cause inconsistencies between the game states observed by each of the players. Different genres of online games present specific latency requirements, depending on the game dynamics, their characteristics, and the level of interaction between the players. First-Person Shooters are considered as the game genre with the tightest latency limits, but other genres also have interactivity requirements, so a “delay sensitivity” can be defined for each of them.

Online games have employed different architectures, being client–server the one employed in the vast majority of the cases. It has many advantages, such as the existence of a central authority, cheating prevention, an easier implementation of billing functionalities, etc.

Online games (except those known as “cloud games”) generate low-bandwidth traffic flows, in order to allow a high packet rate and constant updates of the client movements and the game status. Different protocols (UDP or TCP) are employed at transport level, and this is somewhat related with the game genre and its interactivity requirements.

The different techniques used for estimating the user’s Quality of Experience from network parameters always consider latency as the most important parameter. They provide a weighting factor measuring the importance of each network impairment, such as delay, packet loss and delay variation (jitter). Some factors may increase or reduce the effect of each parameter: for example, in TCP-based games, packet loss is not a problem by itself, since lost packets are retransmitted, but it is translated into an additional latency. Additionally, some First-Person Shooter games are very robust against packet loss, but have more sensitivity to jitter.

Finally, different QoE-enhancing mechanisms are used by game providers, such as client-side prediction, which try to hide latency not only of the local client but also of the other participating players. Another mechanism is server delay compensation. When employed, the server goes back in time before calculating the results of a player action, taking into account his/her associated latency. Other scalability-related techniques are the use of an Area of Interest, which means that only the actions of the avatars near the user are sent to his/her client. Different methods are used in order to distribute the workload between different game servers, such as creating multiple instances (*sharding*) of the virtual world, dividing it into zones, or parallelizing game sessions corresponding to densely populated zones.

Recommended Reading

R. Adams, Active queue management: a survey. IEEE Commun. Surv. Tutorials **15**(3), 1425–1476 (2013). doi:10.1109/SURV.2012.082212.00018, Third Quarter 2013

S.H. Batool, K. Mahmood, Entertainment, communication or academic use? A survey of Internet cafe users in Lahore, Pakistan. Inf. Dev. **26**, 141–147 (2010). doi:10.1177/0266666910366650

Y. Bernier, Latency compensating methods in client/server in-game protocol design and optimization, in *Proceedings of the Game Developers Conference*, vol. 98033, no. 425 (San Jose, 2001)

P.A. Branch, A.L. Cricenti, G.J. Armitage, An ARMA(1,1) prediction model of first person shooter game traffic, in *Proceedings IEEE 10th Workshop on Multimedia Signal Processing (MMSP’08)*, (Cairns, 2008), pp. 736–741. doi:10.1109/MMSP.2008.4665172

O. Brun, C. Bockstal, J. Garcia, A simple formula for end-to-end jitter estimation in packet-switching networks, in *Networking, International Conference on Systems and International*

Conference on Mobile Communications and Learning Technologies (2006), Morne, Mauritius p. 14. doi:10.1109/ICNICONSMCL.2006.34

M. Cajada, VFC-RTS: vector-field consistency para real-time-strategy multiplayer games. Master of Science Dissertation (2012), Available online: http://www.gsd.inesc-id.pt/~lveiga/prosopon/pubs/35-msc-cajada.pdf

C. Chambers, W.C. Feng, S. Sahu, D. Saha, Measurement-based characterization of a collection of on-line games, in *Proceedings of the 5th ACM SIGCOMM Conference on Internet Measurement,* Berkeley, CA (USENIX Association, 2005), p. 1-1

K. Chen, P. Huang, L. Chin-Luang, How sensitive are online gamers to network quality? Commun. ACM **49**(11), 34–38 (2006a). doi:10.1145/1167838.1167859

K.T. Chen, C.Y. Huang, P. Huang, C.L. Lei, An empirical evaluation of TCP performance in online games, in *Proceedings of the ACM SIGCHI international conference on advances in computer entertainment technology (ACE'06)*, (Hollywood, CA, USA 2006b). doi:10.1145/1178823.1178830

J. Chen, Flow in games (and everything else). Commun. ACM **50**(4), 31–34 (2007). doi:10.1145/1232743.1232769

M. Claypool, The effect of latency on user performance in real-time strategy games. Comput. Netw. **49**(1), 52–70 (2005). doi:10.1016/j.comnet.2005.04.008

M. Claypool, K. Claypool, Latency and player actions in online games. Commun. ACM **49**(11), 40–45 (2006). doi:10.1145/1167838.1167860

M. Claypool, D. LaPoint, J. Winslow, Network analysis of counter-strike and Starcraft, in *Proceedings of the 22nd IEEE International Performance, Computing, and Communications Conference (IPCCC)*, (Phoenix, Arizona, US 2003), pp. 261–268. doi:10.1109/PCCC.2003.1203707

M. Claypool, D. Finkel, A. Grant, M. Solano, Thin to win? Network performance analysis of the OnLive thin client game system, in *IEEE 11th Annual Workshop on Network and Systems Support for Games (NetGames)*, Venice (2012), pp. 1–6

R.G. Cole, J.H. Rosenbluth, Voice over IP performance monitoring. SIGCOMM Comput. Commun. Rev. **31**(2), 9–24 (2001). doi:10.1145/505666.505669

T. Debeauvais, B. Nardi, A qualitative study of Ragnarök online private servers: in-game sociological issues, in *Proceedings of the Fifth International Conference on the Foundations of Digital Games*, Monterey, California, USA (ACM, 2010). doi:10.1145/1822348.1822355

E.L. Deci, R.M. Ryan, *Intrinsic motivation and self-determination in human behavior (Perspectives in social psychology)* (Plenum Press, New York/London, 1985)

M. Dick, O. Wellnitz, L. Wolf, Analysis of factors affecting players' performance and perception in multiplayer games, in *Proceedings of 4th ACM SIGCOMM Workshop on Network and System Support for Games*, (Hawthorne, NY, USA 2005), pp. 1–7. doi:10.1145/1103599.1103624

W.C. Feng, F. Chang, W. Feng, J. Walpole, A traffic characterization of popular on-line games. IEEE/ACM Trans. Networking **13**(3), 488–500 (2005). doi:10.1109/TNET.2005.850221

M. Ford, Workshop report: reducing internet latency, 2013. ACM SIGCOMM Comput. Commun. Rev. **44**(2), 80–86 (2014). doi:10.1145/2602204.2602218

B. Furuholt, S. Kristiansen, F. Wahid, Gaming or gaining? Comparing the use of internet cafes in Indonesia and Tanzania. Int. Inf. Libr. Rev. **40**(2), 129–139 (2008). doi:10.1016/j.iilr.2008.02.001

L. Gautier, C. Diot, Design and evaluation of mimaze a multi-player game on the internet, in *Proceedings IEEE international conference on multimedia computing and systems, Austin, TX 1998*, (1998). doi:10.1109/MMCS.1998.693647

J. Gettys, K. Nichols, Bufferbloat: dark duffers in the internet. Queue **9**(11), 40 (2011). doi:10.1145/2063166.2071893. 15 p

M. Gurol, T. Sevindik, Profile of internet cafe users in Turkey. Telematics Inform. **24**(1), 59–68 (2007). doi:10.1016/j.tele.2005.12.004

T. Henderson, S. Bhatti, Networked games: a QoS-sensitive application for QoS-insensitive users?, in *Proceedings of the ACM SIGCOMM workshop on Revisiting IP QoS: What have* *we learned, why do we care?*, (Bolton Landing, NY, USA 2003), pp. 141–147. doi:10.1145/944592.944601

C.Y. Huang, C.H. Hsu, Y.C. Chang, K.T. Chen, GamingAnywhere: an open cloud gaming system, in *Proceedings of the 4th ACM Multimedia Systems Conference (MMSys '13)*, (ACM, New York, 2013), pp. 36–47. doi:10.1145/2483977.2483981

A. Kaiser, D. Maggiorini, K. Boussetta, N Achir, On the objective evaluation of real-time networked games, in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM 2009)*, (Honolulu, HI 2009). doi:10.1109/GLOCOM.2009.5426032

T. Lang, P. Branch, G. Armitage, A synthetic traffic model for Quake3, in *Proceedings of the ACM SIGCHI International Conference on Advances in Computer Entertainment Technology (ACE '04)*, (Singapore, 2005), pp. 233–238. doi:10.1145/1067343.1067373

C.S. Lee, The revolution of StarCraft network traffic, in *Proceedings of the 11th Annual Workshop on Network and Systems Support for Games NetGames 2012*, (2012), Available online: http://dl.acm.org/citation.cfm?id=2501583

W. Lin, C.C. Jay Kuo, Perceptual visual quality metrics: a survey. J. Vis. Commun. Image Represent. **22**(4), 297–312 (2011). doi:10.1016/j.jvcir.2011.01.005

J.L. Miller, J. Crowcroft, The near-term feasibility of P2P MMOG's, *IEEE 9th Annual Workshop on Network and Systems Support for Games (NetGames)*, (Taipei, Taiwan 2010), Article 5, 6 p

J. Muller, H. Metzen, A. Ploss, M. Schellmann, S. Gorlatch, Rokkatan: scaling an RTS game design to the massively multiplayer realm. Comput. Entertain. **4**(3) (2006), Article 11. doi:10.1145/1146816.1146833

M. Oliveira, T. Henderson, What online gamers really think of the Internet?, in *Proceedings of the 2nd Workshop on Network and System Support for Games (NetGames '03)*, (ACM, New York, 2003), pp. 185–193. doi:10.1145/963900.963918

Qualinet White Paper on Definitions of Quality of Experience (2013), European Network on Quality of Experience in Multimedia Systems and Services (COST Action IC 1003), ed. by P. Le Callet, S. Möller, A. Perkis (Lausanne, Switzerland, Version 1.2, Mar 2013)

S. Ratti, B. Hariri, S. Shirmohammadi, A survey of first-person shooter gaming traffic on the internet. IEEE Internet Comput **14**(5), 60–69 (2010). doi:10.1109/MIC.2010.57

K. Rehman-Laghari, N. Crespi, B. Molina, C.E. Palau, QoE aware service delivery in distributed environment. Advanced Information Networking and Applications (WAINA), in *2011 I.E. Workshops of International Conference on*, (Biopolis, Singapore 2011), pp. 837–842, 22–25. doi:10.1109/WAINA.2011.58

M. Ries, P. Svoboda, M. Rupp, Empirical study of subjective quality for massive multiplayer games, in *Proceedings of the 15th International Conference on Systems, Signals and Image Processing*, (Bratislava, Slovakia 2008), pp. 181–184. doi:10.1109/IWSSIP.2008.4604397

J. Saldana, J. Fernandez-Navajas, J. Ruiz-Mas, D. Wing, M.A.M. Perumal, M. Ramalho, G. Camarillo, F. Pascual, D.R. Lopez, M. Nunez, D. Florez, J.A. Castell, T. de Cola, M. Berioli, Emerging real-time services: optimizing traffic by smart cooperation in the network. IEEE Commun. Mag. **51**(11), 127–136 (2013). doi:10.1109/MCOM.2013.6658664

M. Suznjevic, M. Matijasevic, Player behavior and traffic characterization for MMORPGs: a survey. Multimedia Systems **19**(3), 199–220 (2012). doi:10.1007/s00530-012-0270-4

M. Suznjevic, O. Dobrijevic, M. Matijasevic, MMORPG player actions: network performance, session patterns and latency requirements analysis. Multimedia Tools Appl **45**(1–3), 191–214 (2009). doi:10.1007/s11042-009-0300-1

M. Suznjevic, L. Skorin-Kapov, M. Matijasevic, The impact of user, system, and context factors on gaming QoE: a case study involving MMORPGs, in *Network and Systems Support for Games (NetGames), 12th Annual Workshop on. IEEE*, (Denver, CO, USA 2013). doi:10.1109/NetGames.2013.6820606

M. Suznjevic, J. Saldana, M. Matijasevic, J. Fernandez-Navajas, J Ruiz-Mas, Analyzing the effect of TCP and server population on massively multiplayer games. Int. J. Comput. Games Technol. (2014), Article ID 602403, 17 p. doi:10.1155/2014/602403

P. Svoboda, W. Karner, M. Rupp, Traffic analysis and modeling for world of warcraft. ICC '07, in *IEEE International Conference on Communications*, (Glasgow, UK 2007), pp. 1612–1617. doi:10.1109/ICC.2007.270

P.Y. Tarng, K.T. Chen, P. Huang, An analysis of WoW players' game hours, in *Proceedings of the 7th ACM SIGCOMM Workshop on Network and System Support for Games*, (Worcester, MA, USA ACM, 2008). doi:10.1145/1517494.1517504

B. Thompson, T. Koren, D. Wing D, IETF RFC 4170: tunneling multiplexed compressed RTP (TCRTP) (2005)

Ubicom White Paper: OPScore, or online playability score: a metric for playability of online games with network impairments (2005), Available online at http://www.kevingee.biz/wp-content/uploads/2011/04/IP3K-DWP-OPSCORE-10.pdf

A.F. Wattimena, R.E. Kooij, J.M. van Vugt, O.K. Ahmed, Predicting the perceived quality of a first person shooter: the Quake IV G-model, in *Proceedings of 5th ACM SIGCOMM workshop on Network and system support for games (NetGames '06)*, (ACM, New York, 2006), Article 42. doi:10.1145/1230040.1230052

C.C. Wu, K.T. Chen, C.M. Chen, P. Huang, C.L. Lei, On the challenge and design of transport protocols for MMORPGs. Multimedia Tools Appl **45**(1), 7–32 (2009). doi:10.1007/s11042-009-0297-5

S. Zander, G. Armitage, Empirically measuring the QoS sensitivity of interactive online game players, in *Proceedings of the Australian Telecommunications Networks & Applications Conference (ATNAC 2004)*, (Sydney, Dec 2004)

Video Gaming on Ad Hoc Networks: Challenges and Solutions

24

Nadjib Achir and Khaled Boussetta

Contents

Abstract

This chapter is devoted to video game support over wireless ad hoc networks. The main issue is to understand the real correlation between network conditions and players expected performance. Thus, we started this chapter by presenting an objective assessment methodology. The experimental results show clearly that from the player perspective the most important metric is the delay fairness and from the network perspective, the most important metric is the energy consumption. Considering that, we focus on the improvement of the gameplay fairness and energy efficiency, and we present two approaches. One first approach at the application level according to a new packetization technique and the second approach according to a multi-metric routing protocol based on energy consumption and end-to-end delay. Finally, we concentrate on the problem of game disconnection due to nodes mobility. We present a multipath

N. Achir (✉)
L2TI – Institut Galilée, University Paris 13, Sorbone Paris Cité, Villetaneuse, France
e-mail: nadjib.achir@univ-paris13.fr

K. Boussetta
Urbanet, CITI Insa Lyon / INRIA Grenoble Rhône-Alpes, CITI lab, Villeurbanne, France
e-mail: khaled.boussetta@univ-paris13.fr

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_36

OLSR-based routing protocol that computes two paths to reach the destination. In addition, the traffic generated by each node is balanced between the two paths according the quality of each path.

Keywords
Video games • First Pearson Shooters (FPS) • Lag techniques • Mobile ad hoc network (MANET) • Mobility management • Routing protocol • Multipath routing protocol • Multi-metric routing protocol

Introduction

Over the past 15 years, the video games market has grown at an exceptional rate, and this progression has been sustained by the innovation and the creativity that play such a vital role in this industry. Indeed, the continuous technological advances, in terms of both software and hardware, have offered game designers new opportunities to create innovative concepts for playing. In particular, the increasing processing capacities of graphics cards, the miniaturization and the portability of devices, the sophistication of games engines, with enhanced AI (Artificial Intelligence), 3D rendering and physics modules, have all contributed to making playing video games a captivating experience. A key element among the basic ingredients in recent successful games is interactivity supported by network communications, and most of today's popular games integrate a multiplayer mode.

Up to now, the multiplayer mode has generally been supported by wired communication infrastructures. Wireless links, based on cellular technology (e.g., UMTS or WiFi), can also be used but only as a last link in the transmission path. This is typically the case with laptops, PDAs or recent consoles such as Nintendo Wii, Nintendo DS/3DS, Sony's Portable PlayStation (PSP)/PS Vita, and Microsoft Xbox/Xbox One, all of which could either offer a built-in WiFi interface or be equipped with a USB WiFi connector. Thus, in the last generation of consoles, LAN multiplayer mode based on WiFi is now available. However, only one single hop is supported. In other words, in order to play together, all players must be located within the same radio coverage. Multi-hop radio communications using a MANET (Mobile Ad hoc NETwork) (IETF MANET WG) technology is not yet supported by the current generation of consoles. MANETs are an ultimate generalization of wireless networks. Indeed, ad hoc network technology allows the spontaneous creation of data networks by exploiting the capacity of mobile terminals to communicate directly between each other without the need for a centralized infrastructure. Terminals forming the ad hoc network should play the role of routers in order to ensure the gradual transfer of data packets. Moreover, the network management tasks (auto-configuration, security, etc.) are carried out by the terminals themselves, in an entirely distributed way.

Taking into consideration the popularity of LAN multiplayer games and the tremendous success of wireless networks, we claim that future game engines will benefit from the ability to support the MANET mode. In addition to the social and cultural interest of playing with geographically close persons, MANET technology could offer the comfort of mobility to a much higher number of potential players than is currently supported by the latest generation of commercialized consoles. Other advantages of MANET technology for LAN multiparty gaming are facility of deployment, autonomic reconfiguration, self-healing, and the possibility of increasing capacity without preliminary planning. All these elements make us firmly believe in the suitability of integrating of MANET technology in future game terminals. Various scenarios will be possible, like group made up of varying number of kids equipped with portable consoles supporting MANET services and wishing to play a WLAN party in a playground (without the need for any particular fixed network infrastructure).

In order to make such scenario a common reality in the near future, several issues still have to be addressed by the research community. A number of difficulties come from the inherent constraints that characterize the MANET technology. In particular, the mobility of the core networks, energy limitations, and variations in terminal resources sharing. All these problems lead to variable connectivity and a potentially high fluctuation in the Quality of Service provided, as multiparty games are very sensitive to resource availability and fluctuation. For example, *Game play* could suffer from significant degradation in the case of variable delays, which is typically the case in MANETs. Hence, certain technological issues remain to be solved in order to achieve a real deployment of such MANETs in portable game terminals.

In this chapter, we focus on the video game support over wireless networks. First of all a brief outline of engines architecture for FPS games is presented. Thereafter, focus on understanding the real impact of wireless network parameters (delay, jitter, and losses) on the game play. Afterward, we concentrate on how to improve the gameplay fairness and energy efficiency for multiplayers games in MANET. Then, we address the disconnections problem due to nodes' mobility. Finally, we conclude the chapter.

Video Game Engines Architecture

A game engine is software that aims to provide all the functionalities needed to play video games using a hardware terminal (e.g., a console or a computer). It includes several components, the most important ones are:

- The *2D/3D graphic* render, which is responsible for all visual game rendering such as avatars and background details. The majority of state-of-the-art graphics engines adopt an object-based representation of the virtual game environment. This representation provides simplified game design and can be used efficiently for large virtual worlds.

- The *physics simulation* engine that takes in charge the visual coherence of avatars' and objects' movements within the game. Its basic functionality is to simulate the effects of physical behavior of objects, like collisions.
- The *Artificial Intelligence* (AI) module aims to provide rich and realistic interactions between players and environment. More precisely, it produces the illusion of natural/intelligent behavior of Nonplaying Characters (NPCs) by simulating actions and movements as a human-controlled avatar would perform.
- The *network* module is of particular importance when the multiplayer mode is proposed. Basically, the networking subsystem is in charge of transmitting game's data between all entities involved in the game.

Nowadays, most online multiplayer games use centralized client–server architecture, in which, one part of the game software is located in each client and the other part in a central server, which acts as a coordinator. With this architecture, the gameplay is managed by having continuous communication between each user's client and the server. The clients are responsible for collecting and sending to the server all the primitive commands related to the players' actions within the game, such as avatar control commands (i.e., movements and shooting). For its part, the server is responsible for computing the game actions received from each client, in addition to other global information, such as gravity and battlefield map, to compute the new overall game state. The server transmits this new game state back to all the clients. Because of the centralized nature of the client–server architecture, the server can easily maintain a consistent game state.

Beside the obvious problem of scalability, the most important requirement necessary to achieve good performances is to maintain a low round-trip delay between clients and the game server. If the sum of round-trip delay over the network and computational time on the server is greater than the state update interval for a client, the player will experience poor performance. When subject to excessive delay, the player's vision of the virtual world will not be in synch with the server and actions will be misplaced in the next state computation.

Video Games over Ad Hoc Networks

Objective Evaluation of Real-Time Networked Games

In order to achieve a better understating of the real correlation between network conditions and players' expected performance, we start this chapter by an objective assessment methodology for video games when experiencing important end-to-end delay. In the literature, either simulations or real tests are used to evaluate the real correlation between the network conditions and the players' expected performances. Several papers have used network simulators, like NS2, to assess the performance of network mechanisms supporting games traffic. Several models for game traffic have been studied and proposed (Cricenti et al. 2007; Feng et al. 2002; Branch et al. 2005). However, with simulations, only network performance metrics could be evaluated. Gameplay metrics, like the *score* or *players*' *ranking*, could not be considered. Simulations are not able to capture the inference between the players' behavior when network conditions vary.

On the other hand, some of the existing contributions considered real game traces. In Armitage et al. (2003), and later in Armitage et al. (2004), the authors considered a real Quake III online game server and drew some conclusions about delay and jitter based on user's behavior. In the former work (Armitage et al. 2003), only delay information is considered; the general conclusion is that the lower the observed delay the longer a player will stay connected, at the same time the number of frags – i.e., temporary kills, using a military slang – does not seems to be very significant. In the latter contribution (Armitage et al. 2004), instead, jitter is also considered; authors come to the conclusion that for the game to be considered playable by users jitter should be limited to one fifth of the delay. Moreover, from the reports, there is also evidence that frag rate is related to jitter level. The general conclusion is that the lower the observed delay is the longer a player will stay connected to the game server.

Unfortunately, due to the extremely dynamic activity of online games and the short life of gaming sessions, it is generally difficult to get accurate measurements. Indeed, with the aforementioned methodology it is hard to understand bounds for delay and jitter because there is no reliable way to get feedback from the players on their reasons for leaving the game. For this reason, other experiments have been performed using a smaller population of players. For this reason in other cases, experiments have been performed using a smaller population and later questioning the players.

In Quax et al. (2004) and Beigbeder et al. (2003) experiments have been carried out using a lab testbed with controlled network conditions and using real players. In both cases, the analysis was performed using Unreal Tournament 2003 and network delay has been artificially simulated. In Quax et al. (2004) quality of gaming is assessed by a survey submitted from the players. From this survey, authors try to figure out if a player understands when he/she has been penalized or not. Results pointed out that a delay greater than 60 ms is disturbing for the user and jitter does not play a prominent role. A comparison between the subjective survey results and an analysis of game logs leaded authors to divide the players into two groups: "complainers" and "optimists," which perceive the game quality differently and provide the general bias of this approach. In Beigbeder et al. (2003), instead, players have been able to notice latencies as low as 75 ms and found gameplay less enjoyable at latencies over 100 ms. However, the delay did not seem to affect the game scores in the tests. From analysis of the game logs authors suggested to keep the delay below 150 ms to have a playable match.

Another contribution in Dick et al. (2005) focuses mainly on perceptual quality. A number of different games have been proposed to real players and feedbacks collected by means of a survey. Authors come to the consideration that players tend to underestimate their maximum tolerance to network latency, being a little too picky. Moreover, delay and jitter affect differently the proposed games; this is supposedly due to specific lag compensation algorithms; Unreal Tournament 2004 proved to have the strongest correlation between subjective perception and the objective outcome of the game. In Wattimena et al. (2006), Quake 4 is used to develop an end-to-end quality measurement method which has been validated against subjective experiments. By applying this method, based on delay and jitter values, authors report that the mean opinion score shows a very high correlation with subjective data.

The main weakness of the above studies is related to the fact that none of them are reproducible. Indeed, video games are not deterministic. Many events occur randomly (the spawn position of a player in the game, the availability of an item, etc.) and affect the players' way of playing. In addition, to obtain exploitable and smooth results, we need to run exhaustive experimentations (over several hours), and it is clearly not reasonable to expect people to play a game for hours on end. Moreover, real players tend to adapt to network conditions. Therefore, the results obtained would not be exploitable because they would be biased by the players' skills.

Video Game Emulator Testbed

In order to be able to reproduce experiments in a controllable network environment, and also to run exhaustive gaming sessions, we decided in Kaiser et al. (2009a) to develop a *Video Game Emulator Testbed*. This platform is composed of a number of clients, one server, and one scheduler. To run completely objective experiments, all the clients need to have the same hardware and software configurations (i.e., operating system, processor, memory, graphic, and network cards). In order to emulate network delay and jitter, we use the Linux traffic control (TC) subsystem.

For the video game, we adopted Quake III. Quake III is a well-known online multiplayer *First Pearson Shooters* (FPS) which is, nowadays, still largely played over the Internet. This game has been released to the open source as IOQuake III (http://ioquake3.org/ 2005). To be able to reproduce the experiments, we used autonomous bots as acting players. The kind of bot we adopted is a client-side one: unlike the classical bots provided by the Quake III engine, it is run on a client and mimic a real-user behavior. Parameters of the bots have been set in a way to behave as an "*average good player*" and we applied the same bot configuration on all clients. Moreover, in all the experiments, the same map has been used and the game mode has been set to "*death match*"; which means, the first player who manages to reach a certain score, by killing other players, wins the match. In each experiment, 100 consecutive matches have been played. The duration of a match varies usually between 8 and 9 min; which means approximately 15 h for each experiment. Finally, to avoid learning effects in the bots' behavior, we decided to kill all client processes and start them again on every match. This has been done because we observed an unfair behavior in score distribution when performing very long experiments. Results after adopting this strategy showed to be fairer, even on the long run.

The automation of the testbed is done using scripts on a dedicated machine playing the role of a testbed scheduler. We use a configuration file on the scheduler to define the parameters of the experimentations. Precisely, client agents receive

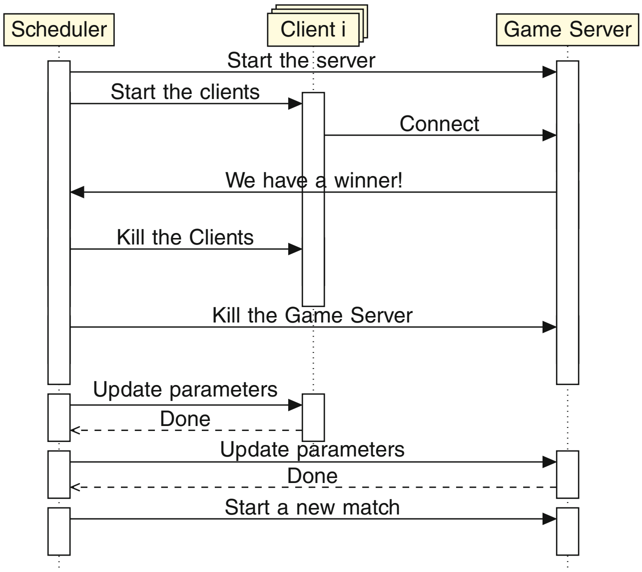
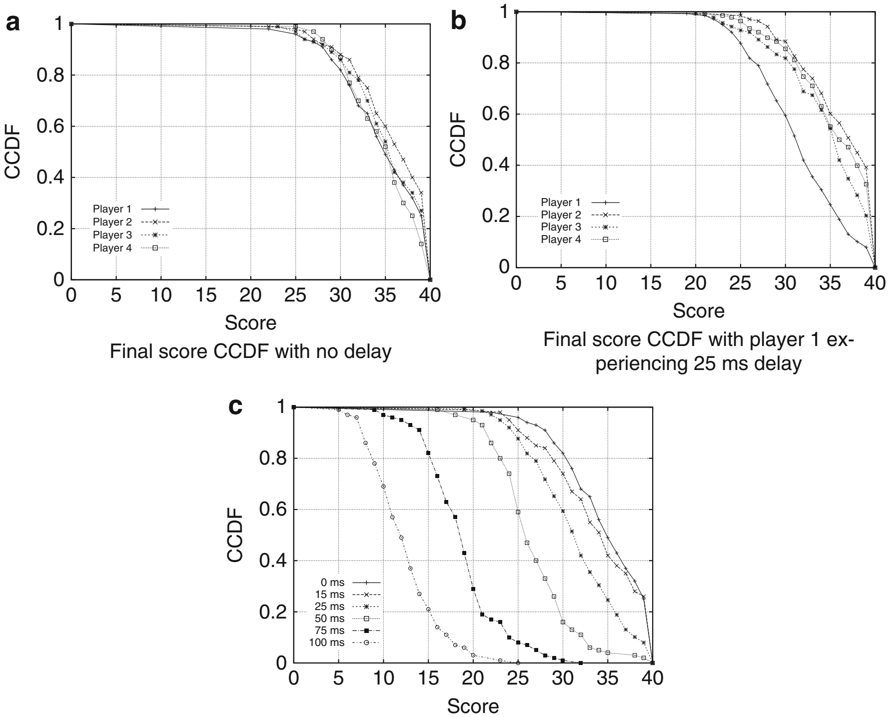


Fig. 1 Video game emulator testbed: messages diagram

messages from the scheduler and the server and interact with the Quake client to obtain match synchronization and perform software restart, as illustrated in Fig. 1. In addition to the end-to-end delay impact we also considered the impact of jitter and packet losses. Experiments have been run with penalizing a single player. Delays have been changed between 0, 25, 50, 75, and 100 ms while jitter values have been 0, 10, and 20 ms. Regarding packet loss, a uniform loss probability of 0, 10, 20, and 30 % have been applied.

Score at a match end has always been considered a good metric to understand if a player is well placed in the arena or, more in general, if he/she is playing in a satisfactory way. Figure 2a outlines the score CCDF when no player is subject to any delay; as can it be observed, the match is fair because there is an extreme high probability for each player to score at least 25 but there is small evidence of an effective balance of skills in the range from 30 to 40 due to visible oscillations. These oscillations are due to the fact that there is only one winner for a match – the one reaching score 40; the other players are left behind by one or two points, tagging them as *losers* even if the quality of their match has been the same as the winner's one.

Penalizing one of the players brings the obvious result as presented in Fig. 2b. Despite the fact that commercial FPSs claim to be playable with up to 150 ms delay on the round-trip, we can clearly observe a performance degradation for one player with just 25 ms (50 ms round-trip delay) In a real-life environment, this situation could be mistakenly classified as "fair" based on the fact that the probability for a player to score at least 25 is still very high despite his/her changes to win have actually decreased. An interesting consideration comes from the observation that we are not actually able to detect small changes in behavior: a small penalization



a Final score CCDF with no delay

b Final score CCDF with player 1 experiencing 25 ms delay

c Score performance degradation based on player's delay

Fig. 2 Score

may get confused with the aforementioned oscillations and ignored. This is not really a problem in the current scenario but it might be in other games, depending on the performed activities.

Performance degradation based on delay seems to follow an exponential behavior, as depicted in Fig. 2c. Starting from 50 ms delay – which is close to the threshold where an FPS starts to be less playable – we can actually see a significative decrease for the player to reach score 25; hence, a human player will actually start to complain.

During experiment we found out that rate of actions – or rate of frags, in this specific case – can also be used very efficiently to estimate a user's satisfaction during online gaming. In some cases, as we are going to show, this metric gave also a better understanding of each player's status. In Fig. 3a–c we plot the CCDF of inter-time between frags. In this case, we can observe a much steadier behavior than before making it possible to detect small penalties. Moreover, frags rate can be computed in real time: it is no longer mandatory to wait for the game to end, and we can detect a player not having fun in the middle of the match.

The experimental results from this testbed show the important sensitivity of the gameplay (especially regarding the fairness among players) of the delays over the

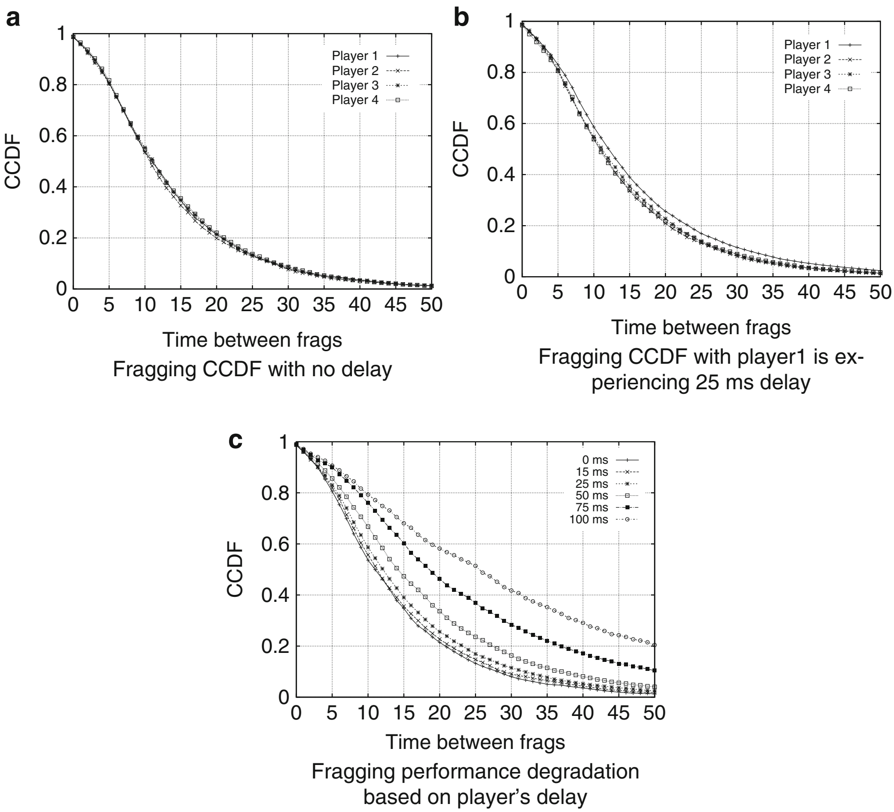


Fig. 3 Time between frags

jitter or packet losses. These experiments have shown that the game may be sensibly unfair for a player experiencing just 50 ms delay more than other players. These results will be useful for a game provider to perform preemptive traffic engineering, evaluation of network infrastructures, and true-skill matching between players taking network handicap into account.

After this preliminary evaluation, we evolved the platform in order to support "*Mobile* Ad Hoc *Network*" (MANET). The main objective is to evaluate the effects of MANET routing on time sensitive networked video gaming experience. In the first version of the emulator, all nodes are connected to each other using their Ethernet interface to a dedicated LAN switch. To emulate a multi-hop network we use the Linux *iptables* tool to filter incoming packets. Basically, each node drops all packets, based on the source MAC address, which come from nodes that are not in its neighborhood. More precisely, each node drops all packets that it should not hear in a real ad hoc network. Moreover, to support the nodes mobility, we implement a limited random waypoint mobility model. Our mobility model limits the movement of nodes in such a way that the network remains connected. Keeping the network

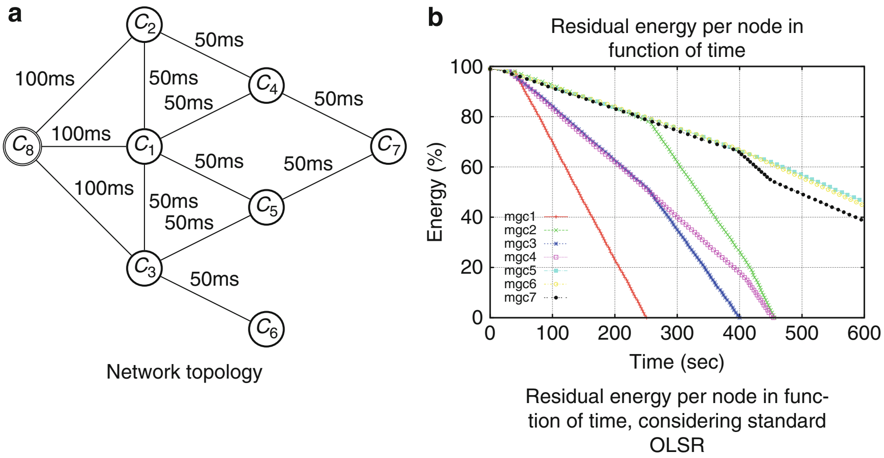


Fig. 4 Energy consumption

connected ensures that the players always have at least one path to reach the game server and thus ensures that they continue to play. During the experiments, the scheduler sends to each node its current position and informs them about their actual neighbor nodes. Nodes can then apply the corresponding *iptables* filters to emulate the multi-hop ad hoc network.

In addition to the mobility management, since our computers are directly connected to AC power, we also added an emulation of the battery consumption. We focus on energy consumption due to network activity. Hence, we do not take into consideration the energy consumed by local CPU computing (e.g., 3D graphics display). In the beginning of the experimentation, each node has 100 % of available energy, except the game server, which is supposed to be plugged to an AC power (so that it will not fail before any other client). Each time a node sends or receives a packet, we decrease its available energy according to the linear energy consumption model presented in Feeney (2001). More precisely, we use the *libpcap* API to capture incoming and outgoing packets and decrease the available energy according to the size of the packets, using the energy consumption model. Finally, as routing protocol we use the well-known proactive OLSR protocol (Jacquet et al. 2001). We use the OLSR daemon (OLSRd) version 0.5.5, and we modify it to fit our needs.

Using this platform, we were able to analyze the impact of MANET (Kaiser et al. 2009b) routing issue on multiplayers networked games. More precisely, we concentrated our analysis on energy consumption and delay on the gameplay. We focused on static topology, as illustrated in Fig. 4a. The node S is the game server and nodes C_i, $i \in \{1, 2, ..7\}$, are only game players. We fix the topology in such way that for some nodes we can have until three hops to reach the server and also to have several different paths to reach the leaf nodes. With this topology, we can easily observe OLSR behavior. As expected, a client who has a great end-to-end delay with the game server has a bad game quality. On the contrary, a client who is close to the game server in terms of end-to-end delay has a better game quality, but his energy decreases dramatically faster since he has to forward the game traffic coming from the other players.

Fairness Improvement

In the last section, we conclude that the end-to-end delay is a critical metric that can bring unfairness among players and even more if we consider FPS games. Moreover, if we consider video games on MANET, energy consumption should also be considered. Indeed, typical FPS game traffic is characterized by small packets size (few dozens of bytes) sent at low time intervals (few ms). Without any specific solution, the radio interface will generate at low time intervals PHY frames with very large headers in comparison with the small payload they are carrying. This important overhead could be considered as an energy waste. It can consequently lead into faster battery consumption and could induce brutal disconnection of some players. To overcome this problem, we present in the following two approaches, the first approach at application level and second approach at the network level.

Application-Level Solutions

The objective here is to increase the fairness between the clients at the application level (mainly at the server side). The classical approaches are the Δ-*causality control* (Ishibashi et al. 2007; Ikedo et al. 2006), the *dead-reckoning* (Aggarwal et al. 2004, 2005), and the *timewarp* (Mauve et al. 2000). The objective of the first approach is to reduce the impact of the heterogeneous end-to-end delays between players and the server by adding an artificial playback buffer of a fixed period at the server side. In this case, each game terminal has to input the information about the player's event (such as position) at regular intervals and send it to the game server with its timestamp, in order to denote the generation time of the event. At the server, the received information is saved in the playback buffer until the generation time plus A seconds ($\Delta \geq 0$). By doing that, the server computes the players' events according to their generation time instead of their reception time. The main weakness of this methods is related to the playback buffer period is fix. In order to resolve this problem, works like Ishibashi et al. (2007); Kusunose et al. (2010); and Zander et al. (2005) propose an adaptive Δ-*causality control*. The main objective is to reduce/increase the playback buffer period according to the end-to-end delay variations.

The objective of the second approach, *dead-reckoning* (IEEE standard 2012; Singhal et al. 1994; Duncan et al. 2003), is to permit to a client to not be blocked and to continue playing even if it is not receiving the data in case of a short-term disconnection. More precisely, server and clients have to compute the current game's state (such as players' avatar position) by using a previously determined position and advancing that position based upon known or estimated speeds. Unfortunately, *dead-reckoning* is subject to cumulative prediction errors. This problem becomes worse in wireless networks because of their higher delays as compared to wired networks.

The last approach, *time warp* (Jefferson 1985; Lin et al. 1991; Mauve et al. 2004) is a synchronization mechanism. This approach allows the game server to execute events without the guarantee of a causally consistent execution. Upon the detection of a causality violation, *rollback* procedures are executed to recover the game state to a coherent state. Unfortunately, this approach suffers from two main problems. First, the game server has to store previous states, for which memory is needed. Second, *rollbacks* demand processing power which can be a costly resource on some limited power terminals. Finally, these three methods are eventually combined at the server side for what is known as a **lag compensation technique**. Unfortunately, as we are considering wireless ad hoc network environment, all the approaches presented earlier could not be considered since we face the very classical energy problem.

As introduced earlier, packets exchanged between players and the game server are very small and sent very frequently. Consequently, this kind of traffic leads to an important overhead, and thus to inefficient energy consumption. To improve the energy efficiency, a possible approach is to concatenate, at the client side, several of these small game data packets into longer packets (Kaiser et al. 2010a). This simple concatenation procedure will decrease the overhead and diminish the number of packets exchanged between players and the game server, which will lead to reduce the energy consumption. Unfortunately, packets concatenation will necessarily lead to increase packets reception delay by the server, which will impact gameplay fairness.

With the classical method, a client who has a short delay with the server could see his delay artificially increased by the server, so that the gap with other long delay players is diminished. However, in our case, instead of adding these artificial delays at the server side for all players experiencing short delays, we propose to add these delays at the client side. While equalizing the delays of different players at the client side, we will exploit the added delay time to concatenate game packets and form a longer packet, which will be sent at the appropriate time. Precisely, we propose to concatenate packets only during a certain period of time, based on values of end-to-end delay of each player.

To illustrate this approach, let us consider in Fig. 5a the topology we used in our experimentations. We defined a static topology such that we have multihop routes and all nodes have several routes to reach the game server. The value on a link represents the end-to-end delay value of this link. These values are chosen by considering the number of neighbors a node has. For example, node C4 has four neighbors, thus its end-to-end delay from it to C1, C2, C5, and C7 is 10 ms*4 = 40 ms.

In the first experimentation, we use a lag compensator technique on the game server side coupled with the energy-enhanced version of OLSR. We downloaded and installed the package for IOQuake 3 Arena which contains the server side lag compensator. This lag compensator uses the timewarp technique. It keeps in memory the game state of the last 2 s. Therefore, it can go back in time to process packets it receives delayed from players. On the other hand, in the second

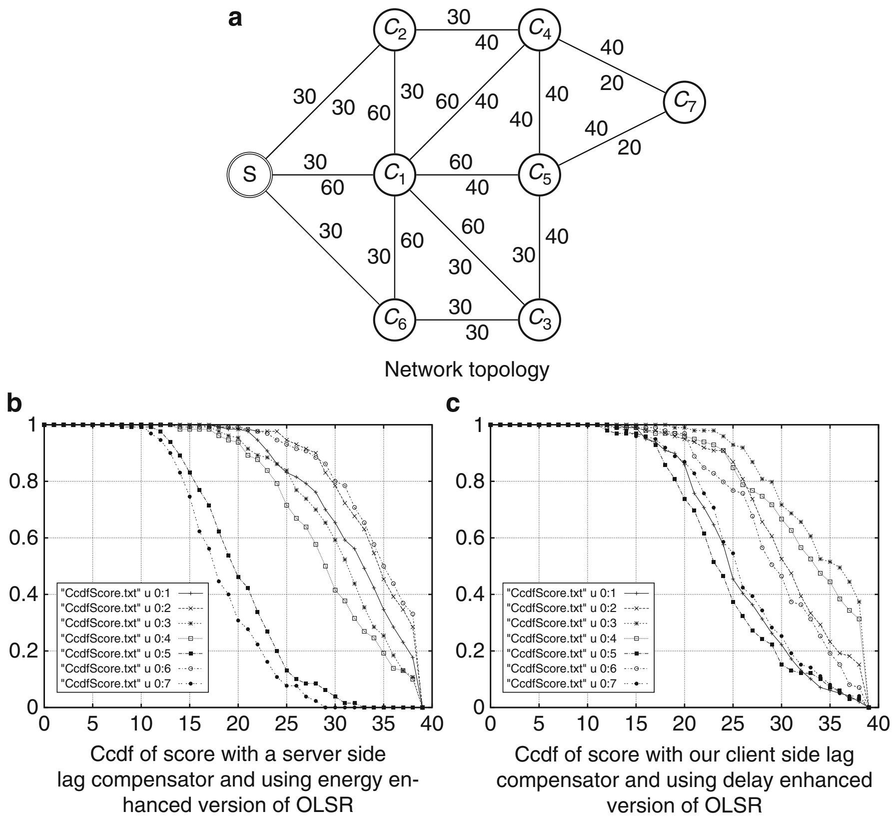
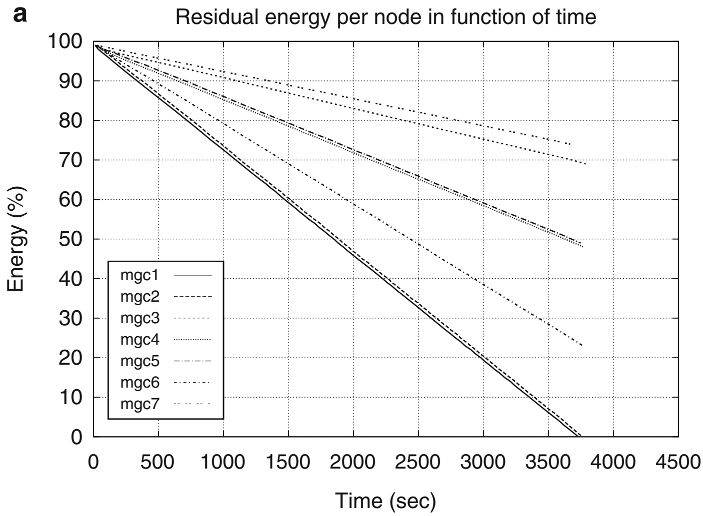


Fig. 5 Lag compensation

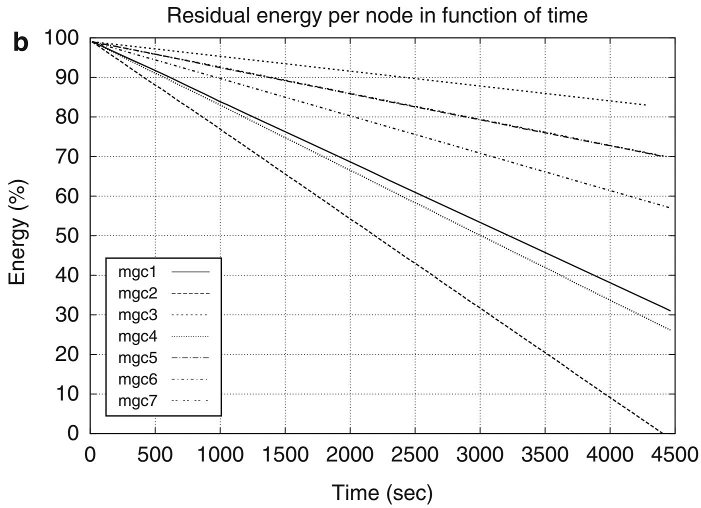
experimentation, we test our proposition. We implement our lag compensator on each computer and coupled it with the delay enhanced version of OLSR. We compare results from this experimentation with the first one.

Figure 5b, c show the complementary cumulative distribution function (CCDF) of score of each player for the first and the second experimentation, respectively . Results show that this solution tends to bring together curves of all players by increasing the mean score of players who are the most penalized. Indeed, if we look at curves of player 7, its chance to reach a score greater than or equal to 20 at the end of a game is about 30 % in the first experimentation. With our proposition, this percentage increases until 85 %. A similar improvement is also observed for player 5. It chances to reach a score greater than or equal to 20 increases from 45 % to 75 %. Therefore, this shows that fairness between players is enhanced.

Finally, the last two figures illustrate the energy consumption. Figure 6a shows the residual energy of each node for the first experimentation, and Fig. 6b shows the residual energy of each node for the second experimentation. We stop the



Energy consumption of nodes versus time with the server side lag compensator and the energy enhanced version of OLSR



Energy consumption of nodes versus time with our client side lag compensator and the delay enhanced version of OLSR

Fig. 6 Energy consumption

experimentations when one node of the network reaches the value of 0 %. We can see in Fig. 6a that after about 3,750 s, a player has lost all its energy. The experimentation then stops. In Figure 6b, the first node loses all its energy after about 4,400 s. Our proposition has thus increased the game session duration by about 17 %.

Network Level Solutions

The objective of network level approaches is to consider routing protocols in order to increase the fairness between the clients and to optimize the energy consumption. Early MANET routing protocols focused only on finding a feasible route between a source and a destination, without any application's QoS considerations. To support QoS, the essential problem is to find a route with sufficient available resources to meet the QoS constraints and possibly to incorporate optimizations, such as finding the lowest cost or most stable of the routes that meet the QoS constraints. To address these routing issue, several researchers addressed the general problem of QoS in MANETs (Mohapatra et al. 2003; Zhu et al. 2004; Reddy et al. 2006). In our work, we focus mainly on the OLSR protocol (Clausen et al. 2003).

If we focus only on the energy consumption, many works were done in the literature to combine the OLSR protocol with the energy metric, such as Kunz et al. (2008); Mahfoudh et al. (2008); Benslimane et al. (2006); and Rango et al. (2008). In Kunz et al. (2008), instead of the number of hops, the authors consider the residual energy of nodes to compute routes. Thus, they avoid nodes with low remaining energy, leading to increase the network lifetime. In Mahfoudh et al. (2008), the authors go further and also consider the energy consumed at the 1-hop and 2-hop neighbors. Indeed, as we are in a wireless environment, when a node sends a message, all its neighbors receive it, even if they are not concerned by it, leading to consume energy. In Benslimane et al. (2006), the authors base their route computing on a min-max energy algorithm. In each available route to reach a destination, they look at the node, which has the lowest remaining energy. They choose the route whose lowest remaining energy node has the greatest value. Finally, in Rango et al. (2008), the authors exploit the "Willingness" field of the HELLO messages to compute their link cost. The Willingness is a value that informs if a node will forward data or not. When a node has a low remaining energy, it sets its Willingness to 0 in order to inform the other nodes that it would not forward traffic anymore.

On the other hand, other works focus only on the delay as routing metric, such as Badis et al. (2003) and Meraihi et al. (2003). In Badis et al. (2003), the authors assume that nodes are synchronized and compute the delay by timestamping the HELLO signaling messages. In this case, nodes receiving a HELLO message can calculate the delay with their neighbors. This delay metric is then used in the Dijkstra's shortest path algorithm to compute routes. In Meraihi et al. (2003), the authors add three new signaling messages in the OLSR protocol. These messages enable to compute the end-to-end delay value between two nodes in the network, without the need of a synchronized network.

Finally, some works addressed multi-metric routing and combine energy and delay metrics. In Sanchez-Miquel et al. (2005), the authors propose EDC-AODV (Energy Delay Constrained AODV), a modified version of the Ad hoc On-demand Distant Vector (AODV) (Perkins et al. 1999) routing protocol. Each node takes into consideration the current size of its buffer (which informs about the delay and congestion at this node) as a first metric, and its residual energy as a second metric.

The cost of a link between two neighboring nodes is processed using a balancing equation with a parameter a in order to strike a balance between the two metrics. In Regatte et al. (2005), the authors propose OEDR (Optimized Energy-Delay Routing), an ad hoc routing protocol considering energy and delay metrics and OLSR. They consider the delay between two neighbors, measured thanks to timestamped HELLO messages, as the delay metric. As for the energy metric, they consider the residual energy of the nodes. They use the HELLO messages to transmit the energy and the delay values. Then, the cost of a link is the multiplication of the delay and the energy values. The Dijkstra's shortest path algorithm is finally processed using this cost to find the best routes. Unfortunately, the above proposals were not proposed specifically to the videogame applications.

In this case, one possible approach is to use multi-metric OLSR-based routing protocol (Kaiser et al. 2012). As energy metric, we consider the residual energy of the nodes, and as delay metric we consider the delay measured on a link between two neighbor nodes by time stamping the OLSR HELLO signaling messages. Finally, both delay and energy metrics are then combined to compute the cost of each link. More precisely, we propose that each node start by computing its shortest path to reach the game server by considering only the delay metric. It is clear that players having a great delay path costs should not consider the energy metric, but only focus on the delay metric. To this aim, we define a *delay threshold*, from which we consider that the players are too far from the game server. In this case, all the nodes having a delay path cost greater than the *delay threshold* only consider the delay metric to compute their paths. Otherwise, all the other nodes consider both energy and delay metrics to compute the link cost according to a balancing equation with a parameter a. This balancing factor determines the weight of each metric. Finally, we decide to dynamically modify the value of the balancing parameter depending on the distance, in terms of end-to-end delay, that separates a player and the game server. More precisely, each node computes its own balancing factor a as the ratio between the path delay and the delay threshold. If a player has a great end-to-end delay with the game server, he gives priority to the delay metric. On the contrary, a player who has a smaller end-to-end delay with the game server favors the energy metric.

To illustrate this approach, let consider in Fig. 7a the topology we used for experimentations. In Figure 7b, we present the energy consumption versus time while using the delay-efficient OLSR. According to the topology presented in Fig. 7a, C_2, and C_3 are the 1-hop neighbors of the game server which belong to the shortest paths, in terms of delay, of nodes C_4, C_5, C_6, and C_7. Thus, they forward their game traffic; that is why their residual energy decreases faster. Indeed, C_3 forwards the traffic of C_5 and C_6, whereas C_2 forwards the traffic of C_4 and C_7. Until 3,000 s., we can distinguish three groups of nodes: C_2 and C_3 which forward two traffics each, C_4 which forwards one traffic (the one of C_7) and leaf nodes (C_1, C_5, C_6 and C_7) which do not forward any traffic. After 3,000 s., C_2 and C_3 reach 0 % of residual energy and shut down. Then, C_1 takes over and forward the traffics of C_4, C, and C_7. As a consequence, its residual energy decreases faster. The only path for C_6 to reach the game server is through C_3. However, as this node is out of energy,

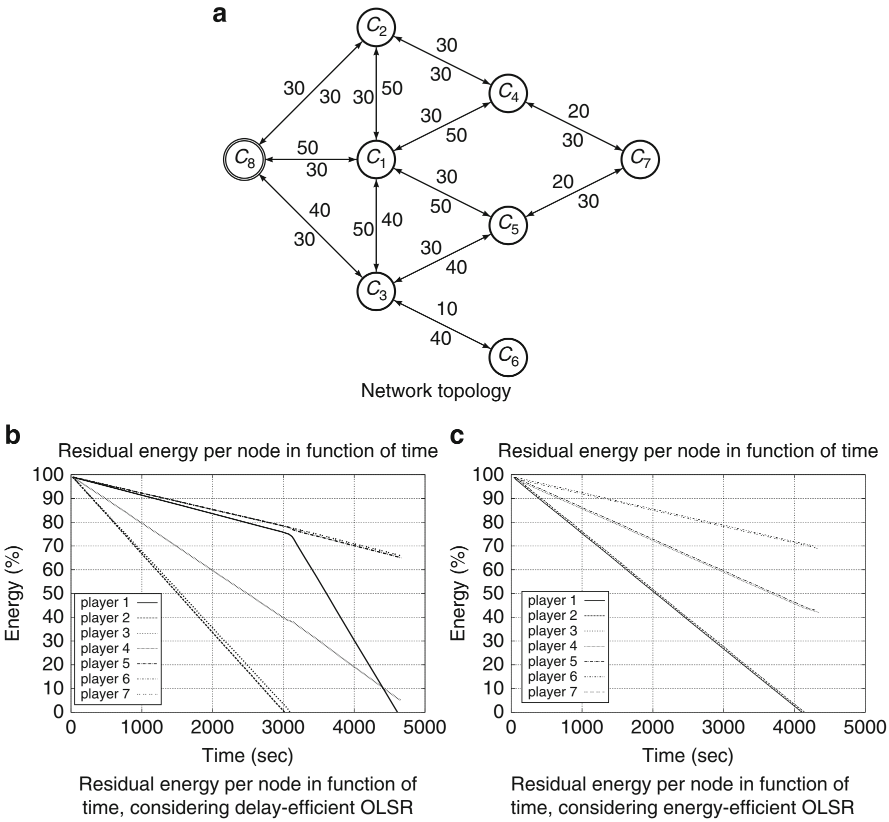


Fig. 7 Energy consumption: delay-efficient OLSR vs energy-efficient OLSR

C_6 cannot reach the game server anymore and has to leave the game. Finally, after 4,600 s., C_1 is out of energy and no other node can reach the game server. These results show that finding routes for traffic according to only the delay metric seriously reduce the game session.

In Figure 7c, we present the energy consumption versus time while using the energy-efficient OLSR. Routes are now computed according to only the energy metric. We can clearly distinguish three groups of curves in this figure. Members of these groups are in fact determined by their position compared with the game server. Indeed, the first three curves are the ones of nodes C_1, C_2, and C_3, which are all 1-hop far from the game server. These three nodes forward game traffic from and to nodes C_4, C_5, C_6, and C_7, that is why their residual energy decreases faster than the other nodes. The second group of curves is composed of nodes C_4 and C_5, which both are 2-hops far from the game server. They fairly forward traffic from and to C_7. Finally, curves of C_6 and C_7 represent the third group. Those two nodes are leaf nodes and do not forward any other traffic than their self. Thus, their residual energy decreases slower than all other nodes. This figure shows well that

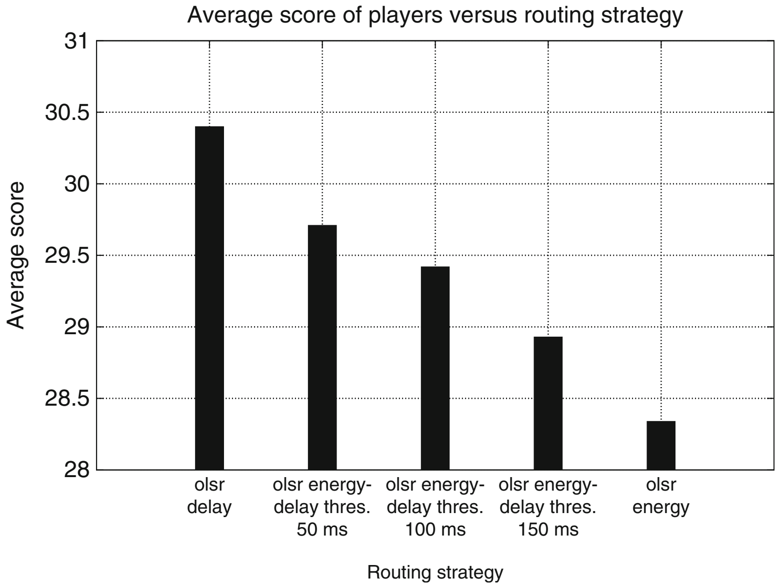


Fig. 8 Average score of player versus routing strategy

the traffic load is fairly distributed among the nodes of a same group to avoid excessive energy consumption of only one node. Indeed, all the players can play together up to only 3,000 s. against 4,000 s. for energy-efficient routing. Using the energy metric rather than the delay metric increased the game session from 3,000 to 4,000 s.

Finally, Fig. 8 depicts the average score of the players in function of the routing strategy chosen. Results show that increasing the *delay_threshold* value leads to slightly decreasing the average score. Increasing the *delay_threshold* value improves lifetime of game session and fairness among players but at the cost of some gaming quality (in particular for players which are close to the game server). On the contrary, reducing the *delay_threshold* value decreases the lifetime of game session and the fairness among the players but improves the game quality.

Mobility Management

As introduced earlier, the last issue is related to nodes' mobility. Indeed, if we consider MANET environment, due to nodes movements, disconnections between a game client and the game server might occur. From the player's standpoint, network disconnection leads to the blocking of the input and output processes (i.e., freezing screen). If this disconnection situation continues for a while, the server could reject the player. Otherwise, if the routing protocol is able to repair the broken path or establish an alternative path to reach the game server then the player can resume its party. However, during the disruption period, the server and other clients often consider the player's avatar as active. Thus, it might be an easy target for other gamers. When the player resumes the party, there is a high probability that its avatar has been penalized (e.g., fragged in a FPS game).

To overcome this situation, one possible solution is to use multipath routing protocol. Many works concerning multipath exist in the literature (Khan et al. 2006; Wang et al. 2001; Yi et al. 2008, 2009; Mueller et al. 2004) and state that multipath routing is a suitable technique in MANET. However, performances of multipath routing protocols depend strongly on how they are configured. For instance, how many different paths should be used? Are paths nodes disjointed? etc. In our opinion, it is quite clear that multipath routing protocol configuration has to be done according to the application constraints.

In Kaiser et al. (2010b) the objective is to use a multipath routing protocol to maintain the connection between each client and the server by building up two paths.

More precisely, we compute two possible routes; the first route is selected on the end-to-end delay metric, since the delay is highly impacting the gameplay and fairness between the players. Moreover, and in order to reduce the impact of mobility disconnection, we compute a second route by removing from the network topology, the link(s) having the worst value of ETX (i.e., the less stable link in the route) and belonging to the first route. To be sure that the traffic generated by each node is using only one of the two routes and in order to avoid routing loops or intermediates node routing change, we use the source routing technique. In fact, each node will add in the header of its packets the whole path that these packets must use to reach their destination.

After computing and memorizing the two paths to reach the destination, each node balances its traffic between the two paths in function of the ETX value of the worst link found in the first path. More the ETX value is worse, higher is the percentage of traffic sent using the second path. Concretely, if the ETX value is equal to 1 (i.e., optimum link quality), then all the packets are sent using the first path. The latter being the shortest end-to-end delay route to the server. However, if the ETX value is greater than 1, the traffic is balanced according to given thresholds. The thresholds have to be configured accordingly to the game QoS constraints.

This threshold-based traffic balancing strategy is motivated by the fact that in many time sensitive games, such as FPS, the traffic generated by the players is composed of a lot of very small packets sent at very high frequency. Thus, if the game server receives only three packets out of four coming from one player, the player can still play in acceptable conditions. By balancing traffic between two paths, we can ensure that an amount of packets still arrive at the game server even if the first path encounter problems. Therefore, if the first path finally breaks, packets send on the second path ensure that the player is still connected to the game server until the routing protocol reacts to the link failure and find a new shortest path.

To illustrate this approach, we carried out a set of experimentations using the emulation platform. In Figure 9a, we show the mean number of disconnections a player may experience during one game session. We can clearly see that the multipath OLSR source routing protocol is reducing the number of disconnections

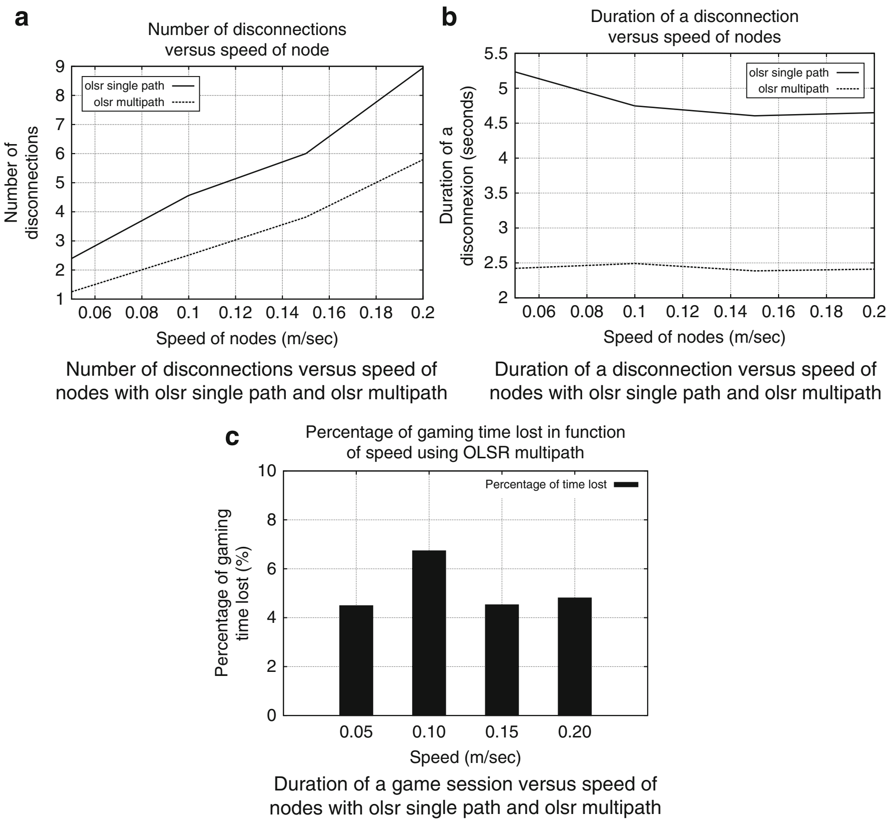


Fig. 9 Mobility Management: single path vs multipath olsr

compared to the classical OLSR routing protocol. This will result in an improvement of the gameplay. We can also notice that even if we provide more than one route to reach the destination the number of disconnections is not equal to zero. This can be explained by the fact that, in some cases, some nodes have only one available path to reach the game server. In this case, it is impossible for those nodes to use an alternative path. Thus, if one of the path's links fails the player will be disconnected from the server. Finally, we can also observe that the number of disconnections is increasing when the speed of nodes increases. That is due to the fact that more the nodes are moving fast, higher is the probability that a link fails.

In Figure 9b, we plot the disconnections duration, in seconds, for both multipath OLSR source routing protocol and classical single path OLSR routing protocol. Typically, this duration is expressing the amount of time a player sees his screen freezing and cannot play anymore. Once again, we can clearly see that our multipath algorithm is reducing the disconnection period by a factor of two. This is justified by the fact that, when a path breaks, a period of time is necessary for the routing protocol to detect the path failure and to compute a new path; in the case of multipath algorithm,

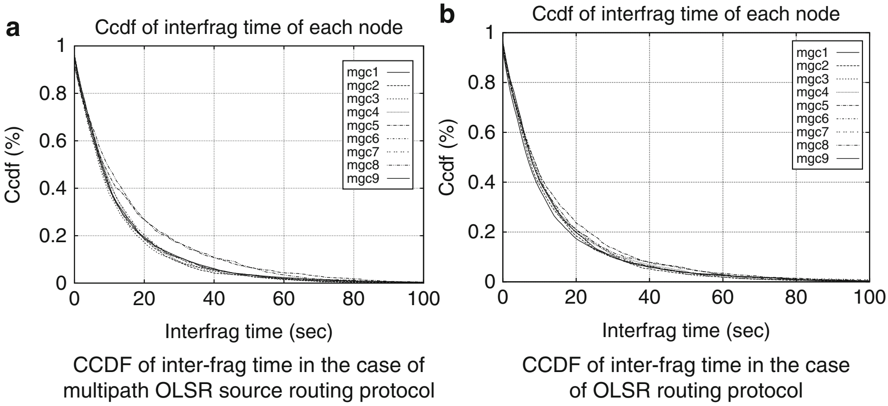


Fig. 10 Inter-frag time: single path vs multipath olsr

observing the ETX values leads to the anticipation of paths failure and then the reduction of the disconnections period. In this case, each node is more able to predict a disconnection and can switch quickly the traffic on another path, if available.

In Figure 9c, we plot the average percentage of time a game session is reduced when using the multipath OLSR source routing algorithm compared to the OLSR routing algorithm. In our experimentations, we consider that a game session ends when the first node leaves the network due to the consumption of all its energy. It was expected that the use of multipath routing reduces the game session's duration. This is due to the fact that the multipath algorithm is using the source routing strategy. That means that, each packet needs to enclose in its header all the path that it should follow to reach the destination. However, we can notice that this reduction is slightly small, between 4 % and 6 %, compared to the gain obtained in terms of mean number of disconnections experienced.

Finally, we plot in Fig. 10a, b the CCDF of inter-frag time for multipath OLSR source routing protocol and OLSR routing protocol in the case of node mobility equal to 0.1 m/s, respectively. Comparing these results we can observe that multipath strategy is slightly impacting the game fairness and penalizing some players. This is caused by the end-to-end delay variation due to the traffic balancing between the paths. But it is clear that most of the players prefer to go through this penalty and continue playing than experiencing high number and longer disconnections.

Summary of Results

As a conclusion, in this chapter we focus on the networked video game support over wireless ad hoc networks. We start this chapter by objective assessment methodology in order to understand the real correlation between network conditions (i.e., delays, jitter, and packet losses) and player's expected performance. The evaluation was conducted experimentally using an FPS game setup running autonomously and simulating the realistic behavior of players, over an emulated network. The experimental results show clearly that considering the score, as an objective metric, we can capture the fairness a party is experiencing. Moreover, the experiments have also derived interesting characteristics of other metrics, e.g., the exponential distribution nature of the inter-fragging times. Finally, the results have proven the important sensitivity of the gameplay (especially regarding the fairness among players) of the delays over the jitter or packet losses.

Thereafter, we evaluated the effects of MANET routing protocols. We have concentrated our analysis on two metrics: (1) energy consumption and (2) impact of the delay on the gameplay. As above, the evaluation was conducted experimentally over an emulated ad hoc network. We choose to compare two versions of OLSR protocol: the native one and an energy-efficient version proposed in the literature, which we have implemented. As expected, the experimental results proved the superiority of the energy-enhanced version of OLSR over native OLSR, in term of network life time duration. Unfortunately, even if the network life duration is increased, the fairness between the players is still missing.

To overcome this limitation, we explore two approaches, one at the application level and one at the network level. Using gameplay objective metrics and extensive experimentations, we show that both approaches can significantly increase the network life duration, while improving the game fairness among players.

Finally, we narrow our attention on the problem of game disconnections due to node's mobility. In this case, we explore a multipath path based on OLSR protocol that computes two paths to reach a destination. A first path is computed using the end-to-end delay metric between the game client and the server. A second path is computed by removing from the network topology, the link(s) having the worst value of ETX and belonging to the first route. Finally, network traffic is balanced between the two paths. The obtained results show that the number and the duration of the disconnections that can experience players are highly reduced.

Recommended Reading

S. Aggarwal, H. Banavar, A. Khandelwal, S. Mukherjee, S. Rangarajan, Accuracy in dead-reckoning based distributed multiplayers games, in *Proceedings of ACM NetGames,* 2004

S. Aggarwal, H. Banavar, S. Mukherjee, S. Rangarajan, Fairness in dead-reckoning based distributed multiplayer games, in *Proceedings of ACM NetGames,* 2005

G. Armitage, An experimental estimation of latency sensitivity in multiplayer quake 3, in *Proceedings of 11th IEEE International Conference on Networks/ICON 2003,* 2003

G. Armitage, L. Stewart, Limitations of using real-world, public servers to estimate jitter tolerance of first person shooter games, in *Proceedings of the 2004 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology,* 2004

H. Badis, A. Munaretto, K.A. Agha, G. Pujolle, Qos for ad hoc networking based on multiple metrics: bandwidth and delay, in *Proceedings of the* 5^t*h IFIP-TC6 Internation Conference,* 2003

T. Beigbeder, R. Coughlan, C. Lusher, J. Plunkett, E. Agu, M. Claypool, The effects of loss and latency on user performance in unreal tournament 2003, in *Proceedings of the 3rd ACM SIGCOMM Workshop on Network and System Support for Games,* 2004

A. Benslimane, R.E. Khoury, R.E. Azouzi, S. Pierre, Energy poweraware routing in olsr protocol, in *Mobile Computing and Wireless Communication International Conference (MCWC),* 2006

P.A. Branch, G.J. Armitage, Towards a general model of first person shooter game traffic, in *CAIA Technical Report,* 2005

T. Clausen, P. Jacquet (eds.), Optimized link state routing protocol (olsr) (RFC Editor, 2003)

A. Cricenti, P. Branch, Arma(1,1) modeling of quake 4 server to client game traffic. in *Proceedings of the 6th ACM SIGCOMM Workshop on Network and System Support for Games (NetGames),* 2007

M. Dick, O. Wellnitz, L. Wolf, Analysis of factors affecting players' performance and perception in multiplayer games, in *Proceedings of the 4th ACM SIGCOMM Workshop on Network and System Support for Games,* 2005

T.P. Duncan, D. Gracanin, Algorithms and analyses: pre-reckoning algorithm for distributed virtual environments, in *Proceedings of the 35th Conference on Winter Simulation: Driving Innovation,* ser. WSC'03. Winter Simulation Conference, 2003, pp. 1086–1093. [Online]. Available: http://dl.acm.org/citation.cfm?id=1030818.1030962

L.M. Feeney, An energy consumption model for performance analysis of routing protocols for mobile ad hoc networks. Mobile Netw. Appl. **6**(3), 239–249 (2001)

W. Feng, F. Chang, W. Feng, J. Walpole, Provisioning on-line games: a traffic analysis of a busy counter-strike server, in *Technical Report,* 2002

http://ioquake3.org/, 2005

IEEE standard for distributed interactive simulation – application protocols, *IEEE Std 1278.1-1995*, 2012, pp. 1–144

IETF MANET WG (mobile ad hoc network). IETF, www.ietf.ora/html.charters/manet-charter.html

T. Ikedo, Y. Ishibashi, An adaptive scheme for consistency among players in networked racing games, in *Mobile Data Management, 2006. MDM 2006. 7th International Conference on,* 2006, pp. 116–116

Y. Ishibashi, Y. Hashimoto, T. Ikedo, S. Sugawara, Adaptive deltacausality control with adaptive dead-reckoning in networked games, in *Proceedings of ACM NetGames,* 2007

P. Jacquet, P. Muhlethaler, T. Clausen, A. Laouiti, A. Qayyum, L. Viennot, Optimized link state routing protocol for ad hoc networks, in *Proceedings of IEEE International Multitopic Conference (INMIC 2001),* 2001

D.R. Jefferson, Virtual time. ACM Trans. Progr. Lang. Syst. **7**, 404–425 (1985)

A. Kaiser, D. Maggiorini, N. Achir, K. Boussetta, On the objective evaluation of real-time networked games, in *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*, 2009a, pp. 1–5

A. Kaiser, N. Achir, K. Boussetta, Multiplayer games over wireless ad hoc networks: energy and delay analysis, in *International Workshop on Ubiquitous Multimedia Systems and Applications (UMSA),* 2009b

A. Kaiser, N. Achir, K. Boussetta, Improving energy efficiency and gameplay fairness for time sensitive multiplayers games in manet, in *Energy Efficiency in Wireless Networks & Wireless Networks for Energy Efficiency Workshop (E2Nets),* 2010

A. Kaiser, N. Achir, K. Boussetta, A multipath traffic balancing proposal to reduce gaming disconnections in manet, in *Wireless Days (WD), 2010 IFIP,* 2010, pp. 1–5

A. Kaiser, K. Boussetta, N. Achir, An energy-delay routing protocol for video games over multihops ad hoc networks, in *Ad Hoc Networks*, ed. by D. Simplot-Ryl, M. Dias de Amorim, S. Giordano, A. Helmy. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol. 89 (Springer, Berlin/Heidelberg, 2012), pp. 193–208

A. Khan, T. Suzuki, M. Kobayashi, M. Morita, A multipath approach for fast re-routing in olsr," in *IEIC Technical Report,* 2006

T. Kunz, Energy-efficient variations of olsr, in *Proceedings of the International Wireless Communications and Mobile Computing Conference (IWCMC),* 2008

Y. Kusunose, Y. Ishibashi, N. Fukushima, S. Sugawara, Qoe assessment in networked air hockey game with haptic media, in *Network and Systems Support for Games (NetGames), 2010 9th Annual Workshop on,* 2010, pp. 1–2

Y.-B. Lin, E.D. Lazowska, A study of time warp rollback mechanisms. ACM Trans. Model. Comput. Simul. **1**(1), 51–72 (1991). [Online]. Available: http://doi.acm.org/10.1145/102810.102813

S. Mahfoudh, P. Minet, An energy efficient routing based on olsr in wireless ad hoc and sensor networks, in *Advanced Information Networking and Applications – Workshops (AINAW),* 2008

M. Mauve, How to keep a dead man from shooting, in *Proceedings of the 7th International Workshop on Interactive Distributed Multimedia Systems and Telecommunication Services (IDMS)*, 2000

M. Mauve, J. Vogel, V. Hilt, W. Effelsberg, Local-lag and timewarp: providing consistency for replicated continuous applications. Multimedia IEEE Trans. **6**(1), 47–57 (2004)

A.N. Meraihi, P. Jacquet, Le côntrole du delai dans le protocole olsr, in *Research Report,* 2003

P. Mohapatra, J. Li, C. Gui, Qos in mobile a hoc networks. Wirel. Commun. IEEE **10**(3), 44–52 (2003)

S. Mueller, R.P. Tsang, D. Ghosal, Multipath routing in mobile ad hoc networks: issues and challenges, in *Performance Tools and Applications to Networked Systems, volume 2965 of LNCS,* 2004

C.E. Perkins, E.M. Royer, Adhoc on-demand distance vector routing, in *Second IEEE Workshop on Mobile Computer Systems and Applications (WMCSA'99),* 1999

P. Quax, P. Monsieurs, W. Lamotte, D.D. Vleeschauwer, N. Degrande, Objective and subjective evaluation of the influence of small amounts of delay and jitter on a recent first person shooter game, in *Proceedings of the 3'd ACM SIGCOMM Workshop on Network and System Support for Games,* 2004

F.D. Rango, M. Fotino, S. Marano, Ee-olsr: energy efficient olsr routing protocol for mobile ad-hoc networks," in *Military Communications Conference (MILCOM),* 2008

T.B. Reddy, I. Karthigeyan, B.S. Manoj, C.S.R. Murthy, Quality of service provisioning in ad hoc wireless networks: a survey of issues and solutions, *Ad Hoc Netw.* **4**(1), 83–124 (2006). [Online]. Available: http://dx.doi.org/10.1016/j.adhoc.2004.04.008

N. Regatte, J. Sarangapani, Optimized energy-delay routing in ad hoc wireless networks, in *Proceedings of World Wireless Conference*, 2005

L. Sanchez-Miquel, N. Vesselinova-Vassileva, F. Barcelo, P. Carbajo-Flores, Energy and delay-constrained routing in mobile ad hoc networks: an initial approach, in *Proceedings of the 2nd ACM International Workshop on Performance Evaluation of Wireless Ad hoc, Sensor, and Ubiquitous Networks (PE-WASUN),* 2005

S.K. Singhal, D.R. Cheriton, Using a position history-based protocol for distributed object visualization, (Tech. Rep., Stanford, 1994)

L. Wang, Y. Shu, M. Dong, L. Zhang, O.W.W. Yang, Adaptive multipath source routing in ad hoc networks, in *IEEE ICC,* 2001

A.F. Wattimena, R.E. Kooij, J.M. van Vugt, O.K. Ahmed, Predicting the perceived quality of a first person shooter: the quake iv g-model, in *Proceedings of the 5th ACM SIGCOMM Workshop on Network and System Support for Games,* 2006

J. Yi, E. Cizeron, S. Hamma, B. Parrein, Simulation and performance analysis of mp-olsr for mobile ad hoc networks, in *IEEE WCNC,* 2008

J. Yi, S. David, A. Adnane, B. Parrein, X. Lecourtier, Multipath olsr: simulation and testbed, in *5th OLSR Interop/Workshop,* 2009

S. Zander, I. Leeder, G. Armitage, Achieving fairness in multiplayer network games through automated latency balancing, in *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology,* ser. ACE'05. (ACM, New York, 2005), pp. 117–124. [Online]. Available: http://doi.acm.org/10.1145/1178477.1178493

H. Zhu, M. Li, I. Chlamtac, B. Prabhakaran, A survey of quality of service in ieee 802.11 networks. Wirel. Commun. IEEE **11**(4), 6–14 (2004)

Part VI

Serious Games

A Tangible Serious Game Approach to Science, Technology, Engineering, and Mathematics (STEM) Education

25

Riccardo Berta, Francesco Bellotti, Erik van der Spek, and Thomas Winkler

Contents

R. Berta (✉) • F. Bellotti
University of Genoa, Genoa, Italy
e-mail: berta@elios.unige.it; franz@elios.unige.it

E. van der Spek
Technische Universiteit Eindhoven, Eindhoven, The Netherlands
e-mail: e.d.v.d.spek@tue.nl

T. Winkler
University of Luebeck, Lübeck, Germany
e-mail: winkler@imis.uni-luebeck.de

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_32

Abstract

The chapter's idea stems from the observation that technology is difficult to learn in an abstract way (books, lectures, etc.) and practical activities are needed not only to apply concepts but also to help learning itself. This is particularly challenging in a younger age where this concern is often neglected frequently, leading to poor instruction, if any. However, science, technology, engineering, and mathematics (STEM) topics typically involve facts and concepts that could be effectively implemented and/or shown through smart objects according to the Internet-of-Things paradigm. Such objects, we called "iBlocks," could be manipulated by young learners to study various types of phenomena/artifacts and compose new aggregations (reflecting – at a younger age – the experience of the "Makers" now successful in world-leading technological universities). The presented idea is to build an environment consisting of physical objects enhanced with sensing, computing, and communicating capabilities in order to support advanced and multimodal/multisensory interaction. An important aspect is that the environment supports the definition of game rules, so that users are stimulated and invited in educational paths involving guided exploration, competition, and collaboration.

Keywords

Serious games • Tangible games • Education

Introduction

Despite the fundamental importance of science, technology, engineering, and math (STEM) education, diffusion of knowledge about technology itself – which is important to engage in a knowledge-based economy – is very problematic.

We propose an approach based on a tangible smart gaming experience (tangible serious game) as a promising approach to allow children to have a multisensorial experience of concepts and facts which are difficult to learn with standard learning tools (books, lessons, multimedia contents, etc.) (Arnab et al. 2012), in particular related to STEM concepts.

The development of such an approach involves a multidisciplinary perspective in order to gather the latest advances in learning theories and educational psychology; moreover, it needs to exploit state-of-the-art electronic devices and gaming technology to provide users with the best learning and interaction support.

The proposed approach arises from the observation that we are progressing into a new technological and innovation driven world that demands a change in the way coming generations learn and think. Being an active participant and contributor to innovation necessitates a solid understanding of STEM fields. However, several topics (e.g., concepts related to technology) are difficult to learn in an abstract way (books, lectures, etc.) and practical activities are needed not only to apply concepts but also to help learning itself. This is particularly challenging in a younger age, where this concern is often neglected, frequently leading to poor instruction, if any.

However, these topics typically involve facts and concepts that could be effectively implemented and/or shown through smart objects, according to the Tangible Bits paradigm. Such objects could be manipulated by young learners to study various types of phenomena/artifacts and compose new aggregations.

The state of the art in the field involves projects like LittleBits, Squishy Circuits, and Electroninks and start-up products, like Sifteo cubes. In this chapter we propose to go beyond these attempts and to design and implement a new approach based on tangible intelligent things (we called "iBlocks") equipped with sensing/computational/communication features and linked to an advanced game engine to define different sets of (programmable) rules for combining them to obtain different behaviors. We borrow the concept of iBlock from the well-known children's games (e.g., Lego bricks), which children can freely compose to build complex objects and scenarios. An important aspect of the proposed approach is that it can also support the definition of game rules, so that users are stimulated and invited in educational paths involving guided exploration, competition, and collaboration following a serious game (SG) approach (Bellotti 2010).

Technology Literacy and STEM Education

As information and communication technologies have become ubiquitous, society has transformed greatly. More information than can humanly be processed is always available at our fingertips; the citizens of the world who are connected in overlapping social networks and accelerating technological breakthroughs afford rapid and unpredictable innovation. As such, the demands of becoming productive and leading a self-actualized life have also changed over the course of the last decades. This has precipitated a change of thinking about education, leading to the formulation of twenty-first-century skills (Dede et al. 2010). One of the central concepts in this twenty-first-century paradigm on education is that children have to become literate in information and communications technologies. This in turn means being able to learn and appropriate knowledge and concepts from logical thinking and algorithms (Levin et al. 2000), more in general from science, technology, engineering, and mathematics (STEM) areas (Trilling and Fadel 2009), being able to analyze and explore complex systems (Zimmerman 2009), as well as being able to use and create technological content in collaborative settings (Jenkins 2009).

The knowledge, skills, and attitudes intrinsic to technological literacy, which are required to lead a self-actualized and productive life in the twenty-first century, however seem to be inappropriately learned through conventional class-room instruction. Technological literacy demands an integrative approach, with hands-on application of knowledge, psychomotor learning elements, and interacting with dynamic processes (International Technology Education Association 2003). In order to reach all of these learning goals effectively, already from an early age, we contend that interactive learning environments should be created that revolve around play. Playful settings engender a safe environment in which concepts can be engagingly explored (Corbeil 1999), and opportunities are provided for simulation and experimentation (Jenkins 2009). Subsequently, there seems to be promise in teaching an integrative approach to STEM education with serious games (Mayo 2009).

Playful Learning

For young children, playing is a learning tool. If the element of enjoyment is nonexistent, the child will not continue playing and will choose a different activity. The element of enjoyment is an emotional component, which drives the child to continue and advance in his or her learning. While playing, the children experience a dynamic process of intriguing occurrences in different fields of interest. The learner is not concerned with failure, and since he or she is merely playing a game, the inpour of knowledge is not overwhelming. The important properties during such a game are a sociocultural environment and the sensation of the learner while playing (Bruce 2011).

The contribution of “free play” is in the development of the child’s imagination, focusing on an object of his or her own choosing, and performing the various stages of the game in his or her own pace. In a “structured (or extended) play” game, children must achieve a set goal while performing stages of inquiry, which include making hypotheses and formulating solutions, among others. In a structured game, the child is exposed to a collection of rules that must be followed and has to formulate additional ones. In this type of game, communal learning takes place, as the work is done in groups and divided between the group members. This enables sociocommunicative skills to be developed.

In physical, sensory motor games, children develop both coarse and fine motor skills, spatial orientation, and systems thinking. The students improve their expressional abilities in the studied field and increasingly use professional terms in place of everyday language. They learn to persevere in order to perform the task to its completion, improving self-efficacy with the subject matter (Van der Spek 2012).

In addition, these games have the ability to create authentic settings for creative collaboration. These settings stimulate identification and improve participatory behavior in technology (Squire 2008).

Serious Games

Serious games (SGs) are games designed for a goal different from pure entertainment (Prensky 2003; Zyda 2005; Bellotti et al. 2010), which are receiving a growing interest for education (De Grove et al. 2010). Exploiting the latest simulation and visualization technologies, SGs are able to contextualize the player’s experience in challenging, realistic environments (situated cognition) (Van Eck 2006). Play supports players in exercising freedom that can complement formal learning by encouraging learners to explore various situations (Klopfer et al. 2009), with no (or lower) barriers of space and time.

SGs can be multiplayer, favoring team-building and collaboration/cooperation in facing issues (Connolly et al. 2007; Islas Sedano et al. 2013; Wendel et al. 2013).

The widespread diffusion of mobile gaming is opening further perspectives also for learning and online socialization (e.g., Parsons et al. 2012; Chi-Husiung Liang 2012). Furthermore, a large and growing population is increasingly familiar with playing games. SGs do not target exclusively power-gamers (typically young males fond of First Person 3D immersive experiences). Power-gamers represent just 11 % of the gaming community, while other types of players (e.g., social, leisure, occasional) are increasing in number. This audience enlargement has been enabled by new typologies of games (e.g., brain training games, intellectual challenges) and by new modalities of interactions (e.g., online collaboration, verbal commands, gesture-based control, social environments, family gaming).

Business data argue for a favorable trend for SGs. IDATE (2012) estimates the current global market of SGs at 2.35 billion euros, with steady growth and huge potential. Positive trends for SGs are also provided by other market and expert analysis and surveys (Pew 2012; Deloitte 2012; Gartner 2011). Performing a survey with e-learning professionals and experts and comparing outcomes with literature review (De Grove et al. 2010) stresses a "positive view," as SGs are perceived as "de facto effective learning environments (Mitchell and Savall-Smith 2004) because games challenge and support players to approach, explore, and overcome problems."

Use of SGs for education and training involves also some concerns that are to be considered. Clark (2003) argues that intended learning outcomes and game objectives might conflict and that games can distract from the learning content. The "suspension of disbelief," typically required in a game, can negatively influence the learning processes. Certain sociodemographic groups may be excluded and "hijacking" gaming as an educational technology might result in aversion towards the medium. Also, some game features (e.g., difficulty level, duration, aesthetic, interaction modalities) might endanger the instructional goal (Mitchell and Savall-Smith 2004). Tashiro (2009) states that "there is insufficient evidence to know if extant serious games improve healthcare education but there is evidence that they may inculcate inadequate clinical pattern recognition."

SGs were initially conceived to train people for tasks in particular jobs, such as training army personnel (Zyda et al. 2003) or salesmen. These tasks were characterized by their specificity and applicability for particular work-related purposes and are typically targeted at a captive audience. With the diffusion of "non-hard" gaming and of new devices (e.g., smartphones, tablets, various types of consoles), a variety of SGs has rapidly established, for different types of users (students, adults, workers, etc.), applications/goals (instruction, training, advertising, politics, etc.), and implementing different genres (arcade, first person shooter, etc.). Klopfer et al. (2009) discuss a number of examples of learning games, showing that uses of SGs now "span everything from advancing social causes to promoting better health to marketing. [. . .] Advergaming is a popular form of advertisement that delivers commercial messages through games."

In order to develop and deploy effective tools for learning, it is necessary to consider all the stakeholders (users, educators, families, researchers, developers/industries) and the whole cycle from research to market and vice-versa. From a scientific point of view, this needs considering a complex mix of disciplines and technologies, such as Artificial Intelligence (AI), Human-Computer Interaction (HCI), networking, computer graphics and architecture, signal processing, web-distributed computing, and neurosciences. These technologies are to be developed and exploited in a target-oriented multidisciplinary approach that puts the user benefits at the centre of the process. Given the instructional goal, SGs should provide quality contents on the target domain and their development should be strongly grounded in proper educational foundations (Greitzer et al. 2007).

Games are "per se" motivating and can be successfully employed with this aim. However, instructional effectiveness involves other aspects that can be analyzed in the light of the pedagogical theories.

The design and use of digital serious games has a certain theoretical foundation in the constructivist learning theories that stress that knowledge is created through experience while exploring the world and performing activities (e.g., Dewey 1933; Kolb 1984). Implications on game design involve the creation of virtual environments, typically 3D, where the player can gain knowledge through exploration and by practice (e.g., manipulating objects), possibly in collaboration with other people.

Constructivism stresses the importance of the learner to build his or her own knowledge. However, Kirschner et al. (2006) argue for the importance of guidance, in particular for novices. They refer to the cognitive load theory (CLT) (Sweller 1988), stressing the need of explicit teaching because of the limitations of the working memory. Analyzing user experience of Crystal Island, Rowe et al. (2010) report that "high-achieving science students tended to demonstrate greater problem-solving efficiency, reported higher levels of interest and presence in the narrative environment, and demonstrated an increased focus on information gathering and information organization gameplay activities." However, "lower-achieving microbiology students gravitated toward novel gameplay elements, such as conversations with non-player characters and the use of laboratory testing equipment." Observing the gameplay, the authors noticed that "high-achieving students tended to utilize more traditional science resources such as textbooks and worksheets while attempting to solve the presented mystery. In contrast, low-achieving students employed the help of expert nonplayer characters and virtual lab equipment to aid in their quest." These observations seem to stress the fact that learning is a complex activity that requires graduality and needs several steps that have to be supported by various tools (e.g., paper and digital, reading and writing) and generally have to be guided, possibly by a real adult, in order to be meaningful/compelling for the learner and not to waste time and energies. Maps, landmarks, contextualized helps, and objectives' lists with status information are game elements that could be employed in order to support avoidance of player's cognitive overloading.

Another important theory is flow, based on Csikszentmihalyi's foundational concepts (Csikszentmihalyi 1990). Flow was first employed to measure engagement in an educational game in Chen and Johnson (2004). Sweetser and Wyeth (2005) drew together various heuristics present in the literature into a concise model, the

GameFlow, consisting of eight elements: concentration, challenge, skills, control, clear goals, feedback, immersion, and social interaction. As the result of more than three decades of commercial competition, most of today's video games deliberately include and leverage the eight components of Flow (Chen 2007).

O' Broin (2011) highlights that most games adequately meet two primary elements of flow – clear goal and feedback – but the balancing of game challenges and player skills is often lacking. In order to maintain a user's Flow experience, the game's activity must balance the inherent challenge and the player's ability to address and overcome it (Chen 2007). If the challenge is beyond that ability, the activity becomes so overwhelming that it generates anxiety. On the other hand, if the challenge fails to engage the player, the player loses interest and quickly tends to get bored. However, designing such a balance becomes a greater challenge as the size of the potential audience grows, which is the typical case of video games.

SGs are being used in a variety of training and educational settings ranging from elementary schools (e.g., with games for mathematics and foreign languages) to universities (e.g., in particular with games for business management, logistics, and manufacturing, e.g., Oliveira et al. 2013). Extensively gamified courses are also being experimented (Bellotti et al. 2013).

Tangible Playful Learning

Results for primary and secondary STEM education through serious games so far have been decidedly mixed (Young et al. 2012). Serious games are often designed as a solitary immersion into a virtual world, whereas games that offer collaboration show better learning (Wouters et al. 2013). In addition, the virtual world may obfuscate real world interactions such as friction and torque, and the strictly preprogrammed nature of many video games may hinder student-driven experimentation (Young et al. 2012).

Furthermore, serious games are often designed and tested with an adolescent or young adult demographic in mind, but this may be too late. Already from a very early age, technology literacy is determined by external factors such as sociocultural background and gender socialization, where for instance girls are less likely to experience self-efficacy with computers and subsequently are less likely to choose a career in STEM areas (Vekiri and Chronaki 2008). Therefore, this project targets children in elementary and early secondary school age in learning central STEM and systems concepts in playful settings.

Tangible user interfaces seem to engage children more than screen-based interfaces when it comes to science learning and learning how to program, an effect that is especially strong for young girls (Horn et al. 2009; Sapounidis and Demetriadis 2013). More importantly, an embodied approach to learning with tangible objects can help children understand the physical properties related to STEM subjects (Young et al. 2012) and interact with difficult problems (Bakker et al. 2011) in systemic settings, as the tangibles can act as distributed cognition (Hornecker and Buur 2006). An embodied approach is crucial in understanding the evolution of simulated systems over time and its configurations in space, thus improving systems thinking as core concept in STEM education and ICT literacy learning (Clark 1999).

According to Papert's constructionist approach, the learner must create or construct an Object-to-think-with: a tangible product that serves as a focal point for thinking, linking sensory perception with abstract thinking, and linking the personal world to the social world. This tangible product can take the form of a LEGO structure, a robot, a computer program, or any other artifact, as long as it has meaning for the learner that has created it, is situated in the public space, and is accessible for examination, evaluation, and reflection by the learner or others (Kafai and Resnick 1996; Papert 1993). Additional perspectives of this approach can be found in modern educational approaches, which complete Papert's approach and implement the use of technological objects as focal points for thinking. Mitchell Resnick, the head of MIT Media lab (Resnick et al. 1998), proposed upgrading and replacing traditional building blocks, such as LEGO bricks, with similar objects that will also contain integrated technological abilities of communication and data processing (Digital manipulatives). The computerized elements allow manipulations with complex objects, which can assist the learner in developing an understanding of dynamic systems. These manipulations are performed by the learner, who uses these complex objects to create a product, which he or she will program to behave as required. Additional researchers (Milrad 2002) propose equipping the learner with interactive models based on modeling tools, construction kits, and simulations, which are used as focal points for thinking for the creation of physical and mental representations of objects and phenomena in the learner's world. This occupation will involve the learner in psychomotor, social, and cognitive processes, which will promote communal knowledge construction.

Having demonstrated the need for a mixed model approach of using tangible objects with the power of serious games in order to develop technological literacy, we propose the concept of a tangible serious game framework. The framework focuses on the technology that is necessary to teach the required STEM concepts through smart tangible objects. That is, designing context-aware blocks with embedded sensors and actuators (called iBlocks) that can be mixed and matched and a tangible serious game development and assessment tools (Bellotti 2013).

Tangible Objects

The use of smart tangible objects, often metaphorically described as "bricks," has been around for some time. Early progenitors such as graspable user interfaces (Fitzmaurice et al. 1995) and the BUILD-IT system (Rauterberg et al. 1998) used physical objects in combination with projected imagery to create mixed reality tangible interfaces. As technology progressed over the years, increasing sensing and actuating capabilities in more diminutive forms, the operating intelligence moved from external computers to inside the tangible objects. Tangible objects as a way of interfacing with a computer gave way to interfacing (interacting) with the objects themselves, leading to, e.g., applications in collocated collaborative work and education (Ullmer and Ishii 2000).

A well-known tangible interface for learning is the MIT Logo Turtle, which teaches children basic imperative programming principles (McNerney 2004). Recently, the commercial LittleBits were introduced, with which children can learn electronics by mixing and matching blocks of electronics. Tangible user interfaces for learning developed at project partner Eindhoven University of Technology include, among others, tangible objects to stimulate learning social behavior for autistic children (Barakova et al. 2007), speech impediments (Hengeveld et al. 2008), communication (Soute et al. 2010), and programming.

However, most of the previous solutions feature tangible objects without any intelligence or can only be used for a single purpose. A problem with current open-ended playful blocks that do have some modular capabilities (e.g., LittleBits; Wyeth 2008; Soute et al. 2010) is that they conversely lack the needed scaffolding to teach the relevant concepts in a structured and engaging manner (i.e., as with serious games). Therefore we propose a framework where the iBlocks feed into an over-arching game system that contains scaffolding and challenge-based user guidance.

Some tangibles include game mechanics in a distributed fashion. Closest to our approach are the Sifteo cubes (Merrill et al. 2012b) which have been extended with games by Pillias and colleagues (Pillias et al. 2014) and more generally by the University of Luebeck in an educational gaming framework called Tangicons 3.0 (Scharf et al. 2012; Winkler et al. 2014b) to teach children algorithmic reasoning. However, Sifteo cubes are not primarily designed as an interactive learning environment. They are only motion aware and cannot be easily extended. Required sensors for the achievement of learning objectives, as intended by us, are missing. As of version two of the Sifteo cubes, games bought in the Sifteo store were only playable through a special Sifteo control box and longer directly from the PC, diminishing their potential for serious game developers. In addition, educators and even serious game developers are not supported in creating educational games with, e.g., tried and tested game mechanics and templates. Moreover, the development, sale, and support of Sifteo cubes were stopped.

Consequently, our concept is significantly more advanced than the state of the art by providing an open system that allows a large number of developers to develop pedagogically relevant cooperative learning games directly to the server side of the iBlocks system. Similar to the Network Environment for Multimedia Objects (NEMO) that holds semantically annotated information (Web 3.0) and which was developed by the project partner University of Luebeck (Lob et al. 2010) (Winkler et al. 2014), the system will be web based, i.e., a server allows the authentication and delivers always new games via a central game engine. This setup should empower serious game developers, educators both in formal and informal education, and researchers in developing the best educational instruments to develop technology literacy in children and young people.

The Tangible Serious Game (TSG) Concept

The main concept (see Fig. 1) of TSG is to build a smart environment consisting of physical objects enhanced with sensing, computing, and communicating capabilities (iBlocks) in order to support advanced and multimodal/multisensory interaction. The TSG approach will enhance a common learning environment (i.e., a classroom, an exhibition, the user home, etc.) with a system composed of a large shared screen or an interactive table controlled by a game engine connected to the Internet and a set of iBlocks that users can manipulate in order to play a game. A game engine IDE allows game designers and serious game companies to use iBlocks-like game objects. A game object is any object in a game that the user can see and interact with (e.g., power-ups, enemies, and walls are all game objects). The iBlock becomes a new type of game object to which the developer can attach aspects, events, rules, interaction modalities, etc. Games will be connected online through an Internet connection in order to allow players to interact remotely with pairs.

The iBlock Device

We define an iBlock as a cyber-physical object featuring a tangible aspect, and a set of programmable logic connection rules that specify how it can combine with other iBlocks when a user manipulates it. An iBlock will be a single device with a cubic shape and a size comparable with the hand of a young student. Users can grasp and manipulate iBlocks, move them in the space, and bring them closer/away from other iBlocks in order to follow the game rules provided through the shared screen/table. The movement of each iBlock and the composition of iBlocks have an effect also in the game virtual world controlled by the game engine. The real-world iBlock is

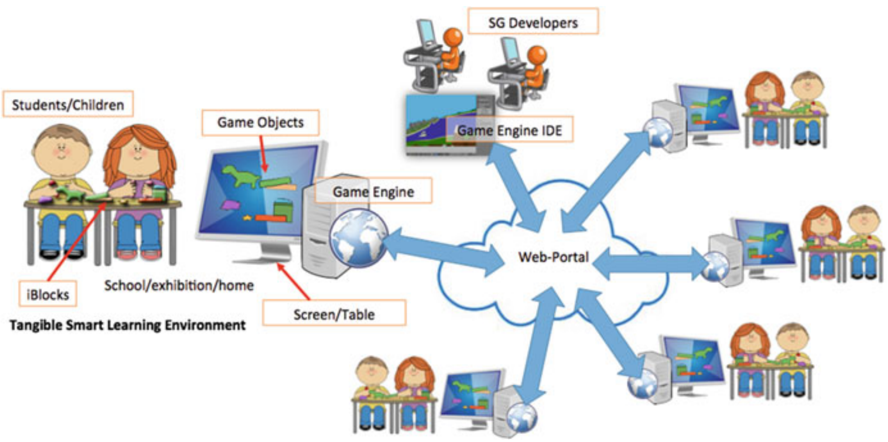
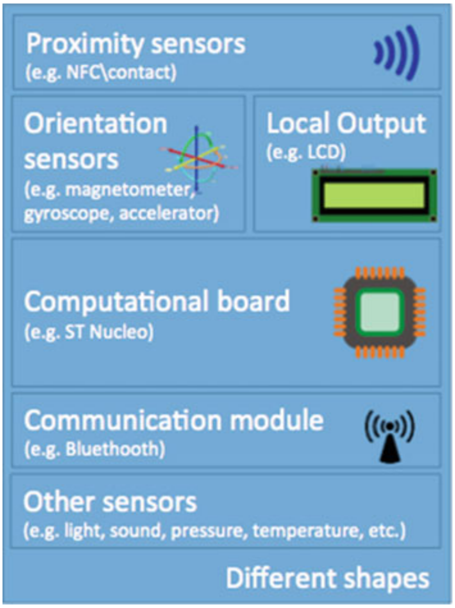


Fig. 1 Tangible serious game architecture

Fig. 2 The iBlock device



related to a virtual content (e.g., 3D shapes, 2D icons, texts, links) that can be accessed online through an extended game engine. Users can observe these effects in a big screen or interactive table that shows the virtual game world.

The virtual representation of an iBlock can be expressed by:

- Elementary geometric components (such as cubes, pyramids, spheres) that can be combined to realize more complex objects
- Elementary electronics sensors (e.g., to measure temperature, sound, lights) or functions (e.g., amplifying a signal, activating a function in response to an event)
- Complex predefined 3D objects, representing natural elements (e.g., trees, greens, flowers, sky) or human artifacts (e.g., buildings, cars, furniture elements)

The development of tangible games is based on a game engine IDE to connect iBlocks and game objects inside the virtual world and decide the game events in response to the player's real actions.

Inside an iBlock, we will have a set of sensing/computational/communication modules (see Fig. 2):

- Computational board (like an Arduino board)
- Positioning module (like an inertial measurement unit combining accelerometers, gyroscopes, and magnetometers), in order to understand its orientation in the space
- Proximity module (like an infrared short range sensor), in order to know the distance of others iBlock in the surroundings

- Wireless connection (like a Bluetooth module), in order to communicate with the game engine
- Other sensors that can also be available for some specialized iBlocks (such as ambient light sensor, temperature sensor, touchless gesture sensor)

Today, with standard boards and shields, a real implementation of an iBlock would cost about 40 euros for a standard version (with only computing, communication, proximity, and positioning capabilities) and about 65 euros for a more advanced version (with light, pressure, touch, sound, etc., sensing and I/O capabilities). This is an encouraging starting point for our concept, and – with the advent of the Internet of Things (28 billion of "things" connected by 2020) – we will see an even larger drop in the cost of sensors, processing power, and bandwidth (e.g., sensor prices have dropped to an average of 60 cents from 1.30 euros in the past 10 years, source Goldman Sachs, "IoT primer," September 3, 2014). On the basis of this preliminary evidence, we can foresee that over the next 5 years, the cost of the implementation of a real iBlock will fall on the order of 10 euros, making the described approach economically feasible.

Game Mechanics and Social Support

A special game engine allows the developer to connect to the iBlock device with the game virtual world and to use the iBlock as a standard game object inside his or her game. The game engine features an IDE with a visual editor through which a developer will be able to exploit iBlocks inside his or her own games. For example, the developer will be able to bind a game object (e.g., a 3D shape developed inside a game engine) to an iBlock by dragging and dropping the iBlock icon inside the interface of the game object in the game engine.

Game Mechanics will be the TSG instrument through which the SG developer can exploit iBlocks in order to create specific instances of these mechanics with own content and parameter values. A mechanic does not define the contents of a game, but the structure and the interaction modalities between players, standard game objects, and iBlocks. The mechanics will combine, in several meaningful ways, sensors and features of iBlocks. For example, an interaction mechanic could involve a 3D shape (a virtual object) inside the virtual game world attached to a specific iBlock in the real world. The player can control the virtual object by moving the related iBlock. Another example, a trigger mechanic could involve a condition in the real world (i.e., the contact between two iBlocks) with an event triggered in the virtual world. Mechanics would offer the access to all the parameters needed to manage the interaction inside a game engine. A game designer could choose a mechanic from the list available in the game engine IDE and then use the IDE to instantiate his or her own tangible game implementation by specifying the content (e.g., the virtual object and the events related). The game engine will feature an IDE with a visual editor through which a developer will be able to choose a mechanic from an existing list and insert his or her own content and parameters.

A web-portal provides interested users (developers, students, teachers, trainers) a list of tangible games made available to customers by game developers. Each game within the web-portal typically includes a description/categorization of the game, the requirements in terms of iBlock devices (type of sensors it needs), the subscription costs (if any), and how to download and deploy the game. Any interested user will be able to go to the web-portal to search for a specific game, see description and details before starting using it.

The IDE involves also collaborative features (e.g., sharing game with others, publishing results on social networks) in order to support collaboration among designers in building shared contents. The web-portal will also involve a forum where developers will be able to propose best practices and comment their experience with the system.

Early Evidence

For more than 8 years, researchers at the University of Luebeck worked on the concept of Tangicons, a noncompetitive learning game with tangible programming cubes (Winkler 2014). The aim of the game is to confront children of pre- and primary schools with logical-abstract problems in the context of fine and gross motor activities.

Tangicons has been developed together with children during codesign processes. So far we used self-made cubes with sensors and actuators, Sifteo cubes (Merrill et al. 2012), or recently also smartphones or mini-tablets as hardware.

In the game, each child is responsible for a physical programming block, which is part of a program line, moving a virtual object. Collectively mastering the situation leads to deep communication of the learners. A further important element of the learning game is that the children have to move also gross motor with the programming blocks in physical space.

Evaluation results indicate that children with an age of 5 years can create simple lines of code using Tangicons. They further show that children, who play in a group and must run in between, have more enjoyment of the game and are faster and livelier when programming.

With the third generation of Tangicons, we are working now on prototypes, where the physical programming blocks (or other small devices) are linked to a web-based semantic cloud system, including a narrator. Here the respective challenges in the learning game adapt to the individual abilities of a single child as well as the group of learners. This becomes possible by detecting the game history of each learner and generating appropriate narrative structures with this information for the tangible learning game.

While research results with the highly advanced software are promising, appropriate hardware is still missing to optimize the prototypical variants of the game for a real use in the context of pre-school and school.

Learning Scenarios

Students explores the concepts by interacting and enhancing their environments (that can be deployed in different contexts such as classrooms and dedicated exhibitions) through the creation of scenarios, which they built by manipulating/combining the iBlocks in order to discover and become familiar with STEM concepts. The platform, through the gaming interface, manages such sandbox physical smart learning environments, also providing assessment, feedback, and coaching. We show the kind of interaction between students and the iBlock device in the context of two examples about logical thinking and basic mathematical concepts.

Logical Thinking Example

The game design example adopts a strategy based on simple electronic circuits in order to introduce the player to logical concepts. The learner will be challenged to discover the right configuration of iBlocks (representing different logic operations) in order to solve quizzes and puzzles. Therefore, the learner might play the following roles: given a behavior, build the circuit implementing it; given a circuit that is hidden, find the stimuli that activate the response; and build your own circuits for others to solve. The logic circuit building will be only the language to specify the logic condition of the behavior rule, ignoring the electronic side. Higher-order negotiation and strategic learning goals can be added by setting up collaborative or competitive game goals, for instance, if two children are given different sensor and actuator iBlocks that can be placed in sequence and need to work together to solve a problem or to negotiate multiple end goals.

Basic Mathematics Example

This tangible game example revolves around learning basic mathematical and programming concepts and skills in order for children to create melodies and experiment with tempo changes. Main learning components are learning mathematical concepts pertaining to music and being able to apply these in the creation of new art. As an example, the game will allow the player to choose an instrument he or she wants to play (piano, violin, trumpet, flute, guitar, drums, etc.). Each iBlock will represent an instrument and the different players play different instruments but the same melody. The speed with which the iBlock is moved by the player dictates the tempo. A player chooses a tempo using his or her iBlock and registers it. The other players have to play the same tempo with their instruments (iBlocks). To check if the melody is in the same tempo, the one register and the one played by the user will be listening in the same time. The player could then try to fit the first melody. Thus, the player can test music notes moving in the space and the significance of the tempo. Other applications of the game could be to discover the physical basis for an instrument to play music. What is going on when you touch a cord? When you blow in a flute? What is important in the shape of an instrument to have better sounds?

Potential Impact

One of the potential impacts of the proposed approach is the creation of new business opportunities, in particular in the direction of improving the quality of learning impact of SGs in education through the creation of smart, tangible learning environments based on iBlock devices. This is expected to attract the interest of traditional toys companies that will find in the SG market an outlet of ever-increasing importance. From a business perspective, the game engine and the social support will bring to the world of tangible SGs the idea of the app stores, which have opened and are nurturing the flourishing market of applications in the mobile sector.

Traditional toys and games industries, the research community, and other technological developers (companies and practitioners) will have the opportunity to enhance their current product (e.g., toys and educational computer games) and findings (e.g., sensors) and sell them into a new market. In this way, SG developers will have key technological tools to transform the market potential (the educational potential of SGs is widely acknowledged) into a real growing business. Such business concepts are expected to concretely increase the exchange and collaboration between researchers, developers, and users/stakeholders in the related fields.

The iBlock prototypes, game mechanics, and games developed using the TSG system will be collected in the web-portal allowing teachers/trainers/students to have an authoritative reference point for the tangible sector. The availability of the game engine IDE and of the game mechanics will facilitate a widespread uptake of tangible-based game practices by allowing developers to more easily and confidently share contents, games, and proven educational best practices.

Around the iBlock devices, games, and web platforms, a community of developers and end users can be created, which is necessary for knowledge development, transfer, and diffusion about development of SGs based on tangible iBlocks.

In particular, the TSG approach is expected to:

- Support transfer of tools and interest from the toys sector to the increasing market of SGs
- Support production environment features – such as (re)use, maturity gaining, resilience, and scalability – to functionalities and sensors developed by research/industry labs
- Raise the quality and improve the flexibility and the educational value of typically gaming functionalities (giving the possibility of a more natural interaction)

From the point of view of the research, the proposed idea provides impact primarily by allowing definition and development of new devices (iBlocks) able to meet the requirements of a new generation of SGs. The TSG business facilitation tool (the web-portal) aims at enabling the important step from having research prototypes (iBlock devices, game mechanics, game examples) towards having more reliable and exploitable products. The web-portal will be a booster of an organization's research value by offering a more credible and convincing presentation as well as real implementation reference cases (real tangible serious games), supporting the commercialization of research outcomes (serious games and iBlock devices) in form of spin offs.

TSG idea can go beyond the pure STEM education. Core components build the core of all Cyber-Physical Systems (CPS) and Internet-of-Things (IoT) systems. The need of understanding this paradigm not only for children but also for employees, for example, in the logistics and manufacturing industries, is large, and there is a big potential for vocational training in this field. TSG is based upon a cocreative approach, and its components are suitable for being used for cocreative service development. This again will open up new possibilities not only for the toys and iBlock producers but also for a range of other organizations. The TSG concept mediate the IoT concept in a playful way and is therefore in addition also specifically suitable for awareness raising on new opportunities of using modern ICT. The use of games will provide access to user groups (e.g., children at risks, people with cognitive reduced capabilities, or cognitive disabilities) in a more engaging way than traditional teaching has.

The idea presented will implement tools for building a virtuous cycle, establishing strong and forward-looking connections between research and market.

The proposed concept in the long-term aims at a significant societal impact, by favoring a widespread adoption of SGs in learning and training settings (both formal and informal), trying to enhance the effectiveness as learning tool by adopting a smart learning environments approach. Teachers, educators, and trainers will have in their hands a catalog of tangible games that can be searched in terms of contents and parameters (such as sensors requirements) through the web-portal. This will have a significant impact, especially for localized items, such as school-trips and experiments tightly related to the teacher/class experience. This can have a positive influence on improving also adaptive and personalized learning, in order to adapt the environment to an inhomogeneous class with different level of technology confidence.

In a longer term, a further potential benefit is the pressure on toys industry for conceiving and deploying solutions (particularly, iBlock-like devices) progressively more affordable, less expensive, easier to use, and therefore more likely to migrate from a professional usage of educators with specific knowledge on experiential learning, serious games, and Game Based Learning to a wider group of teachers to deploy smart learning environments.

Also in the field of informal training, which often takes place in a different social environment (e.g., science centers), the new system and process can contribute to involvement of the user to a much higher extent than it is currently, especially thanks to the natural way of the tangible interactions style.

In both the formal and informal settings, a key impact is given by the fact that the system supports a variety of approaches in teaching and learning. In fact, developers will be able to choose among a set of different (and appropriately meta-tagged) examples of iBlocks use, supporting different learning theories and approaches. This is fundamental for stimulating interest and importantly ownership with one's own teaching and learning preferences.

A variety of examples of SGs using the tangible interface will be collected on the web-portal, supporting exchange of best practices and developer/user community building.

Barriers and Obstacles

It is possible that the tangible approach can be seen just as a set of technologies for interacting with a game: nothing more than a new interaction tool. The critical (wrong) point in this reasoning is that it confuses the technology (e.g., the iBlock device) with the whole development process that is proposed in the TSG concept. The TSG approach is enabled by the availability of the iBlock-like device and the web-portal but aims at decoupling the game from its current virtual world and move it in a more natural environment (the smart learning environment), in order to allow teachers to focus on the real added value of the game play (e.g., a content presented, a skill showed, a compelling mechanic for teaching a mathematical concept), while using a new multimodal/multisensory interaction technology.

A successful approach adoption requires a cultural shift in the way SGs are designed and developed. The tangible game concept requires a shift from the traditional virtual-world-based style of development SGs to a more real-iteration approach. For the concept success, it is important to analyze and explain the return-of-investment of the paradigm change in order to facilitate the adoption.

A number of issues are to be considered in order to make the TSG idea a real case: What kind of tangible games are developers prepared to pay for? How much would they be willing to pay? What kind of data can the system collect? Will service quality and user charges be monitored and regulated? How could SMEs profit from managing the iBlocks production? What flexible fare schemes could be devised in order to meet different developer requirements and different willingness to pay? Could a freemium approach be possible? What could be offered for free and what should be charged?

Conclusion

The goal of the proposed concept is to drive future research in order to improve the quality of learning impact of SGs in education through the creation of smart, tangible learning based environments. This is expected to attract the interest of traditional toys companies that will find in the SG market an outlet of ever-increasing importance. The chapter defines a new multimodal/multisensory interaction technology that will become the basis for the development of the new tangible serious game concept. From a business perspective, the web-portal will bring to the world of tangible SGs the idea of the app stores, which have opened and are nurturing the flourishing market of applications in the mobile sector.

Traditional toys and games industries, the research community, and other technological developers (companies and practitioners) will have the opportunity to enhance their current product (e.g., toys and educational computer games) and findings (e.g., sensors) and sell them into a new market. In this way, SG developers will have key technological tools to transform the market potential (the educational potential of SGs is widely acknowledged) into a real growing business. Such business concepts are expected to concretely increase the exchange and collaboration between researchers, developers, and users/stakeholders in the related fields.

The iBlock concept, the game mechanics, and the games developed using the system will be collected in the web-portal allowing teachers/trainers/students to have an authoritative reference point for the tangible sector. The availability of the game engine IDE and of the game mechanics will facilitate a widespread uptake of tangible-based games practices by allowing developers to more easily and confidently share contents, games, and proven educational best practices.

Around the iBlock devices, games, and web platforms, a community of developers and end users will be created and nurtured, which is necessary for knowledge development, transfer, and diffusion about development of SGs based on tangible iBlocks.

The TSG approach is expected to raise the quality and improve the flexibility and the educational value of typically gaming functionalities (giving the possibility of a more natural interaction).

In order to demonstrate the impact of the proposed approach, it is needed to validate real implementation of iBlocks in terms of providing educational added-value in authentic contexts of formal and informal learning. Planned future work and improvements include developing a real implementation of the tangible iBlocks-based game concept in several learning scenarios, bridging formal and informal situations. The developed platform will be validated in terms of provided educational benefit in several authentic contexts of use: primary and lower secondary schools. Students will be involved in hands-on activities. In particular, the focus will be on introducing children to basic topics related to basic STEM concepts.

Finally, the TSG concept can go beyond the pure STEM education. TSG core components build the core of all CPS and IoT (Internet-of-Things) systems. The need of understanding this paradigm not only for children but also for employees, for example, in the logistics and manufacturing industries, is large, and there is a big potential for vocational training in this field. TSG is based upon a cocreative approach, and its components are suitable for being used for cocreative service development. The TSG concept mediates the IoT concept in a playful way and is therefore in addition also specifically suitable for awareness raising on new opportunities of using modern ICT. The use of games will provide access to user groups (e.g., children at risks, people with cognitive reduced capabilities, or cognitive disabilities) in a more engaging way than traditional teaching has.

Recommended Reading

S. Arnab, R. Berta, J. Earp, S. de Freitas, M. Popescu, M. Romero, I. Stanescu, M. Usart, Framing the adoption of serious games in formal education. Electron. J. e-Learn. (Special ECGBL Issue) **10**(2), 159–171 (2012)

S. Bakker, E. Van Den Hoven, A.N. Antle, MoSo tangibles: evaluating embodied learning. in *Proceedings of the Fifth International Conference on Tangible, Embedded, and Embodied Interaction* (ACM, 2011), pp. 85–92

E. Barakova, G. van Wanrooij, R. van Limpt, M. Menting,). Using an emergent system concept in designing interactive games for autistic children. in *Proceedings of the 6th International Conference on Interaction Design and Children* (ACM, 2007), pp. 73–76

F. Bellotti, R. Berta, A. De Gloria, Designing effective serious games: opportunities and challenges for research. Int. J. Emerg. Technol. Learn. (IJET) **5**, 22–35 (2010). Special issue: creative learning with serious games

F. Bellotti, R. Berta, A. De Gloria, E. Lavagnino, A. Antonaci, F. Dagnino, M. Ott, A gamified short course for promoting entrepreneurship among ICT engineering students. in *Proceedings of IEEE International Conference on Advanced Learning Technologies (ICALT)* 2013, Bejing (2013a)

F. Bellotti, B. Kapralos, K. Lee, P. Moreno-Ger, R. Berta, Assessment in and of serious games: an overview. Hindawy Adv. Hum. Comput. Interact. **2013** (2013b). doi:10.1155/2013/136864

L. Bloom, J. Perlmutter, L. Burrell, “Let him know we are his friends”: applying constructivism to inclusive Classrooms. Intervention **34**(3), 132–136 (1999)

J. Brophy, T. Good, *Looking in Classrooms*, 9th edn. (Pearson Education, New York, 2003)

T. Bruce, *Learning Through Play: For Babies, Toddlers and Young Children*, 2nd edn. (Hodder Education, London, 2011)

J. Chen, Flow in games (and everything else). Commun. ACM **50**(4), 31–34 (2007)

M. Chen, S. Johnson, Measuring flow in a computer game simulating a foreign language environment (2004). Available online at: http://www.markdangerchen.net/pubs/flow_in_game_simulating_fle.pdf

Chi-Husiung Liang, Solving family communication problems between children and parents by using Mobile Serious Games. in Interactive Collaborative Learning (ICL), 2012 15th International Conference on,pp. 1,6, 26–28 Sept 2012

A. Clark, An embodied cognitive science? Trends Cogn. Sci. **3**(9), 345–351 (1999)

D. Clark, Learning by playing: can computer games and simulations support teaching and learning for post-16 learners in formal, workplace and informal learning contexts? Computer games in education and training. Presentation at LSDA seminar London, Nov 2003

T.M. Connolly, E.A. Boyle, M.H. Stansfield, T. Hainey, The potential of online games as a collaborative learning environment. J. Adv. Technol. Learn. (2007)

P. Corbeil, Learning from the children: practical and theoretical reflections on playing and learning. Simul. Gam. **30**(2), 163–180 (1999)

M. Csikszentmihalyi, *Flow: The Psychology of Optimal Experience* (Harper & Row, New York, 1990)

F. De Grove, P. Mechant, J. Van Looy, Uncharted waters?: exploring experts' opinions on the opportunities and limitations of serious games for foreign language learning. in *Proceedings of the 3rd international Conference on Fun and Games, Leuven*, Sept 2010

C. Dede, Comparing frameworks for 21st century skills. in *21st century Skills: Rethinking How Students Learn*, Solution Tree Press. vol. 20. (2010), pp. 51–76

Deloitte, Technology Trends 2012 – Elevate IT for digital business, online: http://www.deloitte.com/view/en_US/us/Services/consulting/technology-consulting/49ec911905f25310VgnVCM3000001c56f00aRCRD.htm

J. Dewey, *How We Think* (Heath, New York, 1933)

G.W. Fitzmaurice, H. Ishii, W.A. Buxton, Bricks: laying the foundations for graspable user interfaces. in *Proceedings of the SIGCHI conference on Human factors in computing systems* (ACM Press/Addison-Wesley, 1995), pp. 442–449

Gartner, Gartner Predicts Over 70 Percent of Global 2000 Organisations Will Have at Least One Gamified Application by 2014, press release, available online: http://www.gartner.com/it/page.jsp?id=1844115

F.L. Greitzer, O.A. Kuchar, K. Huston, Cognitive science implications for enhancing training effectiveness in a serious gaming context. ACM J. Educ. Res. Comput. **7**(3) (2007)

B. Hengeveld, R. Voort, C. Hummels, J. de Moor, H. van Balkom, K. Overbeeke, A. van der Helm, The development of LinguaBytes: an interactive tangible play and learning system to stimulate the language development of toddlers with multiple disabilities. Adv. Hum. Comput. Interact. **2008**, 1 (2008)

M.S. Horn, E.T. Solovey, R.J. Crouser, R.J. Jacob, Comparing the use of tangible and graphical programming languages for informal science education. in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (ACM, 2009), pp. 975–984

E. Hornecker, J. Buur, Getting a grip on tangible interaction: a framework on physical space and social interaction. in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (ACM, 2006), pp. 437–446

IDATE Serious Games Market Report 2012. http://www.idate.org/en/News/Serious-Games_643.html

International Technology Education Association, *Advancing Excellence in Technological Literacy: Student Assessment, Professional Development, and Program Standards* (International Technology Education Association, Reston, 2003)

C. Islas Sedano, M.B. Carvalho, N. Secco, C.S. Longstreet, Collaborative and cooperative games: facts and assumptions. in *Collaboration Technologies and Systems (CTS), 2013 International Conference on*, pp. 370, 376, 20–24 May 2013. doi:10.1109/CTS.2013.6567257

H. Jenkins, *Confronting the Challenges of Participatory Culture: Media Education for the 21st Century* (MIT Press, Cambridge, MA, 2009)

Y. Kafai, M. Resnick, Introduction, in *Constructionism in Practice: Designing, Thinking, and Learning in a Digital World*, ed. by Y. Kafai, M. Resnick (Lawrence Erlbaum Associates, Mahwah, 1996), pp. 1–8

P.A. Kirschner, J. Sweller, R.E. Clark, Why minimal guidance during instruction does not work: an analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. Educ. Psychol. **41**(2), 75–86 (2006)

E. Klopfer, S. Osterweil, K. Salen, *Moving Learning Games Forward, Obstacles Opportunities & Openness, The Education Arcade*, Massachusetts Institute of Technology, (2009), available online at: http://education.mit.edu/papers/MovingLearningGamesForward_EdArcade.pdf.

D.A. Kolb, *Experiential Learning* (Prentice Hall, New York, 1984)

D.S. Levin, T.K. Ben-Jacob, M.G. Ben-Jacob, The learning environment of the 21st century. AACE J. **1**(13), 8–12 (2000)

S. Lob, J. Cassens, M. Herczeg, J. Stoddart, NEMO: the network environment for multimedia objects. in *Proceedings of the First International Conference on Intelligent Interactive Technologies and Multimedia* (ACM, 2010), pp. 245–249

M.J. Mayo, Video games: a route to large-scale STEM education? Science **323**(5910), 79–82 (2009)

T.S. McNerney, From turtles to tangible programming bricks: explorations in physical language design. Pers. Ubiquit. Comput. **8**(5), 326–337 (2004)

D. Merrill, E. Sun, J. Kalanithi, Sifteo cubes. in *CHI'12 Extended Abstracts on Human Factors in Computing Systems* (ACM, 2012), pp. 1015–1018

M. Milrad, Using construction kits, modeling tools and system dynamics simulations to support collaborative discovery learning. Educ. Technol. Soc. **5**(4), 76–87 (2002)

A. Mitchell, C. Savall-Smith, *The Use of Computer and Video Games for Learning. A Review of the Literature* (The Learning and Skills Development Agency, London, 2004), p. 57

D. O' Broin D, Using a criteria-based user model for facilitating flow in serious games. in *Third International Conference on Games and Virtual Worlds for Serious Applications*, VS-Games 2011

M. Oliveira, G. Cerinsek, H. Duin, M. Taisch, Serious gaming in manufacturing education. in *Serious Games Development and Applications, SGDA 2013, LNCS 8101*, ed. by M. Ma et al. (2013), pp. 130–144

C Pillias, R Robert-Bouchard, G Levieux, Designing tangible video games: lessons learned from the sifteo cubes, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 3163–3166, April 26-May 01, 2014, Toronto, Ontario, Canada.

S. Papert, *The Children's Machine: Rethinking School in the Age of Computer* (Basic Books, New York, 1993)

D. Parsons, K. Petrova, Hokyoung Ryu, Mobile gaming – a serious business!. in *Wireless, Mobile and Ubiquitous Technology in Education (WMUTE), 2012 I.E. Seventh International Conference on*, pp. 17, 24, 27–30 Mar 2012

Pew Research Center, The Future of Gamification http://pewinternet.org/Reports/2012/Future-of-Gamification.aspx

M. Prensky, Digital game-based learning. ACM Comput. Entertain. **1**(1) (2003)

K. Raghavan, M.L. Sartoris, C. Schunn, K. Scott, Middle-school students' perceptions and interpretations of different model types. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Dallas

M. Rauterberg, M. Fjeld, H. Krueger, M. Bichsel, U. Leonhardt, M. Meier, BUILD-IT: a planning tool for construction and design. in *CHI 98 Conference Summary on Human Factors in Computing Systems* (ACM, 1998), pp. 177–178

M. Resnick, F. Martin, R. Berg, R. Borovoy, V. Colella, K. Kramer, B. Silverman, Digital manipulatives: New toys to think with. in *Proceedings of CHI '98, conference on human factors in computing systems,* Los Angeles, 1998, pp. 281–287

J.P. Rowe, L.R. Shores, B.W. Mott, J.C. Lester, Individual differences in gameplay and learning: a narrative-centered learning perspective. in *Proceedings of the Fifth International Conference on the Foundations of Digital Games (FDG)*, Monterey, 2010

T. Sapounidis, S. Demetriadis, Tangible versus graphical user interfaces for robot programming: exploring cross-age children's preferences. Pers. Ubiquit. Comput. **17**(8), 1775–1786 (2013)

F. Scharf, T. Winkler, C. Hahn, C. Wolters, M. Herczeg, Tangicons 3.0: an educational non-competitive collaborative game. in *Proceedings of the 11th International Conference on Interaction Design and Children* (ACM, 2012), pp. 144–151

I. Soute, P. Markopoulos, R. Magielse, Head up games: combining the best of both worlds by merging traditional and digital play. Pers. Ubiquit. Comput. **14**(5), 435–444 (2010)

K. Squire, Open-ended video games: a model for developing learning for the interactive age, in *The John D. and Catherine T. MacArthur Foundation Series on Digital Media and Learning*, ed. by K. Salen (The MIT Press, Cambridge, MA, 2008), pp. 167–198

P. Sweetser, P. Wyeth, GameFlow: a model for evaluating player enjoyment in games. ACM Comput. Entertain.**3**(3) (2005)

J. Sweller, Cognitive load during problem solving: effects on learning. Cogn. Sci. **12**(2) (1988)

J. Tashiro, What really works in serious games for healthcare education, Conference on Future Play on @ GDC Canada, Vancouver, 2009

B. Trilling, C. Fadel, *21st Century Skills: Learning for Life in Our Times: Learning for Life in Our Times*. (John Wiley & Sons, 2009)

B. Ullmer, H. Ishii, Emerging frameworks for tangible user interfaces. IBM Syst. J. **39**(3.4), 915–931 (2000)

E.D. Van Der Spek, Towards designing for competence and engagement in serious games, in *Serious Games Development and Applications* (Springer, Berlin/Heidelberg, 2012), pp. 98–109

R. Van Eck, Digital game-based learning: it's not just the digital natives who are restless. EDUCAUSE Rev. **41**(2) (2006)

I. Vekiri, A. Chronaki, Gender issues in technology use: perceived social support, computer self-efficacy and value beliefs, and computer use beyond school. Comput. Educ. **51**(3), 1392–1404 (2008)

V. Wendel, M. Gutjahr, S. Göbel, R. Steinmetz, Designing collaborative multiplayer serious games. Educ. Inf. Technol. **18**(2) (2013)

T. Winkler, F. Scharf, C. Hahn, C. Wolters, M. Herczeg, Tangicons: a tangible educational game with cognitive, motor and social activities. i-com: J. Interact. Coop. Media **13**, 47–56 (2014a)

T. Winkler, F. Scharf, M. Herczeg, Ambient learning spaces. Informatik Spektrum **37**, 445–448 (2014b). Special issue: interaction beyond the desktop

P. Wouters, C. Van Nimwegen, H. Van Oostendorp, E.D. Van Der Spek, A meta-analysis of the cognitive and motivational effects of serious games. J. Educ. Psychol. **105**(2), 249 (2013)

P. Wyeth, How young children learn to program with sensor, action, and logic blocks. J. Learn. Sci. **17**(4), 517–550 (2008)

M.F. Young, S. Slota, A.B. Cutter, G. Jalette, G. Mullin, B. Lai, M. Yukhymenko, Our princess is in another castle a review of trends in serious gaming for education. Rev. Educ. Res. **82**(1), 61–89 (2012)

E. Zimmerman, Gaming literacy: Game design as a model for literacy in the twenty-first century. Video Game Theory Reader **2**, 23–31 (2009)

M. Zyda, From visual simulation to virtual reality to games. IEEE Comput. (2005)

M. Zyda, J. Hiles, A. Mayberry, C. Wardynski, M. Capps, B. Osborn, R. Shilling, M. Robaszewski, M. Davis, Entertainment R&D for defense. Comput. Graph. Appl. IEEE **23**(1), 28 (2003). 36

Serious Games and Their Application in Creating Corporate Identity

26

Magdalena Bielenia-Grajewska

Contents

Abstract

The aim of this contribution is to investigate the role of serious games in creating corporate identity. The author focuses on showing how serious games determine

M. Bielenia-Grajewska (✉)
Intercultural Communication and Neurolinguistics Laboratory, Department of Translation Studies, Institute of English, Faculty of Languages, University of Gdansk, Gdańsk, Poland
e-mail: angmb@ug.edu.pl

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_53

the way in which organizations behave in the modern reality. To narrow the scope of the research, the author coins the concept of *online ludic corporate identity* to study the role of serious games in creating and maintaining organizational personae on the web and its image in the eyes of the broadly understood online users. The theoretical discussion is supported by a case study on the way in which selected organizations offer serious games at their websites to foster knowledge on monetary issues among different players.

Keywords
Serious games • Corporate identity • Gamification • Education

Introduction

There are different factors that determine the growing importance of games in modern reality. One of them is the increasing role of technology in one's private and professional life that leads not only to more advanced techniques of producing goods and services but also to the creation and application of social online networking tools, websites, and Internet messengers in different aspects of life. The omnipresence of *homo technologicus* in the twenty-first century is also visible in the sphere of games that are used nowadays, available not only for computer users but also for the ones who communicate by smartphones or telephones. Moreover, the application of games in various disciplines has been undergoing key alternations. The disappearing line between serious activities and entertainment is called *entertainment renaissance* by Nakatsu (2010). In the past, games were mainly associated with entertainment, whereas nowadays they are also a part of one's professional life. This change has also led to diversified scopes of game functions, from mainly the entertaining ones to the educational and professional applications. Thus, the *homo ludens* of the twenty-first century is different from the one discussed by Johan Huizinga almost 80 years ago (1938/1950). As has been mentioned, the ludic character of games is not restricted to the pure sphere of entertainment; nowadays, it is also conjoined with "serious" dimensions of life, such as education, profession, health service, or administration. Thus, games are not only associated with spending free time, but they also constitute a tool of creating, disseminating, and gaining knowledge as well as sharing experience. The mentioned changes have not only led to alternations at the individual sphere and one's personal attitude toward games; there are also visible at the level of organizations, making them more recognized on the competitive market, thanks to the application of ludic elements in their performance. The role of games in creating individual identity has been widely known. As an example, participation in games as well as the scoring received in them determines one's position within a group and the aptitude to practice certain skills or acquire new competence. To discuss the role of playfulness from both individual and social perspectives, the place of serious games in the life of modern companies is presented, and the way they create the organizational identity in the eyes of the broadly understood stakeholders constitutes the topic of investigation.

Games: Basic Notions

Taking into account the categorization by Caillois (2001), such notions as *free*, *separate*, *uncertain*, *unproductive*, *governed by rules*, and *make believe* can be used to discuss games. The first feature is called *free*; it means that playing is not obligatory. The second aspect is named *separate*, which refers to games that exist within time and space limits and are defined before one starts playing. The third element is called *uncertain*, meaning that the results of playing are not known before games end and players have initiative as well as influence over game results. The fourth notion is called *unproductive,* which is used to denote that nothing is produced in the process of playing. The fifth feature is referred to as *governed by rules*, understood as having own regulations, apart from the ones valid in a given system, applicable for a particular game. The last characteristic notion*, make believe*, is connected with the existence of a second reality (Caillois 2001). Another important feature of games is their *discernibility*; it means that game results are made available for players in a clear way. Another key notion is *integration*, stressing that elements are an integral part of a larger system (Salen and Zimmerman 2004).

The functions of games can be viewed through the notions mentioned by McGonigal (2015): *you are stronger than you know*, *you are surrounded by potential allies*, *you are the hero of your own story* and *you can make the leap from games to gameful*. The first one, *you are stronger than you know*, is associated with alleviating the physical or psychological pain in games and making one more prone to difficulties and problems encountered in private or professional life. The second one – *you are surrounded by potential allies* – stresses that games facilitate the strengthening of human relations, in both virtual and real life. The third one called *you are the hero of your own story* is the inner ability of relying on such heroic features as determination, willpower, or compassion. Finally, games offer the possibility of using game-created and game-exercised features in real-life situations.

Serious Games

Serious games can be defined as the games that possess other purposes than only pure enjoyment, and they are used to serve educational, professional, cultural, or social functions. There are some scholars who do not use the term serious games and propose alternatives such as *games for health*, *game-based learning*, or *persuasive games* (Ritterfeld et al. 2009). Serious games originate from simulators for the military and the creation of the first flight simulator in the 1920s. The next important stage in the development of serious games was the appearance of, e.g., computer chess program created by Claude Shannon in the 1940s. Moreover, the rapid evolution of games also took place in medicine, with the introduction of computer-aided instruction (CAI) in the 1960s. Since the first technological achievements were too expensive for hospitals and medical centers, the National Library of Medicine sponsored in 1972 the creation of a nationally accessible network of CAI. The next important year for the development of medical serious games was the introduction of the IBM PC in 1981. Another application of serious games in the 1980s was the first courseware created by the Harvard's Decision Systems Group called *HeartLab* to instruct cardiac auscultation (Bergeron 2006).

Serious games can be divided into subcategories by taking into account the domain they are applied in. Thus, such serious games are health games, corporate games, educational games, sport games, or foreign language games. Another distinction is connected with the presence or lack of presence of experts who supervise and/or control a game. In the case of serious games used for medical purposes, players may be supported by medical specialists or doctors. Different applications of serious games can be discussed. These include communication, mainly the presentation of complex content, interaction with the aim of supporting an external activity, such as going on a diet, or coaching, and emotional involvement, all of which may be important for social or promotional reasons (Maestri et al. 2015). Their applicability can also be investigated from the perspective of disciplines. Serious games and their applications in different fields of studies have been researched by different scholars. One domain is medicine, in particular the employment of serious games in the sector of care home and hospitals (Pannese and Morosini 2014). An example is the game *Snow World*, developed by the University of Washington to help patients after burn injuries. It has been proved that the mentioned game reduces pain and even facilitates recovery (McGonigal 2015). Serious games are important for education (Egenfeld-Nielsen et al. 2011). Westera (2015) in his contribution discusses whether games influence motivation and foster cognitive flows and learning. The teachers' perspective is represented by studying the relation between games and monitoring performance and their contribution to modern learning and social learning, and it draws the attention of readers to potential dangers connected with overusing games for educational purposes. Application of serious games is also visible in business (Correia and Carrasco 2012; Masci 2012). They are used in academia, providing students with new ways of acquiring and using knowledge in practice, e.g., in managerial simulations. The second way is to teach managers novel approaches and new skills. Apart from the mentioned educational perspectives, serious games may be used as an instrument of creating and sustaining company image or representation. Such an approach offers a complex study on the application of serious games in business settings. Serious games may also facilitate the involvement of individuals and companies in corporate social responsibility. An example may be the study on the game "SimGreen," described by Zhang and Zwolinski (2015), used to stimulate environmental actions of companies.

5 P Model of Studying Serious Games

After studying different serious games and their characteristics, there are some key functions that can be observed. The 5 P model has been created by the current author to discuss them in greater detail. The first letters stand for the functions of serious games related to problem-solving, personae, people, proficiency, and persuasion.

Problem-Solving

The aim of serious games is to enhance people's capacity for making decisions and confronting challenges. Problem-solving is related to key game processes, such as information processing and decision making. Since information flow in organizational gaming is not restricted in terms of roles or corporate hierarchy, users face different tasks and come up with their own solutions. In addition, serious games are often not restricted to internal users, but their task is to stimulate, e.g., certain reactions among external gamers. Thus, the level of knowledge required to successfully use a game should be carefully planned by game designers, taking into account the level of knowledge represented by gamers and their expectations toward the game. Gamers have the option of searching for solutions while playing, and consequently later they may find it easier to face difficulties in real-life situations. Creative ideas often appear in situations not necessarily focused on solving real problems, and thus the way they cope with everyday reality can be more effective.

Personae

Symbols, coined and exchanged in organizational dialogue, form corporate identity. Analyzing the linguistic dimension, they lead to the creation of names, phrases, and slogans characteristic of a given organization. In the long run, communication works as the verbal representation of organizational personae, being recognized by both workers and broadly understood by other key stakeholders. In some companies the use of symbols in serious games is designed and exercised. Gamers are embedded in a given organizational reality and consciously or subconsciously learn about organizations or companies and their market offerings. In addition, serious games may resemble a transitional phase between work and free time. The mentioned function of rites of passage serves different functions. One of them is to link different stakeholders to the performance of organizations in their free time. Secondly, this method of creating company identity turns out to be efficient since information is encoded and delivered in an entertaining way.

People

Games may become a form of ritualized communication in organizations. The activities connected with gaming become an element of organizational routines, making pauses between difficult and exhausting periods of work. Since games are relaxing and hierarchy becomes of secondary importance, strict organizational divisions play a secondary role in interactions. In this case, employers have the chance to meet their workers in an informal way, understanding their fears and problems connected with organizational performance. In games, stress is released and relations are easier formed, on both internal and external dimensions, e.g., between stakeholders and company workers or between external gamers themselves.

Proficiency

Serious games facilitate the creation and learning of new information. In addition, information is often disseminated more efficiently through channels and participants taking part in games. The reasons for this state are as follows. First, games often allow for free interaction and expression of comments; participants in games are more likely to share criticism since a friendly and relaxed atmosphere facilitates exchanging not only supportive comments about the interlocutor(s) but also the ones representing a completely different point of view. In addition, serious games rely more on social networking tools than other forms of organizational learning. Online tools of creating and disseminating knowledge offer information in both synchronous and asynchronous ways. In other words, gamers can choose when they want to play and learn.

The most popular distinction of knowledge is the tacit-explicit dichotomy. Tacit knowledge is defined as the type of knowledge that is difficult to codify and describe, whereas explicit knowledge is connected with formal and often scientifically regulated and/or proved. Since serious games involve both formal, educational knowledge and informal and less codified one, serious games rely on both types of knowledge in creating and sharing data. Holden and Glisby (2010) list, on the basis of selected literature, several modes of social interaction that facilitate the process of transferring tacit knowledge: apprenticeship, brainstorming, coaching/ mentoring, conversation, e-roundtables, dialogues, internalization of practices, metaphorical discourse, organizational languaging, narratives/storytelling, and socialization. Using the mentioned categorization to describe tacit knowledge flows in serious games, these modes include metaphorical discourse, organizational languaging, narratives, storytelling, and socialization. Explicit knowledge in serious games is visible in offering codified knowledge in a more accessible way. In many cases, especially as far as serious games for the general public are concerned, they use simplified terminology or rely on diversified methods of displaying knowledge. Landers and Callan (2011) list the most important ways of learning in casual social games. First, dedicated social networking sites help motivate learners. Secondly, games are well integrated within the social context, being a part of it. Thirdly, motivation is connected with the recognition of rewards. The fourth element is feedback: possibly in the form of promptly granted rewards, provided immediately. The fifth notion is connected with the link between game awards and task difficulty. Moreover, learning goal is more important than performance. In addition, the acceleration rate is important since the easy rewards should evaluate toward more complicated ones. To finish, games should not be boring; thus, practical usability and human factors should be taken into account by game designers.

Persuasion

Some serious games are called persuasion games; their purpose is often to advertise products, evoke certain purchasing behaviors, etc. Analyzing persuasion from the educational perspective, a friendly and relaxed atmosphere facilitates adaptation and eagerness to accept changes. It may persuade gamers/learners to accept novel knowledge or approaches to an, e.g., economic problem.

Linguistic Dimension of Serious Games

The linguistic dimension of serious games can be investigated at different levels. One of them is the function of language in games, focusing on language and its influence on action. Thus, linguistic tools are directed at making players perform tasks while stimulating their active participation and decision-making processes. Moreover, the language of games shows that players are responsible for making decisions. Serious games studied from the linguistic level may be connected with the discussion on micro- or macro-levels. The micro approach is exemplified in focusing on words and their function in creating games as well as their later usage. The discussion may also involve metaphors, of both verbal and pictorial types. The linguistic dimension is also linked with the functions games play, e.g., with the way rules are explained. Effective communication is also connected with the level of immersion and the involvement in game-related activities. In literature, we encounter different studies on the linguistic dimension of serious games. For example, Kirkpatrick states that *games nearly always come wrapped in some fictional scenario, that is, elements of visual narrative will unfold on the screen: there will be a story with characters communicated by the game documentation, and the different elements of the game (sounds, style of visual representation,* etc.*) will be tinged by qualities that converge on this "story" as a setting for our action* (2015: 517).

Corporate Identity

In the literature, different terms are used to denote the image or representation of companies or organizations. When the target audience is treated as the key factor, corporate identity is directed at the external public since it concentrates on how a company is viewed by stakeholders, whereas organizational identity is internally oriented since it relies on employees' opinions (Cornelissen 2008; Hamilton and Gioia 2009; Vella and Melewar 2008; Whetten and Goldfrey 1998). Since the focus on the research is to investigate the role of serious games as an element of creating external image, the concept of corporate identity is used to denote a research perspective. To discuss the relation between corporate identity and serious games, at least two dimensions should be determined. The first is the macro perspective, looking at the coexistence of games and organizations as a complex and multilayer phenomenon. The second one is the micro approach, focusing on individuals and their usage of games. Another approach is the actor-network approach (Bielenia-Grajewska 2011) that highlights the role of living and nonliving entities in the performance of a studied phenomenon. For its part, corporate identity can be examined through different perspectives, including the linguistic one. The latest one includes discussions on symbolic communication in forming and sustaining the linguistic personae of organizations. An example is the investigation of company linguistic identity and its metaphorical dimensions, including purchasers, personnel, and products studied through the perspective of metaphors (Bielenia-Grajewska 2015a). Another approach is the perception of company identity as a multivocal one. An example can be the concept of *heteroglossic linguistic identity* (Bielenia-Grajewska 2013a), showing organizational linguistic representation as languages, dialects, and genres. The next perspective that stresses the coexistence of different elements is hybridism. *Hybrid linguistic identity* is understood as identity that does not only result from other identities but also has its own outstanding characteristics (Bielenia-Grajewska 2010b). A similar approach is presented by the representatives of connectivism:

> strictly speaking, every entity in the world is a distributed entity (save, perhaps, indivisible subatomic particles – and (in my view) these may exist only by virtue of a reverse distribution, consisting entirely of entities that are larger than they are, much like a point in a moiré pattern – but this is very speculative). Every entity is composed of additional entities, and the properties of the entity in question are not all mere reflections of the smaller entities, but rather, unique properties, that come into existence because of the organization of those entities. Thus the same collection of carbon atoms may result in very soft charcoal or a very hard diamond. (Downes 2008)

Ludic Corporate Identity

Corporate identity can be said to possess its own ludic character. The studies on the relation between computer games and human identity have been conducted by de Mul (2005) who elaborates on the concept of ludic identity and its resemblance to narrative identity. Since players' choices in games are grounded in the past and various types of experience, ludic identity is linked with other types of identities. The modern player is called by de Mul (2005) *homo volens* who can autonomously shape his or her life, being represented in, e.g., game interactivity. The concept of ludic identity has been used by the current author to create a term *ludic corporate identity* that reflects the characteristics of identity in organizational settings. It is represented, for example, in the role of play in different areas of organizational performance (Fig. 1).

Ludic corporate identity is shaped by the broadly understood stakeholders (as perceived by the organization). The first determinant encompasses customers. Ludism in organizational settings is connected with the ludic character of purchasing activities. Shopping for many customers is not just an act of buying new things; it is a ritual that should provide pleasure, accompanied by fun and relaxation. Ludism is also often included in products or services themselves that, apart from serving its primary function, should also be the source of fun.

The next group of stakeholders is workers. Ludism among employers is observed by using different types of humor. *Verbal organizational humor* involves telling jokes or funny stories as well as using comical names, whereas *nonverbal organizational humor* includes humorous gadgets in the office (such as mugs with funny slogans).

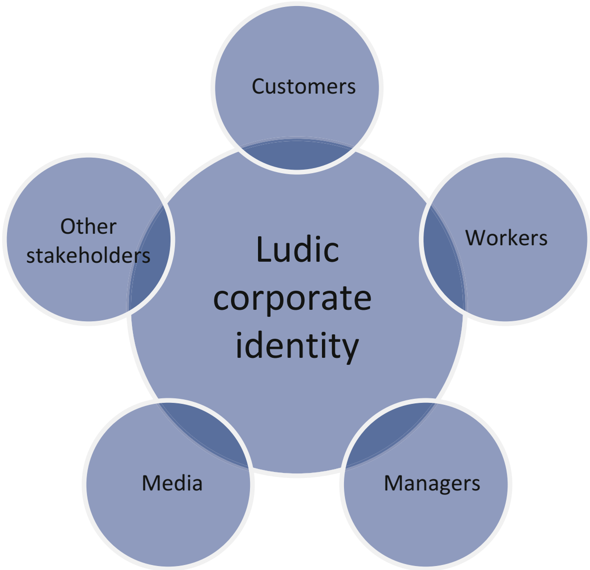


Fig. 1 Ludic corporate identity – stakeholder perspective (Created by the current author)

It should be stressed, however, that ludism is not connected with ridiculous or humiliating instances of humor at work.

The third group includes managers. Humor is used as an instrument of power (Bielenia-Grajewska 2010a) since the way humor is applied depends on one's position in a company and his or her place on the organizational ladder, with workers occupying lower place in hierarchy being less likely to joke about his or her boss. Humor becomes also an element used by managers in difficult situations, as a strategy directed at dealing with stress, risk, or changes in organizations. Managers who use humor at work are often perceived by their workers as more open and flexible.

Media belongs to the fourth group. Taking the organizational perspective into account, advertising in the twenty-first century relies to a large extent on ludism. Modern commercials and advertisement often use puns, humorous stories, funny names, and jokes to make the advertised product more visible on the competitive market. Other stakeholders are the last type of those involved in ludic corporate identity. This group includes those not necessarily directly involved in organizational activities or their recipients, such as local community and nongovernmental organizations. Taking local communities as an example, their festivities, rituals, and other forms of ludic expression may influence organizational performance.

Online Ludic Corporate Identity

There are many reasons why ludic corporate identity is often created and maintained online. The key determinant is technology and the growing role of the Internet in almost all spheres of modern life. To start with, the web offers both synchronous and asynchronous ways of communication. Thus, an organization may select the way it chooses to contact its stakeholders. For example, emails offer interlocutors the time to response, whereas many social networking tools favor immediate interaction. At the same time, organizational stakeholders are free to select the communicative tool that meets their needs and expectations. Moreover, online materials can be relatively quickly changed or updated. This feature is connected with the next important factor of online technology – its low cost in comparison with standard methods of communication, marketing, sales, and distribution. Thus, the Internet is an effective place of creating, using, and modifying serious games. The access to gaming offered online is not restricted; geographical and often economic barriers do not apply. Thus, ludic corporate identity should also take into account cross-cultural differences; what is funny in one culture may be offensive in another one. The relation with interaction with others (including to some extent ludism) in different cultures has been captured in various typologies used in cross-cultural management, such as the ones proposed by Hofstede, Gesteland or Trompenaars, and Hampden-Turner (detailed discussion on the mentioned typologies in, e.g., Holden 2002). Ludism is also an important part of modern online media, treating, e.g., humor as a bargaining card on the competitive media market.

Corporate Ludic Literacy, Corporate Gamer, and Corporate Ludic Environment

The discussion on the role of serious games in modern organizations should focus on three notions that facilitate the understanding how games influence both companies and individuals. The current author has coined the term *corporate ludic literacy* to denote the understanding of games and their functions for the broadly understood stakeholders. The concept stresses different attitudes to gaming that can be observed at both internal and external levels of companies. The proposed approach draws our attention to the fact that both employees and stakeholders differ in their perception of using games in their life. It should be stated that corporate ludic literacy and game-based learning in organizations reflect the situation determining competence required by modern companies. First, skills demanded from workers nowadays differ from the ones required in the past. Nowadays, collaboration, communication, competition, comprehension, adaptation, and innovation play a key role in becoming professionally efficient. Digital (computer and Internet) literacy is indispensable in many professions.

Another notion is called *corporate gamer*. This term is created by the current author to denote a person who uses games offered on the market by organizations. As far as the typology of corporate gamers is concerned, it encompasses both internal and external users as well as categorizes them by the purpose of relying on serious games. Different functions of serious games are presented in the empirical section of this paper.

Apart from the users of serious gaming in organizations and their ability to apply games, the surrounding of organizational game is also crucial in researching the role of serious games for modern organizations. Thus, the notion of *corporate ludic environment* has been created to denote all the features determining the creation, dissemination, and application of games in organizational settings. Among others, such notions as *technology-enhanced learning environment* (TELE) are important determinants shaping the position of games in organizations. In different works devoted to modern learning, the role of TELE in making learning more stimulating is stressed. The same can be observed in organizational settings. When organizations rely on technology in creating and disseminating professional knowledge, the aptitude to rely on serious games is also greater than in companies using standard forms of communication and knowledge sharing. One of the reasons is the positive experience with technological advances and the ease of using novelties.

Corporate Identity and Serious Games

The role of games and game-related elements in modern organizations has been discussed in the literature. For example, some research focuses on the advantages and disadvantages of gamification (e.g., Bielenia-Grajewska 2015b). There are different reasons why serious games can be used to help create effective corporate identity. One argumentation is the interactive character of games, offering stakeholders active participation in what organizations or companies do. Moreover, some authors even highlight the primacy of players over game creators, stressing the formers' effective role in determining the content of games and its popularity. The game environment stimulates competition and cooperation and facilitates group cohesiveness. Moreover, games encourage positive behavior in education, such as gaining goals, working on time, making decisions, and receiving feedback. The mentioned notions also decide for the growing popularity of such methods of acquiring knowledge. Apart from the learner dimension and the mentioned advantages of selecting this type of educational methods, also organizations and institutions benefit from applying games and game-related strategies. One of them is using games as a tool of gaining competitive advantage. The diminishing boundary between private and professional life is another reason for the popularity of applying games in modern education. The lines separating hobbies and work are blurred; professional activities involve elements of entertainment. The popularity of serious games in corporations is connected with different roles that modern companies play nowadays. The first objective is the economic one; the raison d'être of companies is to sell their products or services, satisfy customers, and make profits. Thus, the functions of games are to stimulate customers' interest in products and stimulate purchasing behaviors. As far as other stakeholders are concerned, games are, ideally speaking, an entertaining way of enhancing awareness of products and, in the future, strengthening the projected corporate identity in the eyes of potential stakeholders. The next function is the educational one, making individuals more conscious of a given phenomenon. Game functions may also be studied through the perspective of research levels. For example, micro dimension is connected with the role of serious games is shaping individuals, including employees and customers. The meso level focuses on organizations as such and the way in which gaming determines performance in the marketplace. The macro-sphere is connected with the way serious games influence more complex communities or nations.

Another way of looking at the functionality of serious games in the field of organizational identity is connected with the efficiency of reaching diversified stakeholders. Serious games that can be accessed at websites of organizations do not involve the necessity of intermediaries offering the product. Players can be reached more quickly, more cheaply and often more efficiently. Online versions of games also give organizations the possibility of changing the game content very quickly, adapting it, for example, to the changing economic reality. For companies, this option also offers product placement and advertising, making their services or merchandise more known among the target consumers.

Case Study

In the preceding discussion, the author sought to draw the readers' attention to the diversity of educational tools offered by companies and other organizations operating in different sectors. The mentioned selection of sources arises from technological progress and communication developments, and results in company websites being the most important repository of what a given company or organization does in the market place, including its educational function. As has been discussed above, serious games play an important role in the way companies create their company identity. To narrow the scope of the discussion, attention has been focused on games created by companies operating in the banking sector. Since the issue of monetary union in the European Union is a much discussed topic, the current investigation focuses on the use of games in educating different stakeholders which has been used by different institutions. Since games are the type of activity interesting for different people regardless of their age, institutions provide serious games for different age groups and level of specialized knowledge. To show the diversity of offered games, a corpus of 16 games has been gathered (see Appendix 1). The games selected for the purpose of the study are available in different languages, with some of them offered in multilingual versions. As far as the linguistic sphere of the study is concerned, games in such languages as English, French, German, Italian, and Polish were analyzed. The study was conducted during October, November, and December 2015. The selected games come from countries that are members of the monetary union as well as the ones that do not belong to it. The corporate identity of institutions and organizations in question takes into account the diversity of addressed stakeholders. First, users differ in terms of their age. It should be mentioned that education on the monetary issues is not restricted to adults and this principle is reflected in economic serious games. For example, *Euro Kid Corner*, issued by the European Commission, is aimed to teach children (over 7 years old) important information about the Euro. Apart from the European Union and national banks, such institutions as Il *Museo del Risparmio* in Turin offer the possibility of studying the savings by playing a game. The game *il gioco della formica* can be called the virtual version of snakes and ladders. It checks the knowledge of players on savings, economy, and finance. La *Fondazione per l'Educazione Finanziaria e al Risparmio* from Italy also has an offer on edutainment, dedicated at diversified users of economy. For example, *Economiamo* is a didactic game for children aged between 7 and 12 to gain knowledge on money issues. By contrast, *E-Islands* is dedicated to users from secondary schools who can learn the basics of becoming a successful entrepreneur who makes financial decisions based on proper planning. For its part, the Bank of England offers *Two Point Zero*, aimed at participants aged between 16 and 18. The main task of players is to act as a member of the Bank of England's Monetary Policy Committee and control the rate of inflation. Another game is the one created by *Banco de España*. In *Est@bilo* a player is supposed to check his or her knowledge on the Bank of Spain, by asking questions during the virtual visit in a library. Serious games also take into account the issue of linguistic diversity. For example, the game *Where's the coin from* is offered simultaneously in the 23 languages of the EU. The game does not only stress the euro identity but also the national identity of member countries since the user has to link the given currency with a proper country (one side of a Euro coin has national symbols). Thus, serious games strengthen the complex identity of euro highlighting its national and pan-national elements. For example, in the case of the Economic and Monetary Union (EMU), the role of heteroglossic identity is stressed, highlighting at the same time the single identity and the national identities of member states. Another function of serious games in relation to organizational identity is to strengthen the linguistic identity of organizations, highlighting that organizations cater for diversified linguistic needs. For example, the European Central Bank offers such games as *Top Floor*, *€CONOMIA*, and *Inflation Island*. Analyzing *€CONOMIA* (the monetary policy game) in greater detail, players have the opportunity to learn about the role of interest rates in the monetary policy of the ECB, by establishing themselves key interest rates and observing how they influence inflation rate. The game is available in 23 languages. The mentioned linguistic diversity in offering educational tools is also observable in countries that do not belong to the EMU. For example, the *Swiss National Bank* has created the game called *MOPOS* that is available in English, German, French, and Italian. The mentioned multilingualism of the discussed game is of course connected with the diversity of languages used by the Swiss. In addition, the offered linguistic diversity broadens the group of stakeholders to the ones who speak a given language as a second language. Additionally, it should be noticed that the educational dimension of games is not just restricted to gaining knowledge on monetary matters. The mentioned games created in different linguistic versions can also be used in educating the learners of foreign languages. Thus, another function of educating through organizational identity is advancing linguistic skills. Apart from the educational aspect of organizational identity, visible in fostering monetary or linguistic knowledge (or both), serious games also have social functions, such as forming or strengthening bonds. Attracting more stakeholders is realized by the functions of games; serious games facilitate the formation of communities. Analysis of the organizational dimension makes it possible to widen the group of potential stakeholders. It has also a more compound effect since the bonds that have already been established are strengthened in the course of playing the games. Another social aspect of serious games is connected with offering possibilities for players that are often hard to achieve in the “normal” life. For example, games often stress the change of status; players can get a new job, become promoted, or set up a company. It should be mentioned that serious games used in the monetary sector often fulfill several functions at the same time. For example, the National Bank of Poland offers the decision game “monetary policy.” The aim of the game is to develop competences in the practice of knowledge about monetary policy. The player is the member of the Monetary Policy Council who has, for example, to make decisions that aim to keep inflation close to 2.5 %. It should be stressed that this game serves many purposes. It shows the activity of banks, provides or deepens information on monetary issues, and offers players the opportunity to exercise decision skills.

The game *€CONOMIA* created by the European Central Bank can be presented as an example of the link between serious games and corporate identity. Educational functions are exercised by providing knowledge on the inflation rate that can be supported by reading press articles, monitoring one’s decisions and their impact on the monetary policy, as well as reading the comments made by other team members. The linguistic aspect is represented in many language versions of the mentioned game and specialized terminology embedded in the game, additionally provided in the glossary section which explains the economic terms used in the game. The social dimension is exercised through, e.g., *Hall of Fame* (listing the best players) and the section called *Your Team* where the opinions of other “team members” follow players’ decisions.

In comparison with other tools used to create corporate identity, serious games belong to the most effective ones. The reasons are as follows. First, they often rely on the web that offers diversified ways of creating and maintaining a given function. An example of online tools is websites that simultaneously offer different educational tools. For instance, the National Bank of Austria (Oesterreichische Nationalbank) provides different resources at its website, including simple games, crosswords, and quizzes. Secondly, as discussed, serious games offered online are an economic and efficient way of stressing different functions of organizations at the same time, such as social and educational ones. Thirdly, the multifunctionality of organizational identity is of a hybrid and synergic character, creating an organizational image that is not merely unique in its form, but also strong and effective.

Future of Serious Games and Organizations

De Gloria et al. (2014) underline that although serious games are attractive and motivating sources for learning, there is still much that should be done in order to make this educational instrument a more efficient tool of learning in different contexts. Observing the speed of technological advance, it can be predicted that serious games will become even more complex, offering ingenious ways of acquiring knowledge in different settings. Another important feature is the use of neuroscientific tools to examine behaviors and actions. Such an approach has already been used in analyzing serious games, for example, in the work by Hoffman et al (2011) who with colleagues, using fMRI scans, investigate the reduction of procedural pain among burn patients exposed to virtual reality analgesia. Neuroscience is also used in international neuromanagement, with such areas as international neurobusiness, international neurostrategy, neuromarketing, neuroentrepreneurship, and neuroethics (Bielenia-Grajewska 2013b). The future development of serious games is expected to follow the advancement in the sphere of economics as such. Thus, edutainment will become an important tool of teaching novel economic concepts to the broadly understood stakeholders.

Conclusion

The aim of this contribution was to discuss the role of serious games in creating organizational identity. The author has concentrated on organizations and institutions, namely, banks, that focus in their performance on teaching economic issues. To narrow the scope of the research, attention was focused mainly on the monetary notions. The mentioned serious games create corporate identity in a number of ways. One of them is the educational aspect of companies and organizations, offering not only opportunities to study the nuances of monetary union but also enhancing one's communicative or linguistic skills. Another key dimension is the social one, strengthening relations between organizations and stakeholders. As has been proved in this chapter, there are different reasons why serious games are a powerful instrument of creating and maintaining corporate identity. The most important one is connected with websites, being the place where corporate serious games are offered. As has been discussed, the online character of corporate gaming is responsible for its broad availability and multifunctionality, making serious game a crucial tool of establishing and sustaining dialogues with stakeholders on different levels. It can be predicted that in the future more organizations and institutions will rely on serious games in promoting their products, services, and ideas.

Cross-References

► A Tangible Serious Game Approach to Science, Technology, Engineering, and Mathematics (STEM) Education
► Towards Serious Games for Improved BCI

Recommended Reading

B. Bergeron, *Developing Serious Games* (Charles River Media, Hingham, 2006)
M. Bielenia-Grajewska, Humour in business education. Innov. Manag. J. **5**, 13–21 (2010a)
M. Bielenia-Grajewska, The linguistic dimension of expatriatism – hybrid environment, hybrid linguistic identity. Eur. J. Cross-Cult. Competence Manag. **1**(2/3), 212–231 (2010b).

M. Bielenia-Grajewska, A potential application of Actor Network Theory in organizational studies. The company as an ecosystem and its power relations from the ANT perspective, in *Actor-Network Theory and Technology Innovation: Advancements and New Concepts*, ed. by A. Tatnall (IGI, Hershey, 2011), pp. 247–258

M. Bielenia-Grajewska, Linguistic aspects of informal learning, in *Virtual Professional Development and Informal learning via Social Networks*, ed. by V.P. Dennen, J.B. Myers (IGI, Hershey, 2012), pp. 93–112

M. Bielenia-Grajewska, The heteroglossic linguistic identity of modern companies. Manag. Bus. Adm. Cent. Eur. **21/4**(123), 120–131 (2013a)

M. Bielenia-Grajewska, International neuromanagement: deconstructing international management education with neuroscience, in *The Routledge Companion to International Management Education*, ed. by D. Tsang, H.H. Kazeroony, G. Ellis (Routledge, Abingdon, 2013b), pp. 358–373

M. Bielenia-Grajewska, Company linguistic identity and its metaphorical dimensions. Purchasers, personnel and products through the perspective of metaphors, in *The Routledge Companion to Cross-Cultural Management*, ed. by N. Holden, S. Michailova, S. Tietze (Routledge, Abingdon, 2015a), pp. 170–179

M. Bielenia-Grajewska, Jak „rozegrać" zajęcia, czyli o roli gamifikacji w dydaktyce akademickiej, in *Wyzwania Współczesnej Pedagogiki*, ed. by D. Becker-Pestka, E. Kowalik (Wyższa Szkoła Bankowa, Gdańsk, 2015b), pp. 245–256

R. Caillois, *Man, Play, and Games* (University of Illinois Press, Urbana/Chicago, 2001)

J. Cornelissen, *Corporate Communication: A Guide to Theory and Practice* (Sage, London, 2008)

P. Correia, P. Carrasco, Serious games for serious business: improving management processes, in *Handbook of Research on Serious Games as Educational, Business and Research Tools*, ed. by M. Cruz-Cunha (Information Science Reference, Hershey, 2012), pp. 598–614. doi:10.4018/978-1-4666-0149-9.ch030

S. De Freitas, F. Liarokapis, Serious Games: a new paradigm for education? in *Serious Games and Edutainment Applications*, ed. by M. Ma, A. Oikonomou, L.C. Jain (Springer, London, 2011), pp. 9–23

A. De Gloria, F. Bellotti, R. Berta, E. Lavagnino, Serious Games for education and training. Int. J. Ser. Games. **1**(1) (2014). doi:http://dx.doi.org/10.17083/ijsg.v1i1.11

J. de Mul, The game of life. Narrative and ludic identity formation in computer games, in *Handbook of Computer Games Studies*, ed. by J. Goldstein, J. Raessens (MIT Press, Cambridge, MA, 2005), pp. 251–266

S. Downes, An introduction to connective knowledge, in *Media, Knowledge & Education – Exploring new Spaces, Relations and Dynamics in Digital Media Ecologies. Proceedings of the International Conference held on June 25–26, 2007*, ed. by T. Hug (Innsbruck University Press, Innsbruck, 2008), pp. 77–102

S. Egenfeld-Nielsen, M. Bente, B.H. Sørensen, *Serious Games in Education: A Global Perspective* (Aarhus University Press, Aarhus, 2011)

A. Hamilton, D.A. Gioia, Fostering sustainability-focused organizational identities, in *Exploring Positive Identities and Organizations. Building a Theoretical and Research Foundation*, ed. by L.M. Roberts, J.E. Dutton (Routledge, New York, 2009), pp. 435–460

H.G. Hoffman et al., Virtual reality as an adjunctive non-pharmacologic analgesic for acute burn pain during medical procedures. Ann. Behav. Med. **41**(2), 183–191 (2011)

N. Holden, *Cross-cultural management: a knowledge management perspective* (Pearson Education, Harlow, 2002)

N. Holden, M. Glisby, *Creating Knowledge Advantage. The Tacit Dimensions of International Competition and Cooperation* (Copenhagen Business School Press, Copenhagen, 2010)

J. Huizinga, *Homo Ludens. A Study on the Play-Element in Culture* (The Beacon Press, Boston, 1950)

G. Kirkpatrick, Ludefaction: Fracking of the radical Imaginary. Games Cult. **10**(6), 507–524 (2015), 1–18

R.N. Landers, R.C. Callan, Casual social games as serious games: the psychology of gamification in undergraduate education and employee training, in *Serious Games and Edutainment Applications*, ed. by M. Ma, A. Oikonomou, L.C. Jain (Springer, London, 2011), pp. 399–423

A. Maestri, J. Sassoon, P. Polsinelli, *Giochi da prendere sul serio. Gamification, storytelling e game design per progetti innovativi* (FrancoAngeli, Milan, 2015)

S. Masci, *Giochi nella formazione aziendale. Utilizzo di modelli di counseling integrato in azienda* (FrancoAngeli, Milan, 2012)

J. McGonigal, *SuperBetter: A Revolutionary Approach to Getting Stronger, Happier, Braver and More Resilient–Powered by the Science of Games* (Penguin Press, New York, 2015)

R. Nakatsu, Entertainment and its future, in *ECS 2010, IFIP AICT 333*, eds. by R. Nakatsu et al. (Springer, Heidelberg and Berlin, 2010), pp. 233–242.

L. Pannese, D. Morosini, Serious Games to support reflection in the healthcare sector. Int. J. Ser. Games **1**(3), 5–14 (2014)

U. Ritterfeld, M. Cody, P. Vorderer, *Serious Games: Mechanisms and Effects* (Routledge, New York, 2009)

K. Salen, E. Zimmerman, *Rules of Play: Game Design Fundamentals* (The MIT Press, Cambridge, 2004)

K.J. Vella, T.C. Melewar, Explicating the relationship between identity and culture. A multi-perspective conceptual model, in *Facets of Corporate Identity, Communication and Reputation*, ed. by T.C. Melewar (Routledge, Abingdon, 2008), pp. 3–34

W. Westera, Games are motivating, aren't they? Disputing the arguments for digital game-based learning. Int. J. Ser. Games. **2**(2) (2015)

D.A. Whetten, P.C. Goldfrey, *Identity in Organizations: Building Theory Through Conversations* (Sage, Thousand Oaks, 1998)

F. Zhang, P. Zwolinski, SimGreen: a serious game to learn how to improve environmental integration into companies. Proc. CIRP **29**, 281–286 (2015)

Appendix 1: Research Materials (Serious Games and Educational Tools Available Online)

€conomia. https://www.ecb.europa.eu/ecb/educational/economia/html/index.en.html. Accessed Oct-Dec 2015

Economiamo. Fondazione per l'Educazione Finanziaria e al Risparmio. http://www.feduf.it/container/scuole/economiamo. Accessed Oct–Dec 2015

E-Islands. Fondazione per l'Educazione Finanziaria e al Risparmio. http://www.feduf.it/feduf-content/eislands/gioco.php. Accessed Oct–Dec 2015

Estabilo. http://aulavirtual.bde.es/wav/html/estabilo/estabilo.html. Accessed Oct–Dec 2015

Euro Kids Corner. Available at: http://ec.europa.eu/economy_finance/netstartsearch/euro/kids/index_en.htm. Accessed Oct–Dec 2015

Il gioco della formica. Available at: http://ilgiocodellaformica.museodelrisparmio.it/. Accessed Oct–Dec 2015

Inflation Island. Available at: http://www.ecb.europa.eu/ecb/educational/inflationisland/html/index.en.html. Accessed Oct–Dec 2015

Inflationscockpit. Available at: https://www.oenb.at/docroot/inflationscockpit/gewinnspiel.html. Access: Oct–Dec 2015

MOPOS. https://www.iconomix.ch/en/resources/m04/. Accessed Oct–Dec 2015

Österreichische Nationalbank (OeNB). Available at: https://www.oenb.at/Ueber-Uns/Geldmuseum/Museumspaedagogik/Lehrmaterial.html. Accessed Oct–Dec 2015

Polityka Pieniężna. Narodowy bank Polski. Available at: http://www.nbportal.pl/rozrywka/gry/polityka-pieniezna/gra. Accessed Oct–Dec 2015

Sterowiec Inflacyjny. Narodowy Bank Polski. Available at: http://www.nbportal.pl/rozrywka/gry/sterowiec-inflacyjny/gra. Accessed Oct–Dec 2015

Target Two Point Zero. Available at: http://www.bankofengland.co.uk/education/pages/targettwopointzero/default.aspx. Accessed Oct–Dec 2015

Top Floor. European Central Bank. Available at: http://www.ecb.europa.eu/ecb/educational/topfloor/html/index.en.html. Accessed Oct–Dec 2015

Where's the coin from. Available at: http://www.new-euro-banknotes.eu/Educational-Publications/Where's-the-coin-from. Accessed Oct–Dec 2015

Wykrywacz Kłamstw. Narodowy Bank Polski. Available at: http://www.nbportal.pl/rozrywka/gry/wykrywacz-klamstw/gra. Accessed Oct–Dec 2015

Ethics, Privacy, and Trust in Serious Games 27

Rod McCall and Lynne Baillie

Contents

R. McCall (✉)
Environmental Research and Innovation, Luxembourg Institute of Science and Technology, Esch-sur-Alzette, Luxembourg
e-mail: roderick.mccall@list.lu

L. Baillie
Department of Mathematical and Computer Science, Heriot-Watt University, Edinburgh, UK
e-mail: l.baillie@hw.ac.uk

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_37

Abstract

This chapter presents a set of ethical concepts and related guidelines starting from a historical review perspective leading through to present day work on computer ethics, privacy, and trust. The review forms a basis for a discussion on ethics within serious games including related types such as persuasive games, gamification, and pervasive games. The objective being is to identify common issues and areas of interest where the boundaries between these types of game blur. The result is a set of concepts which specifically explore the gaps which have arisen, for example, in areas such competing values between stakeholders, the effect of serious pervasive games on non-participants, rules of play, transparency, consent, and autonomy. The guidelines contain sets of issues, questions, and examples that aim to alert researchers and practitioners to key concerns.

Keywords

Ethics • Privacy • Trust • Concepts • Guidelines

Introduction

Serious games are those which have a purpose beyond simply entertainment, and they are used across a variety of contexts ranging from traffic management (McCall and Koenig 2012), location-based historical games (Blum et al. 2012), cultural heritage (Chochliouros et al. 2013), and health. Quite often these games either seek to change behavior or teach people about a specific domain. However, serious games raise a set of unique challenges as they are no longer simply about entertaining users within a predefined gaming context but rather are focused on how they can affect the everyday life of their users, for example, influencing health, transportation, or educational choices. In turn such possible effects are directly impacted by the ethical, privacy, and trust aspects which are either explicitly identified by the developers or which may arise indirectly while playing the game. This in turn gives rise to professional, legal, moral, social, and other concerns (Floridi and Sanders 2002). They are also impacted upon by the nature of the game, for example, whether it is within a constrained classroom-type environment or whether it takes place in a more pervasive environment where limits and boundaries are less clear, e.g., on the road. However, regardless of the particular type of game, there is a need to consider the impact on the users and wider society as they increase in popularity and availability.

The chapter takes a predominantly practical position in that it seeks to provide researchers, practitioners, and developers with an overview of ethics, privacy, and trust issues and how these impact upon the area of serious games; it does not seek to debate the relative merits of different models of ethics (Floridi and Sanders 2002;

Versteeg 2013), or to assess the precise moral problems posed from any philosophical view point. Instead it seeks to inform and highlight issues that are relevant to the field. The chapter expands on discussions relating to data privacy and trustworthiness as these are of increasing importance and have an impact on ethics. As a result it recognizes that while the field of computer ethics has been emerging for some decades, there are some unique problems with serious games, and that while this chapter will alert the reader to potential issues, the list is (1) by no means exhaustive (2) that there are many other authors who have contributed concepts and guidelines and that it is (3) open to debate. However, in all cases these issues should be considered during the design, deployment, and testing of any serious game. The chapter begins by discussing the generic challenges proposed by serious games, and it then explores the background to ethics and generally accepted practice drawn from a range of backgrounds including medical and computer ethics. This is followed by a review of different game types and how ethical issues impact upon each one. The chapter closes with a set of ethical concepts and guidelines and a conclusion.

Background

The Ethics of Serious Games

Serious games are a growing area of work, and for many they represent a good way to improve the lives of people either by educating them, encouraging them to change their health and fitness patterns, or through improving their workplace performance (and hence job prospects). They use fun elements to influence either knowledge or behavior outside of the game itself. For example, a school child learning about simple arithmetic through a game should be able to benefit from the knowledge and skills learned in the game so that when they sit their end-of-year exam, their score improves. Likewise the person who is not exercising enough may use a persuasive game to encourage them to start jogging, while teams of employees in a call center may use gamification approaches to improve call-handling efficiency. There would, for example, be little point in a serious game that resulted in the child only being able to answer mathematical problems in a gamelike environment or a gamified application that sowed workplace discord even if corporate productivity improved as employees cheated or perceived it as a monitoring service. These are perhaps overly simplified examples, yet it is precisely these kinds of problems which ultimately must force us to consider a number of key issues. Prior work by Koenig et al. (2012) and McCall et al. (2012) succinctly summarized the core challenges with respect to pervasive traffic games and helped to start a discussion on three key concerns:

- **Ethics**: including aspects such as consent, risk, etc.
- **Privacy**: including aspects such as what data is collected, how it is stored, and how it is shared

- **Trust**: the end users' perception of whether they trust the system

While on the face of it, ethics may seem different from privacy, but the two are linked. For example, as noted later, a lack of transparency can lead to users being unable to see what data is being collected on them or how it is being used. Likewise users may not trust a system if they are unsure what they are consenting to, either in terms of direct use, what it may lead to later, or how their data may be shared.

A strict definition of serious games would probably mean that emerging areas such as gamification, persuasive, or pervasive games should be ignored; however, many of the challenges that exist within serious games are equally relevant to gamified applications. Furthermore, serious games are increasingly incorporating elements drawn from pervasive and persuasive games. As a result, this chapter will explore a plethora of serious games and related game "types" such as gamification, pervasive, and persuasive gaming. The basic premise is that serious games should be viewed as an umbrella term that can include pervasive and persuasive games as well as gamified applications. Indeed the view here is that any game or gamelike environment that results in a serious external benefit (to, e.g., health, education, or workplace behavior) should be considered within the category of serious games. There are, however, some limitations and caveats to this position, for example, not all pervasive games are serious games. Also, while it is debatable whether gamified applications are serious games, their end objectives are often quite similar, i.e., to influence behavior in an environment external to the game itself. The chapters draw on work mainly from psychology and related areas, primarily because these appear to be well developed, and also in many cases it is the duplicitous nature of serious games which leaves them open to debate in terms of ethics, privacy, and trust. A serious game may have the desired effect of improving a particular skill or physical problem. For example, in the case of exergaming, this may be hidden from the end user, and the effect of using the game may not be clear until much later either in terms of both learning that skill or the impact on their future lives.

Background

Ethics

This section will use as its starting point the ethical codes drawn from the APA (American Psychological Association) and the quotation later perhaps best illustrates the diverse areas which although drawn from psychology are often inextricably linked to serious games. The primary reason for starting from this angle is that such an approach initially focuses on the effect on humans, rather than the underlying technology. However, as Brey (2000) points out, this approach is not without its limitations; as in effect we are forcing moral and ethical codes from non-computer science areas onto computer science when in reality, wider society may have no clear understanding about the morality issues relating to a particular technology. In essence the ethical and moral approach in this sense relies on prior

Table 1 Overview of ethical sections proposed by the APA (2010) (*left*) and Brey (2000) (*right*)

Principle A: beneficence and nonmaleficence	Justice
Principle B: fidelity and responsibility	Autonomy and freedom
Principle C: integrity	Democracy
Principle D: justice	Privacy
Principle E: respect for people's rights and dignity	

experience of non-computer science areas simply because we have experience of them, and as a result we may end up being unable to see potential ethical issues in future. A good example would be the use of data collected by Google, Yahoo, and Microsoft by State Security Agencies such as the NSA (USA), GCHQ (UK), and the German Federal Security Ministry; many end users of these services would probably never have knowingly or explicitly agreed to such use. As a result Brey argues the technology is not a neutral part of the equation and indeed may be core to the effect such systems have on users. Nonetheless and despite these limitations, it is important at least at this stage to take a step back from the technology and focus on basic ethical issues.

> Areas covered include but are not limited to the clinical, counselling and school practice of psychology; research; teaching; supervision of trainees; public service; policy development; social intervention; development of assessment instruments; conducting assessments; educational counselling; organizational consulting; forensic activities; program design and evaluation; and administration. This Ethics Code applies to these activities across a variety of contexts, such as in person, postal, telephone, internet and other electronic transmissions. (Extract from APA Ethical Guidelines (American Psychological Association 2010))

In the quotation earlier, it can be seen that a range of end-user domains are highly relevant, e.g., conducting assessments. For example, serious games for use in education would fall under the above set of regulations as would any mobile game which uses the Internet. In Table 1 below, the ethical guidelines from the APA are listed against those of computer ethicist, Brey.

Readers are advised to consult the relevant section on the APA Website, and additionally the European Association of Psychological associations provides guidance (European Federation of Psychologists 2005) although covering similar topics, it reduces the principles to four key areas: (1) respect for a person's rights and dignity, (2) competence, (3) responsibility, and (4) integrity.

Broadly speaking, beneficence and nonmalfeance mean that the study should not harm the participants and that the end objective should be to benefit those with whom they work or, as Beuchamp (2013) said, beneficence is "all forms of action intended to benefit or promote the good of others" ((Beauchamp 2013), p. 1). As a starting point, this should provide a strong basis for any work in serious games; however, it should be noted that when commercial objectives come into play, there is a risk that the benefit is not for the user but for the company, which in turn creates ethical problems.

Fidelity and responsibility cover dealing with people, in particular, how (in this case) the psychologist views the impact on the participants and wider society and importantly to undertake such work for little or no personal or financial gain for themselves. Integrity relates to the need to behave in a way that is open, transparent, and honest. Importantly it recognizes that when deception is used that the benefits for the participants should outweigh the potential harm, indeed, where possible, there should be no harm. If there is a risk of harm, then the psychologist must correct any issues that could or may arise after the study.

Justice relates to the concepts of providing fairness and justice to all those involved in such work, in particular, ensuring that all have access to the benefits of psychology as a whole and the work undertaken. In the APA example, justice extends to the benefits of the studies being fairly distributed in the sense that the benefits accrue in a way that is fair to the participant, psychologists, and ultimately wider society. A good example here would be how to balance the effects of traffic congestion games to benefit both the individual driver and not just the state or wider society. From Brey's perspective, justice is based on the distribution of social goods such as freedom, democracy, and privacy. For example, where a system distributes such basic elements (or indeed others) unfairly or redistributes them in a way that is unfair, this is ethically and morally undesirable.

In a free society, people are often assumed to have the autonomy to decide for themselves what they wish to do; indeed as noted later this should not be coercive. However, as Brey (2000) notes, technology itself can impact on the ability of people to act autonomously and exercise a free choice. For example, addiction to computer games is a widely debated topic (Griffiths and Hunt 1998; Grüsser et al. 2007), and within this context a serious game could create a dependency culture in terms of how people exist both outside the game (e.g., what decisions they take) and also for that matter whether the game itself creates social and other benefits which result in some form of dependency.

The final general principle explores "respect for the rights and dignity" of the participants and starts to highlight the importance of issues such as privacy, confidentiality, and determination. It also explores welfare issues and specifically addresses aspects such as age, gender, and race that may impact on the study or indeed be subjects to bias from the psychologist. This overarching concept can also be extended to include Brey's last two points: democracy and privacy. In terms of democracy, he identifies that systems have the potential to redistribute power from those who should have control over the system to those who control the system. The concept of privacy is also relevant within the field of rights, and Brey describes privacy in two ways, *informational* (*data*) and *relational* (*to others and our environment*). Taking a purely information technology perspective, it is clearly desirable to allow people to control what they share and who they share it with. However, increasingly, as pervasive games and technologies become popular, there is a need to ensure that the other side of his argument is adequately considered, e.g., relational privacy. This is essentially the freedom to be left alone by others and in turn be free from observation and interference. In the age of social networking and the increasing posting of pictures online, it could be argued that this right has been eroded. Also in terms of impact, people who play serious games may be reducing the relational privacy of others even if those others are not aware of such impacts (see discussion on pervasive games in this chapter).

Floridi and Sanders (2002) provides a review and discussion of different approaches to computer ethics. At the lowest level, the NA (no resolution approach) sees attempting to define an area of computer ethics as pointless and without any convincing conceptual model. A more pragmatic approach would seem to be the PA (professional approach) which seeks to instill a set of values, rules, and judgments based on professional standards and a concern for the user of the computing artifact (Gotterbarn 1991 sourced from (Floridi and Sanders 2002)). The two contrasting ones are the radical approach (RA) and the conservative approach (CA). The former views computer ethics as nothing particularly unique and that certain classic metaethical issues are merely transformed into the ICT domain. The innovative approach (IA) seeks to build a middle ground between CA and RA by rejecting a purely metaethical approach from (CA) while adding to it the concepts from a more radical approach. In essence it recognizes the uniqueness of the ICT domain but seeks to build on existing ethical practice.

As can be seen later in this chapter,both existing ethical practices from the APA and the emergent ideas of Brey point to the need not only to base practice on externally already known challenges but increasingly the need to explore the specific new challenges posed by the various sub-types of serious games. As a starting point, it is worth considering the impact of such technologies on ethical areas such as consent, risk and harm, and coercion.

Consent

Many researchers and practitioners derive their underlying historical position of consent from the Nuremberg code, which followed on from the trials in the post-World War 2 in Germany. Although as noted by Vollmann and Winnua (1996), the notion of experimental ethics in medicine pre-dates the Nuremberg code with the Prussian Ministry of the Interior issuing guidance as far back as 1891. These historical issues aside, the Code from 1947 primarily arose as a direct result of the medical and other experiments which were conducted on human subjects, without their consent and often with horrific consequences. Examples of brutal testing practices included the experiments on twins undertaken by Joseph Mengele in Auschwitz. His scant regard for ethics included acting as the doctor who screened patients for fitness; those deemed unfit were sent to the gas chambers. Additionally and without consent, he would often use these screening to assess whether particular people could be used as test subjects. His overall scientific ethos broadly speaking was to prove the supremacy of the heredity over environmental influence on humans, in essence, to prove the supremacy of the Aryan race. In essence he was using science to prove a particular point without considering the effects of his work in the individuals concerned. He also did not even entertain the concept of consent either in form of even making people aware of the effects of the screening but also in alerting them to the inherent risks and harm.

While it is clear that his research resulted in acts of horrendous human suffering, they had the effect of alerting the wider population post-World War 2 to the need for strong ethical principles. As a result the code enshrined the principles that consent must be entirely *voluntary*, the person must have the mental *competency* to take the decision, must be able to *informed any risks to them, and must be able to comprehend* any impact on them (see Nuremberg code 1947). In essence, if Mengele had been alive today, he would have broken every key aspect of the Nuremberg code and more importantly would have deliberately harmed the participants.

> The growth of ubiquitous computing also poses some unique problems with respect to consent for example "if we continue to transpose existing models of consent to ubicomp, there is a danger that without attention, the design of pervasive systems will continue without adequate support for user consent." (Luger et al. 2013, Ubicomp, p. 5)

Luger and Rodden (Luger et al. 2013) identify consent within ubicomp systems as being particularly problematic as until now it has been given only passing consideration, which is in stark contrast to fields such as psychology or medicine (as can be seen from the Nuremberg code). Furthermore, within these more traditional fields, the notion of what people are consenting to is perhaps clearer due to the concept of consent being better understood, for example, the study may be of fixed length, and to explore card-sorting skills. In ubicomp, while lab studies during the early stages may limit some potential risks, for the system to be properly tested, it ultimately has to be used in the wild, with the net result that the parameters of the test and possible effect on participants and other people becomes more difficult to quantify. Thus the notion of what people are consenting to and the impact becomes vaguer, in particular, in the longer term when users' notion of consent may vary over time (Luger et al. 2013). The variation of consent over time is an increasing problem within the area of cloud computing services (which are often also used within location-aware games). For example, companies such as Google and Facebook require people to provide explicit agreement to use their services. However, over time, new services are added which make use of the data previously shared. Furthermore, the terms and conditions are also often modified over time. The effect being that users are essentially signing up to an open-ended consent agreement where both what data will be used for and for how long it is used remains extremely fluid. In many cases, EULAs are used to provide consent, and as Luger et al. (2013) notes these are often opaque and difficult to understand, while Luger and Rodden (2013) note that many go unread. Therefore while the act of giving consent to users is commendable, this is often undone by the sheer complexity of such documents, the net effect being that people sign up to something which they are failing to understand, thus breaking one of the key rules, i.e., comprehensibility. Furthermore, they are often little more than contracts outlining data protection issues from the perspective of what the service supplied desires (in particular protection from legal actions), which, although relevant to privacy, are largely ignoring ethics.

Risk and Harm

As noted under consent, a key issue is being able to quantify the risk to participants and making them aware of this. In essence identifying risks is about being aware of the potential harm that any serious game (or study) may result in. Traditionally harm prevention has tended to focus on the physical body, for example, through injury due to the test procedure. However, the definition of harm extends far beyond the physical body to anything that may have a negative impact on a person. As noted by Shahri et al. (2014), people can perceive enterprise gamification applications as having negative impacts ranging from spying through to increasing workplace tensions. Furthermore, they could be used to inform decisions about hiring, firing, and promotion. Thus even a simple serious game that aims to improve communication between employees could have detrimental effects, even if neither the designers nor employers intended this at the outset. Even cultural differences can play a role in the potential negative effects of serious games. In one anonymous example mentioned during the CHI 2013 conference in Paris, a Japanese company had implemented a gamified application, and many of its junior employees did very well and collected more points than the senior managers. However, after a short while, it became clear that this was unacceptable in the organization, and the junior employees started to "lose" more frequently in order to allow their managers to remain at the top of the scoring tables. While it is not clear what if any harm was done in this case, there is a risk that the introduction of such games could undermine workplace coherence and can be used as methods to single out employees that could be seen as a threat to their managers.

In 2009 serious games for health were estimated to be generating sales of $7bn (García Sánchez et al. 2012). These games or gamified applications can, for example, encourage people to jog more frequently, change their diets, or engage in alternative activities in order to remain active. In each case, however, and unlike close-contact studies (e.g., in a lab), the publisher is not in a position to assess the inherent risks to a particular individual if they undertake certain actions. Indeed the most they can do is point to the user in the direction of their doctor for advice and resort to overly complex EULAs.

Coercion

With respect to coercion, it is important to consider both indirect (even if this is subtle) and direct coercive approaches. As an example of indirect coercion, imagine an office-based serious game that is designed to train people to optimize their daily productivity around aspects such as completion time and success rates. The game itself may have nothing to do with their actual work (i.e., working in a financial office) but it would be designed to improve cognitive abilities relating to particular work tasks. The employees could also be totally free to choose whether or not to play this game. The employees are also informed about what the game will do, i.e., improve their cognitive skills, and that there is no risk to their health – they are even invited to sign an informed consent form. In every sense of the word, this particular serious game is complying with ethical requirements and what is more a data protection policy is even provided. However, as many employees play the game, there has arisen a small competitive element in the office, and those who are perceived as not playing the game face some social pressure to take part. Here, although there is no particular threat made to employees who perform poorly or who refuse to take part, there is a clear "implication" in the air of the office that they are expected to do so. Also it is unclear if the management makes any particular use of the data obtained when undertaking staffing decisions. In this case the coercion is there but more subtle than actively forcing people to take part. In contrast direct coercion is perhaps clearer to define and focuses on threats made against the person or persons playing the game in order to force them to take part. This could include sacking people if they do not take part or denying them certain employment opportunities through to clear any specific threats against them in other ways.

Privacy

There are many issues to consider regarding privacy and serious games especially given the recent growth in the concept of medically orientated games (Uzor and Baillie 2013; Alankus et al. 2010). The issue of privacy is suddenly much more acute and important to maintain in comparison to other game genres as the designer is now manipulating patients. However, the information must be shared when it comes to these games, otherwise the researchers and medical staff will not be able to assess whether or not the serious game has had the desired effect, e.g., made the person physically or mentally better. As is stated in the introduction, serious games can "often" obscure from the end user the actual underlying and overarching goals of the game. For example, Uzor and Baillie (2014) provided exergames to older adults at risk of falls, and the actual games mimicked the exercises in the British National Health Service falls booklet but this may have not been obvious to the fall patients. For example, in one game the player had to assume the role of a horse racing jockey and make their horse jump and clear ten hurdles in order to win the game. It was not obvious to the patient that what he or she is doing was in fact a series of ten exercise squats and that their performance data was being tracked. In this example, there is an issue with obfuscation to the end user along with problems relating to tracking and logging user data and how it is stored, manipulated, and shared. With respect to more traditional games, e.g., Assassin's Creed, it would not be such a major concern, but it does become an issue as regards serious games as the data becomes much more than merely a game score but an actual indicator of health, potential education achievement, antienvironmental, or antisocial behavior. The data collected, therefore, by these games suddenly becomes interesting to a much wider number of interested parties, e.g., police, health practitioners, insurers (car and health), local and environmental bodies, etc. Given this obfuscation from the end user, is it even possible for the end user to give "informed consent" in the way that was outlined earlier? In a recent study by Morrison et al. (2014), they found that even though they had obtained the consent from their mobile game users when they downloaded and installed the app which explained in detail what the app would collect, problems arose when an update after a week of use indicated what had been collected on them. The result was that many users expressed outrage and deleted the app. Therefore, even when users seemingly agree to data on them being collected and used, problems remain if their expectations do not match what they later perceive to be reality.

As a result there is a need to move to a model in which consent is not just collected at the outset but also during the entire use cycle. For example, providing a copy of the data collected to the game participants with examples of how the information could be visualized and used. At this point consent could be requested again. However, in the short term, the challenge is to provide as much information to the user as possible regarding data collection and use without giving away too much of the underlying game logic so that they can undermine the operating of the game.

Trust

At the moment the vast majority of research which has been published as regards games, serious games, and trust has focused on education (Anwar et al. 2006; Jerman-Blažič and Klobučar 2005; Malheiros et al. 2011) and in particular on training. As serious games evolve into other spheres, e.g., rehabilitation, transport, mental health, and so on, it will be interesting to see what the differences are as regards to trust and the end users' perception.

As a consequence of this evolution, there is a need to develop ethical practices in order that users have the correct expectations regarding the collection of their data and how it could be used. Prior work has shown that when users do not understand how a technology is tracking them, a breakdown and reluctance to participate can occur (Baillie and Morton 2009). In order to develop this understanding, studies must be undertaken whenever a serious game is deployed by an organization that collects or stores personal data. Research suggests that trust can be an important factor influencing user acceptance (Anwar et al. 2006; Riegelsberger et al. 2005; Adams 2001). When concerns around the collection of data are not addressed, the potential consequences include low morale, chilling effects, reduced commitment to the goal of the game or the organization deploying it, and lack of interest in playing the game (Fairweather 1999; Ariss 2002; Snyder 2010; Chen and Sanders 2007).

Although some research has been carried out on the impact of privacy issues on educational serious games (Anwar et al. 2006; Jerman-Blažič and Klobučar 2005; Nejdl and Wolpers 2000), it has focused on either types of data that have low sharing concerns, e.g., reading materials to high sharing of data, e.g., individuals' scores, date of birth, etc. There is currently a gap in the literature regarding trust issues of learning systems: (1) how learner users perceive and react to different data practices and interactions with other stakeholders and (2) what impact this can have on system acceptance and effectiveness. In conclusion there is a lack of methods to help developers incorporate trust considerations into the design of their serious games.

One way of starting to incorporate concepts of trust when designing is to use the model provided by Malheiros et al. (2011) when creating the structure of the game and then to think about using one of the models outlined in the transparency section (see later in this chapter).

Games

According to Sicart, "a game is a formal set of rules that project a fictional world that a player has to experience (Sicart 2005, p. 14)" and importantly "a game only exists when played, even though it is possible to describe its rules (Sicart 2005, p. 14)." Furthermore, drawing on Juul (2003), it should be noted that the game actions are voluntary and that the game outcomes should depend on the actions (this is related to the concept of autonomy discussed later). Huizinga (1955) provides a more formal definition arguing that games exist within a Magic Circle and that such a circle consists of a range of facets outlined below. However, as noted later in this chapter, both the views of Sicart and Huizinga are debatable when employed to serious, pervasive, or persuasive games.

- Huizinga's Magic Circle:
  - Rules (they only apply within the game space)
  - Objects (assume a context and range of effects within the game itself)
  - Location (a game of bowls is limited to a particular bowling alley)
  - People (namely, the players)
  - Time period (the game ends at a given point when a set of conditions are met)

Furthermore, Sicart notes that game designers often provide affordances within games with the sole intention of shaping behavior. These affordances can also pose ethical problems both in terms of how the game is played (e.g., discouraging certain actions or players from taking part). Although from traditional games, he pointed out that players often twist rules, and indeed in World of Warcraft the designers had to provide new affordances in order to overcome this problem. It is worth therefore considering the different types of rules that may exist and that are outlined below (see Salen and Zimmerman 2004):

- Operational
- Constituative
- Implicit

Operational rules are those which are provided and that must be adhered to in order to play a game, for example, rolling two dice and moving clockwise around a Monopoly board. Constituative rules are those that exist below the level of the operational rules and are embedded in the game itself (e.g., the range of possible moves in a game of chess). Implicit rules are those that arise during a game and those that reflect good sportsmanship or desirable game behavior. As Salen and Zimmermann point out, these can change between different game instances and between different sets of players, regardless of which type of rule they should exist to allow players to achieve the game objectives, even if they act in a way to slow down such progress (e.g., rolling dice and getting a low number).

The ethical nature of play is also shaped by the types of players; a naive assessment may lead us to conclude that there are players who adhere to rules and those who do not. In this simple ideal world, the objective would be to stop the “bad” players from undertaking improper activities. However, such an approach ignores the wider spectrum of player types that impact on ethical play within games. As noted below, Salen and Zimmerman identify a range of player types, each with a unique set of goals and objectives, but also how they interpret and implement their adherence to rules.

- Standard
- Dedicated
- Unsportsmanlike
- Cheater
- Spoil sport

The standard player follows the rules and has a genuine, although perhaps not passionate, interest in the game being played. This contrasts with the dedicated player who could be considered a hard-core gamer who is dedicated to winning and possesses a strong desire to win. Both of these types of players also pose a lusory attitude; this means the player is willing to adopt the game rules in pursuit of an objective. In the middle of the player group are unsportsman-like players who follow the letter of the rule of book but twist them in pursuit of winning.

People play computer games whether serious or for entertainment purposes due to a range of intrinsic and extrinsic motivations, in essence, for their own benefit with no external rewards (intrinsic) or when there are specific external rewards (extrinsic). The former could simply accept and complete a challenge and fulfill particular personal goals, while the latter could include external financial benefits or even social recognition among peers when a high score is obtained. Both intrinsic and extrinsic incentives have certain limitations, for example, over time, players may become less interested in the intrinsic benefits of such a game, thereby reducing their levels of participation. Extrinsic incentives also have to be decided carefully, for example, ensuring that they are not simply given out for undertaking activities where the same degree of participation would be obtained purely through intrinsic motivation. They also suffer from an erosion effect over time, in particular, when people come to expect the use of external incentives in order to take part. Regardless of the type of serious game, the form motivation can raise serious ethical issues, for example, disproportionate rewards could lead to radical changes in both in-game and outside-game behavior that may have long-term negative effects. They may also attract people to take part who are not the target group and act as a deterrent to some people. Also extrinsic motivational factors such as peer pressure may come into effect that could be psychologically harmful for the players. Therefore, from a purely ethical point of view, the incentive has to be considered as much as the desired game actions.

In addition to basic game mechanics and motivations, one major aspect of game design is to maintain a sense of being immersed in the experience, often referred to as flow (Csikszentmihalyi and Csikszentmihalyi 1988). Flow is the total absorption into a given task as a result of being motivated to partake in it. The concept of flow was extended by Sweetser and Wyeth (2005) to computer games and is deemed to arise through a variety of factors: concentration, challenge, skills, control, clear goals, feedback, immersion, and social interaction. A good example would be a person playing a 3D Shoot 'Em Up and being almost unaware of the surroundings or those around them. A large part of being in a state of flow relates to the underlying personality traits of the individuals; this can extend from the propensity to feel present in mediated experiences (Witmer and Singer 1998), and it could be argued it depends on how the particular gaming experience fulfills the personal objectives of those playing it, in particular, whether the goals of the game match the goals of the players, either in terms of play (e.g., completing a level) or higher-level constructs such as self-fulfillment or fulfillment with respect to wider society. Within the self-determination theory (SDT), Ryan et al. (Ryan and Deci 2002) indicate that human beings have basic underlying needs such as competence, relatedness, and autonomy, and when these are met, people will gravitate toward experiences that provide them. Competency is when people feel they are affecting the outcome within the wider social environment and are given opportunities to exercise and express their preferences; in a gaming context, players seek challenges in order to feel a sense of success. It could therefore be argued that a sense of challenge is a fundamental part of human development, and when this is provided in computer games, a basic need is being fulfilled which may in turn lead to players experiencing a feeling of flow. Importantly Ryan indicates that competence is not an attained skill but an ongoing desire. Thus, completing a level and feeling competent will not be an end in itself. Relatedness is connected to feeling part of a group, while autonomy relates to the feeling of being in control of one's own behavior, and as noted earlier this is also a key ethical requirement. Ryan et al. do, however, point out that their definition of needs and motives is quite different from personal motives or desires. For example, basic survival needs are not part of personal motivation and desires but importantly if these basic needs are not fulfilled, people may turn to other motivations and desires as a form of substitution. Furthermore, such personal motives and desires may change over time. It should be noted that Ryan and Rigby et al. also found that the needs approach from SDT applied also within single- and multiplayer games, although to a more limited degree within single-player games (Ryan et al. 2006).

To summarize, a game is something that exists beyond the actual underlying mechanics (such as rules, people, objects, and locations) and is instead something that exists within a given context. The desire for those to take part in the game and achieve the desired goals is determined by their underlying needs, motivations (intrinsic and extrinsic), and personality traits. A successful game should bring together all these elements in order to encourage participation and ultimately (if desired) a feeling of flow. The situation is, however, complicated by issues such as the nature of the rules (both operational and implicit) and also the range of players who take part from the cheater through to the ideal player. From an ethical perspective, this will give rise to a series of problems ranging from not misusing an understanding of the players underlying needs or motivations and how to ensure (especially in more pervasive and social games) that the players are aware of any implicit rules. Furthermore, while autonomy is a desirable feature both in terms of making ethical and in-game choices, it remains problematic in situations where the degree of free choice within a serious game may be limited by work or social context.

Types of Games

Serious Games

An early example of a non-computer-based serious game is Kriegspiel (war game); it dates from 1812 and was used for training the Prussian and German Armies. Military training using games (for a discussion see: (Smith 2010)) has also been extensively used. Within this context, the issue of ethics and safety impacts upon both the players and potentially those who the player may encounter in actual war zones. Outwith war, serious games are often used to teach people about issues that can have long-term effects on their lives. In essence they use the fun element of games to achieve a serious benefit for the player and other stakeholders. For example, serious games could in theory be used as a way to assess people, therefore having long-term consequences for their future careers and increasingly their health.

Pervasive Games

Pervasive games on the other hand are not bound by as many constraints as traditional serious or normal games. According to Montola (2005), they are not spatially, temporally, or socially bounded by Huizinga's Magic Circle (see earlier). For example, more traditional games such as geocaching can be played in any location where a cache can be left and/or found by players. Also the number of players may vary during the game, and there is no time limit on the total game play and in theory the game could last forever. Within the serious gaming context, such examples include Time Warp which was a mobile mixed reality game played in Cologne, Germany (Blum et al. 2012). Here players would walk around the Rhine bank area in central Cologne looking for characters and objects within different time periods, including the future. Although the future aspect had no educational value, the underlying historical elements aimed to make people aware of some aspects of the city's past (it should be noted the players were not assessed directly for their knowledge). Although the game took place in a relatively spatially defined area, it did alert the developers to some challenges involving players and non-players within complex urban environments.

In common with other products and services, serious games are also increasingly being played on mobile devices such as tablets, mobile phones, and augmented reality visors. These pervasive games bring with them a set of ethical challenges (as noted earlier) as they often seek to extend gameplay beyond the magic circle (see earlier work by Huizinga (1955)). While these facets may apply within a pervasive game, it is no longer as Montola describes spatially, temporally, or socially bounded. Taking the last aspect, the lack of social bounds means that serious game designer must start considering the role of others within not only the design of the game but also from the perspective of ethics, privacy, and trust. Take, for example, the list of player and non-player types that may either be impacted by or impact a pervasive game (Montola 2005):

- Player: someone who influences the game.
- Non-player participant: lack of personal game goal, e.g., referee.
- Spectator: are aware of game and can be actively involved, but cannot directly influence game perhaps only indirectly.
- Bystander: unaware of ongoing game and have no ability to participate in it. They are also insulated from taking part in it.

As can be seen above, there are varying degrees of participation within pervasive games; while player and non-player participants probably provide some form of consent to take part, it is debatable whether any consideration has been given to spectators and bystanders. Indeed, it is also difficult to assess the precise impacts and risks that particular games will have on them and consequently what degree of consent may be required. Take, for example, a game where one challenge is to locate a stranger in the street and photograph them after they answer a question about the locality. They collect no points and derive little, if any, benefit from the game itself but at the lowest level have their day impacted upon by taking part (although perhaps not in a negative way). However, it is unlikely that they are asked to sign any agreement or indeed that they are fully aware of how their picture or data could be used in future.

In addition to the different range of people who may in some way be connected to the experience, there are also (and in common with other game types) a range of players which apply, for example, those who adhere strictly to the rules but maintain the spirit of the game through to those who adhere strictly to the rules but not the spirit of the game finally to those who simply break them (see earlier summary by Salen and Zimmerman (2004)). Within a traditional (i.e., not a serious game) Magic Circle-bounded game, the potential impacts are mostly confined to the game itself, although some players may be angry or frustrated with those who have cheated after the game. In contrast, in a pervasive game where the authors may have limited control over where it is played, the effect of either breaking the rules or not playing in the spirit of the game could be harmful (e.g., in the case of traffic games). At the lower end of the spectrum, as the game is perhaps less time bounded, the effects of "foul play" could last significantly longer than within a game which can be easily ended due to a time limit. In both cases the potential health and social impacts could extend far beyond the actual playing time.

Persuasive Games and Gamification

Persuasive games draw on the wider area of persuasive technologies which "attempt to shape, reinforce, or change behaviors, feelings, or thoughts about an issue, object, or action (Fogg 2002)," for example, reducing smoking or other health improvements. It should be noted that persuasive games need to not use any technology; however, for the purposes of this chapter, the focus will be on games delivered via some forms of computer mediation. Although there are differences from gamification (discussed later), there are a number of similarities, namely, (1) they use technology (2) aims at affecting the attitudes of people such that (3) their behavior is altered. According to Hamari et al. (2014), the topic area emerged in the 1990s with the main writings arising from 2005 onwards, with 2013 being abundant in terms of publications. This is perhaps best emphasized with the emergence of the *persuasive* conferences which specifically tackle the topic from a range of perspectives. Therefore, although older than gamification, it is still a relative new field both technically and from the perspective of ethics, privacy, and trust. Oinas-Kukkonen et al. (Oinas-kukkonen and Harjumaa 2009) define a number of design principles for persuasive systems (see below); however, the primary focus in this chapter is on aspects relating to credibility.

1. Trustworthiness
2. Expertise
3. Surface credibility
4. Real-world feel
5. Authority
6. Third-party endorsements

The concept of trustworthiness was discussed earlier; however, generally the more trustworthy the system, the more persuasive it will be. Expertise is also a key component as the system on the face of it should possess knowledge and experience and competence when attempting to persuade the user. Surface credibility is also important as people tend to make quick so-called first impressions when deciding on a system. Real-world feel relates in some ways to the transparency, in that a system that makes the organization and people behind it clear should have more credibility. Authority primarily relates to whether the system has clear outside authority given to it, e.g., from a government agency. Third-party endorsements relate to others endorsing a pervasive game and thus boosting its perceived credibility. Collectively these elements should allow for persuasive system to be more effective, in turn, though the use of certain aspects may also create ethical issues, for example, a trusted company endorsing a product that would otherwise be considered ethically dubious by the public.

Gamification

In many ways gamified applications should be viewed as an extension to persuasive games in that their explicit intention is to change behavior whether by making this clear or through more covert means. Therefore, there is a strong overlap in the ethical problems that may arise in pure serious games and within gamified applications. Gamifications' growing popularity, clearly persuasive objectives combined with relative immaturity, should lead to concerns with regard to ethics privacy and trust.

The term gamification has come into widespread use in recent years with it broadly speaking being defined by Deterding et al. (2011). Gamification can be thought of as the application of gamelike elements (e.g., points, leaderboards, and badges) to non-game environments such as health and traffic. There are many contrasting views on the merits of gamification with its father figure Deterding offering no perceived judgments regarding its value although it would be hoped that a system would bring about benefits for its users. This is in stark contrast to Zichermann and Cunningham (2008) who see it as little more than a way for a company to achieve its corporate profit-driven policies. In order to illustrate this point, take two gamified applications Leave Now and Foursquare. The former is an academic research tool designed to reduce traffic congestion; while the users may derive benefits such as reduced time spent in traffic beyond the data collected, the researchers involved and the institution will make no profit. Indeed even the data collected will only be used to improve future versions; again there is no personal profit motivation involved. In contrast Foursquare lets users check in at locations, receive points, and perhaps also discounts from the various sites. It is collecting such data to share with its business partners who may be able to offer incentives to their potential customers. Therefore, while the user may derive some fun from using the application, the real benefits go to the businesses involved. This is perhaps not immediately obvious for end users. While there is significant research being undertaken into gamification, the significant interest that it has received from business is a clear indication of the benefits that they perceive from it either in the form of selling more products and services or improving employee performance. In the case of the latter, Shahri et al. (2014) noted that gamification can bring excessive tension and pressure to the workplace.

As Versteeg (2013) argues, it is essentially the fun factor which is used to tempt people to change their behavior, and that gamification has a number of facets which make it explicitly persuasive in nature and that these elements give rise to ethical concerns. Drawing on the work of Fogg (2002), Versteeg argues that gamification can often exhibit the following facets:

- Reduction: persuasion through simplification of the task
- Tunneling: persuasion via leading people through a task
- Tailoring: providing user-specific information to increase chances of persuasion
- Suggestion: suggesting an action based on underlying motivations
- Self-monitoring: allowing the user to view their progress toward a goal
- Surveillance: making user aware they are being watched and providing options based on this
- Conditioning: providing rewards for an action(s)

A critical point of the above list is that while these techniques may be used frequently by the designers of gamified applications, the users may not be aware they are being used. Indeed this may be a specific desire by the designers. As a result these comparatively subtle approaches may lead to a situation where behavior change is encouraged to meet the goals of another party and not that of the player; a good example would be suggestions in online shopping applications or where surveillance is used to reduce behavior of benefit to the players but which does not meet corporate objectives.

A number of authors have developed ethical guidelines for gamified applications (Versteeg 2013; Berdichevsky and Neuenschwander 1999; Marczewski 2013). Zichermann provides three foundational concepts which can broadly speaking be defined as (1) designing systems which help individuals/organizations reach their goals by aligning with their values, (2) making the objectives and purpose transparent, and (3) Making results available (where legally possible) in order to promote best practice. Marczweski arguably goes further and again defines three major areas: (1) honesty, (2) transparency with regard to data collection, aims etc., and (3) provide best service quality possible. Marczweski extends honesty to include making potential customers and users aware of the limitations of gamification and more importantly to state that gamified applications should not collect data that users would not otherwise choose to share. Both Marczweski and Zichermann provide useful sets of guidance; however, although they deal with issues such as transparency, they are limited and not in sufficient detail. Instead they are perhaps best thought of as a set of overarching guiding principles.

Berdichevsky and Neuenschwander (1999) and Versteeg (2013) provide a far greater set of ethical guidelines. Although in the case of Versteeg, greater effort is made to take into account the holistic process from design and implementation to testing. Berdichevsky shares some similarities with the earlier sets of ethical guidelines. However, a fundamental change is that they place responsibility for all possible outcomes of using the persuasive technology on the designers of the system and that the designer would also be prepared to use the same technology. Berdichevsky further adds that the use of the technology should not in itself be the reason for undertaking a particular form of persuasion and that persuasive technology should only be used if the same actions would take place without the technology. In essence these three points move the focus from simply undertaking gamification to one which is more based on underlying ethical concepts, e.g., only do to others what you would do to yourself, and that it is irrelevant how such persuasion is undertaken but that fact that it is being undertaken. While these are beneficial aspects, the idea that the designer should be responsible for all possible forseen and unforseen outcomes is perhaps at best impractical. Indeed Shahri et al. (2014) noted that workplace gamification applications can result in increased divisions between staff and hence tension, can be perceived as a method of monitoring, can erode privacy, and may be seen as a way of exploiting employees. While there are no doubt some employers have these specific goals, it should be clear that for the most part such by-products of use are likely to be harmful on the users and ultimately wider society (or the workplace).

Versteeg (2013) proposes a comprehensive approach which places emphasis on the entire process of developing gamified (persuasive) applications. The approach aims to bridge what Versteeg sees as the gaps identified in the other guidelines with a thorough grounding different ethical approaches (see Versteeg for a full review). Versteeg's framework consists of four main elements: (1) moral principles and values, (2) conceptual investigation, (3) involvement of stakeholders, and (4) evaluation and iteration. The work draws on aspects such as participatory design and traditional ethics; in particular the proposed frameworks' main strength is that it includes all those affected by such a system, not just the end users. The moral principles and values component shares certain aspects of previous ethical guidelines; however, it also includes aspects such as (1) taking people toward a path of virtue, not one of vice, and (2) all users should be treated equally (nondiscrimination). As acknowledged by Versteeg, it is not an exhaustive list of concepts or a complete framework.

Discussion

While there are differences between game types, e.g., classic serious, persuasive, pervasive games, and gamified applications, it is increasingly the case that the boundaries between them are becoming blurred. For example, serious games may include pervasive elements, and often the desired outcome of a serious game is behavior change (or learning of a new skill) outside of the game itself. Therefore, there is a need to explore overarching ethical concerns across the different game types from the perspective of core gaming elements such as rules, people, objects, locations, and time periods (Huizinga 1955). By using these elements as a way to define the core gaming elements or indeed, gamified apps, it is then possible to see how they affect underlying ethical issues,' impact on all people (not just the players), and minimize negative impacts. Therefore, in the following section, a set of overarching concepts and guidelines in the form of questions is presented, and, in common with Fogg (2002) and Versteeg (2013), the involvement of stakeholders is an important part of the process. Underpinning this is a return to the basic moral concepts such as consent, autonomy, beneficence, fidelity and responsibility, integrity, justice and respect for rights, and dignity. As an example, taking informed consent as a core part, a serious game should support *voluntary participation*, *competency to decide to take part*, *informed of all aspects*, and *comprehensible information*. A typical example would be within the people category and concept of consent, i.e., while a player may consent to take part the wider population or indeed others who are not directly playing may not have done so but still be impacted by the results of the game. However, it should be noted that obtaining consent from a possible unlimited number of people would be problematic.

The guidelines as they are presented specifically build on both the underlying positions relating to ethics, privacy, and trust as well as the more applied ideas presented by the various computer, persuasive, and gamification ethicists discussed earlier. However, in contrast with higher-level approaches, the intention in the remaining section is to ask specific questions which open up the ethical debate during the design process so that any potential ethical issues can hopefully be resolved. Thus ethics, privacy, and trust are embedded in the concepts rather than always being explicitly stated, and also specific technical aspects such as encryption or data formats are not discussed. Furthermore, the process of obtaining requirements and feedback from players is not discussed. The concepts are not mutually exclusive, and similar issues will often be addressed from multiple perspectives. Also not all concepts and guidelines will be relevant to all serious game types.

Identify Individual, Social, and Corporate Objectives

Serious games will have some set of overall objectives, whether this is to enhance employee performance or persuade people to undertake different commuting behaviors. Often these objectives can be contradictory, for example, a company may wish to improve employee performance without paying extra salaries, while the employee may view the game as a way to obtain promotion or a pay rise. On a social level, a game may seek to improve communication between people in an office.

- What are the objectives?
- Who has these objectives?
- How does the game align with pre-existing objectives of the individual and the organization (if required)?
- What contradictions are there?
  - Are the corporate objectives different from those of the individual? How is this situation resolved (if at all)?

Benefit and Risk Impacts

The game should first and foremost provide benefits to those who take part; additionally if there are wider societal or corporate benefits, these should be taken into account. There will be some link between the objectives and the benefits; however, there are some subtle differences. Benefits should be considered across all stakeholders from players to the organization, developers, and third parties (e.g., work colleagues).

- What are the benefits?
  - Will these benefits have a direct or indirect negative impact on anyone?
- Who benefits (individual, company, society)?
- What are the balance of benefits between stakeholders?
- What competing benefits are there in the system?
- What are the risks?
- Who experiences these risks (individual, company, society)?
- What are the balance of risks between stakeholders?
- What happens in the event of failure?

Rules and Styles of Play

As noted earlier, games have a variety of player types from those who adhere to the rules and to those who cheat. Increasingly, within gamified applications, there is a risk that the behavior of players who adhere to the rules of a game but not its spirit may have an impact on others. For this the designer should identify a set of impacts.

- Are players aware of the different rules both those explicitly stated and those that are implied within the context of use?
- Is it possible to play by the rules but also play against the spirit of the game?
  - What impact will this have on the overall game and other players?
- What impact will different player types have on the game objectives and other players?
- Are rules and other mechanisms used to change the users' playing habit in a way that encourages them to follow a path that they would find not to be beneficial to the non-game environment?
- Are the rules used going to aid the players in achieving their particular goal? If not, is the restriction harmful (note: rules can be used as a way to increase difficulty; hence, they should also reflect the objectives of the game and the underlying motivations and needs of the players)?

Transparency

Transparency relates to a number of key facts extending from benefits, risk, and harm to the use of data. A key problem arises when certain aspects, of the system may need to be hidden from the user; in particular its real objectives, related to this, are problems with deception of players, etc. Where possible, such aspects should be avoided but there are times when they may be necessary. Furthermore, transparency also relates to players being aware of the impacts of the game on others, or indeed their awareness of the range of rules that may apply.

- Are the goals of the system clearly stated?
- Are the benefits and risks clearly stated?
- Are all rules clearly stated?
- Are players aware of the impact on non-players and the wider environment?
- Where deception is used can this be avoided or mitigated in some way?
- Are data and privacy issues clearly stated?

Temporality and Use

Increasingly games are played over a longer period of time, and certain pervasive and gamified applications may have no particular temporal limit. In this context, the impacts of the serious game need to be considered over a longer period of time. In particular benefits external to the game environment could outlive the time of play, while negative impacts should be avoided. Negative effects could include social and peer pressure issues. There are also issues to do with updating or changing a game and whether the player has consented to take part. Also, when playing a game, care should be taken to avoid any indirect negative impacts.

- If the game has a finite end point, are any benefits lasting beyond this point?
  - Consequently are negative impacts contained to the time of play? If not, what effect will they have and for how long?
- For the duration of play (use), are there negative impacts on aspects of the players work or personal life?
- For the duration of play (use), are there negative impacts on aspects of the wider environment, e.g., locations or other people?
- Are gaming elements, privacy, and trust aspects liable to change over time? If this is the case, what degree of awareness and consent are provided to end users?

Motivations, Incentives, and Flow

The underlying needs and motivations of players will make a significant impact on the use, adoption, and success of serious games (across all types). However, while fulfilling needs and allowing games to match the motivations of the players may be perceived as a good thing, it is not without some inherent risks, in particular, if an individual has an addictive personality which could result in them being overly manipulated by a serious game.

- Are methods provided to ensure that competence is rewarded?
- What are the motivations of the users to play the game both extrinsic and intrinsic?
  - How are these motivations manipulated or utilized within the game, for example, through incentives?

- Are the benefits of using the system the main reason for its use as opposed to extrinsic incentives awarded? For example, if a cash payment were to stop, would the game or specific outcome still be seen as beneficial?
- Is there a risk that heightened attentional focus on the game (sense of flow) could result in the players playing the game for its' own sake rather than for the specific objectives/benefits it sets out to achieve? Here the objective is to avoid addiction to a particular game or situations where judgment with respect to the desired real world outcome becomes clouded due to particular game structures.
- Are the incentives provided necessary for the player to maintain the social, personal, or work status? If these were not provided, would any harm to them arise? For example, lack of points in a workplace game may result in social exclusion or bullying.
- In the event of termination of the game, have the impacts on players, data policy, and wider society been considered?

Autonomy

Players should be free not only to choose how to take part but also how to undertake actions within any serious game, and particular attention should be paid to items which are used to influence behavior such as tunneling, surveillance, or suggestion. For example, if surveillance is being used, are players genuinely able to opt out or behave in ways that they would choose normally to do?

- Can players excise a right to leave the game without negative impacts?
  - If not, what impacts may occur?
- Are players free to decide on their own course of action (including leaving)? In essence are forms of direct and indirect coercion avoided?
- When methods are used specifically to channel or modify behavior (e.g., tunneling), is this approach necessary, appropriate, not harmful, and in the players interest?

Privacy

This should relate to all parties, not just players, for example, bystanders or spectators. It extends from if data is stored on people to how it should be used and any sharing or destruction policy. A significant amount of the data privacy issues are also related to transparency and trust aspects.

- Which groups of people are having data stored directly or indirectly on them?
- Have users and others expressed explicit consent for data to be stored on them?
- Have users and others expressed explicit consent on how such data will be used?
- What unnecessary information is stored? Only data which is required to allow the smooth running of the game should be stored.

- Is information only stored for the minimum period of time necessary?
- Is there a data destruction policy?
- Where data is stored anonymously, how easy is it to still decode the data such that players and others may be identified?
- Will data sharing and use change over the lifetime of the game and is this clearly articulated to the user?
- What policies are in place to ensure that non-players' privacy is maintained? This is more relevant to pervasive games.
- What policies are in place to ensure that the desire for privacy in certain locations is maintained while a game is being played?.
- Where sensitive information (see next section) may be stored or used, is its' storage and use minimized? If possible, no such information should be stored or used.

Rights and Respect for Players

Players' background such as gender, race, religion, politics, and sexual orientation should not be misused when designing serious games. However, it should be noted that certain serious games may target a particular group of people, for example, in order to educate them about a particular topic. Where this is the case, the benefit should be clear and should not be to the detriment of them or other people. In this context, the term "profiling" is meant with respect to selecting a group of users based on their backgrounds so that they take part in a game or that the game responds to such information.

- Where profiling is used, is this related to providing a benefit for that particular target group?
- Where profiling is used, are negative effects for those not selected removed? For example, if the particular background of an individual provides them with an advantage, this should not be at the expense of those without such a background.
- Are issues related to player background which may have negative impacts on players avoided? An example here would be in an office where some new employees do not speak the national language fluently. They are disadvantaged when playing the game which may in turn have negative impacts on their careers.
- Are all game design concepts, data privacy standards, and incentives easily comprehensible and beneficial for all potential players?
- To whom does liability fall in the event of an ethical, privacy, or trust problem? There should always be someone who is responsible.
- To what extent is the responsible party (parties) awares of the ethical, privacy and trust issues?
- What procedures are in place to ensure that the developer/project instigator is held responsible for actions both directly and indirectly resulting from playing the game?

Trust and Credibility

Trust is not only whether a system is trustworthy but whether perceived trust can act as a way to manipulate players. For example, XYZ is a well-known and respectable brand using deceptive tactics in order to change behavior or collect excess data. Trust is also related to aspects such as transparency that were outlined earlier.

- What and how are external methods such as branding or third-party endorsements used to improve the perceived degree of trust?
- Would the system be trusted by players *without* external trust approaches?
- Are third-party and external endorsements being used to manipulate the players' behavior?

Relationship Between the Gaming Context and the Technology

Serious games are ultimately about bringing a benefit to the player in their everyday life, whether this is for work, sport, or pleasure. Also as noted by Brey (2000), the technology cannot be isolated from the wider impact and should also be considered.

- Does the game support a benefit outwith the game environment?
- Would the benefit exist without the game? If not, to what extent is this new benefit desirable?
- What technologies are being used to collect data and how might these impact on user behavior?
- Are the technologies used likely to have an impact on those not playing the game?
- Are specific interaction approaches being used which are designed to shape behavior, and what is their impact?
- Does the game use coercive approaches to restrict or eliminate player choice? For example, if a benefit is withdrawn will this have a negative impact on the player or is a specific interface/interaction type used to coerce players?

The previous concepts and guidelines provide a list of issues to explore when designing serious games and also evaluating those which are either under development or already being used. As noted by Versteeg (2013), an iterative development model is highly relevant and beneficial in this case, in particular, when used with approaches to correctly identify users, context of use, and requirements. It should be stressed that it is not intended as a tick-box-style assessment or a method to provide an ethical compliance score. Finally and indeed importantly, would those responsible for the design and rollout of a serious game use it themselves?

Conclusion

This chapter started with a discussion on ethical issues initially outside the domain of serious games and then within and across various types of serious games. It highlighted problem areas through an examination of various types of serious game with an emphasis on looking at common issues that may arise, or where differences may also result in problems. In doing so, it sought to specifically address the earlier high-level identification of the issues highlighted by McCall et al. (2012) and Koenig et al. (2012). The discussion section provided a (non-exhaustive) list of concepts and guidelines in the form of questions that are intended to provide a way to examine designs and currently implemented systems. In what is perhaps a shift from existing approaches, the ethical discussion and subsequent context also included unintended impacts on third parties (i.e., non-players), through to how consent varies over time. The concepts and guidelines do not aim to provide an assessment mechanism but rather a method by which practitioners and researchers can consider ethical issues and start the discussion on potential weaknesses in their own systems. They are also intended to provoke discussion and provide a starting point upon which they can be refined and developed.

Recommended Reading

A. Adams, Users' perception of privacy in multimedia communication, in *CHI'99 Extended Abstracts on Human Factors in Computing Systems* (ACM, Pittsburgh, 1999), pp. 53–54.

G. Alankus et al., Towards customizable games for stroke rehabilitation, in *Proceedings of the 28th International Conference on Human Factors in Computing Systems CHI 10*, Vancouver (2010), pp. 2113–2122

American Psychological Association. *Ethical Principles of Psychologists and Code of Conduct*. (2010), Available at: http://www.apa.org/ethics/code/. Accessed 23 Jan 2015

M.M. Anwar, J. Greer, C.A. Brooks, Privacy enhanced personalization in e-learning, in *Proceedings of the 2006 International Conference on Privacy Security and Trust Bridge the Gap Between PST Technologies and Business Services PST 06* (2006), p. 1. Available at: http://portal.acm.org/citation.cfm?doid=1501434.1501485

S.S. Ariss, Computer monitoring: benefits and pitfalls facing management. Inf. Manag. **39**, 553–558 (2002)

L. Baillie, L. Morton, Chaps Final Report. Paths to Health (2009), Available at: http://www.pathsforall.org.uk/component/option,com_docman/Itemid,166/gid,153/task,cat_view. Accessed 30 Jan 2015

T. Beauchamp, The principle of beneficence in applied ethics, in *The Stanford Encyclopedia of Philosophy*, Winter 2013 edn, ed. by E.N. Zalta (2013). http://plato.stanford.edu/archives/win2013/entries/principle-beneficence

D. Berdichevsky, E. Neuenschwander, Toward an ethics of persuasive technology. Commun. ACM **42**, 51–58 (1999)

L. Blum et al., The final TimeWarp: using form and content to support player experience and presence when designing location-aware mobile augmented reality games, in *Proceedings of*

the Designing Interactive Systems Conference, DIS'12 (2012), pp.711–720. Available at: http://www.scopus.com/inward/record.url?eid=2-s2.0-84864773076%26partnerID=40%26md5=12db25ea432cb500fa65d8ca5ed455e5
P. Brey, Method in computer ethics: towards a multi-level interdisciplinary approach. Ethics Inf. Technol. **2**, 125–129 (2000). Available at: http://link.springer.com/article/10.1023/A:1010076000182
R. Chen, G.L. Sanders, Electronic monitoring in workplace: synthesis and theorizing, in *Proceedings of the Americas Conference on Information Systems; 3; 1900–1907 Americas Conference on Information Systems AMCIS 13th, Americas Conference on Information Systems AMCIS*, Keystone (AMCIS, 2007), p. 8
I. Chochliouros et al., (Semi-) pervasive gaming educational and entertainment facilities via interactive video-to-video communication over the internet, for museum exhibits, in *Artificial Intelligence Applications and Innovations SE – 48. IFIP Advances in Information and Communication Technology*, ed. by H. Papadopoulos et al. (Springer, Berlin/Heidelberg, 2013), pp. 474–485. Available at: http://dx.doi.org/10.1007/978-3-642-41142-7_48
M. Csikszentmihalyi, I.S. Csikszentmihalyi (eds.), Optimal experience: psychological studies of flow in consciousness, in *Optimal Experience: Psychological Studies of Flow in Consciousness* (1988), Available at: http://ovidsp.ovid.com/ovidweb.cgi?T=JS%26PAGE=reference%26D=psyc3%26NEWS=N%26AN=1988-98551-000
S. Deterding, R. Khaled, L.E. Nacke, D. Dixon, Gamification: toward a definition, in *CHI 2011 Gamification Workshop Proceedings*, Vancouver (2011), pp. 12–15
European Federation of Psychologists, *Meta-Code of Ethics* (2005), Available at: http://ethics.efpa.eu/meta-code/
N. Fairweather, Surveillance in employment: The case of teleworking. J. Bus. Ethics **22**(1), 39–49 (1999)
L. Floridi, J.W. Sanders, Mapping the foundationalist debate in computer ethics. Ethics Inf. Technol. **4**, 1–9 (2002). Available at: http://dx.doi.org/10.1023/A:1015209807065
B.J. Fogg, *Persuasive Technology: Using Computers to Change What We Think and Do* (Morgan Kaufmann, San Francisco, 2002)
R. García Sánchez et al., Value propositions for serious games in health and well-being, in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Springer, Berlin/Heidelberg, 2012)
M.D. Griffiths, N. Hunt, Dependence on computer games by adolescents. Psychol. Rep. **82**, 475–480 (1998)
S.M. Grüsser, R. Thalemann, M.D. Griffiths, Excessive computer game playing: evidence for addiction and aggression? Cyberpsychology Behav.: Impact Internet, Multimed. Virtual Real. Behav. Soc. **10**, 290–292 (2007)
J. Hamari, J. Koivisito, T. Pakkanen, Do persuasive technologies persuade? – a review of empirical studies, in *Persuasive Technology,* eds. by A. Spagnolli, L. Chittaro, L. Gamberini (Springer International Publishing, Berlin/Heidelberg, 2014)
J. Huizinga, *Homo Ludens: A Study of the Play Element in Culture* (Beacon, Boston, 1955)
B. Jerman-Blažič, T. Klobučar, Privacy provision in e-learning standardized systems: status and improvements. Comput. Stand. Interfaces **27**, 561–578 (2005)
J. Juul, The game, the player, the world: looking for a heart of gameness, in *Level Up: Digital Games Research Conference Proceedings* (2003), pp. 30–45. Available at: http://www.jesperjuul.net/text/gameplayerworld/
V. Koenig, F. Boehm, R. McCall, Pervasive gaming as a potential solution to traffic congestion: new challenges regarding ethics, privacy and trust, in *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* Springer Berlin Heidelberg, (2012), pp. 586–593
E. Luger, T. Rodden, An informed view on consent for UbiComp, in *Proceedings of the 2013 ACM Conference on Ubquitous Computing UBI* (ACM, Zurich 2013), pp. 529–538. Available at: http://dl.acm.org/citation.cfm?id=2493446

E. Luger, S. Moran, T. Rodden, Consent for all: revealing the hidden complexity of terms and conditions, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (ACM, 2013), pp. 2687–2696

M. Malheiros et al., Trusting to learn: trust and privacy issues in serious games, in *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Springer Berlin Heidelberg, 2011), pp. 116–130

A. Marczewski, Gamification code of ethics. *Andrey's Blog* (2013). Available at: http://marczewski.me.uk/resources/gamification-code-of-ethics/

R. McCall, V. Koenig, Gaming concepts and incentives to change driver behaviour, in *2012 The 11th Annual Mediterranean Ad Hoc Networking Workshop (Med-Hoc-Net)* (IEEE, 2012), pp. 146–151. Available at: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6257115

R. McCall et al., Workshop on: exploring the challenges of ethics, privacy and trust in serious gaming, in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Springer International Publishing, Heidelberg, 2012), Heidelberg, pp. 584–585

M. Montola, *Exploring the Edge of the Magic Circle: Defining Pervasive Games* (2005), Available at: http://remotedevice.net/main/cmap/exploringtheedge.pdf. Accessed 02 May 2012

A. Morrison, D. McMillan, M. Chalmers, Improving consent in large scale mobile HCI through personalised representations of data, in *NordiCHI'14 Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational* (ACM Press, Helsinki, 2014), pp. 471–480

W. Nejdl, M. Wolpers, European E-learning: important research issues and application scenarios restructuring and integrating European research on E-Learning Academy and Competence Centre in the PROLEARN Context Research Issues for E-Learning, in *Learning Technology*, **54**, 1137–1144 (2000). Available at: http://www.ncbi.nlm.nih.gov/pubmed/21145810

H. Oinas-kukkonen, M. Harjumaa, Persuasive systems design: key issues, process model, and system features. Communications of the Association for Information Systems **24**(1), 485–500 (2009)

J. Riegelsberger, M.A. Sasse, J.D. McCarthy, The mechanics of trust: a framework for research and design. Int. J. Hum. Comput. Stud. **62**, 381–422 (2005)

R.M. Ryan, E.L. Deci, Self-determination theory: an organismic dialectical perspective, in *Handbook of Self-Determination Research* (University Rochester Press, 2002), pp. 3–34

R.M. Ryan, C.S. Rigby, A. Przybylski, The motivational pull of video games: a self-determination theory approach. Motiv. Emot. **30**, 347–363 (2006)

K. Salen, E. Zimmerman, *Rules of Play: Game Design Fundamentals*, (MIT Press, Cambridge, MA, 2004)

A. Shahri et al., Towards a code of ethics for gamification at enterprise, in *Lecture Notes in Business Information Processing*, ed. by U. Frank et al. (Springer, Berlin/Heidelberg, 2014), pp. 234–245

M. Sicart, Game, player, ethics: a virtue ethics approach to computer games. Int. Rev. Inf. Ethics **4**, 13–18 (2005). Available at: http://www.i-r-i-e.net/inhalt/004/Sicart.pdf

R. Smith, The long history of gaming in military training. Simul. Gaming **41**(1), 6–19 (2010)

J.L. Snyder, E-mail privacy in the workplace: a boundary regulation perspective. J. Bus. Commun. **47**, 266–294 (2010)

P. Sweetser, P. Wyeth, GameFlow. Comput. Entertain. **3**, 3 (2005)

S. Uzor, L. Baillie, Exploring & designing tools to enhance falls rehabilitation in the home, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems – CHI'13* (ACM, New York, 2013), p. 1233–1242. Available at: http://www.scopus.com/inward/record.url?eid=2-s2.0-84877933397%26partnerID=tZOtx3y1

C. Versteeg, Ethics & Gamification design: a moral framework for taking responsibility. Masters thesis, Utrech University, Netherlands, 2013

J. Vollmann, R. Winau, NUREMBERG DOCTORS' TRIAL Informed consent in human experimentation before the Nuremberg code I i _., 313 (December). (1996), pp. 1445–1449

B.G. Witmer, M.J. Singer, Measuring presence in virtual environments: a presence questionnaire. Presence: Teleoperators Virtual Environ. **7**, 225–240 (1998)

G. Zichermann, C. Cunningham, *Gamification By Design* (2008), Available at: http://medcontent.metapress.com/index/A65RM03P4874243N.pdf

Part VII

Art and Entertainment

Computer Music Languages and Systems: The Synergy Between Technology and Creativity

28

Hiroki Nishino and Ryohei Nakatsu

Contents

Abstract

This chapter briefly overviews the history of computer music languages and related systems, mainly focusing on those developed in the research community (hence, less focus is put on those commercial computer music software such as digital audio workstation (DAW) software or sound editor software). As is often seen in other surveys of computer music history, the historical development of computer music languages and systems is divided into several overlapping eras in this chapter. The division between the eras of non-real-time computer music systems and real-time computer music systems is particularly emphasized, as it gave a significant

H. Nishino (✉)
NUS Graduate School for Integrative Sciences and Engineering, National University of Singapore and Graduate School of Media Design, Keio University, Singapore, Singapore
e-mail: hiroki.nishino@acm.org

R. Nakatsu
Design School, Kyoto University, Kyoto, Japan
e-mail: nakatsu.ryohei@gmail.com

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_19

impact on both creative practices by artists and musicians and the design of computer music languages and systems by researchers and engineers.

While the evolution of computer music languages has been largely supported by the advance of computer technology and the achievement of the related research in computer science and audio engineering, it should be also noted that issues found in creative practices also have given significant influences to the development of computer music languages and systems throughout its history. Along with the technical advancement, the synergy between technology and creativity in computer music is also highlighted when appropriate in this chapter, as such a perspective can be beneficial to reconsider the relationship between computer technology and artistic creativity in our decades.

Keywords

Computer music • Sound synthesis • Arts and creativity • Music technology • Programming language

Introduction

Ada Lovelace (1815–1852), the only child of the poet Lord Byron, is often referred as the world's first computer programmer (Fuegi and Francis 2003). She worked with Babbage, who designed the Analytical Engine ("a mechanical precursor of the digital computer" (Risset 1994)) in the nineteenth century.

Indeed, there exists a debate whether Lovelace was really the first computer programmer in history. Bromley discusses as follows: "all but one of the programs cited in her notes had been prepared by Babbage from 3 to 7 years earlier. The exception was prepared by Babbage for her, although she did detect a 'bug' in it. Not only is there no evidence that Ada ever prepared a program for the Analytical Engine, but her correspondence with Babbage shows that she did not have the knowledge to do so" (Bromley 1990). On the contrary, Kim and Toole argue that she "was certainly capable of writing the program herself given the proper formula" and it "is clear from her depth of understanding regarding the process of programming and from her improvements on Babbage's programming notation" (Kim and Toole 1999).

However, regardless of such a debate, she is likely at least the first person to propose the idea to use computational automata for musical creation. Around 1840, when Ada was still working with Babbage, she left some notes as below.

> (The Analytical Engine's) operating mechanism might act upon other things besides numbers, where objects found whose mutual fundamental relations could be expressed by those of the abstract science of operations, and which should be also susceptible of adaptations to the action of the operating notation and mechanism of the engine. Supposing, for instance, that the fundamental relations of pitched sounds in the signs of harmony and of musical composition were susceptible of such expressions and adaptations, the engine might compose elaborate and scientific pieces of music of any degree of complexity or extent. (Bowles 1970, p. 4)

Such an idea that computers can be utilized for musical creation was brought into practices in the later decade when the digital computer was invented. Even in the early days of computing when a computer was so huge that it even occupied a whole room and accessible only to specialists in research institutions, composers of the time showed significant interests in this new territory of creative practices in music exploited by computer technology.

Along with the advance of technology, the new domains for creative expressions in sound and music have been enlarged to be explored by artists and musicians until today. Yet, it should be also noted that the historical development of computer systems and languages was not just supported by the advance of technology but also by the necessity for creative explorations among artists and computer musicians; throughout the history of computer music, there can be observed the mutual influences between technology and creativity.

The following sections give the historical development of computer music systems and languages with a certain emphasis on such synergy between technology and creativity.

The Early Era Before Digital Sound Synthesis

The utilization of digital computers in the early era of computing first began as the exploration of algorithmic composition for musical instruments, as digital sound synthesis was not popularized yet; as described in the next section, digital sound synthesis was not realized until 1957, and it took some more time before computer music software became more accessible to composers.

In 1956, Hiller and Isaacson composed *ILLIAC Suites* for string quartet, which is known as the first algorithmic composition that involves a computer (ILLIAC I at the University of Illinois at Urbana-Champaign). Hiller also developed MUSICOMP for the ILLIAC (and later rewritten for the IBM-7090 computer), which was a library of subroutines for automated compositions, together with Baker around the early 1960s. MUSICOMP is also known to be utilized for Hiller's composition, *Computer Cantata* (1963) (Hiller and Baker 1964).

Such an interest to use computer programs for computer-assisted algorithmic composition was also seen in other contemporary music composers of this era. For example, Iannis Xenakis, a Greek-French composer, is another pioneer of algorithmic composition. He developed the Stochastic Music Program in the early 1960s, which is a computer implementation of the method of stochastic music that he had previously developed and used by hand; for instance, *Achorripsis* by Xenakis is known as a stochastic composition where he applied the stochastic method by hand (Childs 2002).

The Stochastic Music Program was used to compose a series of works such as *ST/10, 1–080262* for 10 players (1956–1962), *ST/48, 1–240162* for 48 players (195962), *Atrées* for 10 players (1962), and *Morsima-Amorsima* for piano, violin, cello, and double bass (1962) (Myhill 1978). The detail of the program is informally described in his book, *Formalized Music* (Xenakis 1992).

For another example, Gottfried Michael Koenig, a German-Dutch composer, is also well known for developing programs for computer composition in the 1960s. Koenig describes his program Project 1 (1964–1966) as “a tool for composition – theoretical investigation (rather than for ‘computer-aided composition’)” and sees the program “as a composer’s tool for reflecting the compositional process in music (but not only music) empirically, and for developing a personal theory of the composer’s process” (Koenig 1991).

This program by Koenig was “based on an extension of serial technique which was prevalent in the 1950s” and Project 1 “outputs information for manual transcription concerning the following parameters: timbre, rhythm, pitch class, octave register, and dynamics” (Buxton 1977).

Based on the experience with Project 1, Koenig also developed Project 2 (first version 1965) with more flexibility. Project 2 allowed the user to “specify the selection principle to be used in generating values for a particular parameter in the output score” (Roads 1996, p. 840). Both programs required the user (the composer) to creatively transcribe and interpret the output from the program into musical scores.

Koenig utilized these programs for his computer-aided compositions *Version 1* for 14 instruments (1965–1966) and *Version 3* for 9 instruments (1967), *Segment 1–7* for Piano (1982), *Segment 99–105* for violin and piano (1982), and *3 ASKO Pieces* for chamber orchestra (1982) (all these works were composed with Project 1) (Koenig 1983), and *Übung für Klavier* (1969) (with Project 2) (Ames 1987).

Such interest in algorithmic composition and computer-assisted composition in this era was taken over by composers in the later decades.

The Eras of Non-Real-Time Computer Music Systems and Languages

The CSIR Mk1 computer in Australia, which was later renamed to CSIRAC, is known to be the first computer to generate sound. A loudspeaker was built in the CSIR Mk1 computer as an output device for warning purpose, since the computer lacked any device for visual feedback. Geoff Hill programmed the computer in 1951 so that it can play musical melodies by sending pulses to the loudspeaker (Doornbusch 2004).

It is also reported by Doornbusch that a similar experiment may have been performed on another early computer, the Ferrati Mark I, and that the tape recording seems archived in the British National Sound Archive, while he also denotes that the detailed documentation for this experiment could not be found (Doornbusch 2004).

However, these experiments didn’t give any significant influence to computer music research in the later decades and are rather considered just as one-time experiments today; while these experiments involved computers to control sounds, sound samples were not computed digitally.

The most significant contribution in the early days of computer music history was made at the Bell laboratory by Mathews and his colleagues. The series of

Fig. 1 An IBM 704 computer – installed at NASA in 1957 (This image is public domain. NASA copyright policy states that NASA material is not protected by copyright unless noted)

computer music systems and languages that they built, which is now called MUSIC-N languages, established the basis for computer music systems and languages in the following decades and gave significant influences to the evolution of computer music languages.

Mathews and his colleagues developed MUSIC-I in 1957, which is known today as the first computer program to synthesize sound digitally (Mathews 1961). They developed MUSIC-I on an IBM 704 mainframe computer (Fig. 1), the capability of which is significantly limited in comparison with computers that exist today. The available memory space was just 4,096 words (36 bits per word) in its magnetic core memory and the computational speed was only 0.006-0.04 MIPS (Dean 2009, pp. 47–50); 1 MIPS stands for million instructions per second, which means that an IBM 704 could process only 6000–40,000 instructions per second; as it can be easily inferred from such limitation of the computational resources, computer music systems in this era could perform only non-real-time digital sound synthesis.

As it may be also inferred from the fact that it was the first digital sound synthesis program in the history of computer music, the capability of MUSIC-I was significantly limited. It could generate only a monophonic melody with triangle wave. While MUSIC-I is considered still more like a program for digital sound synthesis rather than a computer music language, it should be noted that MUSIC-I could take an input of a score file, even though only "a patient user could specify notes only in terms of pitch, waveform, and duration" (Roads 1996, p. 87).

It also required a significant amount of effort to synthesize the sound in those days. After Mathews and his colleagues performed the computation for digital sound synthesis on the IBM 740 computer at IBM World Headquarters in New York and stored the result onto a digital magnetic tape, they had to bring it back to Bell Telephone Laboratories in New Jersey so that they can convert the data to actual sound by a 12 bit vacuum tube converter (Roads 1996, p. 87) (Roads and Mathews 1980). Only after such a labor, the digitally generated sound became audible to human ears.

Following the development of MUSIC-I, Mathews and his colleagues then developed MUSIC-II in 1958. As the successor digital sound synthesis program to MUSIC-I, MUSIC-II had better capability in both timbre and the available number of voices. According to Mathews, "MUSIC-II was capable of four independent voices with a choice of 16 waveforms stored in memory" (Roads and Mathews 1980). As such descendants of the MUSIC-I language are normally referred as MUSIC-N languages.

Around 1960, Mathews and his colleagues had an access to IBM 7094, which is "one of the biggest, fastest computers available" around that time, which was "able to add floating numbers at a speed of about 0.35 MIPS" with 32 kilowords of 36-bit-word memory for standard 7094 (Walden et al. 2011). Making the best use of one of the most powerful mainframe computers of their time, Mathews and his colleagues developed MUSIC-III in 1960 (Dean 2009, p. 96; Roads and Mathews 1980).

MUSIC-III is known as the first computer music language that introduced the *unit-generator* concept in its design. The unit-generator concept is an abstraction for digital sound synthesis, which performs "conceptually similar functions to standard electronic equipment used for electronic sound synthesis" (Mathews et al. 1969, p. 15). From the perspective of implementation, A unit-generator is "a software module that emits audio or control signals (envelopes) or modifies these signals" (Roads 1996, p. 787), "which can be interconnected to form synthesis instruments or patches that generate sound signals" (Roads 1996, p. 89).

Figures 2 and 3 show the pictorial representation of an instrument composed of unit-generators and its definition respectively, as described in *An Acoustic Compiler for Music and Psychological Stimuli,* a paper by Mathews published in 1961 (Mathews 1961). The definition of the instrument in Fig. 3 was given to the IBM 7094 computer as punched cards, as it still lacked human-friendly interfaces such as a CRT monitor and a keyboard.

While some may assume that the unit-generator concept is derived or inspired from the modular synthesizers, Mathews states that both the unit-generator concept and the modular synthesizers were developed fairly simultaneously in an interview (Roads and Mathews 1980), and in fact, as noted by Park in another interview to Mathews (Park 2009), MUSIC-III, which introduced the unit-generator concept, was developed in 1960, clearly before Ketoff developed the Synket synthesizer in the early 1960s and also before Buchla and Moog "independently developed the first voltage-controlled modular synthesizers" in 1964 (Pinch 2001); it is fair to consider the unit-generator concept was invented at least simultaneously and independently with the modular synthesizers (There can be some argument if the Synket synthesizer can be considered as an early example of the modular synthesizer, as it was not fully voltage controlled as seen in the Moog synthesizer. However, it was at least a patchable synthesizer, though it was controlled mainly by many buttons and the flexibility in patching was quite

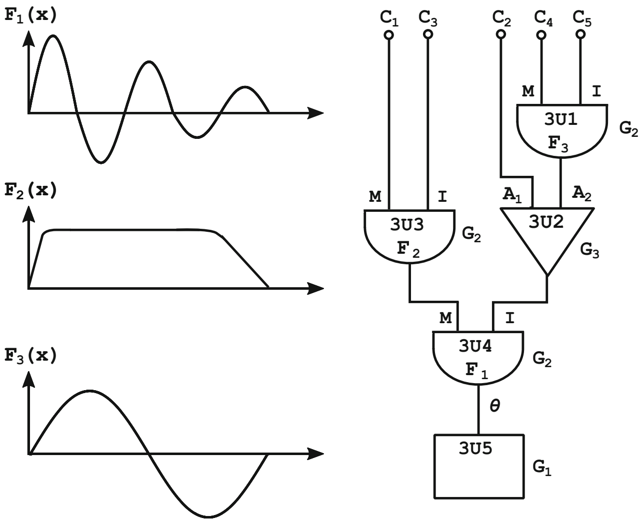


Fig. 2 An Instrument with attack, decay, and vibrato, following (Mathews 1961)

Card Columns		
8910	...	16 ... 72
MAC		G2,3U1,F3,(3U2,A2),x,x
MAC		G3,3U2,(3U4,I),x,x
MAC		G2,3U3,F2,(3U4,M),x,x
MAC		G2,3U4,F1,(3U5,θ),x,x
MAC		G1,3U5
MAC		S1,I3,((3U3,M,C1)(3U3,I,C3)$
ETC		(3U2,A1,C2)(3U1,M,C4)(3U1,I,C5)$
ETC		(3U3,S,PO)

Fig. 3 The definition of the instrument described in Figure 2, following (Mathews 1961). The definition is given to the mainframe computer as punched cards

limited (according to Luigi Pizzaleo, in an email exchange dated September 3, 2015). The year when the Synket synthesizer was developed differs among publications, ranging between 1962 and 1965 (Harrington 2009, p. 251) (Jenkins 2009, p. 21) (Holmes 2012, p. 480) (Eaton 1968). Luigi Pizzaleo remarks that "Ketoff was asked to set up the electronic lab at the American Academy in Rome" in 1962, "but the Synket could not be accomplished before 1964" (in an email exchange dated September 3, 2015). According to John Eaton, the development of the Synket was begun in 1962 and delivered to the American Academy in Rome in 1963. Eaton also notes that he used the Synket in his live performance in 1964, which is considered to be the first live performance with the modular synthesizer in history, but also comments that "it may depend on how one defines modular synthesizer" (in an e-mail exchange dated August 25, 2015), as the flexibility in patching was still significantly limited in the Synket in comparison with the modular synthesizers as we know today; hence, it would be fair to consider the Synket was already developed and playable at least in 1964.).

The unit-generator concept is considered "one of the most significant developments in the design of digital synthesis languages" (Roads 1996, p. 89) and the majority of existing computer music languages still integrate it within the language design as the core abstraction for digital sound synthesis even today. While some recent computer music languages often extend the concept, for instance, by adding the methods (of the object-oriented programming concept) to control the behavior of a unit-generator, the basic concept of unit-generators itself remains almost the same as it was invented in 1960.

Hence, MUSIC-III is considered one of the computer music languages of significant importance in the history of computer music in that it is the language that established two core abstractions in computer music language design. One is the unit-generator concept and the other is the separation between *orchestra* and *score*. The former, *orchestra*, describes the definitions of *instruments* for digital sound synthesis and the latter, *score,* describes the scheduling of musical events to play such instruments, which can also specify parameters for each event independently. Such musical events are often called *notes* in the descendant languages of MUSIC-III.

Considering the aspect that computer music languages are domain-specific languages (Fowler 2010) designed for researchers and musicians, it may be quite impressive that two core abstractions for computer music languages were already established in 1960, in the very early stage of computer music history; it was 1957 when the first commercial release of FORTRAN, which is known as the first high-level programming language in the history of computing (Backus 1978).

This abstraction that separates *orchestra* and *score* is also frequently seen in many computer music languages today, while there also exists a trend to remove such a boundary between sound synthesis and control algorithms, as described in the later section.

It should be noted that the motivation for the invention of such abstractions is derived largely from the domain-specific needs to assist creative explorations in digital audio synthesis. Mathews and his colleagues describe that "the need for a simple, powerful language in which to describe a complex sequence of sound" in (Mathews et al. 1969, p. 34) led to the invention of the unit-generator concept. James C. Tenney, an American composer, describes the benefit of the use of computer score as follows:

> Another obvious advantage in using the computer to generate musical sounds is that the process is automatic. That is, an entire composition can be recorded on the tape with no slicing, editing, timing, mixing, etc. In fact, no manual operation of any kind is required of the composer. His 'score' (in the form of a deck of punched cards) must go through several stages of transformation, using certain special devices in addition to the computer, but the composers is not directly involved in these intermediate operations. The final result is an ordinary magnetic tape recording of the sounds and sound sequences described in the score (Tenney 1963).

Not just Mathews and his colleagues but also other researchers continuously made effort to update MUSIC-N languages. After Mathews and Miller recoded MUSIC-III as MUSIC-IV with a new macro assembly language (Mathews and Miller 1979; Tenney 1963), other researchers began working on its variants. Winham and Howe developed MUSIC-IVB in 1965, which included many additional features to the original MUSIC-IV, and then rewrote it as MUSIC-4BF in FORTRAN II and BAL (Basic Assembly Language) assembler in 1967. Vercoe developed MUSIC 360 for IBM 360 in BAL assembler in 1969. MUS10, which is often also referred as MUSIC-10, was also developed in 1966 by Poole, which emulates the MUSIC IV program on the PDP-10 minicomputer (Roads 1996, p. 789) at Stanford Artificial Intelligence Laboratory (SAIL). While the main part of MUS10 was initially written in the PDP-10 assembly language by Poole, many functions were written also in FORTRAN IV by Chowning.

As seen in the involvement of the FORTRAN language in the development of these MUSIC-N languages listed above, the emergence of high-level programming languages had a significant impact on the development of the later MUSIC-N languages, especially in software portability, and led to the development of machine-independent computer music languages. For instance, MUSIC-V was fully written in FORTRAN IV in 1966 (Mathews et al. 1969).

The development of fast minicomputers and Unix workstations in the 1980s also significantly motivated to port MUSIC-N languages to various environments. Beauchamp developed MUSIC-4C, rewriting MUSIC-IV in the C programming language for DEC VAX-11 (Beauchamp 1993). Around that time, computers became a lot smaller than IBM 7094 (Fig. 1) as shown in Fig. 4 (VAX-11/750). Thanks to the portability of the C programming language, MUSIC-4C was also ported to Unix workstations of the time (Beauchamp 2003).

Among many MUSIC-N family languages, MUSIC-11 was especially noteworthy. While Varcoe and his colleagues first implemented MUSIC-11 in the Macro-11 assembly languages for the PDP-11 minicomputer in 1973, they recoded it in the C programming language for VAX-11 minicomputers in 1986 (Vercoe 1993). This recoded version is known as the first version of Csound (Boulanger 2000), which became one of the most widely used computer music languages in the following decades.

MUSIC-11 is also notable as the first language "designed with two separate levels of processing *control signals* and *audio signals*" (Vercoe 1982) for computational efficiency. Such a design to separate *control rate* and *audio rate* is still frequently seen even among the most recent computer music languages and systems.

The rapid advance of computer technology also released researchers and composers from the tiring procedures. In the earlier days, when punched cards were still required to execute computer music languages, turnaround time for digital sound synthesis was one of the deficits in comparison with analog sound synthesis by electronic devices. Tenney discusses such an issue in his paper from the viewpoint of a composer, while he also found the use of computer score quite beneficial as mentioned earlier in this section (Tenney 1963).

Fig. 4 VAX 11/750 – as exhibited in Vienna Technical Museum (Photo by Dave Fischer. This image is from Wikimedia and licensed under the Creative Commons Attribution-Share Alike 3.0 Unported license)

> Automation brings with in a certain indirection – a time delay (of from a few hours to a few days between the composer's decisions or experimental actions and any auditable results. The standard tape and electronic music techniques thus have a definite advantage over the computer medium, with respect to their immediacy, and the grater ease with which unfamiliar sounds may be tried and tested before actually being used in a composition (Tenney 1963).

In this era, with faster computers and better human-computer interfaces such as CRT terminals (Fig. 5), turnaround time in compositional process of computer music was significantly reduced and creative exploration by composers was facilitated a lot, even though computers were not fast enough to perform digital sound synthesis. MUSIC-11 is also particularly notable for the development of human-computer interfaces for computer music, as two GUI applications for computer music, Musical Score Editor and OEDIT, were developed for MUSIC-11 at MIT (Vercoe 1982) (While the paper by Vercoe (1982), which describes these two GUI applications, was published in 1982, it is described that Steiger and Hale already have developed OEDIT in 1980 in (Puckette 2002)). As shown in Figs. 6 and 7, Musical Score Editor and OEDIT allow the user to directly manipulate a graphically represented score and a unit-generator graph respectively. Musical Score Editor and OEDIT could generate a score file and an orchestra file respectively to render a sound file by MUSIC-11. While there already existed commercial digital synthesizers with GUI interfaces around the same period, such as Fairlight CMI Series II (Fig. 8) and the Qasar M8, the software design of OEDIT clearly is a precursor of graphical computer music programming languages developed in the later decades.

While Music-N languages played a significant role in establishing the foundation of computer music language design, there were other non-real-time computer music languages developed in this era that are worthy to mention.

Cmix was developed by Lansky in 1984 (Lansky 1987). He describes Cmix is "essentially a toolkit for synthesis and analysis" and "differs substantially from

Fig. 5 DEC VT-100 Terminal (Photo by Jason Scott. This image is from Wikimedia and licensed under the Creative Commons BY 2.0 license)

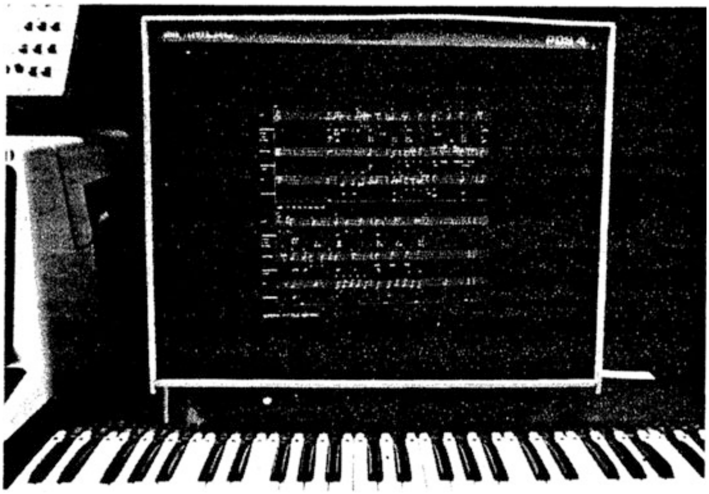


Fig. 6 Musical score editor: manipulating data on a screen in standard music notation (Vercoe 1982). (Courtesy Audio Engineering Society)

most synthesis packages in that it has no scheduler and accumulates mainly by mixing to disk" (Lansky 1990). However, the software package of Cmix included MINC, a tiny programming language that resembles the C language in its syntax so

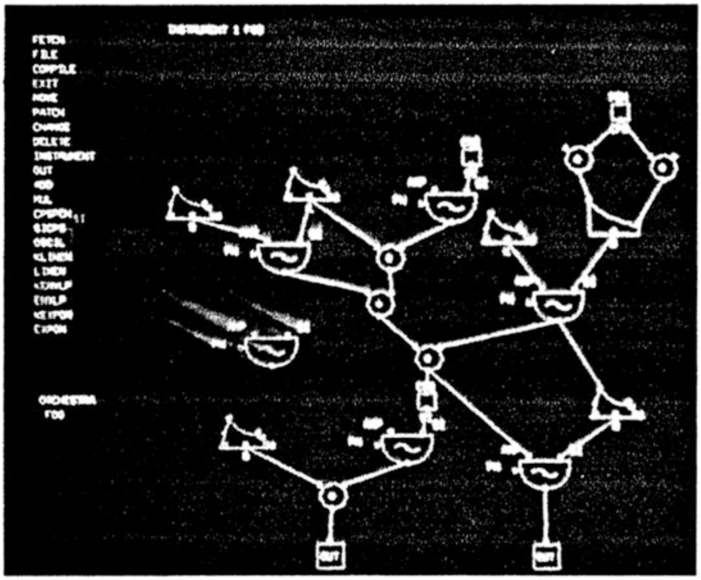


Fig. 7 OEDIT, an orchestra singal-processing network editor: moving a new oscillator into position (Vercoe, 1982). (Courtesy Audio Engineering Society)

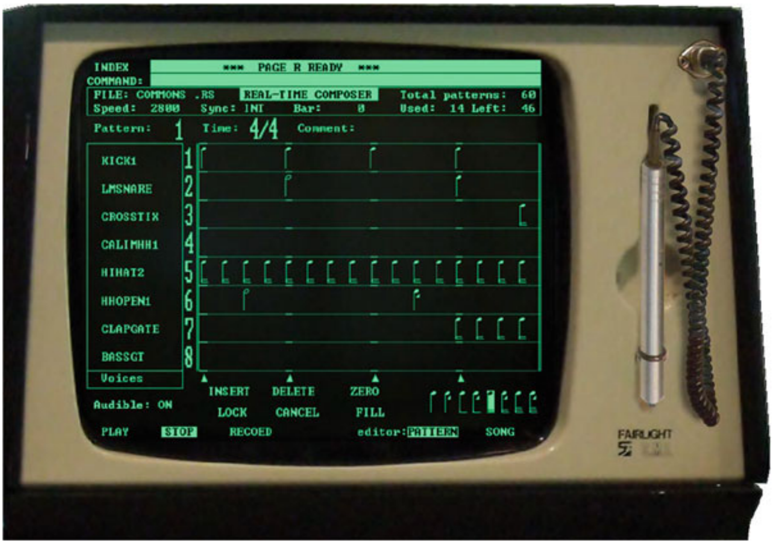


Fig. 8 The "Page R" (Real-Time Composer) on Fairlight CMI II (Photo by Kevan Davis. This image is from Wikimedia and licensed under the Creative Commons Attribution-Share Alike 3.0 Unported)

```
/* START:
p0=start; p1=dur; p2=pitch(oct.pc); p3=fundamental decay time
p4=nyquist decay time; p5 = amp; p6=squish; p7=stereo spread [optional]
p8=flag for deleting pluck arrays (used by FRET, BEND, etc.)[optional]
*/

rtsetparams(44100,2)
load("STRUM")
makegen(2, 2, 7, 7.00, 7.02, 7.05, 7.07, 7.10, 8.00, 8.07)

srand(0)
for (st = 0; st < 15; st = st + 0.1) {
  pind = random() * 7
  pitch = samplefunc(2, pind)
  START(st, 1.0, pitch, 1.0, 0.1, 10000.0, 1, random())
}
```

Fig. 9 A MINC program example (Taken from STRUM1.sco, which is a part of the RTcmix 4.0 package released under GPL license)

that it can facilitate creative exploration. Figure 9 describes a simple example in MINC. While it is not a direct descendant, Cmix seems also significantly influenced by MUSIC-N languages. For instance, the parameters for Cmix's *makegen* function, which can be used to generate amplitude curves, waveform data, and such, are almost the same as MUSIC-N languages.

The CARL system developed by Moore and Loy in 1986 would be another noteworthy non-real-time digital sound synthesis computer music system (Loy 2002) in this era. The CARL system was "a collection of small, command line programs that could send data to each other" (Wang 2008) on a UNIX system. Making the best use of Unix *pipe*, which "is a mechanism to pass the output of one command to the input of another, without creating intermediate files" (Hanson et al. 1984), these small command-line programs in the CARL system, which process digital signals, "may be piped together, which is tantamount to connecting the programs together in a cascade fashion" (Moore 1982).

Figure 10 describes a simple example in CARL by Moore (1982). As shown, these cascaded commands first generate an input score from the file *score.c* by the cscore program and then process it to synthesize the sound output by the cmusic program. The output sound is applied a low-pass filter by the filter program and then sent to the sound output.

Even though non-real-time digital sound synthesis still required a considerable labor, computer music languages of their time already played significant roles in both research and practices. As Risset describes, "the exploration of the musical possibilities of digital synthesis has contributed a lot to the growth of such psycho-acoustic knowledge and knowhow" (Risset 2003).

For example, Moore describes that Mathew's synthesis program contributed to one of the key findings of their time in psychoacoustics that "fairly large-looking differences between two sound waves don't necessarily correspond to large-

```
% cscore score.c | cmusic | filter lowpass | sndout
```

Fig. 10 A command line example in CARL (Moore 1982)

sounding differences, and the opposite is also true" because it "allowed people to specify the physical properties of virtually any musical sound" in (Moore 1996). For another example, Mathews and Pierce investigated musical effects of tones with nonharmonic partials, "to give a better understanding of the subjective basis of harmony and to explore the musical use of sounds with nonharmonic partials" (Mathews and Pierce 1980). The MUSIC-V language was utilized for this experiment, since "by using a computer to generate sounds, one can produce tones which include partials of any chosen frequencies" (Mathews and Pierce 1980).

Not just limited to such a use in psychoacoustic experiments, computer music languages also significantly facilitated creative investigation in timbres and sound textures. The most notable early example would be Risset's *Introductory Catalogue of Computer Synthesized Sounds,* realized by MUSIC-V (Risset 2003). Computer music languages contributed even to the invention of one of the most widely used sound synthesis techniques, frequency modulation synthesis (FM synthesis) (Chowning 1977). John Chowning, an American Compsoer, discovered this new sound synthesis technique in 1967, while he was experimenting on MUS10 (Chowning 2011). This sound synthesis technique was later commercialized by Yamaha as a DX-7 digital sound synthesizer, which is known as one of the most successful digital synthesizers in history.

As computer music languages were already involved in research and creation around this time, the outcome of their research seemed immediately reflected to their creative practices. Risset's composition *Mutation* (1969) uses many sound synthesis techniques that he discovered with the MUSIC-N languages. Chowing composed a series of pieces that utilizes FM sound synthesis techniques such as *Turenas* (1972), *Stria* (1977), and *Phoneé* (1980–1981).

It also should be remarked that the research on computer-assisted composition and algorithmic composition also continued in this era and a number of computer music languages and systems were developed; the list of well-known examples developed in 1970s and 1980s includes TEMPO by Clough (1970), SCORE by Smith (1972), PLAY by Cadabe and Meyers (1978), Tree and Cotree by Roads (Roads and Wieneke 1979), and GGDL by Holtzman (1981).

The Experiments in Musical Intelligence (EMI) software by Cope would be one of the most notable examples of the software for algorithmic composition in the 1980s. Cope began working on the software in 1981, which is an expert system that "creates new but stylistically recognizable music from existing works" (Cope 1991). Cope kept on developing the EMI software for following years and it is reported the resulting software "successfully composed music in the styles of Cope, Mozart, Palestrina, Albinoni, Brahms, Debussy, Bach, Rachmaninoff, Chopin, Stravinsky, and Bartok" (De Mantaras and Arcos 2002; Cope 1987) (Expert systems are “information systems that represent expert knowledge for a particular problem area as a set of rules, and that perform inferences when new data are entered” (Raynor 1999, p. 100)).

Thus, the emergence of digital sound synthesis and computer music languages that facilitates creative explorations contributed to both research and practices in music, while the researchers and composers in this era also took over the research on algorithmic composition and computer-assisted composition.

The Eras of Real-Time Computer Music Systems and Languages

Early Real-Time Computer Music Systems

Even in the early eras before real-time digital sound synthesis is realized, researchers and composers were attracted to the idea to involve computers in live electronic music. The approach taken in this era to realize live electronic music was to control analog sound synthesizers by computers. Buxton noted that “whereas digital synthesis requires a minimum of 32,000 samples per second, hybrid systems only need approximately 100 for each device being controlled” in 1977 (Buxton 1977); while computers were not yet fast enough to perform digital sound synthesis, they became fast enough at least to process such musical events to control analog synthesizers around 1970.

The GROOVE system by Mathews and Moore is one of the early notable examples that took this approach. Mathews and Moore describe the system as “a program to compose, store, and edit functions of time,” which controls analog synthesizers by a small computer (DDP-224). With attached interfaces such as knobs and joysticks as well as a display and a keyboard, the GROOVE systems could perform high-level musical control (Mathews and Moore 1970). This approach to combine a computer with analog synthesizer peripherals was followed by many researchers and engineers and adopted to such live computer music systems as the Hybrid IV system (Fedorkow et al. 1978), PIPER (Gabura and Ciamaga 1967), the Yale synthesizer (Friend 1971), and MUSYS (Grogono 1973) and SYNTOM (Teruggi 2007).

This approach was followed by the emergence of personal computers and MIDI (Musical Instrument Digital Interface) (MIDI Manufacturers Association 1996), replacing minicomputers and analog synthesizers with personal computers and digital synthesizers respectively in the 1980s.

The Emergence of Variable-Function Digital Signal Processors

Digital sound synthesizers became commercially available in the late 1970s. For example, the early commercial digital synthesizers such as New England Digital Corporation-released Synclavier I and Fairlight’s Qasar M8 were already available in the late 1970s (Roads 1996, p. 918; Anonymous 1978). Yet, while the emergence of such commercial digital sound synthesizers made live presentation of digital sound a lot easier to realize, the flexibility in exploring new sound materials was significantly limited in comparison with non-real-time computer music languages with the unit-generator concept, since these early commercial digital synthesizers were normally built upon *fixed-function* digital signal processors (DSPs).

However, the emergence of *variable-function* DSPs finally realized sound synthesis based on the unit-generator concept even in real time. Unlike fixed-function DSPs, variable-function DSPs provide a considerable degree of programmability and such a feature of variable-function DSPs was quite favorable to develop hardware with more flexibility in digital sound synthesis.

The "Samson Box" is one of the noteworthy early examples of such digital synthesizer hardware in history. The hardware was developed by P.R. Samson around 1980 (Samson 1980) and installed at the Center for Computer Research in Music and Acoustics (CCRMA), Stanford University. A series of software were developed for the hardware and used for musical creation, e.g., the Pla language by Schottstaedt (1983) and MUSICBOX by Loy (1981). It should be noted that the Stanford Artificial Intelligence Language (SAIL), which was widely used for the Artificial Intelligence research, was also utilized in the software development for the Samson Box by Schottstaedt and Loy. The language is also used to generate the "note card" images for MUS10 by Chowning for *Stria* (1977) (The readers should note that the same acronym (SAIL) was used for Stanford Artificial Intelligence Laboratory (as described in the earlier section) to avoid confusion).

Di Giugno and his colleagues at IRCAM developed a series of variable-function DSP hardware for digital sound synthesis between the late 1970s and the early 1990s. First, the IRCAM 4A was developed in 1976 (Puckette 1991c). While 4A was not so flexible and just had 256 digital oscillators and matching envelope generators to perform additive synthesis in real time, its successor hardware 4B (Alles and di Giugno 1977) and 4C (Moorer et al. 1979) were developed in the late 1970s, making some extensions to the sound synthesis capability such as FM synthesis.

As seen in the earlier hybrid systems in which computers control analog synthesizer hardware, IRCAM developed control software for these digital synthesizer hardware. Such software for digital synthesizer hardware often designed with the functionality to define how digital sound synthesis be performed, to make the best use of the variable-function DSPs.

The SYN4B language was developed by Rolnick and Prevot for the IRCAM 4B synthesizer. The language could not just play the 4B synthesizer with both real-time control and predefined scores but also could specify oscillator connections to be processed inside 4B. 4CED developed by Abbot is a software package, which contains a unit-generator language and a score language to control an IRCAM 4C digital synthesizer. The 4CED software also supported the interaction by allowing each 4CED score to act as "an independent process that could accept input data and trigger events that caused other scores to start playing" (Roads 1996, p. 806).

4X, the next successor in the series of IRCAM hardware, was developed around 1980 (Di Guigno and Gerzso 1986) and is one of the most significant landmarks in the history of computer music systems (Fig. 11). While 4A, 4B, 4C were "pure

Fig. 11 The IRCAM 4X (the hardware with the cables in the racks in the center) (Photo by Martin Guy, This image is from Wikimedia and licensed under Creative Commons Share-Alike 1.0 License)

synthesizers," 4X was capable of performing "signal processing" of live instruments (Puckette 1991c). Such a capability of real-time signal processing contributed a lot to realize some of the most important contemporary music compositions of the decade. To name some, *Répons* (1981–1984) (Boulez and Gerzso 1988) by Pierre Boulez, *Dé sintégration* (1982–1983) by Tristan Muraiil (Smith 2000), and *Jupiter* (1987, revised 1992) by Philippe Manoury (May 1999) were all composed utilizing the 4X hardware.

The researchers and engineers developed computer music software also for 4X. For instance, Rowe and Koechlin developed 4xy, a compiler of control programs for 4X, and Potacsek developed the 4X patch language, a visual programming language for real-time sound control (Favreau et al. 1986). Yet, the most notable software developed for 4X in this decade would be MAX by Puckette, when considering the influence that the software made to the history of computer music.

According to Puckette, MAX is originated in the m orchestra language for MUSIC 500 system that he developed at MIT (Puckette 1983). The early version of MAX, which was developed in the 4xy language, was "a real-time control system" that configures control processes (or objects) and manages message-based communications between the control processes (Favreau et al. 1986). Puckette describes that the motivation to develop MAX is making it possible "to design elements of a system which can be combined quickly and without changing code" (Favreau et al. 1986). Such an issue was a topic of interest in the decade, since rapid prototyping was one of the most important factors to facilitate the collaboration between composers and engineers.

While the earliest version of MAX was configured by text files, soon after its emergence, Puckette developed the Patcher software, a graphic programming language for the same system (Puckette 1988). While the Patcher program and early versions of MAX could only control MIDI synthesizers or the IRCAM 4X

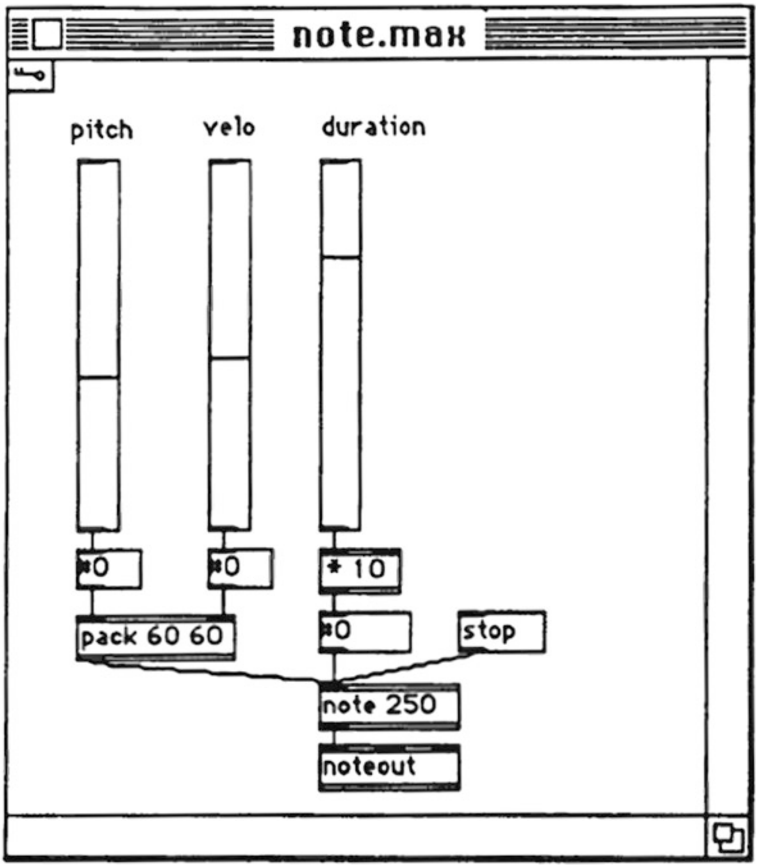


Fig. 12 A Patcher program by Puckette (1988) (Courtesy of Miller Puckette)

hardware by MIDI interfaces, without the capability to perform stand-alone real-time sound synthesis, its design significantly influenced other graphic programming environments to be developed in the later decades. As seen in Fig. 12, the programming environment that the Patcher program provides was already quite similar to what is known as MAX today.

Even though there was some description in (Puckette 1991c) that Di Giugno and his colleagues began prototyping 5A at IRCAM around 1985, with the intention to introduce new functionalities such as floating-point arithmetic and jump instruction, it seems that the 5A project was never completed, as Di Guigno left IRCAM soon in the late 1980s.

Following the success of 4X, IRCAM developed the IRCAM Music Workstation (IMW), which is often referred also as ISPW (IRCAM Signal Processing Workstation) in the late 1980s. Unlike its predecessor IRCAM hardware, an IMW system

Fig. 13 An IRCAM Signal Processing Workstation (A NeXt Cube workstation with three i860 boards installed) (Photo courtesy of Mr. Jean-Bernard Emond)

was composed of "one or more NeXT host computers together with between 2 and 24 i860 coprocessors (CPs) running at 40 MHz, nominally capable of 80 million floating-point operations per second (MFLOPS) apiece. The CoProcessor Operating System (CPOS), has been written specifically to fill the requirements this hardware poses for real time musical synthesis and control" (Puckette 1991b). The pictures of an IRCAM Music Workstation and an i860 board are shown in Figs. 13 and 14 respectively.

Such an integration of digital signal processing hardware within a workstation also made it possible to develop computer music languages that can describe programs for both control algorithms and real-time sound synthesis seamlessly just within one environment. The version of MAX developed for the IMW (MAX/FTS) was indeed a great example of the computer music software of this kind.

In an IMW system, a computer program called FTS ("faster than sound") processed DSP on the CoProcessor Operating System (CPOS), which is an operating system for the i860 board (Puckette 1991b), separately from the NeXT operating system. The IMW version of MAX entirely hides such a boundary between the NeXT operating system and CPOS behind and users could combine both control algorithms and sound synthesis seamlessly within the same graphic programming environment; Fig. 15 shows a MAX/FTS program as Puckette described in (Puckette 1991a). As seen in this example, objects for digital signal processing, such as *osc1∼, line∼*, and *sig∼*, are implemented within one same graphic program with other objects for control algorithms, MIDI I/O, and GUI.

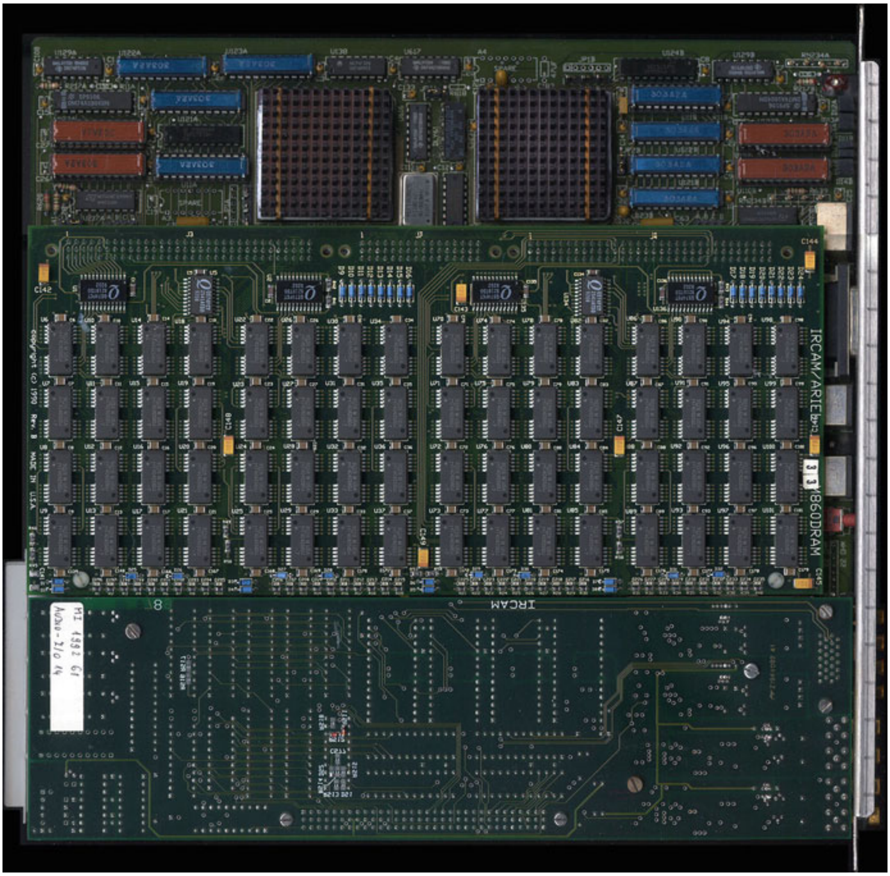


Fig. 14 An i860 board (Photo courtesy of Mr. Jean-Bernard Emond)

Not just IRCAM but also other research institutions and companies developed similar hybrid computer systems composed of a workstation/personal computer and external variable-function DSP hardware. One widely known example of the kind is the KYMA/Platypus Computer Music Workstation by Scaletti and his colleagues. KYMA was a programming environment based on the Smalltalk programming language (Goldberg and Robson 1983) and it ran on Apple Macintosh II to control Platypus, a DSP peripheral (Scaletti and Johnson 1988). The Mars workstation is another example of the same kind, which consisted of an Atari computer and a SM1000 sound generation board, which Di Giugno and his colleagues developed in Italy after he left IRCAM (Armani et al. 1992) (Cavaliere et al. 1992). They also developed its successor system composed of an IBM PC and a NERGAL sound generation board (Paolo et al. 1997). These systems also provide a graphic programming environment for sound synthesis and performance control.

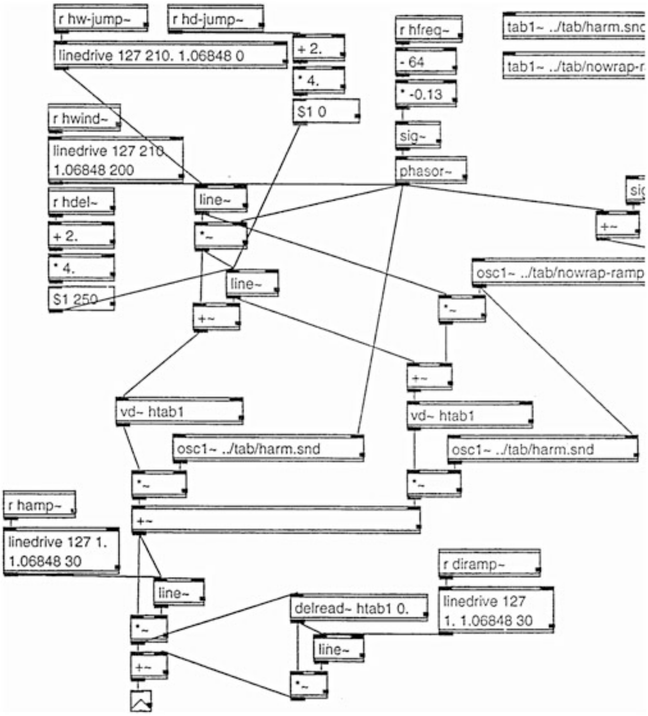


Fig. 15 A Max/FTS patch example by Puckette (Courtesy of Miller Puchette)

Syter, developed by Allouis at GRM around the mid-1980s, is another notable early example of such a system with the custom sound synthesis hardware in the same decade. Syter was a system composed of a PDP 11 minicomputer and a real-time processor called the Syter processor, which has two inputs, eight outputs, and a MIDI interface for control. The Syter system was equipped with graphic screens and mouse for real-time control for mixing/spatialization (Fig. 16) and also provided a number of "basic modular bricks" such as oscillators, envelopes, delays, harmonizers, and noise generators, which can be combined to build more complex structures for sound synthesis (Teruggi 2007).

While these real-time computer music systems developed in this era were still too expensive for personal purchase, the combination of a workstation or high-end personal computer and variable-function DSP hardware significantly reduced the cost for real-time digital signal processing for live computer music; while the market price of the 4X hardware was "beyond the reach of most every computer music studio save IRCAM" (Rowe 2004, p. 213), these systems (such as IMW, KYMA/Platypus, Mars) were affordable for many institutions, resulting in the popularization of real-time sound processing of musical instruments among composers of the time.

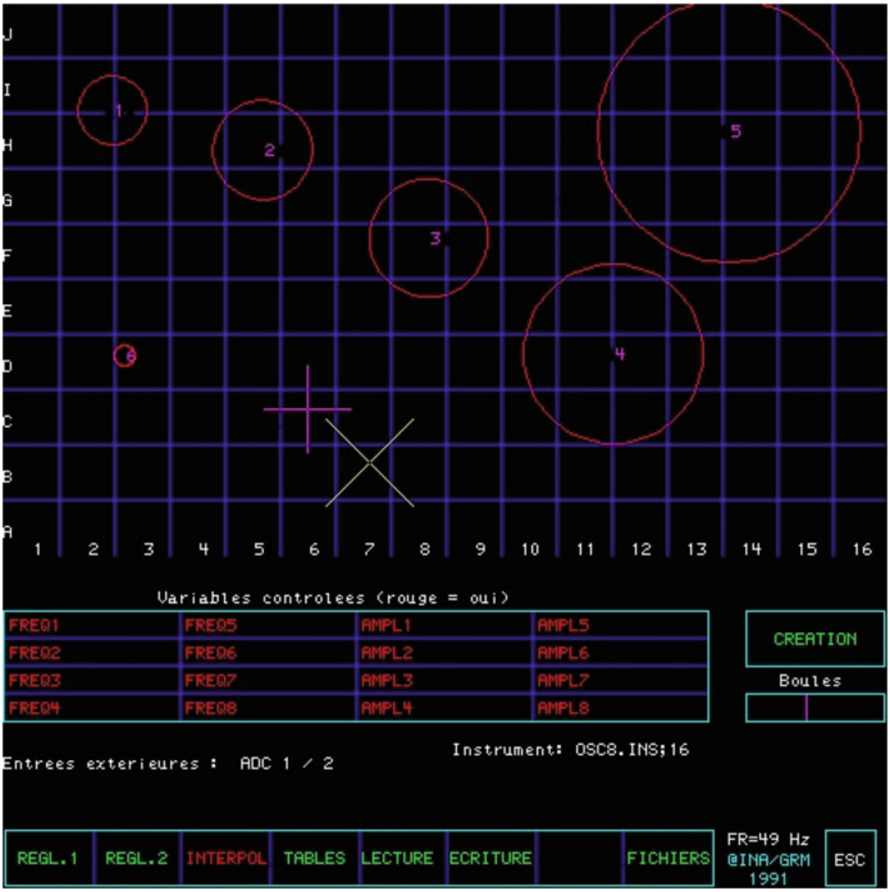


Fig. 16 The interporation screen of Syter (Teruggi 2007) (Courtesy of Daniel Teruggi)

MIDI-Based Computer Music Systems

While real-time sound synthesis still required expensive external hardware in this era, commercial digital synthesizers became rapidly affordable even for personal use. Since personal computers were already fast enough to handle musical events (such as MIDI messages) in real time, composers began experimenting with interactive music with a personal computer and MIDI synthesizer hardware, taking over the interest in algorithmic composition in the previous decade.

Some composers developed their own interactive music software from scratch in general-purpose programming languages. Rowe developed Cypher (Rowe 1992a), an interactive computer music software, for Machintosh around 1990. Cypher could analyze density, attack, loudness, register, duration, and harmony from the streams of MIDI data and interact with the input by transforming MIDI inputs or by

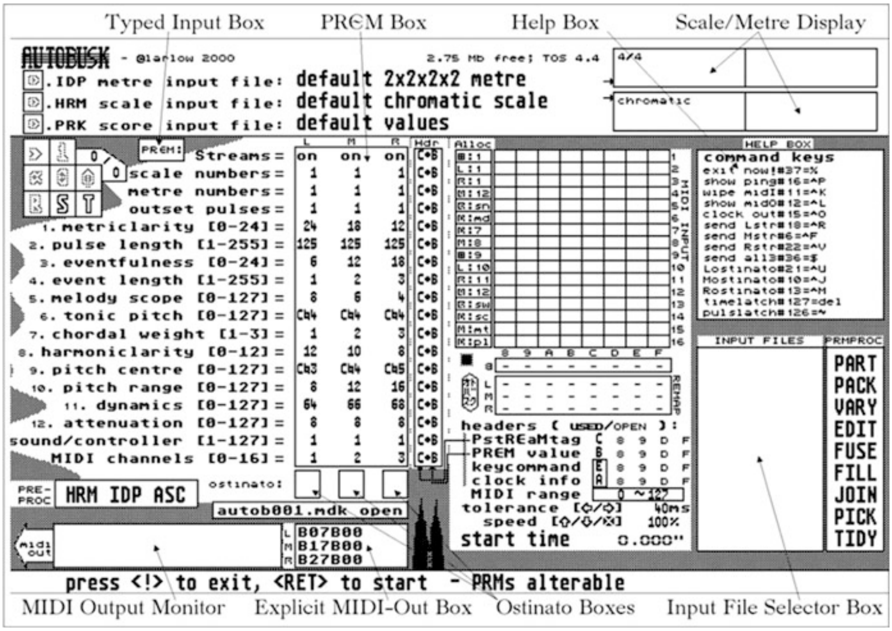


Fig. 17 A screenshot of AUTOBUSK (Barlow, 2000) (Courtesy of Clarence Barlow)

algorithmically generating new musical materials in addition to simply playback from a library of sequences (Rowe 1992b). Cypher was used for Rowe's own interactive music compositions, such as *Flood Gate* for violin, piano (1989), and interactive computer music system, and *Banff Sketches* for solo piano and computer (1989) (Koblyakov and Rowe 1992).

Barlow began developing AUTOBUSK around the mid 1980s. AUTOBUSK was first developed in Pascal on a CP/M operated computer and then ported to the ATARI ST 1040 computer (Barlow 1990, 2000). Barlow describes AUTOBUSK as a "stochastic program," which "generates time-concurrent streams of controlled random rhythmicized MIDI notes based on probabilities calculated from the Harmonicity and Indispensability formulas in the light of a fundamental consideration." Figure 17 is a screenshot of the AUTOBUSK software. Barlow also developed other interactive computer music systems such as TXMS (Textmusic packaged into software), PAPAGEI, Synthrumentator, etc. (Barlow 2011).

Chadabe and Zicarelli developed M (Fig. 18) and Jam Factory for Macintosh. Zicarelli describes each as follows: "both M and Jam Factory are programs that, through the use of graphic and gestural interfaces, provide an envrionment for composing and performing with MID. The resulting envrionment, characterized by controls that offer immediate feed-back is quite analogous to the control panel of an airplane" (Zicarelli 1987).

On the other hand, many computer music languages and software libraries were also developed for such MIDI-based computer music systems. To name some, there

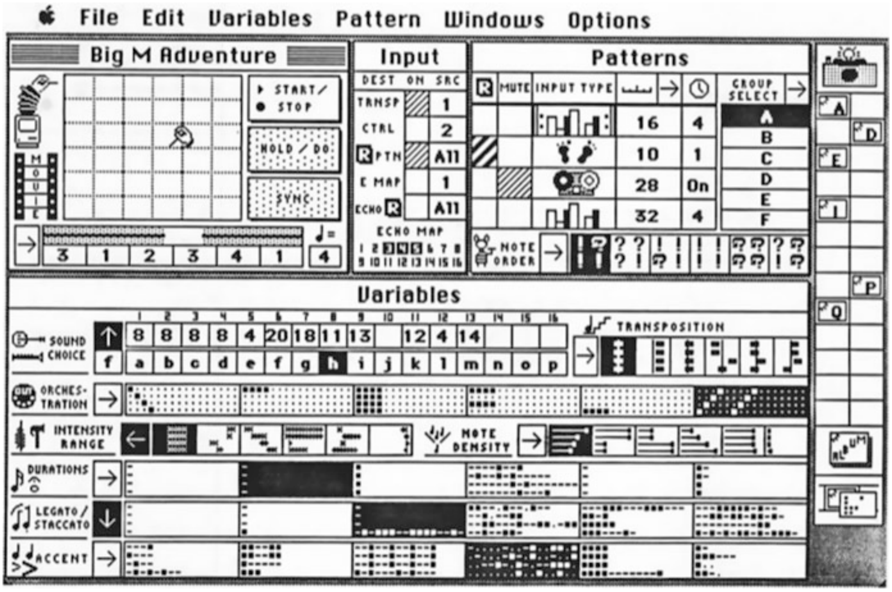


Fig. 18 A screenshot of M (Courtesy of David Zicarelli)

existed MIDI-LISP by Byonton and his colleagues (Boynton 1986), Common Music by Taube (1997), and CMU MIDI Toolkit by Dannenberg (1993). Moxie by Collinge (1984) would be also noteworthy, as it is one of the earliest known examples that present the programming pattern called *temporal recursion,* which is frequently seen in live coding performance (Collins et al. 2003) with the Impromptu computer music language today as described in the later section of this chapter (Soresen and Gardner 2010).

Among those computer music systems/languages, FORMULA (Forth Music Language) (Anderson and Kiuvila 1991) is particularly noteworthy for its software design. FORMULA was based on the FORTH programming language (Kail 1985), yet the entire FORMULA system even included its operating system (Anderson and Kuivila 1990). While computers of this decade were already fast enough to process musical events, there were still problems in synchronizing the timing of musical events on both sides of personal computers and MIDI devices. Internally, FORMULA took an approach to process musical events in *virtual time,* which is a system internal logical time. When FORMULA outputs musical events to the MIDI synthesizer, it buffers the output rather than sending it out immediately so that the events with the same timestamps can be sent to the device without the gap in real time if possible.

Together with the feature of the capability to buffer the MIDI messages with the timestamps on the synthesizer side that some MIDI synthesizers were equipped with, such software design of FORMULA contributed to improve timing precision for live

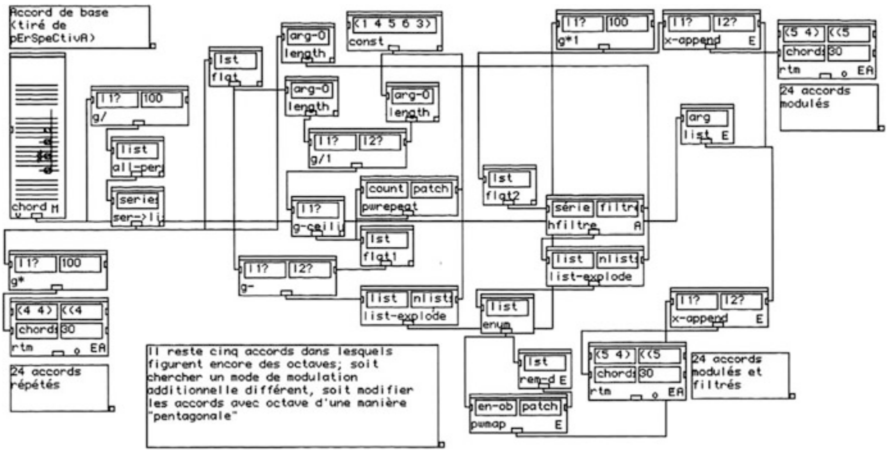


Fig. 19 A Patchwork patch (Assayag et al. 1999) (Courtesy of IRCAM)

computer music presentation; thus, the design of FORMULA represents the challenge in software design for MIDI-based computer music systems in this era and such an approach to separate logical internal time for musical events from real time (with a certain mechanism to coordinate the progress between logical time and real time) is still seen in many computer music systems today.

It should be also emphasized that MAX was licensed for commercialization and ported in 1989 to the Apple Machintosh computer. While it was still MIDI-only version of MAX, this graphic programming environment was sold at an affordable price and made it a lot easier for computer musicians to develop their own interactive music systems, even without expert programming skills.

While it was the software rather for computer-assisted composition than for interactive computer music, Patchwork is another noteworthy example of a graphic computer music language in the same decade (Laurson 1996). The development of Patchwork began in the mid-1980s by Laurson and then he began collaborating with Duthen and Rueda at IRCAM. They describe Patchwork as the "software for computer-assisted composition," which provides "a visual interface to the Lisp Language" (Assayag et al. 1999). Figure 19 shows an example program in Patchwork. While Patchwork never gained much popularity as a graphic computer music language compared to MAX (only 200 registered users around 1999), it is reported that its users included important contemporary music composers of our time (Laurson 1999). For instance, Magnus Lindeberg utilized Patchwork for his composition *Engine* (1996) (Laurson and Kuuskankare 2009).

Thus, while live computer music with real-time digital sound processing was still too expensive for personal use, the technological advance in personal computing in these decades was gradually liberating academic achievements in the previous era, making it available even to independent musicians.

The Emergence of Stand-Alone Real-Time Digital Sound Synthesis

As the computation speed of personal computers became faster, non-real-time computer music languages already began porting to personal computers before the 1990s. For instance, Music 4C was ported to Apple Macintosh and Csound was ported to various platforms such as IBM-PC, Apple Macintosh, and Amiga (Roads 1996, p. 790).

Around the beginning of the 1990s, further advance of computer technology finally made it possible to perform real-time digital signal processing even on personal computers and workstations without any dedicated special DSP hardware. The researchers and engineers began redesigning existing non-real-time computer music languages so that it can perform real-time sound processing on personal computers and workstations. Vercoe and his colleagues developed a real-time variant of Csound, which was already one of the most popular computer music languages then around 1990 (Vercoe and Ellis 1990). Garton and Topper developed RTcmix, which is a real-time version of Cmix, around 1995 (Garton and Topper 1997).

It should be also noted that such an advance contributed to improve not just real-time computer music languages/systems but also non-real-time computer music languages/systems, and the software to perform the CPU intensive sound synthesis techniques became more popularized for personal computers than before. For example, Erbe developed SoundHack (Erbe 1997) for Macintosh, which performs such techniques as convolution, spectral mutations, phase vocoding, and spectral analysis/resynthesis with the user-friendly GUI. For another example, Patchwork was evolved as OpenMusic (Fig. 20) at IRCAM, significantly extending its capability, including the sound processing and analysis feature (Bresson 2006). Such spectral processing/analysis features of OpenMusic were particularly beneficial for spectral music and spectral composers often utilized it for their compositions (Hirs et al. 2009).

The users of OpenMusic were not limited to spectral composers. Other features of OpenMusic, such as constraint programming and aleatoric composition, were also utilized by contemporary music composers with “highly diverse musical and aesthetic backgrounds” and the list of users includes well-known contemporary music composers such as Antoine Bonnet, Brian Ferneyhough, Grard Grisey, Paavo Heininen, Magnus Lindberg, Claudy Malherbe, Tristan Murail, and Kaija Saariaho (Assayag et al. 1999).

MAX, which was already commercialized for the Apple Machintosh computer in 1989, also extended its capability with real-time sound synthesis. Zicarelli developed a significant addition to the original Macintosh version of MAX as MAX/MSP around the late 1990s (Zicarelli 1998). Puckette also developed PureData (Puckette 1996) as open source software, which is a graphic computer music programming language quite similar to MAX, first for SGI IRIX and Windows NT workstations (Puckette 1997), yet was soon ported to other platforms such as Windows, Linux, and Mac OS X.

Not just stand-alone versions of extant computer music languages and systems, researchers and engineers also began developing new languages and systems. SuperCollider and Nyquist are good examples of such languages. SuperCollider (Wilson et al. 2011) was first developed by McCarthy in the mid-1990s as an object-

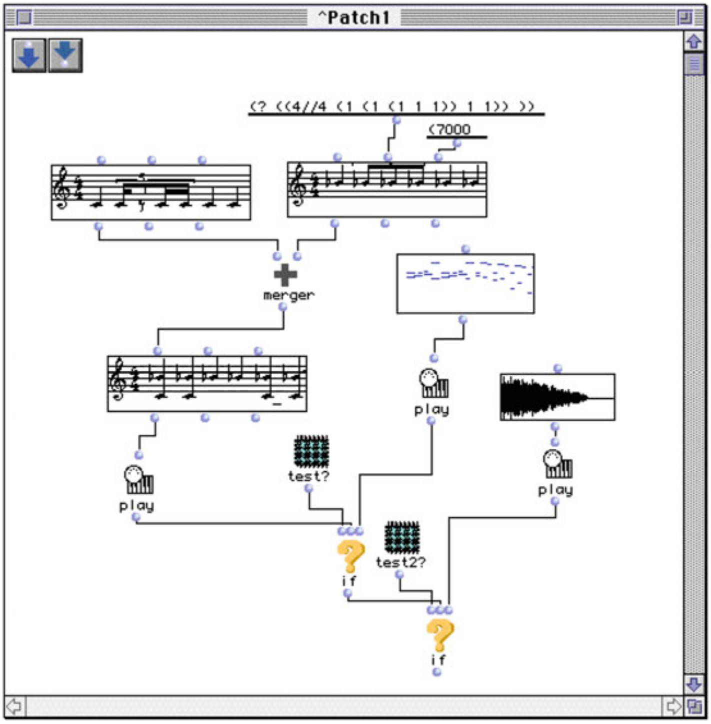


Fig. 20 An OpenMusic patch (Assayag et al. 1999) (Courtesy of IRCAM)

oriented computer music programming language, particularly with a significant influence from the Smalltalk programming language (Goldberg and Robson 1983). In a SuperCollider program, "the users can write both the synthesis and compositional algorithms for their piece in the same high level languages" so that it can support "the creation of synthesis instruments with considerably more flexibility than allowed in lower level synthesis language" (McCartney 1996). Figure 21 shows a screenshot of the programming environment of the first version of SuperCollider.

While the later versions of SuperCollider separated the programming environment to the language interpreter (*sclang*) and the real-time sound synthesis engine *(scserver),* such an idea to design a computer music language that seamlessly integrates compositional algorithms and sound synthesis was of considerable interest in this era. Indeed, as digital sound synthesis can be performed on the same CPU as the language environment unlike the real-time computer music systems of the previous generation with external sound synthesis hardware, it was made significantly easier to remove such a boundary between control algorithms/musical events and sound synthesis.

After designing computer music languages such as Arctic (Dannenberg 1984) and Canon (Dannenberg 1989), Dannenberg developed Nyquist (Dannenberg 1997a, b) in the late 1990s, which is also an unignorable example when reconsidering the removal of such boundary in the traditional score-orchestra

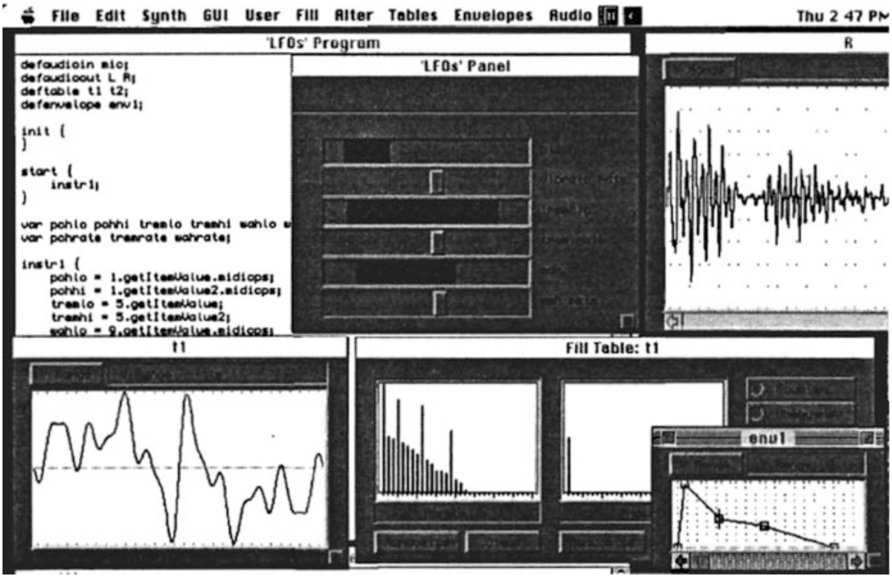


Fig. 21 A screenshot of SuperCollider (McCartney 1996) (Courtesy of James McCartney)

model in computer music language design. While it was first developed as a non-real-time sound synthesis language and later extended for real-time sound synthesis, Nyquist is considered one of "the first computer music programming languages that remove the distinction between the 'orchestra' and the 'score'" (Wang and Cook 2004) along with SuperCollider.

It would be beneficial to discuss about two different kinds of domain-specific languages (DSLs) at this point: internal DSLs and external DSLs. Fowler says an internal DSL "is a DSL represented within the syntax of a general-purpose language" and is "a stylized use of that language for a domain-specific purpose." On the other hand, an external DSL "is a domain-specific language represented in a separate language to the main programming language it's working with" and it "may use a custom syntax, or it may follow the syntax of another representation such as XML" (Fowler 2010, p. 15). According to these definitions, Nyquist is an internal DSL as it is built on the LISP language, using its powerful macro definition feature, whereas SuperCollider is an external DSL in that it has its own syntax, compiler, and runtime environment. Such a difference in the approach to design and develop a computer music language is still frequently seen today as described in the next section.

Some researchers and engineers took a different approach to modularize the real-time DSP functionality as software libraries for general-purpose programming languages. NeXT Sound and Music Kit by Smith and his colleagues (Smith et al. 1989) and Common Lisp Music for NeXT workstations by Schottstaedt (Schottstaedt 1994) are widely known early examples of such an approach in real-time sound synthesis.

Such computer music languages and systems for personal computers that can perform real-time sound processing without external hardware significantly motivated new practices in live computer music performance. Live digital signal processing of instrumental sounds was already realizable even on laptop computers by the end of 1990s, whereas it was only available for the composers with backup from an academic institution with a large budget in the late 1980s.

The popularization of real-time DSP and end-user-friendly computer music language and systems also gave a significant impact even on musicians outside the academia. Techno and noise music composers began exploring new sounds and styles with computer music software developed in academia on their personal computers (Ramsay 2013). Considerable involvement of such technology in the process of musical creations also influenced the investigation of new musical styles and how they consider musical creation as discussed by Cascone in his essay *The Aesthetic of Failure* (Cascone 2000).

New Explorations in Computer Music Languages

In the 2000s, while researchers and engineers are still making effort to develop newer versions of existing languages (e.g., Csound, Max, PureData, SuperCollider, etc.) or successor languages or similar languages to existing languages (e.g., PWGL (Laurson et al. 2009), another successor of Patchwork, or Max-like languages such as Open Sound World (Chaudhary et al. 2000) or jMax (Déchelle et al. 1999)), new computer music languages are also developed.

Wang and his colleagues developed ChucK. The ChucK (Wang 2008) programming language is particularly notable for its proposition of the *strongly timed programming* concept, with a significant focus on precise timing behavior of computer music programs. The concept utilizes logical time internal to the runtime environment and is based on the ideal synchronous hypothesis, in which “all computation and communications are assumed to take zero time (that is, all temporal scopes are executed instantaneously)” (Burns and Wellings 2001, p. 360). As this assumption is clearly unrealizable as it is, in computer music languages and systems, normally it is interpreted to imply that the advance of logical time, which causes the audio computation, is blocked until the system finishes processing all the scheduled tasks scheduled at the current logical time.

Many preceding computer music languages and systems have already taken such an idea into its design. For instance, the concept and utilization of *virtual time* in FORMULA (Anderson and Kiuvila 1991) are quite similar to synchronous programming and PureData also blocks the computation of audio output, which results in the advance of the logical time, until it finishes processing all the tasks scheduled at a certain internal logical time.

However, the strongly timed programming concept in ChucK differs from preceding languages in that it contextualizes this issue of timing precision of a computer music language as a problem in the underlying programming concept. The strongly timed programming concept provides data types directly representing

Fig. 22 A ChucK program to generate a sine wave, changing its frequency of oscillation every 100 milliseconds (Wang 2008, p. 43)

```
01:  // synthesis patch
02:  SinOsc foo => dac;
03:
04:  // infinite time loop
05:  while(true)
06:  {
07:      // randomly choose a frequency
08:      Std.rand2f(30, 1000) => foo.freq;
09:      // advance time
10:      100::ms => now;
11:  }
```

logical time and duration (e.g., *time* and *dur* in ChucK) and the explicit control of the advance of logical time is integrated as a part of the language specification. By such a design, ChucK provides simple and terse means to control timing behavior of a program with sample-rate accuracy in logical time. Figures 22 and 23 illustrate a simple ChucK program as Wang describes in (Wang 2008) and a screenshot of the miniAudicle development environment for ChucK respectively. As shown on line 10 in Fig. 22, the logical time is explicitly advanced by the user program.

Sorensen and his colleagues began developing Impromptu (Soresen and Gardner 2010), which is an internal domain-specific language for computer music built upon the Scheme programming language (Sperber et al. 2007). Impromptu utilizes external software frameworks with respect to sound synthesis, AIME (Sorensen 2005) for the earlier versions and Apple's Audio Units (Adamson et al. 2012) for the later versions and its sound synthesis framework design, its sound synthesis feature may seem less flexible in comparison with unit-generator languages, which allows users to describe various sound synthesis/sound processing algorithms.

However, Impromptu also provides a feature to define a custom Audio Unit and perform the just-in-time compilation of the definition at runtime, and one can utilize this feature to define a desired sound synthesis/sound processing algorithm, as long as one possesses enough programming skill to write one's own code for Audio Unit, while it may require a certain degree of understanding in digital signal processing and more effort in coding. Extempore (Sorensen et al. 2014), the successor language of Impromptu, is being developed also with the extended capability to describe sound synthesis algorithms just within its programming environment, even allowing users to write a function to provide output samples to the audio driver.

Impromptu is also interesting in that it rediscovers two ideas in computer music language design. One is the programming pattern that Sorensen calls as *temporal recursion*. In Impromptu, users often write a callback function that reschedules itself to the underlying scheduler so that the task written in the callee function can be repeatedly invoked by the scheduler. Figure 24 is an example of the temporal recursion programming pattern as Sorensen described in (Soresen and Gardner 2010). As shown, the function *video-player* reschedules itself in lines 08–12.

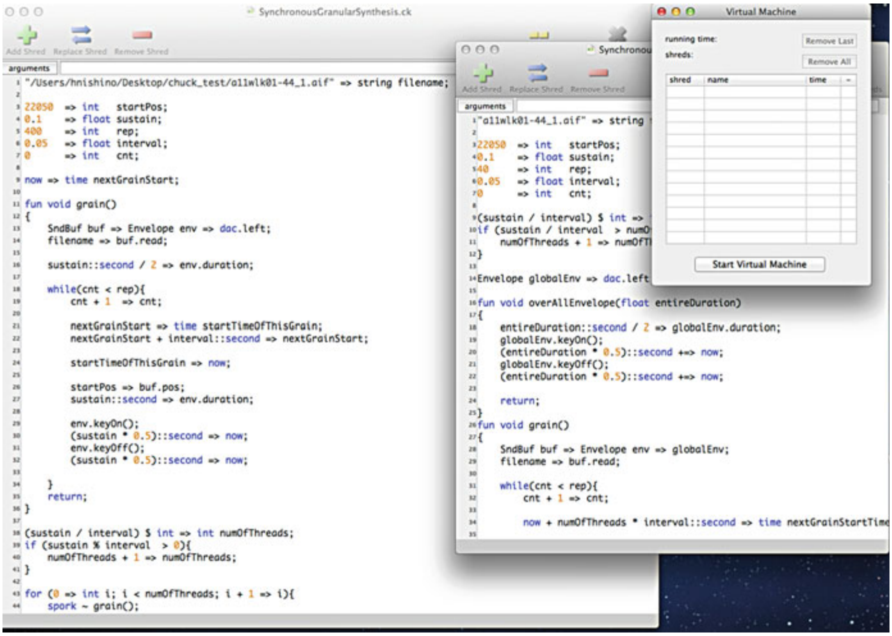


Fig. 23 A screenshot of miniAudicle (Photo by the authors)

The other is the consideration of desirable features regarding time. In (Soresen and Gardner 2010), the designers of Impromptu refer to six desirable features for high-level real-time languages. Lee et al. proposed (Lee et al. 1987) and describe how they took features with respect to time, such as execution time constraint and enforcement of timing constraints, into account in the design of Impromptu. For example, Fig. 24 utilizes the feature of execution time constraint.

Indeed, similar ideas can be seen in earlier computer music languages. While the developer may not have named the programming pattern clearly, Moxie developed by Collinge in the mid-1980s is known as an early example to propose a programming pattern quite similar to temporal recursion (Collinge 1984), and FORMULA by Anderson and his colleagues is designed clearly with the intention to integrate features such as execution time constraint, etc. (Anderson and Kuivila 1990; Anderson and Kiuvila 1991). Yet, the consideration of such features with respect to time seemed less taken into account in computer music languages in the previous decade for some reason; Impromptu may be a good design exemplar to reconsider such ideas.

LuaAV developed by Wakefiled and his colleagues is another noteworthy example among the most recent computer music languages. LuaAV is an internal domain-specific language built upon the LUA programming language (Wakefield and Smith 2007; Wakefield et al. 2010). While LuaAV provides sample-rate accuracy in timing precision as ChucK does, it also provides the capability of

```
;; video player temporal recursion looping
;; at a rate of 1/24th of one second.
;; maximum-execution time of 1/32nd of one second
  (define video-player
     (lambda (time mov position)
        (let ((frame (gfx:get-movie-frame mov position))
           (gfx:draw-image time *canvas* frame 1)
           (schedule (cons (+ (now) (/ *second* 24))
                            (/ *second* 32))
                  video-player
                    (+ time (/ *second* 24))
                    mov (+ position 1/24)))))

;; this call starts the temporal recursion
;; the execution deadline for this first call
;; is the default execution duration of the process
(video-player (now) (gfx:load-movie "/tmp/myfilm.mp4") 0.0)
```

Fig. 24 A screenshot of the programming environmetn of LuaAV (Wakefield et al. 2010) (Courtesy of Graham Wakefield)

image processing both for 2D and 3D. Figure 25 shows a screenshot of the programming environment of LuaAV.

LuaAV also provides sample-rate accurate precise timing behavior in logical time, as in many other languages, yet the strongly timed programming concept is not present in LuaAV; the advance of logical time is made by the call to library functions provided by its software framework and there exists no data type with respect to time. LuaAV also provides the feature of just-in-time compiler for sound synthesis. By compiling a unit-generator graph into native code at runtime, LuaAV can improve the computational efficiency in sound synthesis.

When considering the relationship between creativity of computer musicians and technological research, it should be emphasized that the design and development of these recent computer music languages described above are highly motivated by creative practices of our time. While the emergence of live coding performances (Collins et al. 2003), in which musicians write and modify computer music programs on-the-fly on stage, dates back in the 1980s (Ward et al. 2004), the popularization of interactive computer music languages with the stand-alone real-time sound synthesis capability in the 1990s significantly motivated the rapid growth of live coding community, and researchers and musicians formed TOPLAP (Terrestrial Organisation for the Promotion of Live Algorithm Programming) in 2004 (Toplap 2004).

This motivated researchers and engineers to take live coding as an important design criterion for new computer music languages. All of ChucK, Impromptu, and LuaAV clearly consider the supports for live coding in their language designs.

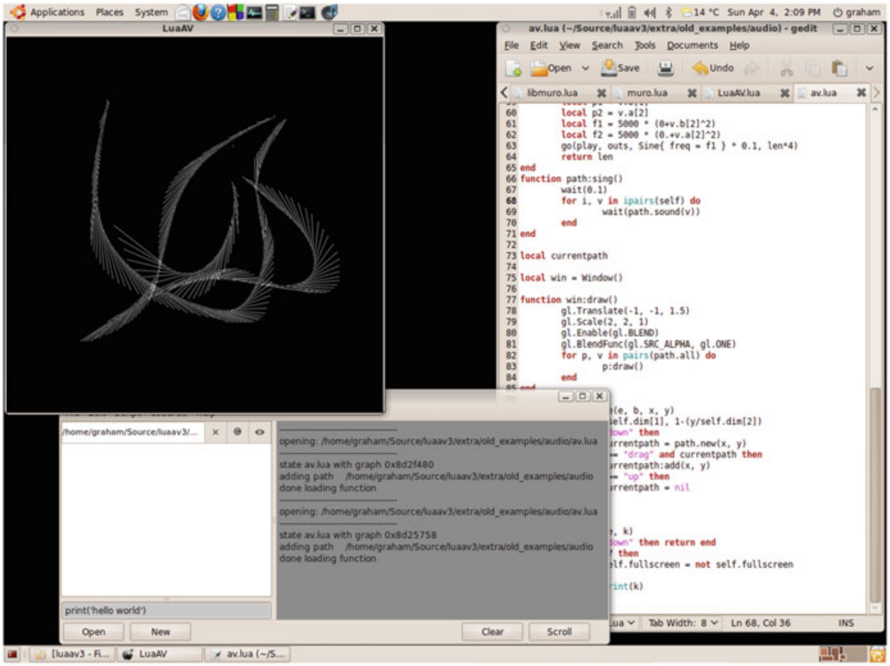


Fig. 25 A screenshot of the programming environment of LuaAV (Wakefield et al. 2010)

Many other recent computer music languages and software (including software frameworks) are also designed with the significant focus on live coding, such as ixi by Magnusson (2011), Tidal and PETROL by Mclean (McLean and Wiggins 2010a, b), sonic Pi and Overtone by Aaron (Aaron and Blackwell 2013), and Conductive by Bell (2011).

Another issue proposed by these languages (except Impromptu) is the necessity of sample-rate accurate timing behavior even for an interactive computer music language. The emergence and popularization of microsound synthesis techniques (Roads 2004), which constitutes the entire sound output by overlapping short-duration sound particles, led to the reconsideration of timing precision of the behavior of a computer music program, as imprecise timing behavior in scheduling such short sound particles (or microsounds) can lead to a theoretically inaccurate sound output, which is often audible to human ears.

Computer music composers also consider that imprecise timing behavior would damage the musical quality of computer music compositions. Lyon discusses such an issue as follows: "even when the amount of deviation from sample accuracy is not clearly noticeable at a rhythmic level, it may still have an undesirable musical effect. For example, a pulsation may feel not quite right when there are a few 10s of milliseconds of inaccuracy in the timing from beat to beat" and "smaller inaccuracies, though rhythmically acceptable, can still cause problems when sequencing sounds with sharp transients, since changes in alignment on the order of a couple of milliseconds will create different comb filtering effects as the transients slightly realign on successive attacks" (Lyon 2006).

Thus, today's creative practices in computer music can require timing precision at both levels of rhythm and sound synthesis to a considerable degree and the support for better timing precision is an important criterion in the design of computer music languages and systems. Along with the desirable features with respect to time, which Sorensen and his colleagues emphasize as important criteria in the design of Impromptu, the issue how computer music programming languages should integrate "time" into its design is still a topic of significant interest even today.

One of the issues in the software and language design that other researchers and developers recently discuss but seems not stated in above three languages is the question if the traditional unit-generator concept is an appropriate abstraction for microsound synthesis techniques. As discussed in section "The Eras of Non-real-time Computer Music Systems and Languages," generally speaking, the abstraction of the unit-generator concept significantly resembles analog sound synthesis by electronic equipment. While this abstraction is still quite beneficial for many sound synthesis techniques, some researchers began investigating if there can be more appropriate abstractions for microsound synthesis techniques.

Brandt argues that the unit-generator concept is "black-box primitive" in computer music language design and that "if a desired operation is not present, and cannot be represented as a composition of primitives, it cannot be realized within the language" and designed Chronic (Brandt 2008), which is an internal DSL for non-real-time sound synthesis built on the OCaml programming language (Leroy et al. 2012). Based on such assessment, Brandt proposed the concept of *temporal type constructor*, by which his Chronic language can describe more "time-structured types" than simple audio streams as seen in the unit-generator concept while recovering the accessibility to low-level data at the same time. Brandt discusses that the concept of temporal type constructors is beneficial as various microsound synthesis techniques can be more tersely described than unit-generator languages in Chronic (Brandt 2008).

While it is not for a computer music language or system by itself, and rather for his own real-time stand-alone software granular synthesizers, Bencina's object-oriented software framework design is another notable example to consider alternative software abstraction, which directly integrates objects for microsounds (Bencina 2006).

Wang and his colleagues also questioned if the unit-generator concept may not be very appropriate for some sound synthesis/analysis techniques in the design of the *unit analyzers* for the ChucK programming language. Chuck's unit analyzers "pass generic data that may be in the form of spectral frames, feature vectors, metadata, or any other (ntermediate) products of analysis" to perform such tasks as domain transformations (e.g., FFT and DWT), feature extractors (e.g., RMS, flux, zero-crossing), and operations (e.g., correlation) (Wang et al. 2007). Similary as in Brandt's discussion on the temporal type constructor concept, they also discuss about the accessibility to low-level data that "the high-level abstraction in the system should expose essential

```
//create a SampleBuffer and fill it with 256 samp sinewave * 4 cycles
var sbuf = new SampleBuffer(1024);
for (var i = 0; i < sbuf.size; i += 1){
  sbuf[i] = Sin(3.14159265359 * 2 *( i * 4.0 / sbuf.size));
}

//create a grain (or a short sound particle).  first
var tmp = sbuf->toSamples();
//apply an envelope to sinewave and resample it to 440 samples.

var win = GenWindow(1024::samp, \hanning);
var grain = tmp->applyEnv(win)->resample(440)->amplify(0.25);

//giving an execution-time constraint of 5 sec.
within(5::second){
  //perform synchronous granular synthesis. schedule short sound
  //particles of 440 samples with a constant interval of 110 samples.
  while(true){
    WriteDAC(grain);
    now += grain.dur / 4;
  }
}
```

Fig. 26 A synchronous granular synthesis example

low-level parameters while doing away with syntactic overhead, thereby providing a highly flexible and open framework that can be easily used for a variety of tasks" when explaining the benefits of the unit-analyzer concept in (Wang et al. 2007).

LC, developed by Nishino in the 2010s, also proposed an alternative design for sound synthesis framework, which directly integrates objects and library functions for microsound synthesis. While LC still can perform sound synthesis by unit-generators, its microsound synthesis framework is highly independent of the unit-generator concept and makes it possible to tersely describe both microsound synthesis and frequency-domain signal processing, without depending on the unit-generator concept. Figure 26 illustrates a simple synchronous granular synthesis example in LC. As shown, there is no unit-generator or data stream involved in this example. Nishino also discusses the benefit of seamless collaboration between this microsound synthesis framework and the unit-generator-based sound synthesis framework in (Nishino 2014).

In the same publication, Nishino also describes the benefits of the adoption of prototype-based programming at both levels of sound synthesis and control algorithms in LC, together with the integration of desirable features with respect to time. The latter includes the mostly strongly timed programming concept, which extends the strongly timed programming concept with the explicit switch between synchronous behavior and asynchronous behavior. Nishino discusses this feature of mostly strongly timed programming as one solution to the problem of temporary suspension in real-time DSP, which can be caused by a time-consuming task in computer music systems and languages designed with the synchronous approach (e.g., ChucK, LuaAV, PureData, etc.), while still maintaining the benefit of precise timing behavior of the synchronous approach.

As above, while computer music languages and systems designed in and after the 2000s took over the achievement by the preceding works in the 1990s, they often focus on the issues that were not taken into much consideration in the 1990s as fundamental design criteria (e.g., live coding, sample-rate accuracy in timing precision, and microsound synthesis), and these issues have a lot to do with creative practices of the decade; thus, even today, the development of new computer music languages was significantly motivated not just by technological advance but also by creative practices, as is often seen in the computer music history.

Other Noteworthy Trends

Along with the development of computer music languages as above, the researchers continuously developed and improved the software libraries and systems through the 2000s. The list of examples includes JSyn (Burk 1998) for Java (written in C++ and Java), Synthesis Toolkit (STK) (Cook and Scavone 1999; Scavone and Cook 2005) for C++, CLAM for C++ (Amatriain 2007; Amatriain et al. 2008), and Marsyas for C++ (Tzanetakis 2007). Not just such new libraries and frameworks, existing languages such as RTcmix and Csound were also integrated into other programming languages (e.g., MAX, PureData, Python, etc.) as library objects or software frameworks (Thompson 2014; Lazzarini et al. 2008); all these frameworks/libraries are designed for stand-alone real-time sound synthesis and processing.

One of the notable trends in these software frameworks and libraries developed around the 2000s in comparison with the 1990s is that they often provide features for physical modeling sound synthesis and music information retrieval, supported by both the further advance of computational speed and the academic achievement in these fields. For instance, STK includes many objects for physical modeling sound synthesis techniques, and CLAM and Marsya include a series of objects for music information retrieval.

It should be also noted that domain-specific languages or programming environments built upon the frameworks/libraries are often provided together; STK and Marsyas provide the SKINI language (Cook and Scavone 1999) and the Marsyas scripting language (Tzanetakis 2007) respectively to facilitate the utilization of the frameworks, and CLAM provides a graphic programming environment called the CLAM Network Editor, in which a user can easily create a patch for synthesis and analysis without expertise in programming (Tzanetakis 2007).

The popularization of real-time sound synthesis in the 1990s also raised remarks among researchers and musicians about "how to articulate computer generated music in a concert setting" (Tanaka 2000) and led to the emergence of the research community that shares the interest in *new interface for music expression* (NIME). While the research on interfaces for computer music is one of the traditional research topics in computer music, the rapid growth of the NIME community was seen in the 2000s. First started as a workshop at the ACM Conference on Human Factors in Computing System (SIGCHI), the NIME conference is held annually every year since 2001 (NIME 2001).

The NIME research and practices in the 2000s involved the creative usages of both gestural input devices and tangible user interfaces built upon the cutting-edge human-computer interface technology of the same decade. For instance, Sensorband (Atau Tanaka, and Edwin van der Heide and Zbigniew Karkowski) are known particularly for their creative uses of various sensors (e.g., biosignal sensors, ultrasound sensors, and invisible infrared beams) in their live computer music performances (Bongers 1998). For another example, The reacTable by Kaltenbrunner and his colleagues is a computer music system that utilizes a tabletop tangible user interface. Performers can dynamically modify the network of sound synthesis/sound processing units in the systems, by directly moving and placing/removing tangible objects on the tabletop interface, each of which represents a sound synthesis/sound processing unit in the system (Jordà et al. 2007). Figures 27 and 28 show the pictures of Sensorband's live performance and the reacTable system respectively.

The growth of the NIME community still continues even in the 2010s, being helped and influenced by the achievement by other movements that reduce the difficulty in making hardware prototypes (e.g., Fab Lab (Walter-Herrmann and Büching 2014) and Maker (Hatch 2014)).

Computer music on mobile platforms is another noteworthy topic raised recently. Since mobile devices such as smart phones and tablets became fast enough to perform real-time sound synthesis, many extant computer music languages began being ported to mobile platforms. Allison and his colleagues reported the Android version of SuperCollider by Shaw in their project (Allison and Dell 2012). PureData were ported to PureData for iOS and Android by Brinkmann and his colleagues (Brinkmann et al. 2011). Wang reported that ChiP, the iPhone version of ChucK, was developed and used for their commercial applications (Wang et al. 2009). RTcmix was ported to iOS by Mailman (2012). The use of PureData and SuperCollider on Raspberry Pi, a low-cost credit-card-size computer that runs Linux, is also reported (Berdahl and Llimona 2013; Bown et al. 2013). As smart phones are normally made with many useful integrated sensors (e.g., compass, GPS, touch screen, accelerometer, and gyro sensor) and wireless networking interface, many NIME projects adopt them for NIME projects.

The rapid growth of the Internet also is giving significant impact on computer music research and practices today. Some researchers note that the origin of the networked music performance can be found in much earlier decades. For instance,

Fig. 27 The live performance by Sensorband: Edwin van der Heide, Zigbniew Karkowski, and Atau Tanaka from *left* to *right* (Photo by Peter Kers)

Fig. 28 A picture of the reacTable (Photo by Thomas Bonte. This image is from Wikimedia and licensed under the Creative Commons Attribution 2.0 Generic license)

Renaud et al. describe that John Cage's *Imaginary Landscape No. 4 for twelve radios* (1951) "is considered to be one of the earliest networked music performance experiment," since the piece utilizes radios as musical instruments, and "although the levels of interactivity were limited to the dialling of radio-stations, gain and tone-colour, the desire to investigate the possibilities of cross-influence in networked instruments is evident in the piece" (Renaud et al. 2007).

The Hub (Gresham-Lancaster 1998), a group that consists of American composers, which is often referred for their early live coding practices (Ward et al. 2004), also provides possibly the earliest example of networked computer music performance. It is reported that the members of the Hub performed in networked distant locations in New York linked by modem (Gann 1987). Pauline Oliveros aslo provides another noteworthy early example of networked music. In 1991, she organized networked improvisation via video telephone transmission and she still frequently experiments in networked improvisation since then, changing the connection from telephone to high-speed Internet (Oliveros 2009).

The experiment in network music performance over the Internet began around the early 1990s. The early noteworthy example is the experiment by Schooler and Touch (1993) in 1993 at the University of Southern California Information Sciences Institute. It is reported that *Haydn Piano Trio No.1 in G, Finale* performed by musicians distributed in L.A. and Boston, to be heard by the audience over the DARTnet network (Sawchuk et al. 2003) (DARTnet (DARPA Testbed Network), was "an experimental network used to learn more about thing like routing protocols," which was founded around the early 1990s, and connected eight key research institutions such as MIT and Xerox PARC (Malamud 1992, p. 313)).

Along with such network music performance experiment with audio streaming, the researchers also investigated on the protocol for musical control events that is suitable for networked music performance. Open Sound Control is possibly the most widely used protocol for networked music performance today, which was developed in the mid-1990s (Wright 1997). It is reported that the Hub utilized Open Sound Control for the wide area network in 1997 (Wright 2005). The Open Sound Control is frequently involved in computer music performances as a communication protocol between computer music software and is still updated since its emergence.

While networked music performance may utilize more general-purpose peer-to-peer Internet broadcasting software such as PeerCast and ShoutCast (Mills 2010), many computer music software were designed and developed for networked music performance until today. The list of well-known software for networked music in the 2000s and 2010s includes FMOL Virtual Music Instrument by Jordà' et al. (Jordà and Barbosa 2001), Public Sound Objects by the Music Technology Group at Pompeu Fabra University (Barbosa and Kaltenbrunner 2002), TransJam by Burk (2000), Quintet.net by Hajdu (2005), SoundJack by Carot et al. (2006), JackTrip by Cáceres and Chafe (2010), and the DIAMOUSES framework by Alexandraki and Akoumianakis (2010).

Thus, networked music performance is still a major research topic under investigation by many research groups (for instance, the research groups including SoundWIRE Research Group at CCRMA, Stanford University, the Sonic Arts Research Centre (SARC) at Queens University Belfast, and Distributed Immersive Performance (DIP) Experiments at the University of Southern California are well-known research groups with the significant focus on the networked music).

It should also be noted that the open source culture (Raymond 2001) began influencing the computer music community. NetJam is a good early example of this kind. It was a computer network that "provides a means for people all over the world to collaborate on musical compositions and otherworks of art" and "participants send Musical Instrument Digital Interface (MIDI) and other data files to each other via electronic mail (e-mail), edit them, and resend them" (Latta 1991). The Free Sound Project is another good example of the kind (Font et al. 2013). Their website provides an enormous number of audio samples under the *Creative Commons* license (http://creativecommons.org/licenses/).

The performance efficiency in real-time sound synthesis and sound processing is still an issue of significant interest even in the 2000s and 2010s. Some computer music languages took the approach to compile a program written in the language into native code so that it can be processed faster.

Faust (Orlarey et al. 2009), developed by Orlarey and his colleagues, is a computer music language of this kind. A program written in Faust is first translated into an equivalent C++ program and then compiled into a native code in various formats such as a stand-alone executable program, libraries, or plug-in for other computer music languages/systems. Faust is also interesting in its syntax that makes it possible tersely to describe low-level block diagrams for digital signal processing. Faust is also interesting in design as a functional programming language and concise syntax with a significant focus on digital signal processing.

Some computer music languages such as Impromptu and LuaAV perform just-in-time compilation for sound synthesis as described in the previous section. In these languages, the user code for sound synthesis is compiled into a native code at runtime. Such an interest is shared by other designers of more recent computer music languages and systems. For instance, the Kronos computer music language by Norilo (2013) also applies just-in-time compilation to sound synthesis algorithms and the recent version of Max integrates the extension called Gen, which allows users to describe low-level sound synthesis algorithms to be compiled into native code (Cycling74 2011).

Another approach to improve the performance efficiency is to utilize *general-purpose computing on graphics processing units* (GPGPUs) for audio computation as seen in (Sosnick and Hsu 2010; Dabrowski et al. 2012; Henry 2011), etc. Graphic processing units (GPUs) are designed to efficiently perform processing floating-point values parallelly, and this characteristic of GPUs is quite favorable to process audio data. It may be possible to see such utilization of GPGPUs as the successor of the research on computer music systems with the variable-function DSP hardware.

Last, but not the least, the research on algorithmic composition and computer-assisted composition is still active. One of the most notable trends emerged around the 2000s would be the application of genetic algorithm to computer music. Such research on *evolutionary computer music* ranges very widely from the sound synthesis and sound design to composition and improvisation (Miranda and Biles 2007).

Conclusion

In this chapter, the history of computer music was briefly reviewed, with a significant focus on computer music languages and related systems/software.

As described, while the evolution of computer music languages and systems has been largely supported by the advance of computer technology and related research, the issues found in creative musical practices also gave significant motivation for the design and development of new computer music languages and systems, since the very early days of computer music.

The research on algorithmic composition and computer-assisted composition is still a topic of significant interest today, and many researchers and composers developed the software for their creative practices even since the very early stage in the history of computer music, and used by many historically important contemporary music composers for their musical creations.

The unit-generator concept was invented as it was necessary to provide “a simple, powerful language” to describe a complex sequence of sound (Mathews et al. 1969), so that researchers and composers can explore the domain of digital sound synthesis, without much difficulty even when they do not possess expert programming skills. The demand for faster sound rendering among computer musicians led to the software design that separates control rate and audio rate as seen in MUSIC-11, which is still often seen among even those most recent computer music languages.

The desire for live computer music presentations in the early era led to the development of a hybrid computer music system composed of a minicomputer and external analog synthesizer hardware. In the later decade, such a hybrid system was replaced by a personal computer with digital MIDI synthesizer hardware, and researchers and composers began working on interactive music compositions on such MIDI-based computer music systems.

In the era when programmable variable-function DSP hardware was developed, researchers began researching on hybrid computer music systems with the capability of real-time sound processing. While the emergence of such computer music systems significantly motivated composers of the time to explore this new territory, the necessity for a development environment that is suitable for rapid prototyping is raised to support the collaboration between engineers and composers.

Also supported by the demand for a more end-user-friendly development environment among computer music composers who wanted to develop their own interactive music systems but didn’t possess expertise in programming, graphic computer music programming environments such as MAX and Kyma/Platpus were developed, considering the better support for rapid prototyping and end-user programming as a significant criterion in computer music programming language design.

The popularization of real-time sound synthesis led to the rapid growth of the research community of NIME, as the development of new musical interfaces to articulate digital sound became an issue of significant interest. Many NIME researchers are now attracted to investigate the utilization of mobile devices as NIME, since mobile devices of our time are already fast enough for real-time sound synthesis while they integrate many useful sensors within. Such an interest led to the development of computer music languages and systems for mobile platforms.

The emergence of more powerful and dynamic computer music languages with the real-time sound synthesis capability in the 1990s also motivated live coding practices. The design of computer music languages in the later decade took the support for live coding as one of their important design criteria. New computer music practices that require more accurate timing precision in sound synthesis and rhythm and the popularization of microsound synthesis techniques cast a question on computer music language design, if there can be an alternative abstraction for digital sound synthesis to the unit-generator concept.

Thus, the view to computer technology that researchers and musicians in the early days of computing had, which foresees the involvement of computers in the process of musical creation further enhancing artistic creativities, has been endorsed by creative practices in the later decades, and is still valid even in our time. Throughout its history, the research on computer music languages and related systems/software has been significantly motivated also by the desire to explore new domains of computer music practices expanded by the advance of computer technology. The synergy between technology and creativity can be observed in every decade in the history of computer music.

References

S. Aaron, A.F. Blackwell, From sonic pi to overtone: creative musical experiences with domain-specific and functional languages, in *Proceedings of the First ACM SIGPLAN Workshop on Functional Art, Music, Modeling & Design* (ACM, Boston, 2013), pp. 35–46

C. Adamson, M. Lee, K. Avila, *Learning Core Audio: A Hands-On Guide to Audio Programming for Mac and iOS* (Addison-Wesley Professional, Upper Saddle River, 2012)

C. Alexandraki, D. Akoumianakis, Exploring new perspectives in network music performance: the diamouses framework. Comput. Music J. **34**(2), 66–83 (2010)

H. Alles, P. di Giugno, A one-card 64 channel digital synthesizer. Comput. Music J. **1**(4), 7–9 (1977)

J. Allison, C. Dell, Aural: a mobile interactive system for geo-locative audio synthesis, in *Proceedings of the 12th Conference on New Interfaces for Musical Expression*, Michigan (2012)

X. Amatriain, Clam: a framework for audio and music application development. IEEE Softw. **24**(1), 82–85 (2007)

X. Amatriain, P. Arumi, D. Garcia, A framework for efficient and rapid development of cross-platform audio applications. Multimedia Syst **14**(1), 15–32 (2008)

C. Ames, Automated composition in retrospect: 1956–1986. Leonardo **20**, 169–185 (1987)

D.P. Anderson, R. Kiuvila, Formula: a programming language for expressive computer music. Computer **24**(7), 12–21 (1991)

D.P. Anderson, R. Kuivila, A system for computer music performance. ACM Trans. Comput. Syst. **8**(1), 56–82 (1990)

Anonymous, Back matter. Comput. Music J. **2**(3), 24–29 (1978)

F. Armani, L. Bizzarri, E. Favreau, A. Paladin, Mars: DSP environment and applications, in *Proceedings of the International Computer Music Conference* (International Computer Music Association, San Jose, 1992), pp. 344–344

G. Assayag, C. Rueda, M. Laurson, C. Agon, O. Delerue, Computer-assisted composition at IRCAM: from patchwork to openmusic. Comput. Music J. **23**(3), 59–72 (1999)

J. Backus, The history of fortran i, ii, and iii, in *History of Programming Languages I* (ACM, New York, 1978), pp. 25–74

Á. Barbosa, M. Kaltenbrunner, Public sound objects: a shared musical space on the web, in *Web Delivering of Music, 2002. WEDELMUSIC 2002. Proceedings. Second International Conference on* (IEEE, Darmstadt, 2002), pp. 9–16

C. Barlow, Autobusk: an algorithmic real-time pitch and rhythm improvisation programme, in *Proceedings of the International Computer Music Conference* (International Computer Music Association, 1990), pp. 166–168

C. Barlow, *AUTOBUSK: A Real-Time Pitch & Rhythm Generator* (University of Mainz, Mainz, 2000)

C. Barlow, Algorithmic composition, illustrated by my own work: a review of the period 1971–2008, in *Proceedings of Korean Electro-Acoustic Music Society's 2011 Annual Conference* (Korean Electro-Acoustic Music Society, Seoul, 2011)

J. Beauchamp, Music 4c introduction, in *Computer Music Project, School of Music, University of Illinois at Urbana-Champaign*, Illinois, vol. 1(99) (1993), p. 3

J.W. Beauchamp, Music 4c, a multi-voiced synthesis program with instruments defined in c. J. Acoust. Soc. Am. **113**, 2215 (2003)

R. Bell, An interface for real-time music using interpreted Haskell, in *Proceedings of LAC 2011*, Maynooth (2011)

R. Bencina, *Audio Anecdotes III, Chapter Implementing Real-Time Granular Synthesis* (A.K Peters, Natick, 2006), pp. 55–83

E. Berdahl, Q. Llimona, Tangible embedded Linux, in *Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction*, (ACM, Barcelona, 2013), pp. 407–410

B. Bongers, An interview with sensorband. Comput. Music J. **22**(1), 13–24 (1998)

R. Boulanger, *The Csound Book: Perspectives in Software Synthesis, Sound Design, Signal Processing, and Programming* (The MIT Press, Cambridge, MA, 2000)

P. Boulez, A. Gerzso, Computers in music. Sci. Am. **258**(4), 44–51 (1988)

E. Bowles, Musickes handmaiden: or technology in the service of the arts, in *The Computer and Music* (Cornell University Press, Ithaca, 1970), pp. 3–20

O. Bown, M. Young, S. Johnson, A java-based remote live coding system for controlling multiple raspberry pi units, in *Proceedings of the 2013 International Computer Music Conference*, Perth (2013), pp. 31–38

L. Boynton, *Midi-Lisp: A Lisp-Based Music Programming Environment for the Macintosh* (MPublishing, University of Michigan Library, Ann Arbor, 1986)

E. Brandt, *Temporal Type Constructors for Computer Music Programming*. Ph.D. thesis, Carnegie Melon University, 2008

J. Bresson, Sound processing in openmusic, in *Proceedings of the International Conference on Digital Audio Effects*, Quebec (2006)

P. Brinkmann, P. Kirn, R. Lawler, C. McCormick, M. Roth, H.-C. Steiner, Embedding pure data with libpd, in *Proceedings of the Pure Data Convention*, Weimar (2011)

A.G. Bromley, Difference and analytical engines, in *Computing Before Computers* (Iowa State University Press, Ames, 1990), pp. 59–98

P. Burk, JSyn–a real-time synthesis api for java, in *Proceedings of the 1998 International Computer Music Conference* (International Computer Music Association, San Francisco, 1998), pp. 252–255

Burk, P. (2000). Jammin'on the web-a new client/server architecture for multi-user musical performance. In *ICMC 2000*. Citeseer.

A. Burns, A.J. Wellings, *Real-Time Systems and Programing Languages: Ada 95, Real-Time Java and Real-Time Posix* (Addison Wesley, Boston, 2001)

W.A. Buxton, A composer's introduction to computer music. J. New Music Res. **6**(2), 57–71 (1977)

J.-P. Cáceres, C. Chafe, Jacktrip: under the hood of an engine for network audio. J. New Music Res. **39**(3), 183–187 (2010)

A. Carôt, A. Renaud, B. Verbrugghe, Network music performance (nmp) with soundjack, in *6th International Conference on New Interfaces for Musical Expression (NIME 06)* (Paris, 2006)

K. Cascone, The aesthetics of failure: 'post-digital' tendencies in contemporary computer music. Comput. Music J. **24**(4), 12–18 (2000)

S. Cavaliere, G. Di Giugno, E. Guarino, Mars: the x20 device and sm100 board, in *Proceedings of the International Computer Music Conference* (International Computer Music Association, San Jose, 1992), pp. 348–348

J. Chadabe, R. Meyers, An introduction to the play program. Comput. Music J. **2**(1), 12–18 (1978)

A. Chaudhary, A. Freed, M. Wright, An open architecture for real-time music software, in *Proceedings of the International Computer Music Conference*, Berlin (2000)

E. Childs, Achorripsis: a sonification of probability distributions, in *Proceedings of the 2002 International Conference on Auditory Display*, Kyoto (2002)

J.M. Chowning, The synthesis of complex audio spectra by means of frequency modulation. Comput. Music J. **1**, 46–54 (1977)

J. Chowning, Turenas: the realization of a dream, in *Proceedings of the 17es Journées dInformatique Musicale* (Saint-Etienne, 2011)

J. Clough, Tempo: a composer's programming language. Perspect. New Music **9**(1), 113–125 (1970)

D. Collinge, Moxie: a language for computer music performance, in *Proceedings of the International Computer Music Conference 1984*, Paris (1984)

N. Collins, A. McLean, J. Rohrhuber, A. Ward, Live coding in laptop performance. Organised Sound **8**(3), 321–330 (2003)

P.R. Cook, G. Scavone, The synthesis toolkit (stk), in *Proceedings of the International Computer Music Conference*, Beijing (1999), pp. 164–166

D. Cope, An expert system for computer-assisted composition. Comput. Music J. **11**, 30–46 (1987)

D. Cope, Recombinant music: using the computer to explore musical style. Computer **24**(7), 22–28 (1991)

Cycling74, *Max 6 Help and Documentation*. Cycling74 (2011)

A. Dabrowski, P. Pawlowski, M. Stankiewicz, F. Misiorek, Fast and accurate digital signal processing realized with GPGPU technology. Przeglad Elektrotechniczny **88**(6), 47–507 (2012)

R.B. Dannenberg, Arctic: a functional language for real-time control, in *Proceedings of the 1984 ACM Symposium on LISP and Functional Programming* (ACM, Texas, 1984), pp. 96–103

R.B. Dannenberg, The canon score language. Comput. Music J. **13**(1), 47–56 (1989)

R.B. Dannenberg, The CMU MIDI toolkit, version 3 (1993)

R.B. Dannenberg, The implementation of nyquist, a sound synthesis language. Comput. Music J. **21**(3), 71–82 (1997a)

R.B. Dannenberg, Machine tongues xix: nyquist, a language for composition and sound synthesis. Comput. Music J. **21**(3), 50–60 (1997b)

R.L. De Mantaras, J.L. Arcos, Ai and music: from composition to expressive performance. AI Mag. **23**(3), 43 (2002)

R.T. Dean (ed.), *The Oxford Handbook of Computer Music* (Oxford University Press, New York, 2009)

F. Déchelle, R. Borghesi, M.D. Cecco, E. Maggi, B. Rovan, N. Schnell, jMax: an environment for real-time musical applications. Comput. Music J. **23**(3), 50–58 (1999)

G. Di Guigno, A. Gerzso, *La station de travail musical 4x*. Technical report (IRCAM Technical Report, Paris, 1986)

C. Dodge, T.A. Jerse, *Computer Music: Synthesis, Composition, and Performance* (Schirmer Books, New York, 1997)

P. Doornbusch, Computer sound synthesis in 1951: the music of CSIRAC. Comput. Music J. **28**(1), 10–25 (2004)

J. Eaton, The humanization of electronic music. Music Educ. J. **55**(3), 101–102 (1968)
T. Erbe, Soundhack: a brief overview. Comput. Music J. **21**, 35–38 (1997)
E. Favreau, M. Fingerhut, O. Koechlin, P. Potacsek, M. Puckette, R. Rowe, Software developments for the 4x real-time system, in *Proceedings of the International Computer Music Conference*, Den Haag (1986), pp. 369–373
G. Fedorkow, W. Buxton, K. Smith, A computer-controlled sound distribution system for the performance of electroacoustic music. Comput. Music J. **2**(3), 33–42 (1978)
F. Font, G. Roma, X. Serra, Freesound technical demo, in *Proceedings of the 21st ACM International Conference on Multimedia* (ACM, Barcelona, 2013), pp. 411–412
M. Fowler, *Domain-Specific Languages* (Addison-Wesley, Upper Saddle River, 2010)
D. Friend, A time-shared hybrid sound synthesizer. J. Audio Eng. Soc. **19**(11), 928–935 (1971)
J. Fuegi, J. Francis, Lovelace & Babbage and the creation of the 1843 ‘notes’. IEEE Ann. Hist. Comput. **25**(4), 16–26 (2003)
J. Gabura, G. Ciamaga, Computer control of sound apparatus for electronic music, in *Audio Engineering Society Convention 33* (1967)
K. Gann, The hub musica telephonica (23 June 1987)
B. Garton, D. Topper, RTcmix–using Cmix in real time, in *Proceedings of the International Computer Music Conference. International Computer Music Association*, Thessaloniki (1997)
A. Goldberg, D. Robson, *Smalltalk-80: The Language and Its Implementation* (Addison-Wesley Longman Publishing Co, Reading, 1983)
S. Gresham-Lancaster, The aesthetics and history of the hub: the effects of changing technology on network computer music. Leonardo Music J. **8**, 39–44 (1998)
P. Grogono, MUSYS: software for an electronic music studio. Softw. Pract. Exp. **3**(4), 369–383 (1973)
G. Hajdu, Quintet. net: an environment for composing and performing music on the internet. Leonardo **38**(1), 23–30 (2005)
S.J. Hanson, R.E. Kraut, J.M. Farber, Interface design and multivariate analysis of UNIX command use. ACM Trans. Inf. Syst. **2**(1), 42–57 (1984)
J.L. Harrington, *Technology and Society* (Jones & Bartlett Learning, Sudbury, 2009)
M. Hatch, *The Maker Movement Manifesto* (McGraw-Hill Education, New York, 2014)
C. Henry, GPU audio signals processing in pure data, and PdCUDA an implementation with the CUDA runtime API, in *Pure Data Convention*, Weimar (2011)
L.A. Hiller, R.A. Baker, Computer cantata: a study in compositional method. Perspect. New Music **3**, 62–90 (1964)
R. Hirs, B. Gilmore, A.H. voor de Kunsten, *Contemporary Compositional Techniques and OpenMusic* (Delatour, 2009)
T. Holmes, *Electronic and Experimental Music: Technology, Music, and Culture* (Routledge, New York, 2012)
S. Holtzman, Using generative grammars for music composition. Comput. Music J. **5**(1), 51–64 (1981)
M. Jenkins, Analog Synthesizers: Understanding, Performing, Buying–From the Legacy of Moog to Software Synthesis (CRC Press, Boca Raton, 2009)
S. Jordà, Á. Barbosa, Computer supported cooperative music: overview of research work and projects at the audiovisual instituteupf’, in *Workshop on Current Research Directions in Computer Music*, Barcelona (2001), pp. 92–96
S. Jordà, G. Geiger, M. Alonso, M. Kaltenbrunner, The reactable: exploring the synergy between live music performance and tabletop tangible interfaces, in *Proceedings of the 1st International Conference on Tangible and Embedded Interaction* (ACM, Baton Rouge, 2007), pp. 139–146
P. Kail, Forth programming language. Softw. World **16**(3), 2–5 (1985)
E.E. Kim, B.A. Toole, Ada and the first computer. Sci. Am. Am. Ed. **280**, 76–81 (1999)
L. Koblyakov, R. Rowe, Score/music orientation: an interview with Robert Rowe. Comput. Music J. **16**, 22–32 (1992)

G.M. Koenig, Aesthetic integration of computer-composed scores. Comput. Music J. **7**, 27–32 (1983)

G.M. Koenig, Working with project 1 my experiences with computer composition. J. New Music Res. **20**(3–4), 175–180 (1991)

P. Lansky, *Cmix. Program Documentation* (Princeton University, Princeton, 1987), http://silvertone.princeton.edu/winham/man

P. Lansky, The architecture and musical logic and Cmix, in *Proceedings of the 2006 International Computer Music Conference*, Glasgow (1990)

C. Latta, Notes from the netjam project. Leonardo Music J. **1**, 103–105 (1991)

M. Laurson, *PATCHWORK: A Visual Programming Language and Some Musical Applications*. Ph.D. thesis, Sibelius Academy Helsinki, 1996

M. Laurson, Recent developments in patchwork: PWConstraints-a rule based approach to complex musical problems, in *Symposium on Systems Research in the Arts*, Baden-Baden, vol. 1 (1999)

M. Laurson, M. Kuuskankare, Two computer-assisted composition case studies. Contemp. Music Rev. **28**(2), 193–203 (2009)

M. Laurson, M. Kuuskankare, V. Norilo, An overview of PWGL, a visual programming environment for music. Comput. Music J. **33**(1), 19–31 (2009)

V. Lazzarini, A. Kirke, E. Miranda, M. Kuuskankare, M. Laurson, F. Thalmann, G. Mazzola, A toolkit for music and audio activities on the XO computer, in *Proceedings of the 2006 International Computer Music Conference*, Abingdon (2008)

I. Lee, S.B. Davidson, V. Fay-Wolfe, *Motivating Time as a First Class Entity*. Technical report (University of Pennsylvania, 1987)

X. Leroy, D. Doligez, A. Frisch, J. Garrigue, D. Rémy, J. Vouillon, The OCaml System release 4.00, in *Documentation and User's Manual* (Projet Gallium, INRIA, Paris, 2012)

D.G. Loy, Notes on the implementation of MUSBOX: a compiler for the systems concepts digital synthesizer. Comput. Music J. **5**(1), 34–50 (1981)

G. Loy, The CARL system: premises, history, and fate. Comput. Music J. **26**(4), 52–60 (2002)

E.A. Lyon, A sample accurate triggering system for pd and max/msp, in *Proceedings of the 2006 International Computer Music Conference*, Karlsruhe (2006)

T. Magnusson, Ixi lang: a supercollider parasite for live coding, in *Proceedings of the International Computer Music Conference* (University of Huddersfield, Huddersfield, 2011)

J.B. Mailman, The fluxations stochastic interactive algorithmic music engine (SIAME) and iphone app, in *Proceedings of the 9th Sound and Music Computing Conference* (Copenhagen, 2012)

C. Malamud, *Exploring the Internet: A Technical Travelogue* (Prentice Hall, Englewood Cliffs, 1992)

M.V. Mathews, An acoustic compiler for music and psychological stimuli. Bell Syst. Tech. J. **40**, 677–694 (1961)

M.V. Mathews, J.E. Miller, *Music IV Programmer's Manual* (Bell Telephone Labs, Murray Hill, 1979)

M.V. Mathews, F.R. Moore, Groove – a program to compose, store, and edit functions of time. Commun. ACM **13**(12), 715–721 (1970)

M.V. Mathews, J.R. Pierce, Harmony and nonharmonic partials. J. Acoust. Soc. Am. **68**(5), 1252–1257 (1980)

M.V. Mathews, J.E. Miller, F.R. Moore, J.R. Pierce, J.-C. Risset, *The Technology of Computer Music* (The MIT Press, Cambridge, 1969)

A. May, Philippe manoury: Jupiter. Comput. Music J. **23**(3), 118–120 (1999)

J. McCartney, Supercollider: a new real time synthesis language, in *Proceedings of the 1996 International Computer Music Conference*, Hong Kong (1996)

A. McLean, G. Wiggins, *Petrol: Reactive Pattern Language for Improvised Music* (MPublishing, University of Michigan Library, Ann Arbor, 2010a)

A. McLean, G. Wiggins, Tidal–pattern language for the live coding of music, in *Proceedings of the 7th Sound and Music Computing Conference*, New York (2010b)

MIDI Manufacturers Association, *The Complete MIDI 1.0 Detailed Specification: Incorporating All Recommended Practices* (MIDI Manufacturers Association, Los Angeles, 1996)

R. Mills, Dislocated sound: a survey of improvisation in networked audio platforms, in *10th International Conference on New Interfaces for Musical Expression (NIME 10)* (University of Technology, Sydney, 2010)

E.R. Miranda, A. Biles, *Evolutionary Computer Music* (Springer, London, 2007)

F.R. Moore, The computer audio research laboratory at UCSD. Comput. Music J. **6**(1), 18–29 (1982)

F.R. Moore, Dreams of computer music: then and now. Comput. Music J. **20**, 25–41 (1996)

J.A. Moorer, A. Chauveau, C. Abbott, P. Eastty, J. Lawson, The 4c machine. Comput. Music J. **3**(3), 16–24 (1979)

J. Myhill, *Some Simplifications and Improvements in the Stochastic Music Program* (MPublishing, University of Michigan Library, Ann Arbor, 1978)

NIME, (NIME, 2001), http://www.nime.org. Accessed 14 Nov 2014

H. Nishino, *LC: A Mostly-Strongly-Timed Prototype-Based Computer Music Programming Language that Integrates Objects and Manipulations for Microsound Synthesis*. Ph.D. thesis, National University of Singapore, 2014

V. Norilo, Recent developments in the Kronos programming language, in *Proceedings of the ICMC2013 International Computer Music Conference*, Perth (2013)

P. Oliveros, From telephone to high speed internet: a brief history of my telemusical performances. in, *Sounding the Margins: Collected Writings 1992–2009* (Deep Listening Publications, Kingston, 2009), pp. 191–194

Y. Orlarey, D. Fober, S. Letz, Faust: an efficient functional approach to DSP programming, in *New Computational Paradigms for Computer Music* (Delatour France, Paris, 2009)

A. Paolo, A. Fabio, B. Renato, P. Andrea, P. Patrizio, P. Angelo, R. Claudio, S. Sylviane, V. Mauro, The new mars workstation, in *Proceedings of the... International Computer Music Conference* (Computer Music Association, Thessaloniki, 1997), p. 215

T.H. Park, An interview with max mathews. Comput. Music J. **33**(3), 9–22 (2009)

T. Pinch, Why you go to a music store to buy a synthesizer: path dependence and the social construction of technology, in *Path Dependence and Creation* (Lawrence Erlbaum Associates, London, 2001), pp. 381–400

M. Puckette, The patcher, in *Proceedings of the International Computer Music Conference*, Cologne (1988)

M. Puckette, Combining event and signal processing in the max graphical programming environment. Comput. Music J. **15**(3), 68–77 (1991a)

M. Puckette, FTS: a real-time monitor for multiprocessor music synthesis. Comput. Music J. **15**(3), 58–67 (1991b)

M. Puckette, Something digital. Comput. Music J. **15**(4), 65–69 (1991c)

M. Puckette, Pure data: another integrated computer music environment, in *Proceedings of the Second Intercollege Computer Music Concerts*, Tokyo (1996), pp. 37–41

M. Puckette, Pure data, in *Proceedings of the International Computer Music Conference*, Thessaloniki (1997)

M. Puckette, Max at seventeen. Comput. Music J. **26**(4), 31–43 (2002)

M. Puckette, MEM Studio, Music 500: a new real-time digital synthesis system, in *Proceedings of the International Computer Music Conference*, Rochester (1983)

B. Ramsay, Social spatialisation: exploring links within contemporary sonic art. *eContact! 14.4-Toronto Electroacoustic Symposium 2011 (TES 2011),* 14(4) (2013)

E.S. Raymond, *The Cathedral & the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary* (O'Reilly Media, Sebastopol, 2001)

W.J. Raynor, *The International Dictionary of Artificial Intelligence* (Global Professional Publishing, Chicago, 1999)

A. Renaud, A. Carôt, P. Rebelo, Networked music performance: state of the art, in *Proceedings of the AES 30th International Conference* (Saariselkä, 2007), p. 16

J.-C. Risset, Sculpting sounds with computers: music, science, technology. Leonardo **27**, 257–261 (1994)
J.-C. Risset, *Computer Music: Why?* (France-University of Texas Institute, Austin, 2003)
C. Roads, *The Computer Music Tutorial* (The MIT Press, Cambridge, MA, 1996)
C. Roads, *Microsound* (The MIT Press, Cambridge, MA, 2004)
C. Roads, P. Wieneke, Grammars as representations for music. Comput. Music J. **3**(1), 48–55 (1979)
C. Roads, M. Mathews, Interview with max Mathews. Comput. Music J. **4**(4), 15–22 (1980)
R. Rowe, *Interactive Music Systems: Machine Listening and Composing* (The MIT Press, Cambridge, MA, 1992a)
R. Rowe, Machine listening and composing with cypher. Comput. Music J. **16**(1), 43–63 (1992b)
R. Rowe, *Machine Musicianship* (The MIT Press, Cambridge, MA, 2004)
P.R. Samson, A general-purpose digital synthesizer. J. Audio Eng. Soc. **28**(3), 106–113 (1980)
A.A. Sawchuk, E. Chew, R. Zimmermann, C. Papadopoulos, C. Kyriakakis, From remote media immersion to distributed immersive performance, in *Proceedings of the 2003 ACM SIGMM Workshop on Experiential telepresence* (ACM, Berkeley, 2003), pp. 110–120
C.A. Scaletti, R.E. Johnson, An interactive environment for object-oriented music composition and sound synthesis, in *Conference Proceedings on Object-Oriented Programming Systems, Languages and Applications*, *OOPSLA '88* (ACM, New York, 1988), pp. 222–233
G. Scavone, P. Cook, RTMidi, RTAudio, and a synthesis toolkit (STK) update, in *Proceedings of the 2005 International Computer Music Conference*, Barcelona (2005)
E. Schooler, J. Touch, *Distributed Music: A Foray into Networked Performance* (International Network Music Festival, Santa Monica, 1993)
B. Schottstaedt, PLA: a composer's idea of a language. Comput. Music J. **7**(1), 11–20 (1983)
B. Schottstaedt, Machine tongues xvii: CLM: music v meets common Lisp. Comput. Music J. **18**(2), 30–37 (1994)
L. Smith, Score-a musician's approach to computer music. J. Audio Eng. Soc. **20**(1), 7–14 (1972)
R.B. Smith, An interview with Tristan Murail. Comput. Music J. **24**(1), 11–19 (2000)
J. Smith, D. Jaffe, L. Boynton, Sound and music on the next computer, in *Audio Engineering Society Conference: 7th International Conference: Audio in Digital Times* (Audio Engineering Society, Toronto, 1989)
A. Sorensen, Impromptu: an interactive programming environment for composition and performance, in *Proceedings of the Australasian Computer Music Conference 2009*, Brisbane (2005)
A. Sorensen, B. Swift, A. Riddell, The many meanings of live coding. Comput. Music J. **38**(1), 65–76 (2014)
A. Soresen, H. Gardner, Programming with time: cyberphysical programming with impromptu, in *Proceedings of the 2010 SPLASH/OOPSLA*, Nevada (2010)
M. Sosnick, W. Hsu, Efficient finite difference-based sound synthesis using GPUs, in *Proceedings of the Sound and Music Computing Conference*, Barcelona (2010)
M. Sperber, R.K. Dybvig, M. Flatt, A. van Straaten, Revised6 report on the algorithmic language scheme (nonnormative appendices) (2007)
A. Tanaka, Musical performance practice on sensor-based instruments. Trends Gestural Control Music **13**, 389–405 (2000)
H. Taube, An introduction to common music. Comput. Music J. **21**(1), 29–34 (1997)
J. Tenney, Sound-generation by means of a digital computer. J. Music Ther. **7**(1), 24–70 (1963)
D. Teruggi, Technology and musique concrète: the technical developments of the groupe de recherches musicales and their implication in musical composition. Organised Sound **12**(3), 213 (2007)
R.S. Thompson, Eric Lyon: designing audio objects for max/msp and pd. Comput. Music J. **38**(2), 80–84 (2014)
Toplap, (Toplap, 2004), http://toplap.org. Accessed 14 Nov 2014
G. Tzanetakis, Marsyas-0.2: a case study in implementing music information retrieval systems. *Intelligent Music Information Systems. IGI Global* (2007)

B. Vercoe, Computer systems and languages for audio research, in *Audio Engineering Society Conference: 1st International Conference: Digital Audio* (Audio Engineering Society, Rye, 1982)

B. Vercoe, *Csound: A Manual for the Audio Processing System and Supporting Programs with Tutorials* (Massachusetts Institute of Technology, Cambridge, MA, 1993)

B. Vercoe, D. Ellis, Real-time Csound: software synthesis with sensing and control, in *Proceedings of the International Computer Music Conference*, Glasgow (1990), pp. 209–211

G. Wakefield, W. Smith, Using LUA for multimedia composition, in *Proceedings of the International Computer Music Conference.* (International Computer Music Association, San Francisco, 2007), pp. 1–4

G. Wakefield, W. Smith, C. Roberts, LuaAV: extensibility and heterogeneity for audiovisual computing, in *Proceedings of the Linux Audio Conference*, Utrecht (2010)

D.C. Walden, T. Van Vleck, F. Corbató, *The Compatible Time Sharing System (1961–1973): Fiftieth Anniversary Commemorative Overview* (IEEE Computer Society, 2011)

J. Walter-Herrmann, C. Büching, *FabLab: Of Machines, Makers and Inventors* (transcript Verlag, Bielefeld, 2014)

G. Wang, P. Cook, Chuck: a programming language for on-the-fly, real-time audio synthesis and multimedia, in *Proceedings of the 12th Annual ACM International Conference on Multimedia* (ACM, New York, 2004), pp. 812–815

G. Wang, R. Fiebrink, P.R. Cook, Combining analysis and synthesis in the chuck programming language, in *Proceedings of the International Computer Music Conference*, Copenhagen (2007), pp. 35–42

G. Wang, *The Chuck Audio Programming Language. A Strongly-Timed and on-the-Fly Environ/Mentality*. Ph.D. thesis, Princeton University, 2008

G. Wang, G. Essl, J. Smith, S. Salazar, P. Cook, R. Hamilton, R. Fiebrink, J. Berger, D. Zhu, M. Ljungstrom, et al., Smule = sonic media: an intersection of the mobile, musical, and social, in *Proceedings of the International Computer Music Conference (ICMC 2009)*, Montreal (2009), pp. 16–21

A. Ward, J. Rohrhuber, F. Olofsson, A. McLean, D. Griffiths, N. Collins, A. Alexander, Live algorithm programming and a temporary organisation for its promotion, in *Proceedings of the README Software Art Conference*, Aarhus (2004)

S. Wilson, D. Cottle, N. Collins, *The SuperCollider Book* (The MIT Press, Cambridge, MA, 2011)

M. Wright, Open sound control-a new protocol for communicating with sound synthesizers, in *Proceedings of the 1997 International Computer Music Conference*, Thessaloniki (1997), pp. 101–104

M. Wright, Open sound control: an enabling technology for musical networking. Organised Sound **10**(03), 193–200 (2005)

I. Xenakis, *Formalized Music: Thought and Mathematics in Composition* (Pendragon Press, Hillsdale, 1992)

D. Zicarelli, M and jam factory. Comput. Music J. **11**, 13–29 (1987)

D. Zicarelli, An extensible real-time signal processing environment for max, in *Proceedings of the 1998 International Computer Music Conference*, Ann Arbor (1998)

Häusliches Glück: A Case Study on Deception in a Mixed Reality Environment

29

Alex Davies and Jeffrey Koh

Contents

Abstract

In this chapter, we present a case study of a mixed reality environment that leverages on concepts inherited from the application of deception as demonstrated by magicians, illusionists, and other practitioners that use deception to develop compelling narratives mapped multimodality. These concepts are demonstrated in the context of a spatial cinematic art installation. From analysis of this case study, we found that these techniques can be effective tools in the creation of convincing mediated experiences. This chapter begins with an overview of the work and development processes. It then examines approaches to illusion in

A. Davies (✉) • J. Koh
Creative Robotics Lab, NIEA, UNSW Art and Design, Sydney, NSW, Australia
e-mail: alex.davies@unsw.edu.au; jeffrey.koh@unsw.edu.au

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_18

terms of physical devices (mechanics) and misdirection (the underpinning psychological principles of conjuring). This is followed by a detailed investigation of the work's two narrative scripts in order to offer insight into the spatial dynamics of the audience experience.

Keywords
Media art • Entertainment • Interactivity • Mixed reality • Deception • Cinema • Scenography • Narrative environment

Introduction

This chapter examines the immersive installation, *Häusliches Glück*, in terms of the creation of illusory presence in mixed reality installations. Drawing from the practices of stage magicians, the work employs a number of techniques to advance the idea that the virtual characters the audience encounters are physically present, rather than media simulations. These are framing context, consistency, continuity, conviction, justification, surprise, and disguise. They are derived from a range of techniques that magicians routinely employ to deceive audiences, commonly referred to under the umbrella term of "misdirection" (Lamont and Wiseman 1999).

You enter the apartment and look around. The rooms appear to still be inhabited. Artifacts from recent human activity are present everywhere: partially prepared food on the kitchen bench top, freshly washed laundry hanging out to dry, an open magazine on the sofa, a television left on in the corner. This place looks lived in. This is an exhibition space, but it feels as if you have walked into someone's home. Where are the inhabitants? What happened here? You hear noises from a back room. As you draw closer, you notice the door is ajar but partially blocked by a cupboard from the other side. You hear voices. A shadow darts across the wall inside the room. You strain to look into the space, but can only see a fragment of the room. You continue to eavesdrop on the conversation that is taking place, and notice that the old television beside you flickers with a grainy black and white image of the room behind the cupboard. A great commotion suddenly occurs from within the room. You hear a crashing sound and, glancing across at the television, see the action unfold within the room. You realize that what you're actually viewing on the screen corresponds with the activity you can hear within the room... A phone rings in the lounge room. Although you are compelled to stay and find out what is happening behind the blocked doorway, you momentary pull yourself away to answer the call. A flustered voice anxiously gives you directions down the phone line. The line goes dead...

Häusliches Glück is an installation in which audience members unwittingly find themselves at the center of a story unfolding around them. The work combines techniques of mixed reality, physical narrative, and cinema. Precomposed sound, video media, and mechanical devices are used in combination with the recreation of a 1950s style working class apartment in Linz, Austria, to form a series of convincing illusions. The installation offers a range of experiences to audience members

Fig. 1 *Häusliches Glück*, installation detail with audience member in a bedroom

drawn from a series of narratives that are presented via a television simulating a live video feed from the adjoining room. The entire apartment is augmented by sound and image technologies that create the illusion of these events occurring "live." Drawn in by the intimate reality of the deserted apartment, visitors transcend the experience of being bystanders, becoming active participants in a layered set of narratives. The documentation of two murders in 1957, unfolding in the present moment, creates a contradiction of the senses (Fig. 1).

Häusliches Glück was exhibited in a dilapidated first floor apartment in Linz from July to September 2009. The work was developed in conjunction with Time's Up, and was presented for the 2009 European Capital of Culture as part of the *Haus Der Geschichten* (House of Stories) project ("Haus Der Geschichten" is translated as "House of Stories" but "Geschicten" can also mean "history"). *Haus Der Geschichten* exhibited the site-specific work of various artists in an old vacant apartment building in downtown Linz throughout 2009. The central idea was to convert disused urban space into "a setting for stories: of wonderment, invention, and drama (Linz09 2009)" (Fig. 2).

The overarching narrative of the installation can be broken down into two distinct acts. The first act is set in 1957 and presented by a narrator who appears to be located in a room adjacent to the audience. The two additional characters in the story, Wilhelm (the victim) and the unnamed killer, play out their roles throughout the apartment via sound, image, and mechanical devices (Fig. 3).

The second act is designed to seamlessly transition from the first and is a counterpoint to the preceding drama; it is firmly placed in the present moment and the context of the exhibition space. Two exhibition attendants enter the room that the narrator previously occupied and proceed to dismantle the installation. Each act employs different framing contexts to examine deception within the mixed reality installation (Fig. 4).

The work was developed on site over 2 months leading up to the public exhibition opening. Two acts combine to an overall duration of approximately 30 min. A script for Act One was developed in the murder mystery genre and shot with three central characters. A loose script was developed for Act Two, defined around key points of

Fig. 2 *Häusliches Glück* – Pfarrplatz 18, Linz, *Haus der Geschichten* location

Fig. 3 *Häusliches Glück*, video still from Act 1, Wilhelm entering the bathroom

drama that enabled the two actors to improvise the scene. The apartment was intricately dressed and artifacts were precisely maintained for the public presentation period so that they would match the precomposed video content. In addition to the

Fig. 4 *Häusliches Glück*, video still from Act 2, Linz09 gallery attendants

Fig. 5 *Häusliches Glück,* entrance/kitchen before and after installation

media content, all these physical elements enabled the audience to construct their own stories based upon the artifacts they encountered and directly reinforced the predetermined narrative in Act One (Fig. 5).

Based on the content of the scripts, actors were shot using CCTV cameras located in the rear room (obstructed by the cupboard) and the bathroom. The rear room was subsequently emptied and used as a machine room (housing the technology) for controlling the rest of the installation space. Props from the shoot (books, furniture, paintings, etc.) are maintained within the narrow sight lines between the cupboard and door jam. However, the rest of the room is completely replaced with the controlling technology.

This is the first instance of deception in the work. The cupboard not only provides a physical space where the virtual characters can exist, it also functions to conceal the underlying technologies. Following is a discussion of these technologies (Fig. 6).

Mechanics

This section examines key technologies that are used to augment the installation environment. Various technologies are used in the work to create the virtual characters and, additionally, to support the physical actions of those characters. In the first act, these technologies convey the story presented by the narrator and in the second act they support the characters' decisions and actions (i.e., the virtual characters' actions are consistent with the environmental stimuli). A range of technologies are implemented to achieve a cohesive illusion throughout the whole apartment. Put simply, everything that is described via the narrator in the story simultaneously occurs spatially in the apartment via sound, video, lighting, and mechanical devices. In such a polysensory environment, it is important that each modality works in unison because, as Anderson contends "when we perceive multi-modally, we seek the invariant properties of an event across modalities (Anderson 1996, p. 86)." *Häusliches Glück* not only utilizes a wide range of technologies but does so in a globally consistent manner to reinforce the illusion.

Audio

Audio is a central device for establishing illusion within the installation. While the television image presents a compelling focal point for the narrative, sound contributes to spatial awareness of virtual characters throughout the apartment.

The placement of sound within the installation is analogous to offscreen sound in cinema. However, rather than just representing action in an abstract location beyond screen space, sound emanates from points in the apartment outside the audience's sight lines, and is linked to specific physical locations within the apartment. Both these applications of sound achieve the same pronounced effect, creating an imaginary presence beyond the user's immediate visual awareness. To implement this sonic landscape, speakers are concealed at key locations in the installation where narrative activity occurs.

Following is a list of the six speakers employed throughout the apartment and their function in terms of supporting the live narrative.

Fig. 6 *Häusliches Glück,* installation floor plan

1. *Radio*: The internal speaker of the radio is used to present diegetic sound for the installation. Although the device gives the impression of an old wireless unit from the turn of the century, it is in fact controlled via the software and wired to the central amplifier point in the control room. Diegetic sound is a particularly useful approach to illusion as, by its very nature, it is associated with a device that is commonly used to generate sound. Familiarity and consistency therefore reduce suspicion. In this instance, using such a device eliminates the need to disguise it, while, at the same time, supporting the framing context of the narrative.

Fig. 7 *Häusliches Glück*, radio

Function: Atmosphere is created through diegetic sound directly related to the narrator's story. As the narrator describes this element of the story, the radio crackles to life and music fades up and continues for the duration of Act One.

The musical accompaniment is drawn from the 1957 Eurovision song contest and creates a whimsical counterpoint to tension developed in the rest of the narrative. Musical accompaniment from the radio also provides continuity for the work as a whole by assisting in fluid transitions between the two distinct acts (Fig. 7).

2. *Kitchen*: A speaker is concealed within the kitchen cupboard behind a range of kitchen implements.
   Function: This speaker is used to represent narrative elements that occur in the region of the entrance and kitchen. Sound is panned from this point to speakers located in the bathroom and bedroom to indicate movement of virtual characters when appropriate (footsteps, for example). Additionally, it is used to represent point sources of sound occurring at that specific location, such as when Wilhelm is rummaging through the drawers to find candles (Fig. 8).
3. *Bathroom*: A speaker is concealed near the ceiling on the top of the water heater in the bathroom behind a rag and stack of old newspapers. A second speaker is concealed underneath the bathtub.
   Function: The speaker located near the ceiling is used momentarily in the first act to represent activity from the upstairs neighbors and, therefore, is very specific in its function. The second speaker underneath the bathtub presents all other activity that occurs within the bathroom (Fig. 9).
4. *Windows*: Speakers are concealed within each of the apartment's three external windows. Shutters and curtains are added to the interior of the window frames to conceal the devices from the audience, and similarly, curtains are used to conceal them from the outside of the apartment.
   Function: All three speakers are used to represent the world beyond the confines of the apartment. This includes environmental elements that occur during the narrative such as rain and human activity from other imagined apartments in the vicinity. For instance, the sound of neighbors arguing can be heard in the distance (Fig. 10).

Fig. 8 *Häusliches Glück*, kitchen speaker (in cupboard)

Fig. 9 *Häusliches Glück,* bathroom speaker (above water heater)

Fig. 10 *Häusliches Glück,* window speaker

Fig. 11 *Häusliches Glück*, narrator speaker located above cupboard

5. *Narrator's Room*: In addition to the window speaker located in the narrator's room, one additional speaker is located centrally within the room. As with all the other mechanical elements present, this is concealed (in this instance, by the cupboard).
   Function: The speaker conveys sound related to any activity that occurs within the room such as the narrator's monologue, footsteps, or simulated glass smashing. As the sound is filtered via the cupboard and indirectly heard by the audience at a distance, the single speaker is capable of producing enough spatial resolution to cater for all these dramatic elements convincingly (Fig. 11).
6. *Phone*: A phone located in the bedroom was rewired and is controlled via computer to ring and, if picked up, to voice a response.
   Function: The phone is used as a plot device in Act Two. It is triggered to ring during the narrative. The virtual characters encourage the audience to assist them (i.e., by answering it). Depending on the audience's actions, the narrative goes off on one of two paths. If an audience member picks up the phone at any other time, they are presented with a dial tone, then the phone seemingly commences an auto dial and, after ringing for a short time, is answered with a randomly selected recording of a confused woman who has been woken from a deep slumber by the call. The woman, with some irritability, quickly terminates the conversation (Fig. 12).

Video

The actors were shot with CCTV cameras located throughout the apartment, one in the rear room, one in the entrance/kitchen, and one in the bathroom. These three cameras remain as props in the installation to reinforce the illusion. When a character can be seen walking into the bathroom, for instance, the aim is that the audience interprets this action as occurring live via the camera they noticed earlier, not prerecorded.

Fig. 12 *Häusliches Glück,* phone in bedroom

Fig. 13 *Häusliches Glück*, television in bedroom displaying "CCTV" video

Video technology is used to convey the illusory sense of virtual presence in several ways. The image on the television forms the central focal point of the illusion, and a range of other techniques are employed to support the image presented on screen (Fig. 13).

The apartment contained three CCTV cameras mounted in discreet locations within the entrance, bathroom, and narrator's room. While the camera's presence is unassuming, there was no attempt to completely conceal them, as the audience must believe these cameras are presenting "live" video feeds to the television.

The low-resolution aesthetic of the image presented on the television is therefore interpreted as normal and consistent with audience's expectations of the visual appearance of CCTV cameras. The presence of cameras, however, is out of character with the rest of the environment's design, as all other artifacts are located within an earlier time period. These disparate objects innately draw attention to themselves in the context of the surrounding environment and aid in creating a false frame of reference, an expectation that the audience is viewing live events on the television, a mediated performance in the first act, and reality in the second act (Fig. 14).

Fig. 14 *Häusliches Glück*, camera location in kitchen (behind lamp)

Fig. 15 *Häusliches Glück*, audience sight lines to the narrator's room (camera *top left*)

The visibility of the camera is particularly important in the narrator's room (Fig. 15), as this is one of several indicators that the characters in the room are being transmitted live to the television. The camera is visible through the narrow gap in the door that is otherwise obscured by the cupboard. The audience's sight lines into the room are heavily controlled and very few artifacts are visible in this section of the room, further drawing attention to the device. The camera is mounted in a prominent position in the corner of the room, so when the audience peeks into the room, they see the CCTV camera, hear the character from within the room, and see the "live" image on the screen. The combination of these elements, in addition to the mechanical devices creates a palpable sense of human presence.

To summarize, video technologies are used in conjunction with other devices as a means to support the two narratives presented and develop the illusion of virtual characters coexisting with the audience. The audience's preconceptions concerning the technologies themselves are exploited and aid in the establishment of appropriate framing contexts required for a successful illusion.

Physical Elements

Häusliches Glück employs a range of mechanical devices activated during specific points in the narrative when it is appropriate for the actions of the virtual characters to be physically manifest within the environment. These devices are used in conjunction with the visual and aural elements to create a more cohesive illusion. The addition of mechanical elements in the work plays a significant role in the perceptual realism of the virtual characters through multimodal reinforcement, the illusory sum being greater than the parts. The mechanical elements also shift the focus from simulating agents via the abstract elements of sound and image, to events that physically occur in the environment. The actions of virtual characters therefore have tangible physical outcomes in the real world. In terms of illusion, perceptual realism will inevitably have a more pronounced and plausible effect if the object itself achieves it, rather than by simulation, no matter how sophisticated. For example, why simulate the sound of a person moving a cupboard when the audience can experience the cupboard physically move? Mechanical devices therefore assist in convincingly bringing the characters from the virtual world to cohabit real physical space with the audience. Additionally, mechanics are employed to develop the *mise-en-scène* of the installation as a whole.

Just as sound plays a crucial role in the spatial representation of the narrative, so too does lighting. Every lighting element in the installation is directly linked to the narrative and controlled via software. Lighting fixtures in the apartment are rewired to a DMX lighting dimmer and controlled in relation to the story elements. In addition to the visible elements in the apartment such as ceiling lights and lamps, lighting is also concealed within the window shutters to simulate lightning, and within the narrator's room for specific effects.

The lighting elements are therefore experienced directly (ceiling fixtures and lamps) or indirectly (obscured lighting events occurring in the narrator's room), analogous to the way audiences perceive onscreen and offscreen elements in cinema, yet, in this instance, occurring in physical space (Fig. 16).

The same lighting control software is used during production and exhibition so that the physical and virtual content can be synchronized. The lighting sequences

Fig. 16 *Häusliches Glück*, bedroom light, rewired to DMX controller

that occur in physical space therefore directly relate to the lighting represented on the television image (as the audience is meant to believe that the cameras portraying live events are within the same physical space).

In addition to lighting that encompasses atmospheric elements inherent in the narrative, lighting also represents actions directly related to the activity of the virtual characters. Three other lighting elements are located in the narrator's room and therefore are not directly visible to the audience. As these lighting effects are experienced emanating from the inaccessible room, they can present an accurate simulation of reality. The audience is not in a position to confirm the presence of individuals within the room controlling the lighting changes, only the results of their actions. These items include a floor lamp, a flashlight, and a lighter, the use of which will be discussed in detail within the narrative analysis section.

While lighting has a pronounced effect on the atmosphere of the installation, and therefore significantly contributes to conviction in the work, by its nature, it is still somewhat intangible, much like the characteristics of sound in the work. Although both highly effective in their own right, these devices do not have the same visceral impact that a physical object would impart.

To further enhance the influence of virtual characters in real physical space, the cupboard is called into play. It is not only used to obstruct audience access to the room but also conveys the actions of the characters by mechanical means through motorized control.

While disguise of visual elements is central to conjuring and is similarly necessary for illusion in mixed realities, disguise is also required to minimize any incongruous audio elements of the work. Just as ill-placed cables have the ability to detract from the illusion, so do unwarranted sounds. A range of motorized devices is used to augment the environment. The goal is that the audience perceives the effect of these devices, but not the device itself, for example, seeing the cupboard shake or the flashlight sweep around the room, without the artificial whirring of motors. Both these elements are soundproofed, the servo mechanism for the flashlight via a transparent dome that transmits light yet no sound, and the cupboard motor in a

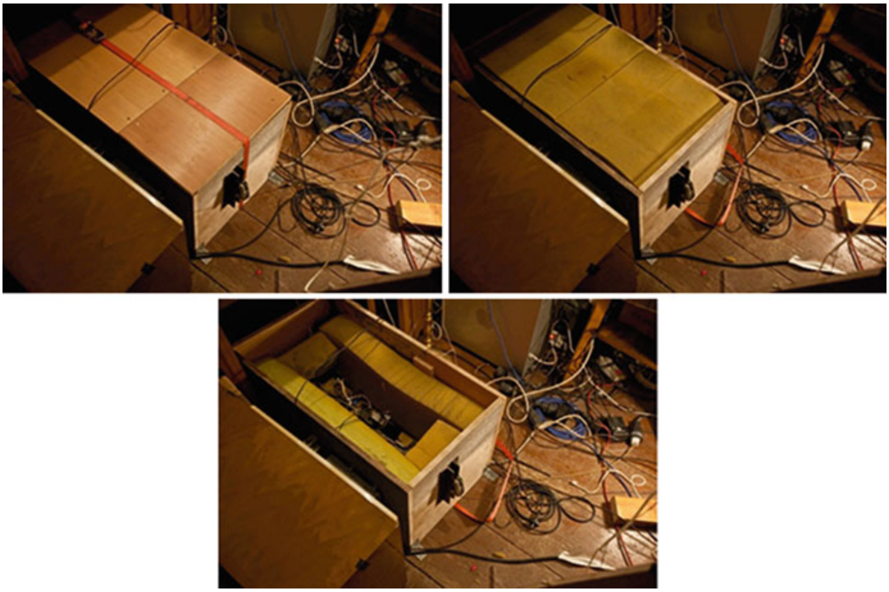


Fig. 17 *Häusliches Glück, s*ound treatment for cupboard shaker motor

Fig. 18 *Häusliches Glück*, sound treatment dome for flashlight servo

sound proof box. The creaks and clunks of the cupboard moving additionally aid in masking the motor sound when activated (Figs. 17 and 18).

Misdirection

This section examines how the psychological principles of deception employed in magic are used in the work and how these techniques contribute to the palpable sense of virtual characters occupying the installation space with the audience.

Fig. 19 *Häusliches Glück,* apartment/installation entrance (*left door*)

Framing Context

Appropriate framing is central to the success of an illusion. *Häusliches Glück* combines several framing contexts to achieve the desired illusory outcomes. The first is inherently part of the gallery-based experience, the subsequent two are systematically constructed via the mechanics and content to establish a deliberately ambiguous context in which the work is interpreted.

The installation was presented over a 3-month period in 2009 as part of the *Haus Der Geschichten* exhibition. In terms of audience experience, the work was therefore framed, firstly, as an artwork presented in the site-specific context of an old apartment building. Beyond this fact however, the audience was not aware of the specific nature of the work. The catalog publication and promotional material were intentionally vague and did not convey any detail as to the type or content of installation. Audiences therefore encountered the work with few preconceptions (Fig. 19).

Upon initially entering the work, the framing context of *Häusliches Glück* becomes even more ambiguous as the narratives consist of several discreet and shifting framing contexts that are revealed as the work evolves.

On one hand, the work is framed as a spatial cinema installation in which the narrative elements are clearly fictional (displaced in time for instance). On the other hand, as the narrative unfolds, certain aspects are framed to give the impression that the characters presented are in fact real and physically present (i.e., actors presenting the work live). This tension between the two forms of illusion (the deception of the conjuring trick and suspension of disbelief) causes the audience to continuously reframe and question their relationship to the media. Act Two, the second narrative component of the work is framed as a counterpoint to the overt drama of the first narrative. Rather than sharing the apartment with a murderer from the 1950s, the audience is again transferred to the present and encounters two gallery attendants coinhabiting the space. The banality of these characters and their actions separate them from the previous content and reframe the experience as present reality. Rather than willing the audience to suspend disbelief, the presence of these virtual characters is articulated as fact with the aim of being interpreted as such.

Consistency

Häusliches Glück, as a narrative driven work, requires consistency in characterization and the broader environment in order to create and maintain a convincing illusionary space. Consistency is addressed within each scene and across the work as a whole.

In the first act, the characters' actions are consistent in a similar manner to the traits of cinematic or theatrical actors. Their motivations and actions are in accordance with the heightened sense of drama presented. In the second act, the characters are designed to display normal and appropriate behavior for gallery attendants. Consistency of character enhances the perceptual realism of the scene and opens the possibility for the deception to occur.

In summary, the actors behave like actors in a fictional drama, and the Linz09 characters behave like gallery attendants. Although each act relies on consistency of characterization to develop the illusion, the first act does so to foster suspension of disbelief, and the second to deceive the audience.

Continuity

Continuity is the sequential flow of narrative events, in other words, the consistent arrangement of events in a logical manner. The principle of continuity in magic and cinema are equally applicable in the domain of mixed reality as a means to maintain audience interest and reduce suspicion through the removal of incongruous elements. Continuity is an integral part of illusion. If the sequence of events defy logic and confuse the audience, the next step of the story in cinema, or the performance in conjuring, will be missed. Interest is reduced when this occurs, suspension of disbelief is diminished in cinema, and deception is jeopardized in conjuring. Disruptions in continuity are specifically avoided in *Häusliches Glück* and influence the way in which the illusion is implemented.

Continuity within *Häusliches Glück* can be examined within each act and also in terms of the dynamic flow of the installation as a whole. Act One demonstrates a linear narrative that follows the fundamental principles of structure found in a magic trick or conventional cinema narrative. The story follows a path of cascading smaller climaxes until reaching the primary climax at the end. This structure keeps the audience engaged with the narrative and encourages suspension of disbelief.

In contrast, Act Two intentionally does not follow these conventions as the aim is to represent reality, and reality is often mundane, lacking the climactic dynamics of fiction. Although there are several minor dramatic climaxes in the scene to sustain audience interest, they are not overt and are in keeping with the deceptive framing context of the act. Overall continuity is maintained through the characters' actions following a logical course.

While continuity is inherent in each act, it is also employed so that the transitions between acts are fluid and logical. The installation moves seamlessly from one dramatic point to the next, looping around itself so that it can be experienced at any given point in time by an audience member. Unlike a conventional cinematic or theatrical experience, the audience is not necessarily required to experience the work between finite start and end times. Act One begins with what could be considered the default environmental state of the installation. The narrator enters and the environmental elements progressively change in accordance to the narrative. For the work to flow, these changes have to be restored in a coherent and logical manner for the work to restart from the beginning without jarring transitions that would disrupt the experience. [Methods used to achieve these transitions are discussed further in the Narrative Analysis section below].

Conviction

As illusion in the work relies heavily on narrative elements to create a compelling experience for the user, conviction is critical. On one level, the audience requires conviction to fundamentally engage with the narrative. If the narrative does not prove to be sufficiently engaging, the audience's interest will waver and the work will not achieve a high degree of presence.

For this reason, the first act is based on what could only be described as a clichéd murder mystery. The first act's overall dramatic structure and content are easy to comprehend; however, it is the spatial approach to narrative that creates a richly compelling scenario. Due to the nature of the exhibition environment, that is, nonfixed audience entry times and variable attendance durations, a simple and well-trodden story framework is appropriate as it falls within a familiar genre. The audience is therefore able to easily engage with the story at any given point of entry. Additionally, if the story is illogical or unconventional, the addition of the wide ranging mediation used to present the drama would more likely create a state of confusion, disrupting suspension of disbelief. As Nelms maintains, when the degree of conviction is high, and the audience is engaged in the dramatic elements of the presentation, they are less likely to seek out the methods of deception (Nelms 1969). Further, a high degree of presence also establishes the possibility for the audience to perceive the simulated characters as possibly existing in the space and presenting the drama as live theater. Without the foundation of conviction, illusory deception would not be possible.

Whereas conviction is achieved predominately through narrative in the areas of cinema, theater, and literature, in this instance the story's unique presentation environment is also a contributing factor. Conviction is shaped through a wide range of other elements in the installation encompassing objects and artifacts, lighting, spatial sound, and olfactory elements. For example, unlike installations presented in relatively sterile "white cube" gallery spaces, *Häusliches Glück* takes place in an old apartment building at Pfarrplatz 18, Linz. The building itself has had over a century to develop its distinct scent. When entering the apartment, a heavy musty smell, mingling with stale tobacco greets the audience. The scent significantly contributes to the sense of immersion within the environment, even prior to the commencement of the electronically mediated elements.

Conviction in the second act serves a different purpose. The audience is not required to suspend disbelief in the drama, they need to be convinced that there was in fact *no drama*, and that the attendants were real and physically present. Conviction in this instance is based on the plausibility of the characters' correspondence to reality. Due to this, dramatic devices are not employed in the second scenario, for example, with lightning. The dramatic elements in Act One are even treated by the attendants as artificial, belonging to a fictional world, rather than the present reality that they inhabit. For example, the attendant pulling a power plug and stopping the music and rain further distances the characters from the previously artificial environment. In doing so, the attendants convey a similar perspective to the media that is also held by the audience, one of clear artifice. Their behavior effectively reveals the trickery of the prior Act. The combination of these actions, and the context in which they are presented, aims to deceive the audience into believing that the virtual characters are not virtual and indeed are physically present.

Justification

Justification is used throughout the work as a concealing device to maintain a fluid narrative and resolve the illusory deception. Characters' actions are required to be justified as in the above case of consistency, but beyond the behavior of the characters themselves, elements of the physical environment are used to justify aspects of the narrative that would otherwise diminish the illusion. In this instance, principles of magic are applied in the work to conceal deficiencies in technology. As it is uncommon for many technologies that developers encounter to behave *precisely* in the required manner, the use of the conjuring principle of justification is an excellent device in maintaining illusion in media arts.

A multifaceted example of the use of justification is evident in the cupboard that partially blocks the door to the rear room, obscuring the audience's view and inhibiting entry to the room beyond. Firstly, this obstruction offers justification for the inaccessible room, a controlled space from where many of the illusory elements occur or are controlled from. Secondly, due to the nature of the mechanics used to move the cupboard, only a close approximation of the virtual characters' actions (on screen) and the resulting movement of the cupboard are possible. Any major movement caused by a character would therefore be interpreted by the audience as unnatural, thus increasing suspicion. This mechanical issue is addressed via two different approaches, each of which is appropriate to the specific context of each Act.

The first act exploits light as a concealing device. When the killer enters the room with the narrator in the last scene, the mechanical cupboard is activated during the struggle in which the narrator is thrown into the object. It is deployed again after the narrator is murdered, when the killer attempts to gain access to the adjacent bedroom where the audience is situated by violently shaking the cupboard in an attempt to dislodge it. Preceding these two actions involving the cupboard, the killer walks through the apartment and systematically turns each light off, casting the audience – and eventually the narrator – into relative darkness. Simultaneously, the lightning simulation increases in a naturally rising progression coinciding with the increasing tension. By the end of the killer's walk, the only light visible in the apartment comes from flashes of lightning through the window shutters, a dim desk lamp in the bedroom with the audience, and a floor lamp obscured behind the cupboard in the narrator's room. This is a complex sequence of events and in the following section I will explain how each of these specific devices are used to justify the presence and absence of other elements, thereby maintaining consistency, continuity, and conviction.

Returning to the struggle in the room, as the narrator grapples with the killer, the audience sees her on the television being swung into the floor lamp, which subsequently topples and smashes on the floor. Synchronized DMX lighting simultaneously disables this lamp, and in conjunction with the television image, sound of the breaking lamp from within the room, and sudden darkness, creates a compelling illusion that the lamp has actually been smashed. This action justifies that the only remaining light in the room occurs only during lightning flashes. When the cupboard is violently shaken, the movement of the cupboard can be physically seen and heard, but the image of the killer's actions behind the cupboard are only seen as flickering fragments on the television due to the intermittent lightning activity. The audience therefore does not perceive discrepancies in synchronization. The fragmented visual information additionally adds to the drama as it increases uncertainty of what is occurring. The systematic lighting design of this final scene is therefore employed as a concealing device. The characteristics of the lighting are justified and do not arouse suspicion as they are appear consistent with the progression of the narrative.

Surprise

Surprise in magic occurs when the framing context is disrupted and an alternate reality is revealed. The element of surprise in *Häusliches Glück* is directly shaped by the installation's narrative structure, when the framing contexts shift between the two Acts. Surprise also operates indirectly at any given point of the work when the audience's perception of the virtual characters shifts either way between reality and virtuality.

In addition to surprise generated though the disruption of framing contexts, surprise also occurs as a dramatic device, primarily evident in the first act. This example to some degree is also linked to a perceived shift in framing. For instance, the audience has certain expectations that they are going to experience a prerecorded video work when they enter the apartment and first see the narrator on the television screen. They are subsequently surprised when the narrative elements such as the storm or the music spatially occur around them. Each new element builds on this surprise, as the situation the audience is presented with is unexpected and no longer conforms to media experiences they are familiar with. Although familiarity with the genre may lead them towards the overall outcome of the narrative, they are unable to predict how this will transpire within the mediated environment. An example of this is evident in the actions of the killer in the last scene. I observed audience members shocked when the cupboard violently moves as the narrator is thrown into it during the final struggle. I also saw audience members leap across the room in shock when the killer begins to furiously shake the cupboard to gain access to the bedroom where audience members are located. This event is wholly unexpected. Even though the audience may frame the incident as fiction presented by mediated characters, the element of surprise still works just as it does in a cinematic thriller, perhaps even more effectively than in a cinema space, due to the high degree of presence the installation environment is capable of producing.

Disguise

Disguise encompasses any number of techniques employed to conceal the method in order to create an effect. As illustrated above, disguise is closely linked to the psychology of justification in magic. Justification is employed in order to disguise devices via dramatic elements. Disguise is also necessary in the installation in order to conceal the wide array of technical elements (the mechanics) required to create the augmented environment. By concealing these elements, the audience experience is not clouded by elements that may indicate pretext, thereby ensuring that they instead focus on the significant factors contributing to the illusion. Conjuring, with its particular context and set of requirements, generally employs a far wider range of disguise including psychological expedients such as attention control, in which the magician may cunningly induce the spectator to focus their gaze on a specific subject to conceal a concurrent maneuver. In the case of the installation, physical disguise is appropriate and is the primary method used in concealing the mechanics of the illusion. Again these methods are closely linked to justification, as any method of disguise that is seen to be incongruous increases suspicion to the detriment of the illusion.

The audience experiences the work throughout the whole apartment, therefore the apartment as a whole and its containing elements could be interpreted as analogous to a device used by magician to execute a trick, the mechanics for illusion. Concealing devices encourage the suspension of disbelief by eliminating these mechanics from the audience's perception with the goal of a transparent medium. The apparent transparency of the mechanics enables the additional deceptive trickery that occurs and supports suspension of disbelief in the narrative. In all instances, disguise is employed to reduce suspicion by developing a natural and familiar environment in keeping with the framing context of the work.

Narrative Analysis

This section examines the two key narrative scripts presented in *Häusliches Glück* to provide background into the design of audience experience, timing and spatial media elements of the work, and to demonstrate how these are combined to advance the overall illusion. In contrast to the preceding general discussion, this section focuses on principles of deception that are applied to precise moments within the narratives.

As discussed previously, much of the success of a magic trick is due to destabilizing an audience's expectations. *Häusliches Glück* uses both mechanical devices and dramatic elements to continuously undermine audience expectations during the work. Although their expectation is primarily that they are experiencing a fictional narrative in the form of precomposed film, the illusory devices consistently indicate otherwise that the work is being played out by actors who are physically present or, indeed, gallery attendants who cohabit the space.

Analyses of the following scripts are provided in italics, capitals define characters, dialogue is indented, and action is indicated by parenthesis. (The following script for Act One was the original shooting script developed for the work. It was subsequently translated by Andreas Mayerhofer at Time's Up to a colloquial form of Austrian German for the production of the work. The choice of language was necessary to logically fit within the framing context of the story and cater for the local audience.)

Act One

Audience members enter the building and ascend stone stairs to the first floor. They notice a dimly lit apartment with the door slightly ajar. Upon entering the old apartment, they are greeted by the musty odor of stale tobacco, old books, and furnishings. Although the apartment appears to be vacant, it eerily shows signs of being recently inhabited. The bed is unmade, cigarettes overflow from ashtrays, food is half prepared, and cloths are strewn on the floor.

A computer vision system determines the presence of the spectator and activates the narrator sequence. The narrator enters the rear room through a door and sits down in a chair.

[The narrator addresses the audience directly. . ..]

NARRATOR Hello, come closer, I'd like to tell you a little story... About an ordinary apartment, with an extraordinary history...

This is timed so that the audience enters the bedroom from the entrance/kitchen approximately the same time as the narrator walks across the room and sits down. The audience can see the narrator on the television screen via what appears to be CCTV and hear her voice coming from within the next room. This split between sound and image is designed to accentuate the fact that although the interior of the blocked room is visible via a camera, the sound is direct from the narrator, not presented via the television. By presenting the sound that corresponds to the visual elements on the screen emanating from the location of the perceived source, the audience is likely to acknowledge that the mediated content is in fact occurring live from the next room rather than pre recorded content. Rather than talking towards the camera, she arches her head towards the bedroom as if addressing the audience directly.

NARRATOR The year is 1957; Bob Martin, Austria's first Eurovision contestant, had just finished in last place with *Wohin Kleines Pony*. This was not the only tragedy to befall Austria on that fateful day...

The radio located near the front entrance crackles to life and the song "Wohin Kleines Pony" can be heard. This song continues in its entirety as background music while the story unfolds. After the song is complete, the software randomly selects other songs that were also performed during 1957 Eurovision Song Contest. The radio is employed as a device for diegetic sound to create atmosphere for the narrative. The random function of the music provides variation within the work for repeat audience visits thus destabilizes expectations.

NARRATOR That night, our friend Wilhelm had just made it home as the storm broke, but not before he was caught in the deluge. He entered the apartment, peeled off his wet coat and momentarily slumped on the chair.

As the narrator describes the deluge, a flash of lightning can be seen though the shutters on the apartment windows, and the sound of rolling thunder and rain fade up through the speakers concealed behind each of the three windows in the apartment. The speakers are oriented so that the environmental sounds are perceived as coming directly from their source. The environmental effects suddenly shift the audience from the position of a passive observer to right in the centre of the fictional world.

The sound of Wilhelm entering can be heard in the kitchen via a speaker concealed within the cupboard. As the audience is currently located in the bedroom watching the narrator on the television, they can only hear this sound coming from behind them in the kitchen, as there are no direct sight lines to the source. The image on the television flickers and cuts from the live narrator camera to the bathroom camera. The audience sees Wilhelm enter the bathroom and hang up his umbrella, soaked and dripping from the rain. Wilhelm is seen onscreen drenched with water to reinforce the physical presence of the storm that the audience is experiencing live within the installation, thus narrowing the schism between the real and virtual.

When the video cuts between the narrator camera and bathroom camera it is prior to the entrance of Wilhelm. The audience therefore sees the bathroom empty for several seconds. This separates the work from film editing conventions that would naturally place the edit on his entrance. Although this is not perceived as a significant fact, due to the audience's familiarity with the smooth flowing edits inherent in cinematic drama, the cut subconsciously gives the impression that they are viewing a somewhat randomly switching CCTV system, not a refined cinematic sequence. All edits within the work are aimed at being slightly off, while at the same time finding a balance of not disrupting the narrative and therefore reducing suspension of disbelief.

[Wilhelm exits the bathroom and shortly after the camera cuts back to the narrator.]

NARRATOR It had been a long day at the Voerst and he was tired and a little drunk (the Voerst Alpina is a major steelworks operating in Linz and was chosen to make the characters' traits plausible and context specific). Prying his body from the chair he walked a little unsteadily into the bathroom to prepare a bath.

The actions of Wilhelm can be heard shifting from the kitchen back to the bathroom via the respective speakers.

WILHEM [Turns on tap] [A thunderclap is heard and the apartment lights flicker.] Mutters to himself cursing the weather

The audience sees Wilhelm turn on the tap in the bathroom via the television, however, also hear the sound emanating from the bathroom. When the apartment lights flicker the CCTV image on the television screen similarly flickers, dropping to black as the lights dip. This is the first instance when a physical occurrence experienced by the audience (the apartment lights) is mirrored in the virtual content (the video), thereby perceptively linking the two as occurring live.

NARRATOR He went over to search for candles just in case the power failed as it often did during these storms. The wiring in the building was old you see, just like most of the inhabitants.

A large burst of thunder can be heard and the lights in the apartment flicker. Wilhelm can be seen exiting the bathroom off camera and is then heard again rummaging through drawers in the kitchen.

NARRATOR Locating a candle, he returned to the bathroom.

[Wilhelm returns to bathroom and prepares for shaving with straight razor]

During this process Wilhelm begins whistling along with the music playing on the radio ("Wohin Klines Pony"). This again provides a correlation between the actions of the virtual character to the music emanating from an object physically present for the purpose of encouraging the audience to infer that both aspects are occurring in the physical environment. Simultaneous to Wilhelm's audio activity, the audience sees the narrator on the television getting out of her seat and rummaging through the room looking for candles. After opening some drawers in the room, she walks over to the cupboard that is blocking the door and opens it. As she does so, the cupboard moves ever so slightly as the motor control is activated. The audience sees this action on the television, hears her movement within the room, and sees the cupboard physically shudder. The narrator's action opening the cupboard is so incidental and matter-of-fact that no weight is placed on this occurrence. The dramatic impact, however, is significant. This is the first instance where the actions of a virtual character physically influence the environment (unlike the apartment lights previously in which a physical event impacted upon the media, that is, virtual to real as opposed to real to virtual.) Though subtle, there is a shift in the perceived framing context at this moment. The audience, initially comfortable with the fictional narrative played out as cinema, potentially questions their interpretation of the events. It is in fact the subtlety of this action that makes it effective. The result is unsettling as the *audience, while watching the television in front of them, notices the movement in the left periphery of their vision. If this movement is perceived, the device shifts the framing context from an immersive cinematic work to a live theatrical production.*

NARRATOR The neighbors were arguing about something trivial yet again, they always argued, it was not the fact that they argued that irked him, but the trivial nature of their conflicts. The whole building was like a giant buzzing hive and there rarely seemed to be any peace. [Sound of neighbors arguing/breaking glass]

WILHELM [Yelling from bathroom] Shut up, it's late for Christ's sake.

[Neighbors fall quiet for a moment, then respond with an antagonistic remark.]

The speaker concealed in the window cavity plays the distant sound of an argument. The sound is equalized and the volume adjusted so that it appears to come from an apartment across the atrium outside. The argument is between a man and woman. The dialog and sound effects for the argument and their subsequent response to Wilhelm's statement are randomized in software. Each argument presented is therefore unique in the instance of repeat audience visits. This scene is an example of the benefit of spatial storytelling and is particularly effective. The narrator recounts the tale from her room, Wilhelm can simultaneously be heard yelling from the bathroom on the other side of the apartment while the sound of the arguing couple is perceived in the distance, beyond the confines of the apartment.

WILHELM [Mutters to himself about neighbors]

NARRATOR Looking in the mirror, he noticed some mustard left over from the würst he consumed on the way home, wiping it off [Wilhelm wipes off mustard with finger and then eats it] he began to prepare for a shave; never a good idea after the consumption of alcohol, but this was the kind of idiocy that Wilhelm was accustomed to, and he was managing rather well under the circumstances. "Could mustard actually be used for shaving cream?" he wondered as he worked the lather onto his face. He was so consumed by the shaving process, and the lingering taste of Estragon, that he did not notice the figure enter though the door [Killer enters through the door from the storm and is wearing a soaking raincoat]. Standing in the doorway water pooled at his feet from the storm, his clothes soaked.

[Narrator pauses after describing the murderer]

NARRATOR I'm sorry... one moment, [clears throat, gets up from the chair and gets a glass of water resting on top of the chest of drawers. Sips water, sits back down and continues...] [Sound of footsteps]. Yes, Wilhelm never noticed him enter...

The pause is added for dramatic tension. It disrupts the flow of the fictional narrative while maintaining the progression of the "live" narrative. When the narrator ceases to recount the story, she directly addresses the audience by apologizing. By engaging with the audience, the narrator places herself in the present, unlike the other characters displaced in time, bringing both parties together and encouraging presence via social interaction. Wilhelm is seen on the television oblivious to the presence of the killer that the audience can hear walking across the entrance room towards the bathroom.

[Moments later the killer enters bathroom, Wilhelm turns from mirror still with shaving cream on face.]

WILHELM What?

[The two struggle and the killer wraps a rope around Wilhelm's neck. As they continue to grapple they orient themselves facing away from the camera. The killer can be seen grabbing Wilhelm's hand holding the razor. As they both spin back around towards the camera Wilhelm's throat is cut. He screams and falls limply to the floor]

The killer and Wilhelm are heard engaged in a violent struggle in the bathroom until a scream is heard and the audience hears a body drop.

NARRATOR As you can see, the situation for our friend Wilhelm was not good, not good at all. The blade had expertly sliced through his carotid artery. Blood began to pool around his motionless body, as it lay twisted on the floor, his life rapidly draining away.

The narrator again addresses the audience with this statement (as she does with the inviting statement at the beginning of the story and the prior apology).

[Thumping on roof coming from above the bathroom. Muffled sounds, "keep it quiet down there"]

A speaker concealed near the ceiling in the bathroom plays equalized audio to give the impression that the thumping and dialogue are coming from apartment above. The killer is seen via the television looking up, responding to the source of the sound. The actions of the character on screen are directly linked to the sound physically emanating from a different room. This links the behavior of what may be perceived as a virtual character (mediated via the television) to events that occur in real physical space.

NARRATOR The rain continued to sheet down outside…the storm's ferocity seemingly mirroring the violence of the attack as if they were intertwined in some form of twisted symbiosis…

The simulated storm outside increases in intensity, lightning and thunder are more frequent as is the resulting erratic power in the building, often submerging the *audience in darkness as the lights flicker and at times, drop out entirely for short moments during the simulated brownouts.*

[Narrator pauses for a few seconds]

[Sound of neighbors arguing again in distance]

During the pause in narration, another randomized argument can be heard in the distance, this time somewhat masked and abstracted by the increased sound of the storm.

NARRATOR First he turns the bathroom light off.

[Narrator pauses for a few seconds]
[Light off]
[Footsteps]

NARRATOR Then the kitchen.

[Narrator pauses for a few seconds]
[Light off]

NARRATOR She hears him crossing the room.

[Sound of shuffling footsteps.]
[Narrator pauses for a few seconds]

NARRATOR Then the lounge room.

[Light off]

As the killer moves through the apartment, his footsteps shift from the bathroom, to the entrance, and then to the bedroom. Each time the narrator describes the killer turn off a light, the audience simultaneously hears a switch flick from that particular room and the ceiling light is turned off via DMX casting the area into darkness. The combination of the sound of the killer's movement and the physical enactment of turning the lights off provide a comprehensive sense of the virtual presence.

At this point the audience is situated in a predominantly dark apartment; the only light source is a dim desk lamp in the bedroom and diffused light coming from a floor lamp in the narrator's room. Momentary bursts of lightning brightly illuminate the apartment.

[Suddenly the footsteps cease. The only sounds that can be heard are the radio and the never-ending rain. After a pause there is a sudden shuffle towards the front door]

NARRATOR She sat waiting and waiting . . . and waiting, silent . . . in the dark, frozen in terror, unable to move.

[Narrator fidgets with paper and mutters something to herself]

The narrator at this moment realizes the outcome of the story and becomes somewhat unsettled. After recounting the tale of a murder that occurred in the 1950s, she suddenly realizes that the script that she is reading now refers to herself as the protagonist. The events are no longer displaced in time but have transitioned into the present. This shift occurs, beginning with the statement, "She sat waiting and waiting…" Her statement effectively reflects precisely what the audience is doing at that moment in time also, waiting in a dark apartment, uncertain of what will occur. It reaffirms that the narrator coexists with the audience in the present, which to some extent is now both "in the story together" and indicates that any further events that transpire in the narrative are now occurring "live."

NARRATOR She never heard him enter because of the commotion of the storm.

[Loud thunderclap and lightning. The fake door opens and the killer emerges]
[Shadow of killer on wall from lamp flash during lightning]

NARRATOR: The killer slowly approaches from behind, his shadow dancing around the walls in the flickering light, she feels his presence as he draws closer, [Shudders] his hands reaching towards her trembling body. He grasps around her neck, she gasps, rising in an attempt to pry his crushing grip from her neck. They struggle, he throws her towards the wall, the lamp smashes and they are thrown into inky darkness,

As the killer approaches, the narrator can be seen describing her imminent demise. Although the scene is charged with tension due to the developing drama and the environmental mechanics, it is simultaneously whimsical and absurd as a counterpoint to the violence. That is, she is narrating her own death as it is occurring live, still desperately clutching the manuscript in an attempt to continue reading. The struggle is visible on the television and a cacophony can be heard from inside the room. When the narrator crashes through the floor lamp and it topples, a sound effect is played to simulate a breaking globe, and the DMX lighting control disables the lamp. The effect from outside the room is that the audience hears a crash, the final light source emanating from the room vanishes, and the simulated CCTV camera feed to the television drops to black giving the impression that there is not enough light available to capture an image. The only time the audience can see activity occurring in the room via the television is when a series of lightning flashes occur. This is used as a concealing device for the following actions.

NARRATOR: She sees his leering face in the flickering light from the window [Lightning Flash]. She tries to scream, but can't, her body recoiling in horror. Her mind feels hazy. The enveloping darkness is interspersed by flashes of bright light [Lightning]. She feels her life slipping away but desperately struggles [Crash into cupboard]. She never even felt the razor effortlessly sli…

In this sequence, the two characters continue to struggle and the narrator is hurled into the cupboard. During the collision, the cupboard violently rattles to simulate the impact of her body. As discussed in the section "Justification," the sparse visual information created by the lightning flashes was used to conceal the synchronized movement of the cupboard to the corresponding television image. Unlike the previous movement of the cupboard when the narrator opens the door looking for candles, this mechanical movement is substantial and impossible to miss, the entire cupboard crashes into the doorframe. The story abruptly ends when the narrator's throat is cut by the razor mid sentence. This concluded the tale and the presence of the narrator in the experience. Her death is only witnessed in fleeting video fragments on the television during lightning flashes. The result is her lifeless body sprawled on the floor. The remaining inhabitants of the apartment are now the killer and the audience. Due to the prevailing darkness in the space and substantial lack of video information on the television (the screen is predominantly black at this point, momentary images flash up only during sporadic lightning activity), the audience becomes more reliant on sound or physical cues within the room to determine the location of the killer.

[Lighting and rain continues and we see the killer at the cupboard during a lighting flash and it shakes]

A sudden flurry of lightning reveals on the television that the killer is behind the cupboard. He is heard straining to move the cupboard in an attempt to gain access to the room in which the audience is located and the cupboard heaves back and forth. Simulated sounds of the strenuous activity undertaken by the killer combine with the actual creaks and groans of the cupboard physically shaking. The moment can simultaneously be viewed on the television to the right of the audience, the commotion behind the cupboard seemingly revealed via the camera during lightning flashes.

[The killer exits, the apartment is still, the sound of rain continues to sheet down outside the windows and the narrator's body can be seen in a contorted position on the floor during occasional lightning flashes, punctuated by thunder.]

While the killer attempts to get unsettling close to the audience, he fails and can be seen departing through the rear door during the lightning flashes. The time that elapses until the commencement of the second Act (approximately 90 s) leaves the audience waiting in relative darkness in the apartment with a degree of uncertainty as to whether the murderer may appear again or if indeed it is the end of the work as a whole. Act Two demonstrates this is not the case; however, the events are most certainly not what they expect.

Act Two

As the narrative in Act One is set in 1957, the audience is clearly aware that the characters occupying the environment are not ghostly apparitions from the past, and indeed that the overtly dramatic scenario presented to them has no connection to their current reality. The narrative aspect of the work is undoubtedly interpreted as an intricate fictional simulation. There is, however, the possibility of a sense that the characters in this story may have been be played by actors that cohabit the space. For example, the narrator might be considered an actor playing the role of a narrator within an environment that has clearly fictitious elements such as a radio operating of its own accord.

As a counterpoint to this fictional scenario, a second narrative sequence was developed that fluidly occurs after the first. The aim of this Act is to firmly root the virtual characters in the present (in a similar manner to *The Black Box Sessions*). Not only do these subsequent characters exist in the present, they are portrayed as plausible in the reframed context of the act. The audience therefore initially experiences a wholly fictional narrative based partly upon suspension of disbelief and is then presented by a factual situation founded solely on deception. Additionally, this sequence is a means to achieve a smooth and elegant transition for resetting the environment to the default state, ready for the commencement of the first Act.

Two actors under direction from Alex Davies and Andreas Mayerhofer improvized the following sequence, therefore, no script is available. Discussion of this sequence is via descriptive analysis of the video content presented in the work.

When the first Act ends, all the lights in the apartment remain out due to the storm. The audience waits in near darkness – other than a diffused glow of the bedside lamp and the occasional flickering of lightning through the shutters. All that can be heard is the constant patter of rain and the radio still playing in the kitchen. The television is black after the lamp was smashed during the struggle (giving the impression that the "live" CCTV camera is unable to capture an image in the darkness). The apartment is eerily still after the dramatic climax of the previous narrative.

In the darkness, sounds of movement are heard from the narrator's room: could this be the actors leaving?

After 30 s the audience hears the sound of a door opening emanating from within the narrator's room. A crash is heard – followed by cursing – as an individual bumps into an object in the darkness. The sounds of a lighter flicking are also heard. A small wavering pinpoint of light is visible in the blackness of the television image and faint flickering light is seen emanating from behind the cupboard in the room. The lighter goes out; the room is cast into darkness again. Then it reappears. The audience continues to hear shuffling and then the actor calls out to her colleague to come and assist with a flashlight. Moments later another individual is heard entering the room and a flashlight is seen sweeping around on the television screen and from within the physical space of the room, darting across the walls. After a brief moment of activity, the second character locates a switch. The whole apartment is cast into light as power is apparently restored to its natural state. Simultaneously, the television flickers and an image appears of two women standing in the room, both dressed in Linz09 Culture Capital t-shirts. The women begin to dismantle the installation, seemingly under instruction from the management, moving artifacts and props while generally engaged in small talk. During the course of this activity, the attendants show signs of awareness of the audience's presence in the adjacent room; however, they continue about their business.

The attendants continue to talk amongst themselves while working when a mobile phone rings and is answered by Attendant 1. The audience hears half the phone conversation (the attendant's perspective) of what appears to be a misunderstanding *about the exhibition de-install. As this discussion proceeds, the character states that her phone is running out of batteries, and indeed, moments later it cuts out. While this conversation is occurring, the other attendant continues to work and can be seen on the television unplugging a power cable that runs along the ground behind the cupboard. The very instant the cable is unplugged the television flickers to black and the sound of the radio and rain cease. The only sound that can be heard in the apartment now is directly related to the perceived actions of the characters in the rear room, no longer the broader simulated environment. The act of unplugging the cable ends the remaining legacy media elements from the previous Act. It is the final transition in framing context from a split fictional environment, dislocated in time, to one centered, in fact, purely in the present.*

The attendants discuss their predicament when suddenly the phone located in the bedroom near the audience begins to ring. The audience hears one of the attendants approach the cupboard and attempt to move it in order to access the phone, making the cupboard rattle. She calls out to the audience to assist her and answer the phone as she is unable to gain access, despite heavy shaking to dislodge the cupboard. If the audience answers the phone, they are presented with a fast talking anxious manager insisting that the attendants must call him back immediately. He promptly hangs up before the audience has the chance to engage in conversation – as the line goes dead after the flurry of spoken directions. The swiftness of this exchange and the context it is presented in removes any need for an intelligent response from the software to audience input. It is a decisive end to the interactive element, designed to sustain the illusion. The previously mentioned narrative element of the second attendant disconnecting the power (and therefore the television) additionally played the role of a concealing device in the work. As the television still appears to be disconnected during the phone sequence, the audience is unable to perceive any synchronization issues when the attendant shakes the cupboard as they can't simultaneously see the action occur on the television (working in a similar manner to the use of lightning as a device to conceal discrepancies in Act One however appropriate to the new framing context).

Realizing an error has occurred, the two attendants proceed to roughly move all the props in the room to their original location, restore power to the television (and therefore CCTV image), and hastily depart. Moments later the first attendant reenters and turns the light off again in the narrator's room (as it had been when they arrived). This was executed to justify the following audio transition that was required to occur in darkness while at the same time appearing plausible. That is, the attendant wished to return the room to the way they found it as closely as possible due to the misunderstanding.

This completes the second Act. Following this Act is a further 30-s transition through the use of sound in which it appears that people are moving objects around in the darkness of the narrator's room. This was employed for continuity, so that when the first Act commences again, if audience members are still present from the previous Act, the contents of the room do not appear to jump between the two scenarios. Darkness was used as a means of disguise so that the audience could not see what was actually occurring within the room (nothing) and were therefore reliant on sound *cues and their imagination to construct these events. Through this process the installation is returned to its ready state for the commencement of Act One.*

Conclusion

Häusliches Glück demonstrates the use of conjuring principles as a means of developing the illusion of perceptively real characters in mixed realities. This in turn propels the narrative. The work achieves this through the tight integration of a wide range of media to create an immersive polysensory environment. *Häusliches Glück* demonstrates how narratives in mixed reality can enhance conviction in the very same manner as magicians employ showmanship, both with the common goal of advancing an illusion. It further demonstrates how multimodal experiences increase presence and cumulatively aid in deception.

The installation illustrates the principle that multiple framing contexts can be integrated into a single cohesive work, thereby dynamically shifting the nature of the illusion and audience experience. On one level of the illusion, the audience is encouraged to suspend disbelief through the cinematic fictional narrative they find themselves immersed in. On another level, the environmental elements indicate that humans physically present in the space are playing out the very same story. The audience is finally presented with the last layer of deception occurring in Act Two whereby they are transported again to the present and encounter a relatively familiar everyday situation.

All these permutations of illusion are achieved through the same methods used to execute conjuring tricks: the unified integration of mechanics and showmanship. Unlike magic, however, where the audience is expecting to be deceived, mixed realities have the potential to occur within a wider range of framing contexts. The audience in these circumstances may not expect deception to transpire. Even when a particular aspect of the deception fails (as it often does in these circumstances), each new illusory element adds to the uncertainty and makes the audience continually question what is in fact real or virtual within their surroundings. It is perhaps this uncertainty above all that creates tension and drives audience interest and ongoing engagement with the work.

Recommended Reading

J. Anderson, *The Reality of Illusion : An Ecological Approach to Cognitive Film Theory* (Southern Illinois University Press, Carbondale, 1996)

P. Lamont, R. Wiseman, *Magic in Theory: An Introduction to the Theoretical and Psychological Elements of Conjuring* (University of Hertfordshire Press, Hatfield, 1999)

Linz09, Haus der Geschichten/House of Stories (European Capital of Culture, 2009), http://www.linz09.at/en/projekt-2205346/haus_der_geschichten.html. Accessed 16 Jan

H. Nelms, *Magic and Showmanship: A Handbook for Conjurers* (Dover Publications, New York, 1969)

Entertainment, Culture, and Media Art 30

Ryohei Nakatsu, Naoko Tosa, Matthias Rauterberg, and Wang Xuan

Contents

R. Nakatsu (✉)
Design School, Kyoto University, Kyoto, Japan
e-mail: nakatsu.ryohei@gmail.com

N. Tosa
Academic Center for Computing and Media Studies, Kyoto University, Kyoto, Japan
e-mail: tosa@media.kyoto-u.ac.jp

M. Rauterberg
Industrial Design, Eindhoven University of Technology, Eindhoven, The Netherlands
e-mail: G.W.M.Rauterberg@tue.nl; ifip.tc14.chair@gmail.com

W. Xuan
Rolls-Royce, Singapore, Singapore
e-mail: wangxuan927@gmail.com

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_16

Abstract

This chapter is dedicated to explore the relationship between entertainment, culture, and media art. Firstly, the positioning of entertainment will be described including the historical point of view and also focusing on recent digital entertainment technologies. It will be clarified that entertainment is closely related to the mental sustainability of human and therefore has a close relationship with culture and art. Then, based on this, further investigation will be carried out to explain how culture and technology could be combined together. Firstly, the relationship between culture and technology will be discussed proposing the concept of "cultural computing." Then three typical research examples of cultural computing, "ALICE," "ZENetic Computer," and "Confucius Chat" will be described emphasizing how cultural issues could be treated using computer both in the Western culture and Eastern culture. As art is the core part of culture, as a next step the role of art, especially new art called media art where technology plays an important role, will be discussed. And, as an example, media art created by an Eastern media artist will be introduced as a good example of such media art treating cultural issues.

Keywords

ALICE • Alice in Wonderland • Cultural computing in the West • Advice from caterpillar, stage • Cheshire cat, stage • Eat Me and Drink Me, stage •

Evaluation • In park, stage • In rabbit hole, stage • Pool of tears, stage • Art • Versus entertainment and culture • Buddhist human recognition model • *Godai* • Personality recognition model • Zen dialogue interaction • Chat. *See* Confucius Chat • Communication • Global Village • Confucius Chat • Eastern (Chinese) view • Evaluation • iSage application • Methodology • System description • Cultural computing • ALICE • Confucius Chat • ZENetic Computer • Cultural heritage • Culture • Cultural heritage • Definition • Intangible culture • Tangible culture • Versus entertainment and art • Western • Entertainment • And human life • Digital games • Origin of • Renaissance • Versus culture and art • Waste of time • Four God Flags • Background Sansui image • Korean dragon • Of Goguryeo Tombs and animation • Oracle bone script animation • Visitor's response • Intangible culture • Content archiving and retrieval • Content learning • Cultural computing • Technospiritual research • Media art • And culture • Cultural issues • Definition • History • Yeosu Expo 2012 • Tangible culture • Western culture • Yeosu Expo 2012 • Expo Digital Gallery (EDG) • Selection process • Zen. *See* ZENetic Computer • ZENetic Computer • Cultural computing in Eastern (Japanese) view • Interaction control via chaos • Interactive Buddhist human recognition model • Interactive story generation • *Sanen* design • *Sansui* painting creation • Story experience within Sansui Space

Introduction

Network, information, and media technologies are rapidly changing our society, including human relationship, lifestyle, and communication. Entertainment is the area where these new technologies have a strong influence (Nakatsu 2009). One good example is games. Playing computer/video games, or in other words digital games, is a common daily activity for people, especially for young generation. In particular, smartphone-based casual games have become very popular in both the United States and Asian countries (Juul 2012). Another good example is communication. People communicate with their families and friends through e-mail, mobile phones, texting, Twitter, and other means. Until the 1980s, communication media such as telephones were mainly used for business communication. Today, however, communication extends beyond business conversations and has become a form of everyday entertainment. It looks that “Global Village” proposed by Marshall McLuhan has been already realized (McLuhan and Powers 1992; McLuhan 2011).

There are many discussions regarding these phenomena, but most of them merely observe what is happening in society and report the fact that more people are spending more time enjoying these new forms of entertainment. Unfortunately, however, there has been little consideration regarding the basic reason why these new forms of entertainment have been accepted by people all over the world. One of the fundamental questions is: what is the main role of entertainment in our society and everyday life (Nakatsu 2009)?

The thought that culture, especially art as a core part of culture, is an essential part of our life has been commonly accepted. As entertainment has generally been understood to mean digital games, however, the relationship between entertainment, culture, and art has little been considered and discussed. As various aspects of entertainment, especially from technological point of view, are discussed and described in other sections of this handbook, mainly in this section, the focus will be put on the relationship between entertainment and art. And in this chapter, culture with its broad meaning including art as its core part will be discussed from a point of entertainment and technologies.

In this chapter, firstly, the relationship between technology and culture will be discussed. As a first step, the definition of culture will be clarified and then how culture could be treated by computer will be discussed. Based on this introductory discussion, a new concept called "cultural computing," which is a new research area trying to treat various cultural issues by computer, will be proposed. Then three typical examples of cultural computing will be described and discussed. One example, ALICE (Hu et al. 2008; Bartneck et al. 2008), was carried out in the Western country, and the other two examples, ZENetic Computer (Tosa et al. 2005; Tosa and Matsuoka 2006) and Confucius Chat (Khoo et al. 2011; Wang et al. 2013), were carried out in the Eastern countries. Therefore by comparing these three researches, it would be interesting to notice the similarities and differences on how cultural issues are treated in the West and East.

Secondly, the relationship between art and culture focusing on how technology would work to connect these two will be discussed. Recently new art called media art, where technology plays an important role, has become an important part of fine art. Although there was a tight relationship between culture and art in old days, recent media art has strongly been influenced by technologies, the expression form of which looks more modern and cool. Therefore the relationship between culture and art has become not so explicit. As a first step, the recent trend of media art will be briefly discussed. Then, an example of media art where cultural issues are richly expressed by using technologies will be described (Tosa et al. 2013).

Entertainment

Are People Accused of by Enjoying Entertainment?

It has been commonly understood, especially by the advance of digital technologies, that entertainment generally means digital games such as video game, computer game, etc. In developed countries, new types of entertainment have emerged, such as chat on mobile phones and games on game machines and PCs. People in those countries are tending to spend more and more time enjoying such forms of entertainment. Established entertainment industry is huge and includes the movie, game, sports, and other businesses. There have been significant concerns and complaints about this trend. The basic logic of such complaints is that compared with other human activities, such as education, business, industrial production, and so on, entertainment is not productive. In other words, the complaints suggest that entertainment is only a waste of time.

On the other hand, at least one billon people currently face starvation all over the world. For these people, it is crucial to obtain food for tomorrow or even today. This problem has been one of the most serious topics at global conferences and meetings such as World Economic Forum (Dutta and Mia 2009).

An important question, however, is whether people who live in developed countries should be accused of by enjoying entertainment neglecting the situation in underdeveloping countries. Another question is why the demand and markets for such entertainment are so huge. The issue of whether entertainment is a waste of time clearly requires care.

Origin of Entertainment

What is happening now is not a totally new phenomenon. In earlier times, human life was simple. Humans farmed or hunted to survive. When people were not occupied with these tasks, they entertained themselves by various means. In other words, it can be said that food is strongly related to our physical sustainability, yet at the same time, it can be pointed out that entertainment is related to our mental and spiritual sustainability.

Then the era of civilization began. People introduced various novel types of activities, such as art, business, learning and teaching, and religion. Because of these activities, entertainment came to be considered as a secondary activity in human life. Although entertainment remained a certain part of human everyday life, it has not been considered an essential part. Figure 1 illustrates these changes in human physical and mental sustainability (Nakatsu 2009).

Human Life and Entertainment

Even now sometimes the fact that entertainment is an important part of our life becomes apparent. Consider passengers on an airplane. During a flight, most people sleep, eat, or entertain themselves by watching movies, reading novels, and so forth. Only a few people work during a flight. This means that in a simple situation, our lifestyle consists of three basic activities: sleeping, eating, and entertaining ourselves (Nakatsu 2009).

What is happening now is, in one sense, an “entertainment renaissance.” The introduction of new technologies, especially interactive technologies, into traditional forms of entertainment has totally renewed and strengthened those forms. People are again noticing the basic strength and meaning of entertainment and recognizing that it is a substantial part of their lives. This is a key point in trying to understand such substantial issues as the role of entertainment and the future direction of entertainment (Rauterberg 2006).

Entertainment, Culture, and Art

As is indicated in Fig. 1, there is a tight connection between entertainment, culture, and art. Although more detailed discussion will be carried out later, culture is easily understood something that covers the whole diagram shown in Fig. 1. Also art is an essential core part of culture. As has been described before, by the recent advances in IT, new types of entertainment such as various types of digital games and new communication forms have appeared. This means that it is important at this point to reconsider the relationship between entertainment, culture, and art. As has been understood, entertainment covers wide area of human everyday activities. Therefore by keeping in mind that entertainment has a tight relationship with culture and art, in the following chapters and sections of this chapter, it will be discussed how culture and art could be combined with technologies.

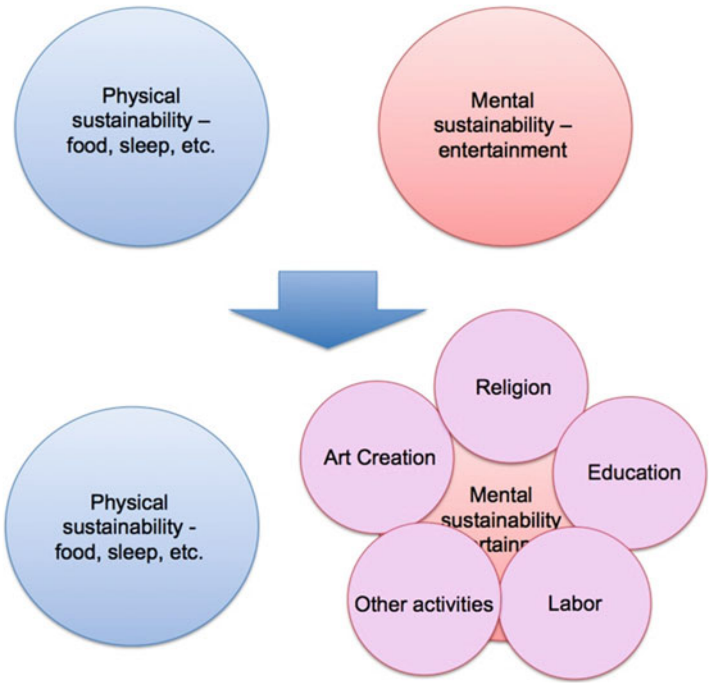


Fig. 1 Changes in the physical and mental sustainability of human beings

Culture

Definition

In many ways, computing has transformed the way people preserve and disseminate culture. Scientific methods have helped in various stages in the identification, preservation, authentication, and retrieval of cultural contents, not only in the form of digital representation and recreation of physical cultural assets as online repositories but also creating new forms of interactions with them. Culture, as a general term, can be divided into two categories: tangible culture and intangible culture. The UNESCO World Heritage Centre defines cultural heritage as "the legacy of physical artifacts and intangible attributes of a group or society that are inherited from past generations, maintained in the present and bestowed for the benefit of future generations" (Eriksen 2001). In this section, the authors provide a review on how computing technology is used for these two types of culture.

Tangible Culture

Tangible culture, or material culture as termed in social sciences, refers to culture that is tangible or touchable, such as sculptures, architectures, paintings, etc. Much of these tangible cultures are considered worthy of preservation for the study and appreciation of human history and society. Computing methods have been utilized to help achieve this goal. One major research area is digital cultural heritage, which specifically concerns with digital recreations of historical sites or artifacts using 2D or 3D modeling techniques or the management of such resources (Cameron and Kenderdine 2007; Koller et al. 2009; Levoy et al. 2000). For example, Okada et al. used advanced analytic techniques to restore the historical scripts and illustrations to its original form (Okada and Shoji 2010), and the Virtual Kyoto project built geotemporal referenced 3D models of the entire city of Kyoto (Nakaya et al. 2010). Some of the models built are accessible online as e-museums (Shiaw et al. 2004), where people can easily browse, retrieve, and explore the tangible culture from a computer interface. Furthermore, interactive installations have been set up in physical museums, allowing easy access to the information in ways such as augmented reality (Choudary et al. 2009), natural gestures (Alisi et al. 2010), etc.

Intangible Culture

Intangible culture, such as oral traditions, customs, ways of life, traditional craftsmanship, performing arts, theater, social practices, etc., represents cultural identities and diversity of mankind and is also an important part of culture. Modern computing and interaction techniques bring new possibilities for its preservation and promotion, allowing not only easier access and retrieval of these intangible cultural contents but also participation and learning through a simulated or recreated experience.

Archiving and Retrieval of Intangible Cultural Contents

Like the work on tangible culture, many of existing works on intangible cultural contents focus on archiving and browsing, where the content structuring and accessibility are the major concerns. For example, Mallik et al. (2011) created a repository that encoded domain knowledge with digitized records of Indian dance and music using an ontology-based approach. A browsing application is developed for semantically access the collection of Indian dance videos. Stavrakis (Stavrakis et al. 2012) created a public digital archive of video recordings and motion captures of Cypriot folk dances. In addition, a 3D video game was created for children to be more aware of their dance heritage. The Bulgarian Folklore Digital Library (Pavlov et al. 2010) offers a web-based platform for the documentation and access to a range of Bulgarian folklore objects. Similarly, Artese built a web-based register for the intangible cultural legacy of the oral history of the Italian Lombardy territory (Artese and Gagliardi 2012).

3D modeling is usually used for virtual recreation of tangible cultural artifacts, but it can also be applied for intangible culture such as traditional festivals and ceremonies. For example, Choi et al. work on the digital reproduction of the Yamahoko parade in the Kyoto Gion Festival through motion data of the oats and crowds (Choi et al. 2010), which allows people to have a realistic view of how the parade is like during the festival.

Learning of Intangible Cultural Content

Hiyama explored new methods for archiving and transferring of craftwork skills. They built a wearable display system that can be used for the skill training on Kamisuki, Japanese traditional papermaking (Hiyama et al. 2011), replacing the traditional skill transfer with word of mouth between a master and apprentices. It provided first-person visual and audio information and biological information (myoelectric signals, breathing, and the eye gaze) of a craftsman. Similarly, calligraphic skills can be trained through system that records an expert's movements and reproduces them through haptic devices (Saga et al. 2005).

Another work on calligraphy is by Xu (Xu et al. 2009), where an intelligent system was developed to learn the style of a particular calligraphist and automatically generate new calligraphy of the particular calligraphist.

Technospiritual Research

Social values and beliefs are also an important component of the intangible culture. People from different cultural backgrounds inherently possess distinct way of thinking, feeling, and worldviews. However, the use of technology for them has been less explored, compared to other cultural contents described above. For example, despite the high-profile calls from Intel anthropologist Genevieve Bell to the HCI community for using technology in spirituality and religion (Bell 2004, 2006), research in this area is still rather sparse. Buie and Blythe (2013) discussed the importance of supporting spirituality and religious practices in the HCI community and proposed several reasons for the lack of research in this area. Although in their paper the term "spiritual" is used in a loose sense, referring to a sense of deep connection to something larger than oneself, whatever that may be (a deity, nature, a cause, the Universe, an intimate group, etc.), and covering all user experiences that can be described as spiritual, religious, transcendent, or numinous, and thus it perhaps does not fully fit into the scope of intangible culture, the authors believe that the spirituality they talked about shares many common characteristics with intangible culture, and one of the important reasons for the lack of attention on them is probably the elusive nature that is hard to quantify and convey digitally and interactively.

Most technospiritual research mainly focuses on facilitating the spiritual practices using technology (e.g., using ICT for communicating religious messages (Bell 2006), using telepresence to help physically isolated people mediate together (Hlubinka et al. 2002), using sacred imagery in a mobile application to support Islamic prayer practices (Wyche et al. 2009), and a special display device to suggest topics for prayers (Gaver et al. 2010)), rather than embedding the cultural content in the interaction itself and creating new experience. The latter approach requires analysis and understanding of the particular cultural contents to reproduce them in the interaction, and this can be called cultural computing research.

Cultural Computing Research

Cultural computing concerns about not only integrating cultural aspects into interaction but also allowing the user to experience the core elements of the culture through the interaction. The target of cultural computing is to extract a basic structure underlying a specific culture and to develop systems that would let people understand this basic structure through interactive experiences. Some work seeks to digitally model the intangible cultural contents and allow people, especially novices, to be able to participate in the creation process of the particular cultural content. For example, Yao et al. (2010) digitized shadow play, a traditional art performance in China, using multipoint interactive control, which can be easily manipulated by a performer with no prior experience. Hitch haiku (Tosa et al. 2008; Xu et al. 2009) is an interactive system that supports the composing of haiku, a Japanese minimal poem form. Using this system, users can choose arbitrary phrases from a repository, and the system generates the haiku based on the essence of words in the phrases. Similarly, Poetry Mix-up (Zhu et al. 2011) is a poetry creation system which models the words and structure of poetry. Users can get the system to generate their own poetry by simply sending an SMS to the system.

Another representative example is the project ALICE, which is an interactive installation that enables users to be immersed in a mixed reality environment featuring the story of the book Alice in Wonderland by Lewis Carroll. Scenarios are selected and implemented as an interactive experience composing six stages, in each of which the user gets immersed in the installation, experiencing and rejecting on the Western culture determinant (Hu et al. 2008; Rauterberg 2006). Another work is the Mysterious Machines, an art installation including three physical robots autonomously talking to each other as well as to the audience about religious issues

(Schonenberg and Bartneck 2010). It uses the New Testament, the Koran, and Rigveda as the knowledge base.

For the Eastern culture, Tosa created the ZENetic Computer (Tosa et al. 2005; Tosa and Matsuoka 2006), which projects the style of Zen communication into computing world, letting people experience Zen Buddhism through Sansui ink painting and construction of one's own virtual space. The system then generates a fragmented Zen story based on user's design of the virtual space and poses Zen questions to the user. Through this interactive experience, users are led to confirm his or her self-awareness, consciousness, and unconscious imagination.

Also about the Eastern culture, Uriu and his colleagues looked at how entertainment computing can be used to support Buddhism practices (Uriu et al. 2012). They use augmented reality technology to help users learn about Todaiji temple, one of the world heritages in Nara, Japan, and a special ritual ceremony held at the temple. Inspired by these works, there is another attempt to further cultural computing research by looking into Chinese philosophy and studying how to provide an interactive experience for the general public (Wang et al. 2013).

In the following chapters, three typical examples of cultural computing research will be described.

ALICE

Cultural Computing: Western View

The word culture (from the Latin "colo, −ere," meaning "to cultivate," "to inhabit," or "to honor") has been defined and used in many ways throughout different contexts. Kroeber and Kluckhohn (1952) compiled a list of more than 156 different definitions for culture. One of the most popular definitions of culture in the field of anthropology is "a complex web of shifting patterns that link people in different locales and that link social formations of different scales." Culture is the integration of human behavior that includes attitudes, norms, values, beliefs, actions, communications, and groups (ethnic, religious, social, etc.). Cultural computing is more than integrating cultural aspects into the interaction. It is about allowing the user to experience an interaction that is closely related to the core aspects of his/her culture, in a way that let him/her engage with an augmented reality using the values and attributes of his/her own culture. As such it is important to understand one's cultural determinants and how to render them during the interaction. In this chapter, the focus will be a cultural computing project from the Western world (prevailing in England).

Over the last 3,000 years, the peoples of four distinct regions of the civilized world created the religious and philosophical traditions that have continued to nourish humanity into the present day: Confucianism and Daoism in China, Hinduism and Buddhism in India, monotheism in the Middle East, and philosophical rationalism in Greece. "Monotheism" and "philosophical rationalism" are the religious and cultural foundation of the occident. Here it is adequate to investigate illustrative stories that are well known, accessible, classical in their culture and relevant from the point of view of cultural computing. It is required to primarily look for narratives that would be helpful in the understanding of the essential aspects of English culture. To this effect, as one example, “Alice in Wonderland” by Lewis Carroll (1865) has been selected and will be described in this chapter. Another example is a story of “Zen Buddhism” attributed to a Bodhidharma (circa 500 AC), which will be treated in the next chapter. Both are examples either to help understand the underlying cultural value (i.e., Zen) or question it (i.e., Alice). For the Eastern and Western culture, the main value dealt with is enlightenment but in different ways. Utilizing on modern technology, Nakatsu et al. (2006) try to give a new direction in form of “Kansei Mediation” to enable societies transforming toward enlightenment (see also Rauterberg 2004). Salem et al. (2005) discuss the relationship of cultural computing and entertainment, and Hu and Bartneck (2005) could conclude that “culture matters.”

Western Culture: Alice in Wonderland

In the West, Kant (1784) gave an answer to the question, “What is enlightenment?” He indicated that the “way out” that characterizes enlightenment in the West is a process that releases us from the status of “immaturity,” and by “immaturity,” he meant a certain state of our will that makes us accept someone else’s authority to lead us in areas where the use of reason is called for. In the Western world, enlightenment is defined by a modification of the preexisting relation linking will, authority, and the use of reason.

Nisbett (Nisbett et al. 2001) can confirm that Westerners are analytic, paying attention primarily to the object and the categories to which it belongs and using rules, including formal logic, to understand its behavior. In contrast East Asians are more holistic, attending to the entire field and assigning causality to it, making relatively little use of categories and formal logic and relying on “dialectical” reasoning. These Western and Eastern types of cognitive processes are embedded in different naive metaphysical systems and tacit epistemologies. Nisbett et al. (2001) speculate that the origin of these differences is traceable to markedly different social systems as part of the underlying cultural determinants.

To address logical reasoning in the Western culture, the most appealing narrative is “Alice in Wonderland” of Lewis Carroll. Charles Lutwidge Dodgson (1832–1898), better known by the pen name Lewis Carroll, was a British author, mathematician, logician, Anglican clergyman, and photographer. His most famous writings are “Alice in Wonderland” and its sequel “Through the Looking-Glass.” His facility at word play, logic, and fantasy has delighted audiences ranging from the most naive to the most sophisticated. He was exceptionally gifted and achievement came easily to him. His works have remained popular since they were published and have influenced not only children’s literature but also a number of major twentieth-century writers such as James Joyce and Jorge Luis Borges. There are societies dedicated to the enjoyment and promotion of Lewis Carroll’s works in many parts of the world including North America, the United Kingdom, and

New Zealand. In this perspective, the book "Alice in Wonderland" can serve as input for a cultural computing project in the West.

The first interactive, but semi-immersive virtual reality system based on parts of "Alice in Wonderland" was developed at the Entertainment Technology Center of Carnegie Mellon University. Pierce (Pierce et al. 1999) created a successful virtual experience based on a head-mounted display to overcome some or all of the following problems: entering a virtual world is a jarring experience, people do not naturally turn their heads or talk to each other while wearing a head-mounted display, putting on the equipment is hard, and people do not realize when the experience is over. In the Electric Garden at SIGGRAPH 97, they presented the Mad Hatter's Tea Party, a shared virtual environment experienced by more than 1,500 SIGGRAPH attendees. They addressed these head-mounted display-related problems with a combination of back story, see-through head-mounted displays, virtual characters, continuity of real and virtual objects, and the layout and setting of the physical and virtual environment.

The cultural computing project ALICE was started as an interactive, entertaining experience (see Nakatsu et al. 2005) inspired from "Alice in Wonderland." In the scope of this project, interactive adventures are experiences provided by an augmented reality (AR) environment based on selected parts from Lewis Carroll's book "Alice in Wonderland." The user assumes the role of Alice and explores this interactive narrative.

ALICE is an exploration of interactive storytelling in AR. By exploiting the unique characteristics of AR compared to established media such as film and interactive media, the project uses AR as a new medium for edutainment and entertainment as a particular carrier for cultural transformations. Innovations include the refashioning of conventions used in film and interactive tools for the development of an AR narrative, and the use of simple artificial virtual and real characters (avatar and robot, respectively) to create an immersive interactive experience.

ALICE is an augmented reality (AR) narrative with intelligent agents acting as characters who lead the user through virtual and real locations, moral choices, and emotional states. The narrative is a surreal quest, sometimes funny, sometimes disturbing. The character White Rabbit (representing the concept of time) introduces him and joins with the user in a series of absurdist challenges. ALICE is an educational journey toward the user's heart's desire, designed to provoke self-reflection on a number of other issues: bullying and trusting others, selfishness and selflessness, enjoying the moment, or sublimating pleasure. The user is given the opportunity to occupy and experience any of these mental and emotional positions. This will be achieved in line with the "Alice in Wonderland" plot (albeit shortened).

Alice in Wonderland can be used to give interesting examples of many of the basic concepts of adolescent psychology. Alice's experiences can be seen as symbolic depictions of important aspects of adolescent development, such as initiation, identity formation, and physical, cognitive, moral, and social developments (Lough 1983). Alice's adventures are DE constructivist in nature and as such are directly challenging the strongly held belief of a linear, single track, and sequential reality.

ALICE: Implementing Cultural Computing in the West

The approach of this project is to create an interactive experience based on the cultural values (e.g., highlighted in the story "Alice in Wonderland"). A Westerner would understand and appreciate the implementation of an interaction inspired from "Alice in Wonderland" adventures or "Le Petit Prince" but will be puzzled if he/she was presented with an interaction inspired from the Eastern "Ox story" or "The Journey to the West." Similarly, an Easterner would appreciate the second set and be confused by the first. Alice's adventures are illustrative of English culture. Indeed English and Western cultures in general are based on monotheist religions (Judaism, Christianity, and Islam) which are concerned with certainty and absolutism, in the sense of absolute truth and certainty. Western culture is also based on Cartesian logic and a linear and constant flow of time. To understand the culture shock that our proposed interactive experience could generate, it is interesting to look at the book "Alice in Wonderland." Alice, who is tired of the rational world she lives in and therefore follows the white rabbit into a hole, leading to a world without rational boundaries, experiences several culture shocks in this new world. Alice's adventures happen in a world of paradox, the absurd and the improbable. The key aspects of Alice in Wonderland can be resumed in the following points: (1) a nonlinear, nonconstant time flow, (2) a distortion of experience in space and with other characters, and (3) a counterintuitive, commonsense defying heuristics.

To be able to investigate the effects of cultural computing on the user's experience, an interactive installation ALICE was build inside of the W-Hal building of the Eindhoven University of Technology. It takes a space of 12 m by 12 m spanning two floors over a total height of 7 m.

The ALICE installation consists of six consecutive stages. The visitor walks by himself/herself through them, one after the other. Given the Western Culture's emphasis on individuality, the installation is designed for individual experiences. Figure 2 shows a schematic drawing of the installation. Notice the panoramic room on top, the spiral rail for the electric seat, and the Cave in front. The function of each part will be explained in more detail within each of the corresponding stages. For each stage, the corresponding section of Carroll's original story is firstly summarized, followed by an impression of the design and the implementation (Table 1).

Stage 1: In the Park

In the first scene of Carroll's original book, Alice is bored and sleepy sitting on a bank with her sister. A white rabbit runs close by her, and while looking at its pocket watch, it cries out "Oh dear! Oh Dear! I shall be late!"

Stage 1 symbolizes the concept of time. The visitor of the ALICE installation enters this stage through two curtains. Inside, a 360° panorama picture of Cambridge surrounds the visitor, the place where Carroll lived and worked. The panorama depicts a location in nature at the River Cam (Fig. 3).

The visitor is told to first fill in a questionnaire at a computer that is located on top of a tree trunk. Afterward, the visitor is left to his/her own devices. The visitor has to wait until he/she is bored and thereby enters a similar state of mind as Alice in

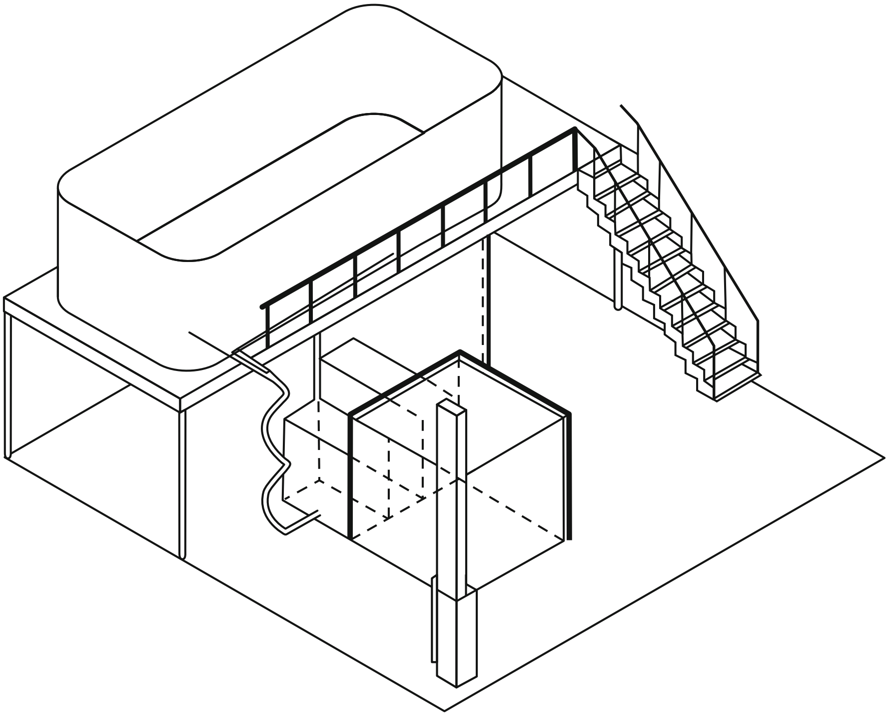


Fig. 2 Schematic drawing of the Alice installation

Table 1 Alice's adventures in six stages

Stage	Theme	Experience	Measurement
In the park	Time	Boredom and curiosity	Retrospective duration
Down the rabbit hole	Environmental space	Disorientation	Presence
Eat Me and Drink Me	Personal space	Shrinking and growing	Presence
The pool of tears	Genesis	Swimming, out of water	
Advice from a caterpillar	Self	Questioning the concept of self	Personality, self-esteem, and self-concept
The Cheshire cat	Logic	Challenged by the logic reasoning	Generating alternatives

the original study. Then, a white rabbit appears from behind the curtain. It has a pocket watch and says through its internal speakers "Oh dear! Oh Dear! I shall be late!" before disappearing into a rabbit hole at the far end of the room. On its way pass the visitor, it tries to seduce the visitor to follow by starting a catch-me-if-you-can type of behavior. The rabbit enters a hidden section inside of the rabbit hole

Fig. 3 A visitor follows the rabbit

before the visitor enters the hole and hence disappears from sight. The visitor enters the rabbit hole after the rabbit and thereby enters stage 2.

The panorama picture was printed on a large fabric. Several lights behind the fabric lighten the room, thereby giving the panorama the brightness of a summer day. The rabbit was build using a carbon fiber fabric. The model is mounted on top of a remote-controlled car. A wide-angle video camera mounted in the ceiling allows the experimenter to see the whole room and thereby enable him to efficiently control the rabbit with the remote control.

Stage 2: Down the Rabbit Hole

Alice falls down the rabbit hole in an unusual fashion. The speed of her fall is very slow and she is able to interact with objects that are attached to the walls of the rabbit hole, such as cupboards and bookshelves. The fall takes such a long time that Alice engages in a conversation with herself before dozing off. At last, she lands on a heap of sticks and dry leaves. Stage 2 can be characterized by environmental space. Alice is discussing whether she can fall right through the earth and come out in New Zealand. She disputes the relationship of herself within the space of the earth.

When entering the rabbit hole in the ALICE installation, the visitor finds an electric seat that is mounted on a rail. Once safely seated, the seat takes the visitor down a tunnel in spiral movement (Fig. 4). Along the walls of the tunnel, cupboards, bookshelves, and lamps are mounted. The speed of the electric seat is slow.

The seat was build using an electric stair lift, commonly used in the houses of elderly to enable them to move from one floor to another. A securely controlled gate prevents the visitor to pass by the seat, thereby endangering him/her to truly fall down the rabbit hole. The seat is only activated once the armrests are put down, which prevents the visitor from falling off while moving. Infrared cameras are placed inside the rabbit hole to allow the experimenter to monitor the visitor.

Stage 3: Eat Me and Drink Me

Alice enters a dark corridor with many doors, which are all locked. She approaches a glass table on which a small golden key lays. She uses the key to open a tiny door

Fig. 4 A visitor descents through the rabbit hole

that leads to a garden, but Alice is too tall to enter. She approaches the table again, and this time she notices a bottle labeled "Drink Me" and later a little cake labeled "Eat Me." By drinking from the bottle, she shrinks, and by eating the cake, she grows. Eventually, she manages to have the appropriate size to enter through the tiny door.

Similar to stage 2, stage 3 is associated with the concept of space. Clearly, the authors could not grow or shrink the visitors directly, but they were able to manipulate their relative size in comparison with the environment by using a cubic Cave. The walls and the ceiling were made of white semitransparent material. The authors used the back-projection method to project a seamless virtual environment onto the walls of the Cave. Figure 5 shows a scale model of the specially designed Cave. A sliding side is connected to the entrance and the exit tunnels, enabling a five-side full projection when the visitor is in the Cave. Both the entrance and the exit tunnels are also build as a three-side projection Cave, but for the time being, only the exit tunnel is used for stage 4 (the pool of tears).

The visitor entered the Cave and had the impression to stand in a virtual room (Fig. 5d). A cookies box labeled "Eat Me" and a bottle labeled "Drink Me" are placed on top of a small table (see Fig. 6). When the visitor drinks from the bottle, the virtual room enlarges, giving the impression that the visitor is shrinking. When eating the cookie, the virtual room shrinks, giving the visitor the impression that he/she is growing.

On one side of the room, a door is shown. Once the visitor reaches the appropriate size, the wall on which the door is shown moves aside and thereby allows the visitor to enter stage 4.

The Vizard VR Toolkit from WorldViz was used to synchronize the five projectors. The model of the virtual room was developed in 3D Studio Max. The floor of the Cave is equipped with pressure sensors that allow us to determine the visitor's position in the Cave. Depending upon his/her location, the perspective of the projection is adjusted to give a true 3D impression of the virtual room. The bottle features touch and tilt sensors to detect drinking. The cookie box is equipped with a microphone that allows us to detect the visitor's chewing sounds when eating the cookie.

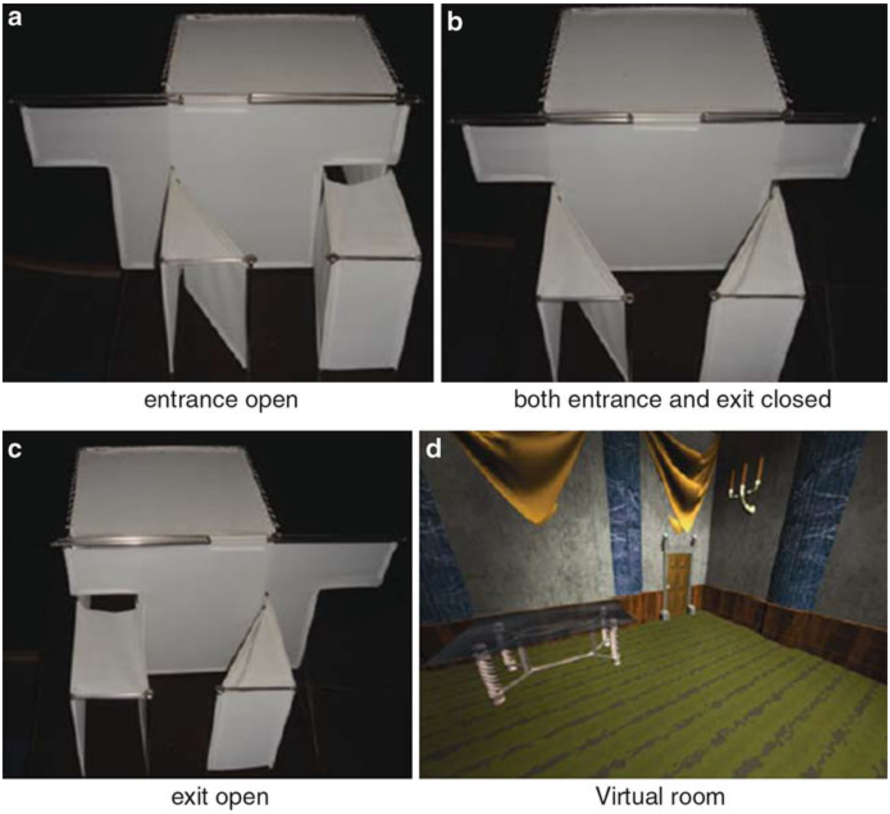


Fig. 5 Scale model of the Cave (**a**–**c**) and the virtual room (**d**)

Stage 4: The Pool of Tears

During her growing and shrinking experience, Alice cries many tears. When she walks through the tiny door, she enters a pool of tears. She talks to a mouse that swims alongside her, and together with some other animals, they finally reach the shore. When the visitor enters stage 4, the sea with a mouse in it is projected on one of the walls (see Fig. 7). In addition, a smoke machine creates an impression of moisture and adds a mystical feeling to this stage. The visitor walks along the projection and thereby enters stage 5.

At the beginning, the authors were discussing building a real pool, but it was not clear what the visitors would do with their clothes. Hence, the authors created a virtual pool of tears with the use of back projection.

Stage 5: Advice from a Caterpillar

The caterpillar engages Alice in discussion about herself. It asks Alice who she really is and how she wants to be. The complete episode is a dialogue, and hence, the authors created a robotic caterpillar to engage the visitor in a similar dialogue (Fig. 8).

Fig. 6 Eat Me and Drink Me (*left*), visitor eating the cookie (*right*)

Fig. 7 Pool of tears

Microphones recorded the utterances of the visitor, and a simple dialogue system manages the conversation. Since most of the questions are metaphysical or mystical, a conclusive dialogue can be created without an extensive artificial intelligence (AI) for the caterpillar.

Fig. 8 A visitor talking to the caterpillar

This stage can be associated with the concept of the self. Throughout history, there have been wide varieties of theories about the self, coming from the fields of philosophy, psychology, and religion. This includes assertions that there is no self; that the idea is a logical, psychological, or grammatical fiction; that the sense of self is properly understood and defined in terms of brain processes; that it is merely a constructed sociological locus or the center of personal and public narratives; or that it belongs in an ineffable category on its own (Gallagher and Shear 2000). There is a significant difference in the construction of the self when comparing European/American culture and the Japanese culture (Kitayama et al. 1997; Nisbett and Masuda 2003). By addressing the Western individual self-concept, Alice's self is challenged in "Advice from a Caterpillar." After she entered the rabbit hole to follow the White Rabbit, she experienced a lot of transformations both physically and mentally. This brought her in an initial state of confusion, which is emphasized in her conversation with the Caterpillar: "Who are YOU?" This challenging attitude of the caterpillar makes Alice uncertain about herself, becoming vulnerable and open for persuasion (Fogg 2003). Such a situation gives the possibility for a confrontation with and stimulates awareness of the self-concept. The character symbolized as a caterpillar is well chosen. One of the most important characteristics of caterpillars and butterflies is their unique life cycle. One of nature's most mysterious metamorphoses occurs when a caterpillar changes from a slow-moving, fat, and ugly creature to a colorfully winged, beautiful butterfly. This metamorphosis happens to a lot of insects, but not as dramatically as it does to a butterfly (Heiligman 1996). In this respect, the "caterpillar" character can unconsciously pursue the message to a human user not to be afraid of a fundamental metamorphosis in his or her self-concept. This symbolic meaning can counterbalances the challenges intended by a conscious dialogue next to it.

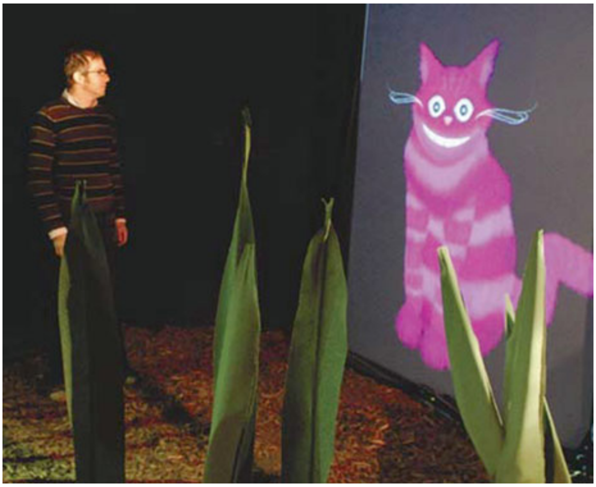


Fig. 9 A visitor talks to the Cheshire cat

Stage 6: The Cheshire Cat

The Cheshire cat involves Alice in a dialogue about logical reasoning and madness. During the dialogue, the cat disappears at times completely and sometimes only its grin remains visible. This episode of Carroll's book is predominantly a dialogue. In contrast to the previous stage, the authors did not use a robotic character, since the transformations of the cat are impossible in reality. Therefore, the authors created a virtual Cheshire cat that is projected on a screen (Fig. 9). When the visitor approaches the cat, it engages him/her in dialogue similar to the one in Carroll's book. A microphone was hidden next to the screen to record the visitor's responses. A simple 2D animation tool is used for the animation of the cat and the dialogue management. Similar to the Caterpillar, most of the questions are metaphysical or mystical, and hence, a conclusive dialogue can be created without an extensive artificial intelligence (AI).

Evaluation

It is likely that the Alice installation will have an effect on its visitors, but it is unclear what effects and how strong they may be. The visitors' experience may range from being mildly entertained to a deeper psychological disturbance. As a first experiment, in the following paragraphs, the intended approach will be described. The experiment will be set up as a repeated measure study. Each participant will be measured before and after visiting the Alice installation. The within-subject setup will decrease the effect of individual differences.

Since the effect of the Alice installation is unknown, it is intended to use several measurement instruments that cover a wide range of concepts. This wide spread of measurements will reduce the risk that the authors may miss an important effect.

In a second experiment, they will then be able to focus on those measurement instruments that proofed to be sufficiently reliable and sensitive. The authors will discuss their choices for measurement instruments in order of the six stages in the Alice installation and their corresponding theme.

After two preliminary studies of single stages (Kooijmans and Rauterberg 2006), the authors are currently planning a first experiment for the complete installation. The experiment will be set up as a repeated measure study. Each participant will be measured before and after visiting the Alice installation. The within-subject setup will decrease the effect of individual differences.

The first stage is associated with the concept of time. Fraisse (1984) and Nakatsu et al. (2006) argued that "the more one pays attention to time, the longer it seems, with the extreme being expectancy which is nothing but expectancy of a desired or feared event. Reciprocally, duration seems short when the task is difficult and/or interesting." Since the authors are not primarily interested in the influence of expectation on the perception of time (e.g., "A watched kettle never boils" Flaherty 1999), the authors will focus on retrospective duration similar to Glicksohn's (1992) study. His study showed that retrospective estimation of the duration to an altered sensory environment was not affected by the environment, but the method of production registered an influence on the participants' cognitive timer.

Stages 2 and 3 deal with space. Stage 2 can be characterized by environmental space, while stage 3 is associated with personal space. In both stages, the relationship between the visitors' and the installation's space is of essence. This relationship can be described within the well-established framework of presence. The Independent Television Commission – Sense of Presence Inventory (ITC-SOPI) questionnaire (Lessiter et al. 2001) is the de facto standard tool to measure presence and has previously demonstrated its usefulness (Hu and Bartneck 2005). The authors therefore included it in our study.

The symbolic interpretation of stage 4 remains difficult. A possible symbolism could be the Christian ritual of baptizing. The authors were not yet able to identify an operationalization of the genesis concept for this stage. This stage is also very short transitional phase in the installation, and hence, they may neglect. The authors will continue our effort and would be grateful for suggestions from the reader.

Stage 5 is associated with the concept of the self. The Alice installation is expected to influence not only the consciousness but also the unconsciousness of its visitors. The authors intend to measure the visitors' conscious perception of the self through a personality test based on the NEO Five-Factor Inventory (NEO-FFI) with 60 items, which is designed to take only 15 min to fill in. For the unconscious self-concept Greenwald, McGhee, and Schwartz (Greenwald et al. 1998) developed the Implicit Association Test to overcome this limitation and demonstrate its effectiveness to measure the self-esteem and self-concept (Greenwald and Farnham 2000).

The necessity of logical reasoning is the topic of the last stage. Generating alternatives is a key component in human reasoning (Newstead et al. 2002), and hence, the authors use the Alternative Generation Task originally proposed by Torrens et al. (1999). Participants are being asked to create as many diagrams as they could think of to visualize a multiple model syllogism. Their responses were scored in terms of the number of different but correct diagrams produced.

ZENetic Computer

Cultural Computing: Eastern (Japanese) View

Since involving various kinds of media technology in everyday lives, people have built a sphere of communication that reaches to all parts of the globe. However, on the other hand, people are starting to feel the danger that, as the communication network expands, the level of personal communication has become shallow.

In this situation, a new communication medium that will convey personal depth of feeling across long distances has become urgently necessary. Within this context, the authors decided with this project to pursue the possibility of a communication medium that incorporates a new kind of interactivity, with editorial engineering (Matsuoka 2001) and art and technology (Tosa 2000) as a foundation and including research on the operation in which interactions of multiple cultures come to fruition and research on the “intelligence” that appears in between the user and the system.

Human communication is originally something cultivated in an environment comprising localities, national customs, and language. Therefore, the fruits of these cultures have strong roots in their unique histories.

However, the media that developed in order to convey these peculiarities across cultures were communication media such as writing, music, and film. Now, as the computer society covers the earth, the task that computers must take on is the clear and accurate intercommunication between local and global cultures. Toward that end, it is first necessary for those involved with computer technology to bring to life local characteristics.

Thus, the authors focused on the cultural roots of their native country, Japan. This includes Buddhist culture, the *kanji* (Chinese character) culture, *haiku* (Japanese short poetry) and other Japanese poetry and song, and traditional Japanese dress (*kimono*). They decided especially to dig into the unique communication space and imagery methods developed in Zen Buddhism and landscape ink painting (*Sansui* painting).

Within the traditional relationship between culture and computers, emphasis has been placed on the preservation of decaying traditional cultures for the sake of future generations, restoration of artifacts, and computer graphics simulations recreating lost relics. However, the authors struck on the possibility of computing the previously unquantifiable essence of culture inherent within people, such as personal subjectivity, feeling, emotion, and cultural personality (Murray 1998).

With this research project, the authors offer the concept and direction of “cultural computing” as above and describe, in simple terms and through the realization of an actual interactive system, a computing method reflecting the differences of emotion, consciousness, and memory that will be indispensable for the future communication abilities of computers.

As cultural computing is a very broad field, in order to produce a specific example, one must pick out a single local culture and use that as a base for building a real system. In this case, the authors chose Zen, a special area of Buddhism, and developed and evaluated ZENetic Computer as a system in which people can experience Zen culture firsthand.

ZENetic Computer: Concept

The authors developed ZENetic Computer as a specific example of cultural computing. They focused on the roots of Japanese culture, including Buddhist culture, *kanji* culture, and *waka* and *haiku* poetry (Matsuoka 2004). The authors decided to especially focus on the unique communication space Zen and *Sansui* (mountain and water) ink painting create.

Below is the explanation of the scenario a user experiences within ZENetic Computer. First, the user builds a three-dimensional *Sansui* ink painting on the display using an intuitive and user-friendly interface, constructing her own virtual space.

These images express the natural world that characterizes the East and Japan and their philosophical concepts, providing the user with a dramatic experience very different from the images seen in modern-day life (Matsuoka 2003). In this way, in the introduction, the system brings about a kind of awakening within the user and encourages their unconscious imagination.

Next, as the system classifies the user's state of consciousness based on the composition of their *Sansui* landscape design, it generates a story appropriate for the user, drawing her through the display into this alternate world.

Within the story are included mechanisms to shake the user's consciousness developed from *haiku* poetry and Zen riddles (*koan*). The story built from these elements is not a complete linear story like those found in movies or novels, but rather a nonlinear collection of short story fragments. A user who experiences these inconclusive story fragments feels a kind of uncertainty and holds an expectation and desire to build a complete story by connecting these fragments. Because of this desire, the user, in being asked questions without a "correct" answer, may hesitate somewhat but cannot help but try to answer these questions.

Through several such triggers lurking within the center of culture, the user connects these stories and builds her own unique narrative. Next, as the user uses a virtual brush, a rake for the rock garden, and images within the screen in response to questions posed by the system via images and voice, she begins to realize that the system is demanding that she meet it face-to-face. This means the door to her "unified consciousness" has begun to open further. As our desire to connect the story fragments mixes with the system's user interface, the distance between our everyday self and our true hidden self begins to shrink.

Fig. 10 ZENetic Computer at SIGGRAPH 2004 emerging technologies

Ma interaction plays an important role in the process of fusing together these two selves. *Ma* is a very Japanese concept; it is one that places a high value on ephemeral events – the here and now – within every experience.

The user, having thus traveled through several stages and several scenes, now coming to the end of the trip, interacts with a bull, which is used in Zen as a metaphor for expressing one's true self. Through this dialogue, the user can experience the process in which the everyday self and the subconscious self fuse together to bring about a unified self-consciousness.

As the surrounding environment plays a very important role in this experience, the authors have made an effort to conjure an Eastern atmosphere for the ZENetic Computer installation (Fig. 10).

Story Generated from Symbols

Creation of a Typical *Sansui* Painting

The authors have divided *Sansui* painting into 12 hieroglyphic characters (rock, mountain, moon, traveler, bridge, bird, tree, flower, wise man, cloud, and water) and made them into icons. The user drags any 2D icon and constructs his or her own 3D *Sansui* painting. Figure 11 is an example user-constructed *Sansui* painting.

Sanen Design

As one can see in the Sesshu painting in Fig. 12, there is a unique method of perspective for *Sansui* paintings called *Sanen*. Within the painting are three perspectives: *Koen*, lying far away with a view from below; *Heien*, with a straight-on view; and *Shinen*, close-up and viewed from above. Depending on the position of the user's icons, graphics corresponding to the *Sanen* area are displayed, increasing the realism of the user-created *Sansui* painting.

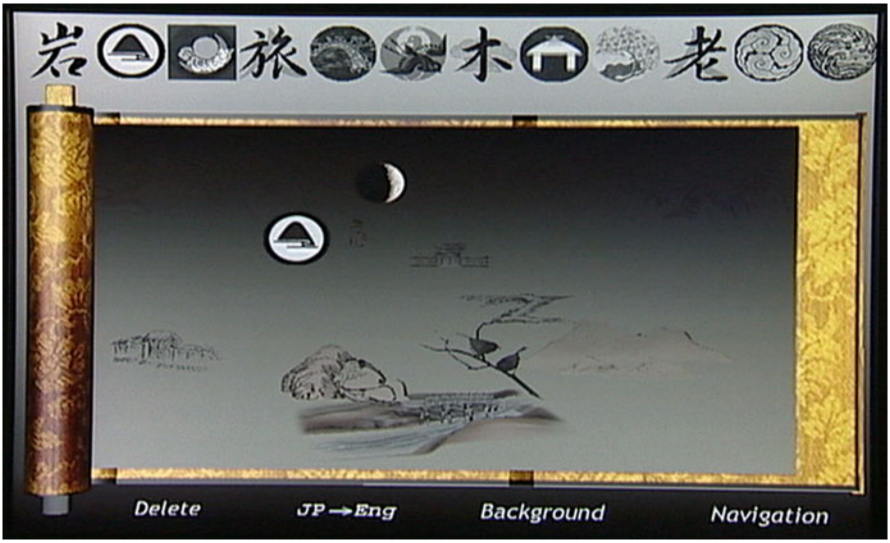


Fig. 11 Making a 3D *Sansui* ink painting

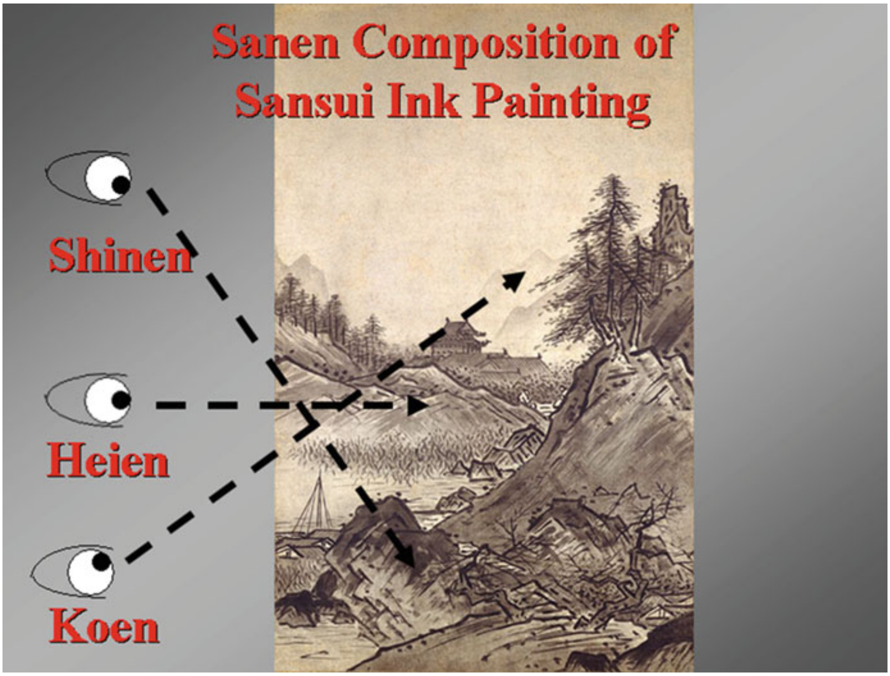


Fig. 12 Composition of *Sanen* perspective in Sesshu's work

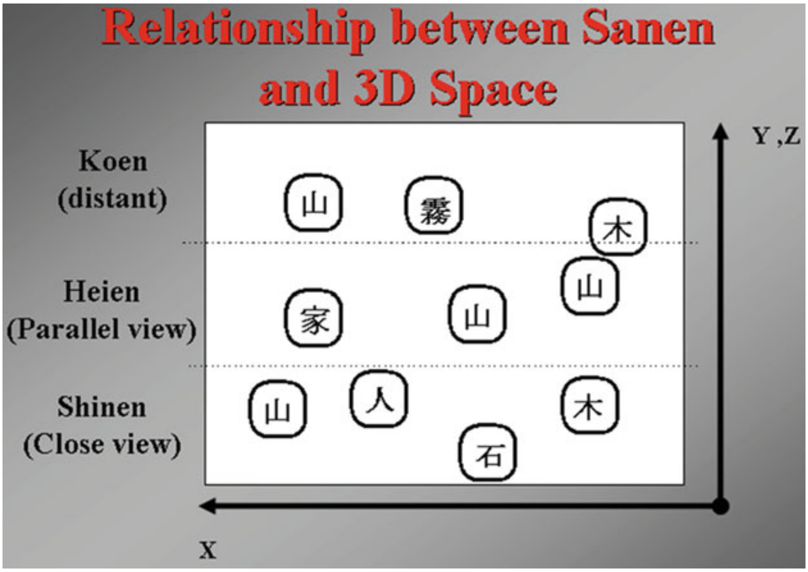


Fig. 13 Composition and distance within 3D space

Interactive Story Generation

When the user finishes creating the *Sansui* painting, she can walk through the three-dimensional *Sansui* space she created by operating the "rock garden interface" containing a touch panel (Fig. 13). By the user's approaching any *Sansui* painting icon within the space, a *haiku* poem or Zen dialogue is output based on the combination of *Sansui* painting icons contained in the framed display (Fig. 14).

Interaction Model Using a Buddhist Human Recognition Model

The authors include the Buddhist communication method between Zen master and pupil, a fashion for the purpose of understanding people, which has been followed for over 2,000 years. This kind of interaction based on the deep understanding of people is a field not yet researched within Western science.

Sansui World Expression Based on World Model "*Godai*" (Sky, Water, Fire, Wind, Earth)

In Buddhism, the directions and the five elements (*godai*) constructing the world are closely related. Upon walking through the *Sansui* painting world, changes in weather based on *godai* appear depending on the direction of movement. For example, weather changes such that if one goes north, it snows; south, a thunderstorm appears; east, it gets foggy; and west, it rains.

Fig. 14 Interaction with the rock garden interface

Classification of User Personality Based on Personality Recognition Model

Goun are the elements that make up the core of the Buddhist thought in which five basic physical and mental elements make up the world; in this interactive system, the authors apply these elements in the classification of personality. The five personality categories are as follows:

色 (Shiki)	How nature and materials actually exist
受 (Jyu)	Intuitive impression
想 (Sou)	Perceived image
行 (Gyou)	Process of mind that activates behavior
識 (Shiki)	Deep mental process reaching beyond the above processes

The authors prepare a two-dimensional *Goun* space made up of ten areas with these values along the vertical axis and their strength (positive or negative) along the horizontal axis. When the user generates a *Sansui* landscape according to her preferences, the system classifies the user's individuality through the combination of *Goun* categories assigned to the icons that make up the landscape. Through this process, the user's individuality is expressed as a *Goun* value, and the initial value is determined as described above.

Zen Dialogue Interaction

When the user approaches a certain object within the *Sansui* painting, a Zen event occurs. Every event is constructed such that one can have an interactive pseudoreal experience with a Zen *koan*. The User, Target, and Zen Master agents exist within each interaction, and the content of the interaction changes based on their interrelationships.

Fig. 15 ZEN dialogue "Dharma Anjin," where the user draws herself using the touch screen

For example, the *koan* "Dharma Anjin" (Fig. 15) is a dialogue where once, in response to a pupil's complaint that his inner spirit is in turmoil even after training, Dharma replied "Alright, then show me your troubled spirit." The authors have translated this into an interaction in which one draws one's inner spirit. The *koan* "The Lotus Smiles" (*Nengemisho*), shown in Fig. 16, holds the meaning of telepathy. In order to express this, they made an interaction like a matching game, hiding Noh Theater masks beneath lotus leaves, such that the leaves change to flower petals when the user finds matching masks. Figure 17 is the *koan* "The Sound of One Hand Clapping," wherein the system judges the calmness of the user's spirit by measuring the regularity of the user's handclapping.

Interaction Control via Chaos

One can think of the interaction for the Zen dialogues as being controlled by a combination of both cooperative and oppositional interactions between three different states: (1) the current state of the user (User), (2) the goal the user should reach (Target), and (3) the Zen master that guides the user (Zen Master). To simulate this process, a model is used such that the reaction of the system during user interactions depends on the interaction of the three elements of User, Target,

Fig. 16 ZEN dialogue "The Lotus Smiles," where the user's *Goun* state increases with successful matching of Noh Theater masks

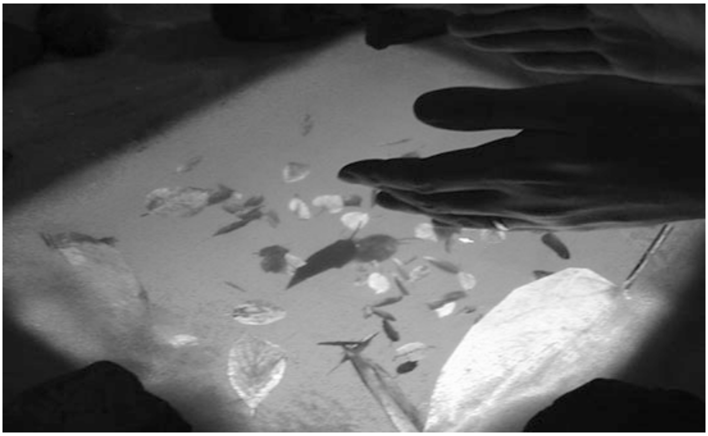


Fig. 17 ZEN dialogue "The Sound of One Hand Clapping"

and Zen Master, which are all expressed as points with *Goun* space. As a method to implement this model, one can think of a way to show the positional relationships between the three elements and the system's reactions as a table. However, because with this method the system's reactions become fixed, one cannot realize a framework allowing the enjoyment of various kinds of interactions spanning several uses. In order to allow many users to experience various interactions each time they interact with the system, it is helpful to introduce an appropriate element of "fluctuation."

The system uses a method for the dual synchronization of chaos to realize this fluctuation (Liu and Davis 2000). The method for dual synchronization of chaos is a model handling the synchronization of three or more chaos states. In this case, the system adopts a model containing three chaos states, corresponding respectively, to the User, Target, and Zen Master. Each chaos state corresponds to a point in *Goun* space. Under the method for dual synchronization of chaos, if one applies an initial value and an appropriate input value, the three chaos states relate to one another, moving through *Goun* space, and generate an output corresponding to their interactions. For the chaos input, the system uses data from the user's interactions. With the basis of Zen, activity, as the axis, the *Goun* value rises (in the plus direction) the more active a user's interaction and falls (in the minus direction) the less active he is. The data output from the chaos model is used after transformation into the system's reaction data for the user. For example, in the *koan* "Dharma Anjin," the position of the Target chaos changes depending on the curvature and density of the drawing the user sketches. The higher the density and curvature are, the better *Goun* state achieved. In order words, the Zen "*Enso*" (circle) is the best. Also, in "The Lotus Smiles," the *Goun* state rises with a higher accuracy in matching images of Noh Theater masks.

The Flow of the Story Experience Within Sansui Space

The story process a user walks through is as follows:

1. Generate a *Sansui* painting.
2. *Haiku* are generated related to the icons on the *Sansui* painting.
3. When the user approaches objects in the *Sansui* painting, associated Zen dialogues appear.
4. Depending on the interaction results from the four Zen dialogues, a form matching the user's personality is determined.
5. In conclusion, the "Ten Ox Story" corresponding to the user's interactions is displayed.

Evaluation and Future Outlook

At the 32nd UNESCO General Conference, the meaning of culture was defined once again. Culture encompasses all of customs, images, expressions, knowledge, skills, and related tools, goods, artificial objects, and cultural spaces. Not only physical cultural relics but also information exchange systems, communal, spiritual, and philosophical systems are included in the definition of culture.

In 2004, ZENetic Computer received second place in the UNESCO-sponsored Digital Storytelling Competition of Intangible Heritage. In the future, as the processing power of computers, high-quality displays, and input devices approach the limits of human perception, it is expected that high technology will enter the spiritual domain. In the West, Japanese Zen is an old and mysterious philosophy. Indeed, although books try their hand at explanation, it is difficult to truly understand Zen by reading alone. ZENetic Computer tries to convey the spirit of a culture through experiences such as participating in Zen dialogues, listening to *haiku*, and exploring *Kimono* patterns.

In the future, there will likely be a strong desire for the thought and design of cultural computing for universal communication, boldly making this kind of cross-cultural connection. This project was planned with this intent in mind, and for its realization, the authors made use of advanced game design, graphics, and interactive displays. The authors are certain that the methods used in ZENetic Computer will flourish in the broad field of education and will make possible experience-based cross-cultural understanding.

Confucius Chat

Cultural Chat: Eastern (Chinese) View

Confucian cultural values and teachings are traditionally preserved in printed media, for example, books. They are often written in an esoteric way which has not only limited appeal to young users, but their sheer volume simply intimidates many. Existing endeavors in making traditional texts friendlier to young readers have rarely gone beyond the hard-core medium of books – the methods they employ include adding modern language interpretations or inserting caricatures to assist understanding. The phenomenal sales for such products indicate a high demand for Confucian knowledge in modern Chinese societies. Others have ventured into the television media by producing cartoon series featuring Confucian figures and their stories. However, this method transforms young people from active readers to passive receivers.

Current projects utilizing the Internet as medium are limited at digitizing ancient texts and creating electronic databases, whose primary purpose is to facilitate academic research. Example of such electronic databases includes the electronic version of Siku Quanshu (the Imperial Collection of the Four Treasuries). However, since such databases are created and managed by academic institutions, which often impose restricted access, people outside the academia are excluded from its usage. Moreover, for users to make meaningful use of such electronic resources, certain knowledge of the entire corpus is necessary before they can make sense of it.

Young people of today possess high level of digital literacy. Through modern networked and social media such as Facebook, Twitter, etc., they make friends, explore new forms of entertainment, seek advices, and expand their knowledge (A Common Sense Media Study 2012; Nielsen Study 2009). It is evident that new media is an appealing and natural-information seeking channel for today's young generation. The authors see cultural computing (Tosa 2010; Rauterberg 2006) as a promising approach to promote traditional culture to the young people today. By embedding the cultural values in the interactive experience, it naturally conveys the information and nurtures cultural understanding. Inspired by the cultural computing project ZENetic Computer (Tosa and Matsuoka 2006), the authors designed Confucius Chat, a system that presents the age-old Confucian philosophy in the new context of interactive media. Here, users are no longer passive receivers listening to lectures of Confucius; instead he/she will be active enquirers engaging in stimulating dialogues with him. By bringing users “face-to-face” with the historical giant who share with them his life experiences and the joys he derived from pursuing the Dao (the Confucian way of moral ethical being), the authors hope to dispel the misconception that the Confucian way of ethical being is no more than an ascetic life of self-restrain. By using interactive media – one of the best parts modernity provides – the authors attempt to revive the appeal of these ancient philosophies and teachings to our young generation and, thereby, do our part in promoting traditional Chinese culture and closing the gap with the older generations.

Methodology

The Confucius Chat project has been a joint effort of an interdisciplinary team consisting of both Confucian scholars and computer scientists. The Confucian scholars collect relevant text materials needed for creating the virtual Confucius, whereas the computer scientists are responsible for designing the algorithms dealing with the materials collected. Both parties work together to discuss how to model Confucian knowledge and present it to users.

The authors create a knowledge base of Confucian statements from four classical texts: the Analects, Confucius sayings in the Mencius, passages directly related to Confucius in the Book of Rites, and the entire Classic of Filial Piety. These four texts constituted a major part of the educational curriculum for children in imperial China (De Bary 1996). They are the classics par excellence in the vast sea of Confucian corpus and textual basis of the culture in which the older generation of Chinese immersed. The authors chose to use James Legge’s (1815–1897) translation for all four texts. Since his translation is one of the earliest English translations and is less literal, the authors hope it can produce a more archaic tone in virtual Confucius’s replies, thereby inducing a more historic experience within users – the feeling that he is actually talking to someone walking out of history. On encountering disputable interpretations, they will consult two other commonly used translations in the field by D. C. Lau and Simon Leys to derive at what they think is an appropriate and more pertinent translation to our modern-day user.

The authors eliminate passages that are too historically specific in nature, out of which no real meaning can be extracted. Also, since many of the passages are lengthy (especially those in the Book of Rites and the Classic of Filial Piety), and are comprised of several parts each with a distinct meaning, instead of transporting the whole paragraph of text into our knowledge base as one entry (i.e., one sentence that can be used as the reply from the virtual Confucius), the authors separate them into short but self-sustained statements. For instance, the opening passage of the Analects becomes three entries in our knowledge base: (1) Is it not pleasant to learn with a constant perseverance and application? (2) Is it not delightful to have friends coming from distant quarters? (3) Is he not a man of complete virtue, who feels no discomposure though men may take no note of him? In this way, they gathered 2069 entries in total.

A topic table is created to describe the contents of the collection through card sorting technique. Topics are classified into eight headings: Action/conduct, virtue/morality, social relations, emotions/feelings, governance, metaphysical concepts, life circumstances and places/areas. Under these headings there are a total of 269 topics. The authors include in our topic table the corresponding meaning of the words as they are used in the classical texts. Word meanings are based on the WordNet (Fellbaum 1998) lexical database. This is necessary, because many of the topics in this domain have special meanings, which makes literal word-level tagging insufficient. For each entry in the knowledge base, 1–3 topics are identified. For sentences containing domain-specific terms (e.g., special concepts, name of historical persons and texts, etc.) that do not have a semantic meaning in the WordNet, a separate database tagged with these terms is created.

In addition to the philosophy knowledge base, the authors also create a knowledge base of answers to questions about Confucius as an individual, for example, his age, date of birth, hometown, etc., as well as information about his disciples, gathered from the earliest reliable historical text of Shiji (Records of the Grand Historian) , by Sima Qian (ca. 110 B.C.). These question and answer pairs are stored as patterns and templates in the Artificial Intelligence Markup Language (AIML) (Wallace 2003). AIML is a commonly used language for building natural language chatterbots. Though it simply relies on pattern matching, it is extremely effective in making the agent seemingly intelligent and human-like.

Furthermore, the authors also prepared a few series of dialogue sequences, which would be initiated by the virtual Confucius asking the user a question. This would make the conversation between the user and virtual Confucius more interactive.

System Description

Building upon our previous web-based proof-of-concept (Khoo et al. 2000, 2011), the authors have designed a system where users can directly have a conversation with a virtual Confucius through a mobile device. Users can think of it as a virtual sage living inside their mobile phones. It allows people to interact with the virtual Confucius anywhere and at any time, thus providing a more convenient way to get exposed to traditional cultural values. In this second iteration of our development, the authors greatly expanded the knowledge base of the system by working with Confucian scholars who gathered a comprehensive collection of Confucius's sayings, and redesigned the algorithm accordingly. This process will be described in this section.

To create the virtual Confucius, the authors first worked with a few Confucian scholars who have the domain knowledge to build a knowledge base containing

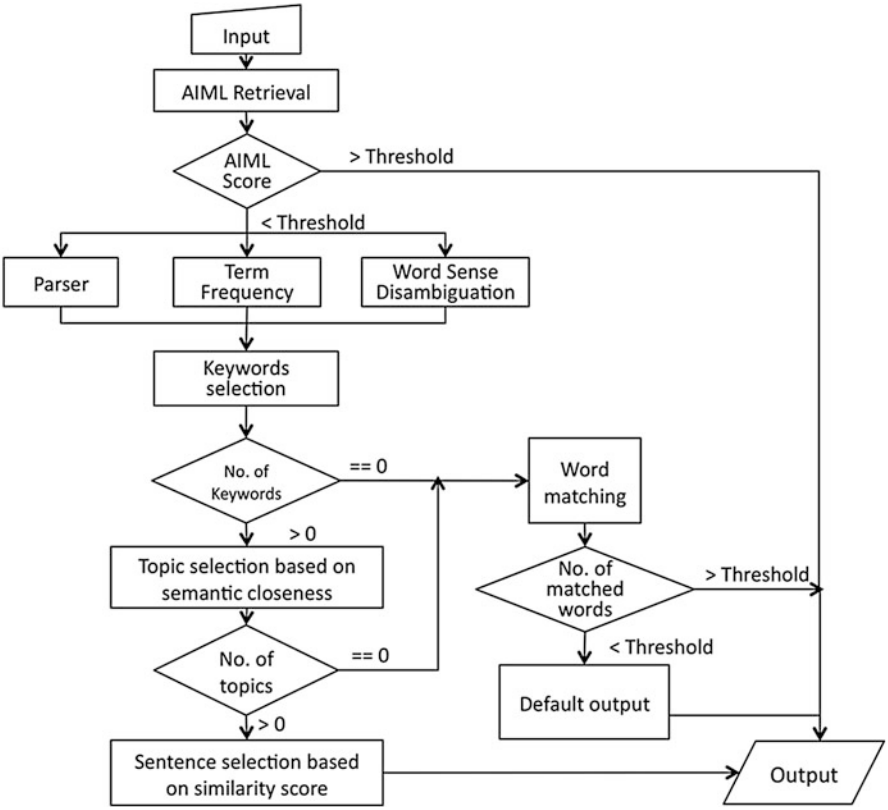


Fig. 18 A flow chart of the iSage thinking engine

sayings from Confucius. The sayings were collected from four classical Confucian texts – the *Analects*, the *Mencius*, the *Book of Rites* and the *Classic of Filial Piety*. The Confucius scholars were asked to manually annotate the knowledge base by labeling each entry with their corresponding topics. They then built a Confucius thinking engine utilizing various natural language processing (NLP) techniques to analyze user inputs and extract answers from the knowledge base.

The authors employ a layered approach for the input analysis and answer generation. Depending on the nature of input questions, three methods are designed to retrieve suitable answers from the knowledge base. Input sentences from users are processed by these methods sequentially, and the system decides if the answer is suitable for output based on certain quality measure. A flow diagram of this process is shown in Fig. 18.

Firstly, an adapted Artificial Linguistic Internet Computer Entity (A.L.I.C.E.) program (Tsai et al. 1996) is used to search for an answer. This method relies on pattern matching to find a suitable reply in prescripted question and answer pairs.

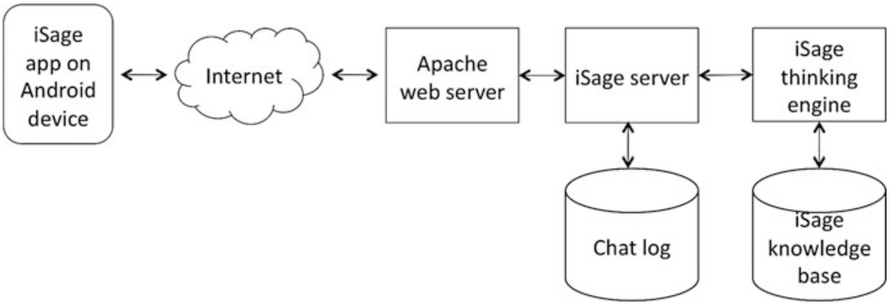


Fig. 19 Architecture of the iSage system

The original scripts were modified by the Confucius scholars, tailored to Confucius's personal information and personality. This method works well for handling small talks, such as saying hello or introducing himself. If the question is out of scope of the adapted A.L.I.C.E. knowledge base, a series of NLP techniques are used to analyze the input sentence and relate it to the sentences in the knowledge base of Confucius's sayings. The authors applied a grammatical parser (Klein and Manning 2003) to understand the structure of the sentence, extract the keywords in the sentence (Huang et al. 2008), and match their semantic closeness (Pedersen 2004; Pedersen and Kolhatkar 2009) with the topics in our knowledge base, with the help of a lexical database called WordNet (Fellbaum 1998). The sentence in the knowledge base with the highest similarity score is selected as the output. This added layer of semantic comparison gives the system the ability to dynamically calculate and respond to unseen inputs, in contrast to other commonly available conversational agents that mostly rely on prescripted answers. In addition, a simple word-matching module was also implemented, in case the previous two methods failed to find an answer. It also has one important function that is necessary in the system – to handle domain-specific terminologies that do not have semantic meanings. Finally, if no suitable output can be found using the abovementioned methods, a default output will be presented, asking the user to rephrase the question.

The iSage thinking engine requires a large comprehensive lexical database and is computationally intensive. Considering the computational power of mobile phones and the wide availability of 3G and Wi-Fi networks today, the authors decided to put the core data processing modules in our server rather than on the mobile phone itself. The mobile application simply accepts the user query and sends it as an HTTP request to the server. The server receives the request, processes it, and sends back the answer retrieved using the iSage thinking engine. The answer is then displayed on the mobile phone as virtual Confucius's reply. Figure 19 shows the architecture of the system. Currently, user's chat history with the virtual Confucius is logged anonymously for data collection and analysis in the future.

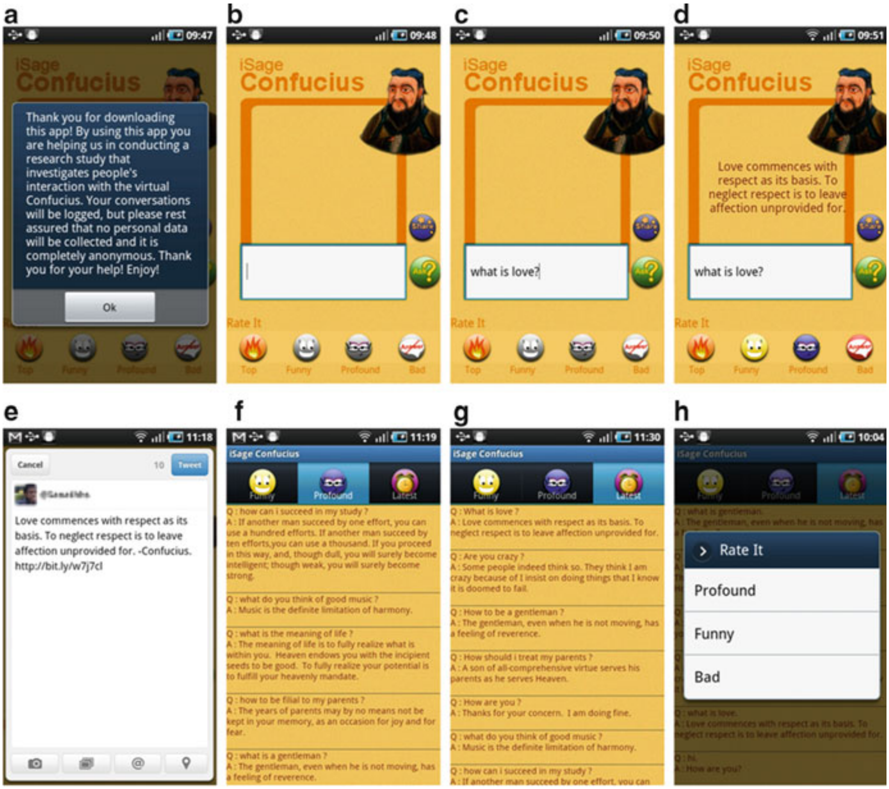


Fig. 20 A screenshot of the iSage Confucius application

Usage

Figure 20a shows a screenshot of the iSage application when it is opened for the first time. Users are informed that this application is part of a research project, and the data will be logged anonymously for research purpose. After they acknowledge it, the main page is shown as in Fig. 20b. To chat with the virtual Confucius, users can type in the question in the text box and press the green "Ask" button next to it (Fig. 20c). The question is then sent to the server, which replies with virtual Confucius's answer after processing. The answer is displayed in the area above the text input, as shown in Fig. 20d. Users can share the answer to their friends to their social networks, such as Facebook and Twitter (Fig. 20e). The authors hope to leverage the power of social networks to spread information about the application and the words from the virtual Confucius, so as to get more people to know about it and thus further promote interaction with traditional Chinese culture.

Users can also rate the given answer if they think the answer is funny, profound, or bad, by pressing the corresponding icons at the bottom of the screen. The "*Bad*" rating works as a feedback for us to find out the problem of the algorithm, so that the authors can improve it further in the future. By tapping the "*Top*" icon, users will be led to another page where they can see the top 10 most funny and most profound question and answer pairs from all users (e.g., Fig. 20f). The latest 10 question and answer pairs can also be viewed (Fig. 20g). This allows users to see what other people are asking the virtual Confucius, so they can get more inspirations or just to have a laugh. Users can also vote for the question and answer pairs here, simply by tapping the particular question and answer (Fig. 20h).

Evaluation

To see people's reactions to this system as a new way to interact with culture and to gather feedbacks to further improve the system, a preliminary user evaluation was conducted with 20 participants. All of the participants are university students and researchers under the age of 35, representing the relatively younger generation who are not so familiar with traditional culture but are frequent users of the Internet and multimedia. All of the participants are Asian except one, but most of them indicated that they are not very familiar with Confucianism.

The participants were first informed that he/she was going to interact with a virtual Confucius, which is a computer program that tries to mimic the real Confucius. They could use this mobile application to converse with Confucius, and they could ask any questions they like. The participants were asked to interact with the system for about 10 min and fill in a questionnaire afterward to report their experience with the application.

Almost all participants indicated that they enjoyed interacting with the application. A majority of them said they became more interested in Confucian culture and would like to know more. It was agreed by most of the participants that it is a good idea to promote ancient culture through this system, and it was more fun to interact with the system than reading a book about Confucian philosophy. However, when asked about the actual effects as compared to book reading, many of them hold a conservative view. Some reported that sometimes it was difficult to understand the answer from the virtual Confucius and suggested to phrase the answer in simpler language and further improve the algorithm to produce more relevant and meaningful answers.

Conclusion

The authors see great potential in combining culture with new media technology. As the first attempt, they have created and published a virtual Confucius that can talk to people. They hope that real users can use our system and enjoy this new experience interacting with ancient philosophy. It is not meant to replace book reading or classroom learning. Instead, by making the classics more approachable, the authors hope to help dispel the common stereotype of a pedantry Confucius who only lectures and preaches and arouse the young generation's interests in Confucian philosophy.

Though the preliminary evaluation showed that the application was successful in capturing user's attention and interests, as users only interacted with the system for a very short period of time, the real benefits of using this system, especially how much users can actually learn by interacting with the application, are still unknown. It also presents great challenges for evaluation in this respect – unlike science or history learning where there can be standard tests, culture is much more abstract and difficult to measure. One possible method is to conduct semistructured interview with the participants before and after longitudinal use of the application. Scenario-based questions can be designed by domain experts and used to assess user's knowledge and understanding of the Confucian culture. The authors will explore this in their future work.

Media Art

Definition

Media art is a new genre of art that started in 1960 and is becoming more and more popular in recent years (Tribe et al. 1995; Shanken 2014). Media art covers a very wide area ranging from film art, photo art, video art, interactive art, virtual reality art, network art, to even game art.

The reason why this new art area is called media art is that, different from conventional media such as papers, marbles, rocks, etc., it has been exhibited in its own deliberation process and exhibition process. New media have been mainly used such as film, video, photo, installation, network, projection, and communication media such as e-mail, SMS, and even online games.

The most distinctive feature that differentiates media art from conventional form of art, usually called "fine art," is that various types of technologies are tightly connected with artistic methodology and plays a substantial role to create new type of art (Candy and Edmonds 2002; Jones et al. 2006). These technologies cover film technologies, video technologies, human-computer interaction technologies, communication technologies, virtual reality technologies, projection technologies, network technologies, communication technologies, game technologies, and so on.

These technologies help artist to consider artistic concept, express artistic concept, create new art form, and create new ways of delivering and exhibiting artworks. One distinguished example is an emergence of "interactive art." Although there has been no direct interaction between viewers of artworks and artworks themselves, by introducing human-computer interaction technologies, it became possible that artworks can change their forms based on the interaction between the viewers and the artworks (Sommerer and Mignonneau 1999). This can open up a totally new type of art creation/anticipation process, although its full realization would still take a long time.

Although there have been a lot of art exhibitions treating media art at museums, galleries all over the world, from business point of view, this new art genre has not yet been fully recognized by people. There are only a few museums and a few art collectors that have collected media art as one of their collections. However in this area, the situation is already changing. Some of video artworks created by well-known vide artists are sold and/or bought by many museums and art collectors.

History

The history of media art traces back its origin to nineteenth century when several types of precinema invention such as zoetrope, praxinoscope, and zoopraxiscope appeared (Leeman 1976). Then the era of cinema came when Auguste Lumiere and Louis Lumiere invented cinematograph (Whiting 2005). Although most of the previous art forms, except performances, were static, cinema introduced moving images as a new expression form of art.

In the mid-twentieth century, various types of kinetic art and light art were developed. Thomas Wilfred's light artworks (Betancourt 2006) and Jean Tinguely's kinetic artworks (Tinguely 1988) should be noted as representative premedia artworks.

Then the era of video artwork came in 1960. In this genre, representative video artists are Nam June Paik and Wolf Vostell (Higgins and Vostell 1971). Nam June Paik's solo exhibition in 1963 is recognized as a first video art exhibition in the world. As video artwork is becoming a target of art collectors' interest recently, the contribution of Nam June Paik and Wolf Vostell should be memorized as a significant mile stone in the history of media art.

Along with the advancement of computer technologies and computer graphics technologies in the 1970s and 1980s, various new types of media art appeared. For example, Ivan Sutherland invented head-mounted display (HMD) and also various kinds of computer graphics algorithms which makes it possible to create various new experiences in computer graphics-based virtual environment. Also new exhibition opportunities for media art were generated during this era. One is Ars Electrônica Center in Australia (Stocker and Leopoldseder 2014), a center focusing on exhibition of media, and another is SIGGRAPH, a conference and exhibition focusing on computer graphics technologies and their applications.

The 1990s is another memorable era for media art. With the advancement of human-computer interactions and real-time technologies, it became possible for computers to react human's action in real time which makes it possible for a new media art called interactive art to emerge. Various types of interactive artworks were created and exhibited as installations. Several representative artists in this area are Roy Ascott (2003), Sommere and Mignonneau (1999), Naoko Tosa (2015), etc. Ars Electrônica and SIGGRAPH (e.g., see ACM 2000) offered good opportunities for these artists and their artworks.

Also new museums, centers dedicated to media art, especially focusing on this new media art of interactive art were built and started their activities. Some of the representative museums and centers are Center for Art and Media (ZKM) in Karlsruhe, Germany (Shwarz 1997), and InterCommunication Center (ICC) in Tokyo, Japan (see http://www.ntticc.or.jp/index_e.html).

Although the basic concept of interactive art which makes possible for interaction between viewers and artworks and through it even interaction between viewers and artists is totally new and is innovative even now (Nakatsu 1998), unfortunately there have been only a few interactive artworks that fully incorporate this concept. Therefore in the 2000s, people's interest in interactive artworks has decreased a bit, although still interactive art exhibitions at SIGGRAPH and Ars Electrônica are the centers of researchers' and artists' attractions who are working in this area.

However in fine art area, the recognition toward media art as a new genre of fine art has been shared among museum curators and art collectors. Therefore in world famous art shows such as Art Basel, Hong Kong Art Show, many new media artworks are exhibited along with traditional fine artworks, and such new media artworks became as a target of art collectors' target.

Art and Culture

From its origin, art has had a close relationship with culture. There are many artworks that record and/or illustrate various kinds of cultural issues such as cultural events, cultural customs, cultural heritages, etc. (Benton and DiYanni 2011; Greenberg 1971). Also there are lots of artworks that illustrate religious issues all over the world, such as various religious events, pictures of famous religious saints, etc. As religion is a key part of culture, this means that art has been closely involved in culture (Plate 2002). Even now especially in underdeveloping countries and in rural areas in developed countries, there are many artists who are trying to draw cultural life of people who live there.

In the case of media art, however, the situation is somewhat different. As was described before, one of the distinguished features of media art is that technologies are tightly connected to various aspects of artworks such as conceptualization, creation, exhibition, deliberation, etc. This means that technologies play an essential role in media art. Technologies have been usually indifferent to cultural issues, and in some sense, based on their futuristic concept, technologies have been trying to change people's everyday customs/habits, which means that technologies have been even destroyers of cultures.

Therefore unfortunately until recently, the relationship between media art and cultures has been weak. The emergence of the concept of cultural computing has been changing this situation. As technologies would make it possible for computers to treat cultural issues, also in the area of media art, it is expected that gradually media artworks would emerge that treat cultural issues.

Here treating cultural issues should be more than simple recording/preserving cultural events, heritages, etc. Three good system examples of cultural computing, ALICE (Hu et al. 2008), ZENetic Computer (Tosa et al. 2005), and Confucius Chat (Wang et al. 2013), were described and discussed previously in this chapter. It should be emphasized that what the researchers of these systems tried to do is, based on the deep understanding of their own cultures, to try to develop systems that make it possible for young generation and people in different cultures to be able to understand longtime existing rich cultures or their own.

Therefore the relationship between media art and culture should also target the same direction. The role of media art when treating cultural issues should not be the simple recording/expressing of cultural issues but should be something deeper. A good example of such integration between media art and cultural issues will be described in next chapter.

Media Art Treating Cultural Issues

Background

Big events such as art exhibitions, technical exhibitions, etc., are good chances to demonstrate the achievements obtained through collaborations among people with different backgrounds. The authors have been carrying out collaborations in the interdisciplinary area between art and technology. One of the authors, Naoko Tosa, is a well-known media artist who has been active in introducing technologies into media art and interactive art she has created. On the other side, other two collaborators, Jong-Il Park and Ryohei Nakatsu, are researchers in the area of communication engineering and have been eager to integrate contents and technologies.

They started their collaboration when they were colleagues at Advanced Telecommunications Research Institute International (ATR) in Japan in the mid-1990s. At present they continue their activities in different countries, so to speak Japan, Korea, and Singapore. These counties basically have different cultures but at the same time share lots of cultures as Asian countries. Therefore based on their collaboration, it is expected their collaboration could create new artworks that could fill the gap of various cultures in Asia and would express something core in the Asian culture.

Based on such basic understanding, they discussed the basic concept of a video artwork the called “Four Gods Flags” that would express the birth, fight, integration, and future of various Asian cultures. Fortunately this concept was accepted by the Yeosu Expo 2012 Committee. Through the creation process of about more than 1 year, they have created the video artwork which was exhibited at the Yeosu Expo during its open period.

This chapter describes this artwork focusing on the relationship between this artwork and the Yeosu Expo, the concept of the artwork, and the creation process of the artwork.

Fig. 21 Bird view of the Yeosu Expo site

Overview of the Yeosu Expo 2012

The Yeosu Expo 2012

The Expo is one of the world's oldest and largest international events sponsored by Bureau International des Expositions (BIE). There are two types of Expos. One is the International Registered Exhibition (World Expo) which takes place every 5 years and lasts for 6 months. The World Expo is characterized by the broad scope of the chosen theme, which must be of universal concern to all of humanity. On the other hand, International Recognized Exhibition (International Expo) is held between two World Expos and its duration is 3 months. The theme of the International Expo must represent, as with the World Expo, a global concern, but it must be more specialized in its scope. For both Expos, participants include states, international organizations, civil society groups, corporations, and citizens.

The Yeosu Expo is one of International Expos and was held at New Port area in Yeosu, Korea, from 12 May 2012 to 12 August 2012 for 3 months. The site of 2,710,000 m^2 consists of exhibition area of 250,000 m^2 and auxiliary facilities. The site has the geographical edge that comes from being adjacent to the ocean and, thus, is ideal for realizing the Expo theme, "The Living Ocean and Coast." The site overlooks Hallyeo Maritime National Park and Odong Island and is surrounded by 317 islets, having the perfect natural conditions to make the Expo's theme to come alive. The whole view of the Yeosu Expo is shown in Fig. 21.

Expo Digital Gallery

Expo Digital Gallery (EDG) is one of the key architectures of the Expo. A large LED screen is installed on the ceiling on the passageway of the International Pavilion and is called EDG. The screen is 218 m long and 30 m wide which is 6,324 units of 60 in. TVs combined together. At the EDG several video works including the author's artwork were displayed.

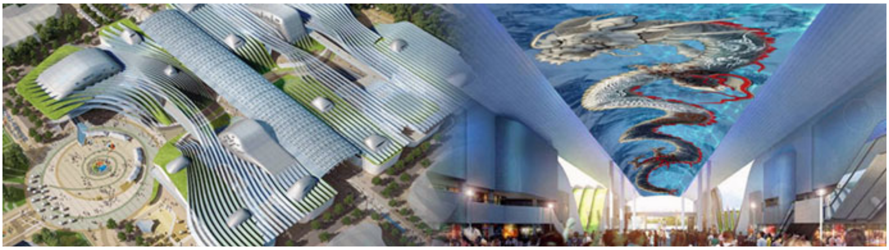


Fig. 22 Expo Digital Gallery and its inside

Unlike other pavilions, it is not necessary for visitors to wait in a long line to watch the EDG, but they can simply look up and see the screen while walking and therefore EDG gathered a lot of attentions during the Expo. The view of EDG and international pavilions are shown in the left side of Fig. 22. Also the right side of Fig. 22 illustrates how EDG looks like.

The Selection Process

The construction company, GL, announced an open call for proposals on the contents to be exhibited at the EDG. Various persons and organizations submitted their proposals, and after their review, our proposal on "Four Gods Flags" has been selected as one of video works to be exhibited at the EDG during the Expo.

Concept of "Four God Flags"

From the ancient time both in Korea, China, and Japan, there has been a legend that four sacred beasts or four gods exist in four directions protecting people. These four gods are the Blue Dragon of the east, the White Tiger of the west, the Red Phoenix of the south, and the Black Turtle-Snake of the north.

This idea has been related to the concept of "fêng shui." Good fêng shui is realized at a place with mountains in the back and water in the front, such as Yeosu city. It has been said that a place with good fêng shui is well protected by the four gods.

In the Expo Digital Gallery, a main street of the Yeosu Expo 2012, the images of these four gods will be shown in a huge ceiling LED display with the size of 218 × 30 m and appealing their dynamic movements to the audience.

The background images express underwater scenes with traditional Asian landscape taste. This is based on the idea of integrating traditional Asian culture and the ocean that is the main theme of the Expo.

At the same time, Expo would be an occasion to show people the vision of our future world. The authors tried to express how knowledge and wisdom came out of old concept of the four gods and also tried to express our future where people all over the world could be united filling the gap of various cultures.

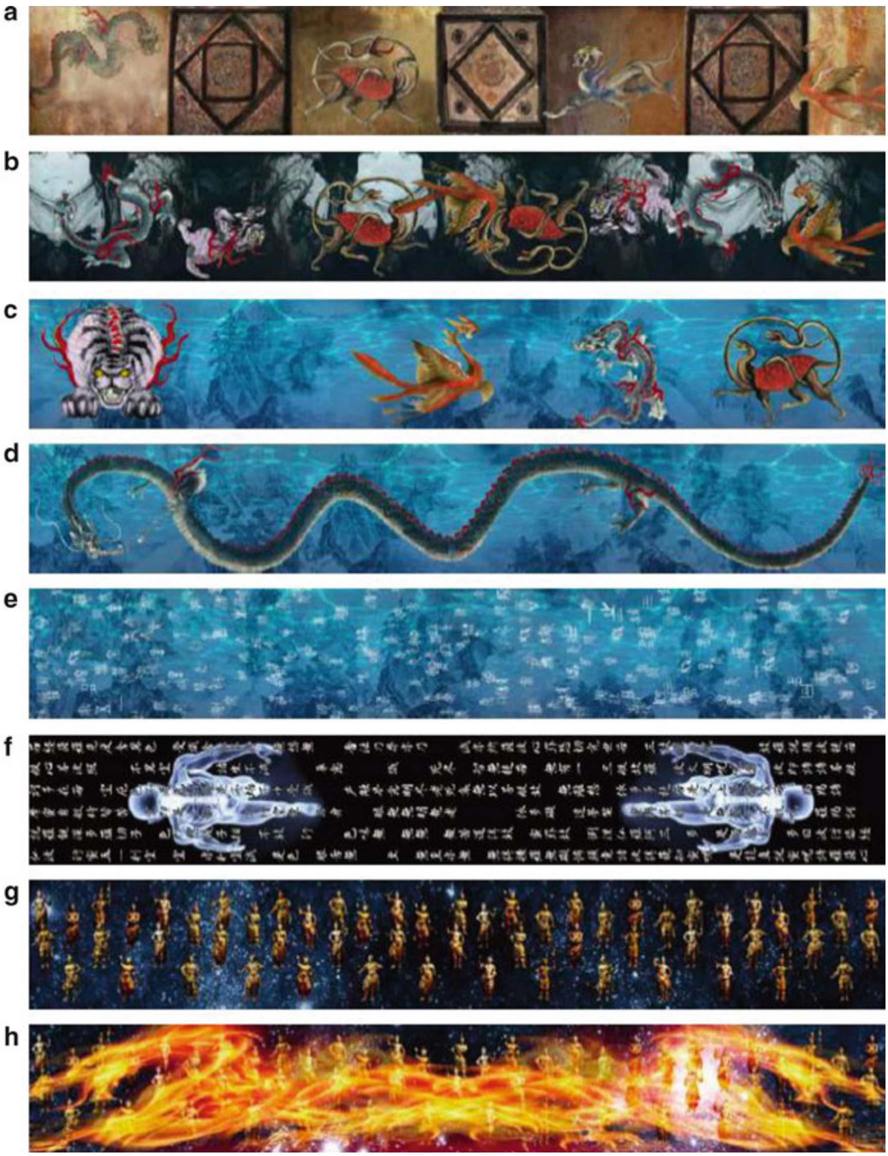


Fig. 23 Screenshot of "Four God Frags"

After the dynamic movements of the four gods, they are merged and are represented by the blue dragon, as 2012 is the year of the dragon (Fig. 23a–d).

Then the process of how human has progressed and obtained knowledge and wisdom throughout the long history is visualized by a group of old Chinese characters that came out of the blue dragon. These old Chinese characters gradually change into the normal Chinese characters and also change into a text (Fig. 23e, f).

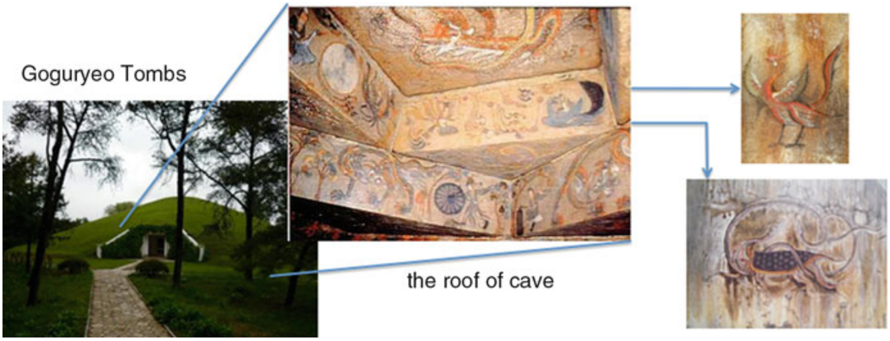


Fig. 24 Four Gods of Goguryeo Tombs

Then the authors tried to express the future when people would have richer wisdom, exchange their wisdom, and share their wisdom filling the gap of countries and cultures through the image that each of these characters would change into a form of Maitreya and they chat each other. Maitreya is a god that appears in the long future and is relevant to express the image of out future. And then in the final stage, the images of Maitreya change into the images of the four gods to express the dynamism human should maintain even in the future regardless of their country, culture, and religion (Fig. 23g, h).

Creation Process

Four Gods of Goguryeo Tombs and Animation

Four gods that are depicted on the ceiling of Goguryeo ancient tomb were used to express their resurrection. The video begins with a scene that these Gods begin to move slowly, directing the movement was solemn. The animation was created with hand-drawn images of all the key frames and then with morphin (Fig. 24).

Dragon with Korean Pattern

Dragon is a sacred animal commonly used in Asian countries. The authors created the dragon image so that it looks like a dragon in Korea. It has big eyes, a feature of the Korean dragon, and its face is not thin like the Japanese dragon. As there are various beautiful patterns of dragon designs in South Korea, these patterns were attached to the body of the dragon. It is very effective and you will see a very beautiful scene of dragon swimming underwater (Fig. 25).

Background Sansui Image and Oracle Bone Script Animation

As the theme of the Expo was "sea," the authors have created a background landscape image expressing an ocean. The landscape image itself is an animation which the authors created showing the old story of the Korean turtle that goes for a journey to get a rabbit heart for its king. In addition, the change from oracle bone

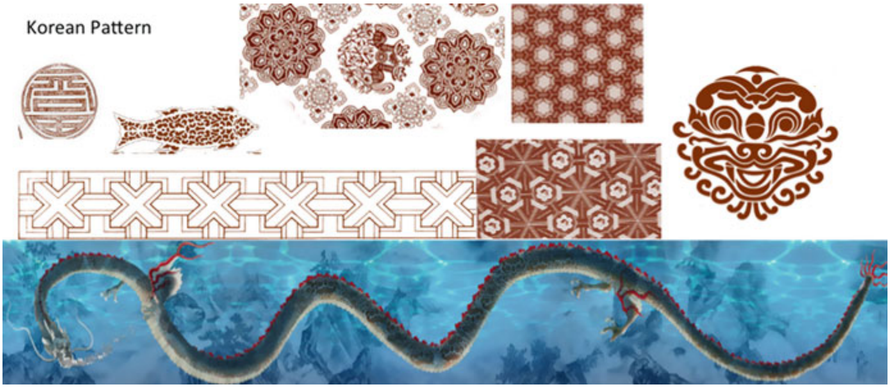


Fig. 25 Dragon with Korean pattern

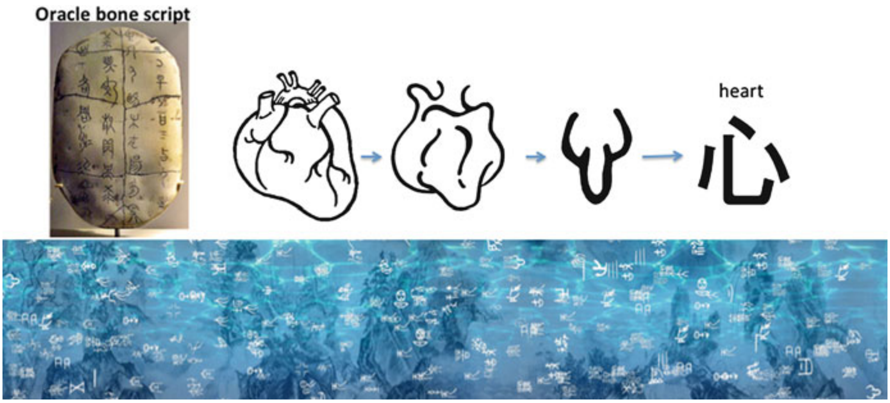


Fig. 26 Underwater Sansui with oracle bone script animation

scripts, which are the roots of Chinese characters, to modern Chinese characters is expressed by animation symbolizing the birth of wisdom. Then these characters change themselves into Heart Sutra symbolizing the Asian philosophy (Fig. 26).

Conclusion

The created video artwork of "Four Gods Flags" was regularly exhibited with other video works in the Expo Digital Gallery (EDG) of the Yeosu Expo 2012 from mid-May 2012 to mid-August and gathered lots of attentions and interest from visitors of the Expo. Especially our artwork was recognized as the most distinguished one among the video works exhibited at the EDG. Based on the authors' contribution to the Expo, one of the authors, Naoko Tosa, was awarded by the Yeosu Expo Committee.

Conclusion

In this chapter, the topic of the relationship among entertainment, culture, and art was treated. As entertainment, especially digital entertainment, has been usually understood to mean games, it has been seldom that entertainment has been discussed in relation with culture and art. Therefore the positioning of this chapter is unique in this handbook and is vulnerable.

Firstly, entertainment was discussed by higher point of view and was concluded that from its origin, entertainment is an essential part of human mental sustainability and therefore should have a close relationship with culture and art.

Then the relationship between culture and technology was discussed and a new interdisciplinary research area of cultural computing was proposed where computers treat various cultural issues. Then three examples of cultural computing research were described to some details so that readers could understand how actually computers can handle cultural issues. Also by reading these three examples of cultural computing, readers would understand how differently cultural issues are treated in the East and West. At the same time, however, it is expected that readers would notice that there are mutual similarities of basic concepts as well as computing methodologies among these three different systems. Most notable similarity is that interactions and dialogues between these systems and users do not aim to give the users right answers that have been memorized in the systems. Instead, through the interactions, all of these systems try to confuse the users, make them to confront their own self, and expect that the users would find out their own self by themselves after their experience. This means that cultural computing could provide totally different experiences to people from those conventional user-computer interactions.

Then as the next step, the relationship between culture and media art was discussed. As a first step, the definition and history of media art were described, making it clear that technologies are closely connected with media art and technologies play an essential role in media art. At the same time, futuristic and material-oriented nature of technologies was pointed out; it has been seldom that cultural issues are treated in media art. As the original relationship between culture and art was tight, however, it was pointed out that media art should treat cultural issues for its sustainability. One good example of media art that treats cultural issues is described to show how media art and cultural issues could be integrated.

Recommended Reading

A Common Sense Media Research Study, *Social Media, Social Life: How Teens View Their Digital Lives* (Common Sense Media, San Francisco, 2012)

ACM, *SIGGRAPH 2000 Conference Proceedings: Computer Graphics Annual Conference Series* (ACM, New York, 2000)

T.M. Alisi, G. D'Amico, A. Ferracani, L. Landucci, N. Torpei, Natural interaction for cultural heritage: the archaeological site of Shawbak, in *Proceedings of the International Conference on Multimedia (MM '10)* (ACM, New York, 2010), pp. 1449–1452

M.T. Artese, I. Gagliardi, Cataloging intangible cultural heritage on the web, in *EuroMed'12: Proceedings of the 4th International Conference on Progress in Cultural Heritage, Springer LNCS 7616* (Springer, Berlin/Heidelberg, 2012), pp. 676–683

R. Ascott, *Telematic Embrace: Visionary Theories of Art, Technology, and Consciousness* (University of California Press, Berkeley, 2003)

B. Salem, M. Rauterberg, R. Nakatsu, "Kansei Mediated Entertainment," Entertainment Computing – ICEC2006, Springer LNCS 4161, (Springer, 2006), pp. 103–116

B. Balazs, *Theory of the Film: Character and Grows of a New Art* (Dover Publications, New York, 1970)

C. Bartneck, J. Hu, B. Salem, R. Cristescu, M. Rauterberg, Applying virtual and augmented reality in cultural computing. Int. J. Virtual Real. **7**(2), 11–18 (2008)

G. Bell, The age of auspicious computing? Interactions **11**(5), 76–77 (2004)

G. Bell, No more SMS from Jesus: ubicomp, religion and techno-spiritual practices, in *UbiComp'06 – Proceedings of the 8th International Conference on Ubiquitous Computing, Springer LNCS 4206* (Springer, Berlin/Heidelberg, 2006), pp. 141–158

J.R. Benton, R.J. DiYanni, *Art and Culture: An Introduction to the Humanities, Combined Volume* (Pearson, 2011)

M. Betancourt, *Thomas Wilfred' Clavilux* (Wildside Press, 2006)

E. Buie, M. Blythe, Spirituality: there's an app for that! (but not a lot of research), in *CHI '13 Extended Abstracts on Human Factors in Computing Systems on – CHI EA '13* (ACM, New York, 2013), pp. 2315–2324

F. Cameron, S. Kenderdine, *Theorizing Digital Cultural Heritage: ACritical Discourse, Volume 59 of Media in transition* (The MIT Press, Cambridge, MA, 2007)

L. Candy, E.A. Edmonds, *Explorations in Art and Technology* (Springer, New York, 2002)

L. Carroll, *Alice's Adventures in Wonderland* (Macmillan, London, 1865)

W. Choi, T. Fukumori, K. Furukawa, K. Hachimura, T. Nishiura, K. Yano, Virtual Yamahoko parade in virtual Kyoto, in *ACM SIGGRAPH 2010 Posters on – SIGGRAPH '10, page Article No.146* (ACM, New York, 2010)

O. Choudary, V. Charvillat, R. Grigoras, P. Gurdjos, MARCH: mobile augmented reality for cultural heritage, in *Proceedings of the 17th ACM International Conference on Multimedia*, vol. 3 (ACM, New York, 2009), pp. 1023–1024

W.M.T. De Bary, Confucian education in premodern East Asia, in *Confucian Traditions in East Asian Modernity*, ed. by T. Wei-ming (Harvard University Press, Cambridge, MA, 1996), pp. 21–38

S. Dutta, I. Mia, *The Global Information Technology Report 2008–2009* (World Economic Forum, Geneva, 2009)

T.H. Eriksen, Between universalism and relativism: a critique of the UNESCO concept of culture, in *Culture and Rights*, ed. by J.K. Cowan et al. (Cambridge University Press, 2001), p. 127

C. Fellbaum, *WordNet: An Electronic Lexical Database* (The MIT Press, Cambridge, MA, 1998)

M.G. Flaherty, *A Watched Pot: How We Experience Time* (New York University Press, New York, 1999)

B.J. Fogg, *Persuasive Technology: Using Computers to Change What We Think and Do* (Morgan Kaufmann, Amsterdam/Boston, 2003)

P. Fraisse, Perception and estimation of time. Annu. Rev. Psychol. **35**, 1–36 (1984)

S. Gallagher, J. Shear (eds.), *Models of the Self* (Imprint Academic, Exeter, 2000)

W. Gaver, M. Blythe, A. Boucher, N. Jarvis, J. Bowers, P. Wright, The prayer companion: openness and specificity, materiality and spirituality, in *Proceedings of the 28th International Conference on Human Factors in Computing Systems – CHI '10* (ACM, New York, 2010), pp. 2055–2064

J. Glicksohn, Subjective time estimation in altered sensory environments. Environ. Behav. **24**, 634–652 (1992)

C. Greenberg, *Art and Culture: Critical Essays* (Beacon Press, Boston, 1971)

A.G. Greenwald, S.D. Farnham, Using the implicit association test to measure self-esteem and self-concept. J. Pers. Soc. Psychol. **79**, 1022–1038 (2000)

A.G. Greenwald, D.E. McGhee, J.K.L. Schwartz, Measuring individual differences in implicit cognition: the implicit association test. J. Pers. Soc. Psychol. **74**, 1464–1480 (1998)

D. Heiligman, *From Caterpillar to Butterfly* (Harper Collins, New York, 1996)

D. Higgins, W. Vostell, *Fantastic Architecture* (Something Else Press, New York, 1971)

A. Hiyama, Y. Doyama, M. Miyashita, E. Ebuchi, M. Seki, M. Hirose, Wearable display system for handing down intangible cultural heritage, in *Proceedings of the 2011 International Conference on Virtual and Mixed Reality: Systems and Applications – Volume Part II* (Springer, Berlin 2011), pp. 158–166

M. Hlubinka, J. Beaudin, E.M. Tapia, S. John, An AltarNation: interface design for meditative communities, in *CHI '02 Extended Abstracts on Human Factors in Computing Systems – CHI '02* (ACM, New York, 2002), pp. 612–613

J. Hu, C. Bartneck, Culture matters – a study on presence in an interactive movie, in *Proceedings of the 8th Annual International Workshop on Presence* (London, 2005) pp. 153–159

J. Hu, C. Bartneck, B. Salem, M. Rauterberg, ALICE's adventures in cultural computing. Int. J. Arts Technol. **1**(1), 102–118 (2008)

Z. Huang, M. Thint, Z. Qin, Question classification using head words and their hypernyms, in *Proceedings of the Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics* (Morristown, 2008), pp. 927–936

K. Isbister, H. Nakanishi, T. Ishida, C. Nass, Helper agent: designing assistant for human-human interaction in a virtual meeting space, in *CHI Letters,* vol. 2, no. 1 (ACM Press, New York, 2000), pp. 57–64

C.A. Jones, B. Aming, M. Briand, MITLCA Center, *Sensorium: Embodied Experience, Technology, and Contemporary Art* (The MIT Press, Boston, 2006)

J. Juul, *A Casual Revolution: Reinventing Video Games and Their Players* (The MIT Press, Boston, 2012)

I. Kant, Beantwortung der Frage: Was ist Aufklärung? Berlinische Monatschrift **2**, 481–494 (1784)

E.T. Khoo, A. Cheok, W. Liu, X. Hu, P. Marini, Confucius computer: bridging intergenerational communication through illogical and cultural computing. Virtual Real. **15**, 249–265 (2000)

E.T. Khoo, A.D. Cheok, W. Liu, X. Hu, P. Marini, V. Saksen, J. Jiang, B.L. Duh, Confucius computer: bridging intergenerational communication through illogical and cultural computing. Virtual Real. **15**(4), 239–265 (2011)

S. Kitayama, H.R. Markus, H. Matsumoto, V. Norasakkunkit, Individual and collective processes in the construction of the self: self-enhancement in the United States and self-criticism in Japan. J. Pers. Soc. Psychol. **72**, 1245–1267 (1997)

D. Klein, C.D. Manning, Accurate unlexicalized parsing, in *Proceedings of the 41st Annual Meeting on Association for Computational Linguistics – ACL'03* (2003), pp. 423–430

D. Koller, B. Frischer, G. Humphreys, Research challenges for digital archives of 3D cultural heritage models. J. Comput. Cult. Herit. **2**(3), 1–17 (2009)

T. Kooijmans, G.W.M. Rauterberg, Advice from a Caterpillar: an application for cultural computing about the self, in *5th International Conference on Entertainment Computing (ICEC)* (Sanda, 2006), pp. 5–8

A.L. Kroeber, C. Kluckhohn, *Culture: A Critical Review of Concepts and Definitions* (Peabody Museum, Cambridge, MA, 1952)

S.K. Lee, S. Rennert, *Nam June Paik* (Tate, 2011)

F. Leeman, *Hidden Images: Games of Perception, Anamorphic Art, Illusion from the Renaissance to the Present* (Harry N. Abrams, 1976)

J. Lessiter, J. Freeman, E. Keogh, J. Davidoff, A cross-media presence questionnaire: the ITC sense of presence inventory. Presence Teleop. Virt. Environ. **10**, 282–297 (2001)

M. Levoy, J. Ginsberg, J. Shade, D. Fulk, K. Pulli, B. Curless, S. Rusinkiewicz, D. Koller, L. Pereira, M. Ginzton, S. Anderson, J. Davis, The digital Michelangelo project, in *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques – SIGGRAPH '00* (ACM Press, 2000), pp. 131–144

Y. Liu, P. Davis, Dual synchronization of chaos. Phys. Rev. E **61**, 2176–2179 (2000)

G.C. Lough, Alice in wonderland and cognitive development: teaching with examples. J. Adolesc. **6**(4), 305–315 (1983)

A. Mallik, S. Chaudhury, H. Ghosh, Nrityakosha: preserving the intangible heritage of Indian classical dance. J. Comput. Cult. Herit. **4**(3), 1–25 (2011)

L. Marc et al., The digital Michelangelo project: 3D scanning of large statues, in *Proceedings of SIGGRAPH 2000* (August 2000), pp. 131–144

S. Matsuoka, *Editorial Science of Intelligence* (Asahi Shimbunsha (in Japanese), 2001)

S. Matsuoka, *Philosophy of Sansui* (Gogatsu Shobo (in Japanese), 2003)

S. Matsuoka, *Science of Flower, Birds, and Moon* (Chuko Bunko (in Japanese), 2004)

M. McLuhan, *The Gutenberg Galaxy* (University of Toronto Press, Toronto, 2011)

M. McLuhan, B.R. Powers, *The Global Village: Transformations in World Life and Media in the 21st Century* (Oxford University Press, New York, 1992)

J.H. Murray, *Hamlet on the Holodeck: The Future of Narrative in Cyberspace* (The MIT Press, Cambridge, MA, 1998)

R. Nakatsu, *Image/Speech Processing Adopting an Artistic Approach – Toward Integration of Art and Technology. in Art@Science* (Springer, Wein/New York, 1998), pp. 38–49

R. Nakatsu, *Integration of Multimedia and Art for New Human-Computer Communications. PRICAI 2002: Trends in Artificial Intelligence* (Springer, 2002), pp. 19–28

R. Nakatsu, M. Rauterberg, P. Vorder, *A New Framework for Entertainment Computing: From Passive to Active Experience. Entertainment Computing – ICEC2005*. Lecture Notes in Computer Science, vol. 3711 (Springer, 2005), pp. 1–12

R. Nakatsu, M. Rauterberg, B. Salem, Forms and theories of communication: from multimedia to Kansei mediation. Multimed Syst **11**(3), 304–312 (2006)

R. Nakatsu, *Logos, Pathos, and Entertainment, Culture and Computing*. Lecture Notes in Computer Science, vol. 6295 (Springer, 2009) pp. 137–1465

R. Nakatsu, C. Edirisinghe, *The Role of Movies and Telephony in the History of Communication Media. Entertainment Computing – ICEC2011*. Lecture Notes in Computer Science, vol. 6972 (Springer, 2011), pp. 448–451

T. Nakaya, K. Yano, Y. Isoda, T. Kawasumi, Virtual Kyoto project: digital diorama of the past, present, and future of the historical city of Kyoto, in *Culture and Computing*, ed. by T. Ishida. Lecture Notes in Computer Science, vol. 6259 (Springer, Berlin/Heidelberg, 2010), pp. 173–187

S.E. Newstead, V.A. Thompson, S.J. Handley, Generating alternatives: a key component in human reasoning? Mem. Cognit. **30**, 129–137 (2002)

Nielsen Study, *How Teens Use Media* (2009)

R.E. Nisbett, K. Peng, I. Choi, A. Norenzayan, Culture and systems of thought: holistic versus analytic cognition. Psychol. Rev. **108**(2), 291–310 (2001)

R.E. Nisbett, T., Masuda, Culture and point of view. Paper presented in the proceedings of the National Academy of Sciences, vol. 10019 (2003), pp. 11163–11170

Y. Okada, T. Shoji, Digital conservation of cultural assets, in *Culture and Computing*. Lecture Notes in Computer Science, vol. 6259 (Springer, Berlin/Heidelberg, 2010), p. 147

R. Pavlov, D. Paneva-Marinova, K. Rangochev, M. Goynov, D. Luchev, Towards online accessibility of valuable phenomena of the Bulgarian Folklore Heritage, in *Proceedings of the 11th International Conference on Computer Systems and Technologies and Workshop for Ph.D. Students in Computing on International Conference on Computer Systems and Technologies – CompSysTech '10* (ACM, New York, 2010), pp. 329–334

T. Pedersen, WordNet:: Similarity – measuring the relatedness of concepts, in *Proceedings of the Nineteenth National Conference on Artificial Intelligence (AAAI-04)* (San Jose, 2004), pp. 1024–1025

T. Pedersen, V. Kolhatkar, WordNet:: SenseRelate:: AllWords: a broad coverage word sense tagger that maximizes semantic relatedness, in *Proceedings of the North American Chapter of the Association for Computational Linguistics – Human Language Technologies 2009 Conference, Association for Computational Linguistics* (Boulder, 2009), pp. 17–20

J.S. Pierce, R. Pausch, C.B. Sturgill, K.D. Christiansen, Designing a successful HMD-based experience. Presence **8**(4), 469–473 (1999)

S.B. Plate, *Religion, Art, and Visual Culture: A Cross-Cultural Reader* (Palgrave Macmillan, New York, 2002)

M. Rauterberg, Positive effects of entertainment technology on human behaviour, in *Building the Information Society*, ed. by R. Jacquart (IFIP/Kluwer, Dordrecht, 2004), pp. 51–58

M. Rauterberg, From personal to cultural computing: how to access a cultural experience, in *uDayIV – Information nutzbar machen*, ed. by G. Kemper, P. von Hollberg (Pabst Science, Lengerich, 2006), pp. 13–21

S. Saga, K. Vlack, H. Kajimoto, S. Tachi, Haptic video, in *ACM SIGGRAPH 2005 Emerging Technologies on – SIGGRAPH '05* (2005)

B. Schonenberg, C. Bartneck, Mysterious machines, in *Proceedings of the 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI '10)* (ACM, New York, 2010), pp. 349–350

E.A. Shanken, *Art and Electronic Media* (Phaidon Press, London, 2014)

H.Y. Shiaw, R.J.K. Jacob, G.R. Crane, The 3D vase museum: a new approach to context in a digital library, in *Proceedings of the 4th ACM/IEEE-CS Joint Conference on Digital Libraries* (ACM, New York, 2004), pp. 125–134

H.P. Shwarz, *Media – Art – History: Media Museum, Zkm, Center for Art and Media Karlsruhe* (Prestel, Munich, 1997)

C. Sommerer, L. Mignonneau, Art as a living system: interactive computer artworks. Leonardo **32**(3), 165–173 (1999)

E. Stavrakis, A. Aristidou, M. Savva, S.L. Himona, Y. Chrysanthou, Digitization of Cypriot folk dances, in *EuroMed'12 – Proceedings of the 4th International Conference on Progress in Cultural Heritage Preservation*. Lecture Notes in Computer Science, vol. 7616 (Springer, Berlin/Heidelberg, 2012), pp. 404–413

G. Stocker, H. Leopoldseder, Art Electronica Center. Edition Lammerhuber; Bilingual edition (2014)

J. Tinguely, *Jean Tinguely: A Magic Stronger Than Death* (Abbeville Press, New York, 1988)

D. Torrens, V.A. Thompson, K.M. Cramer, Individual differences and the belief bias effect: mental models, logical necessity, and abstract reasoning. Think. Reason. **5**, 1–28 (1999)

N. Tosa, *Cross-Cultural Computing: An Artist's Journey*. Series on Cultural Computing. (Springer, 2015)

N. Tosa, Cultural computing – creative power integrating culture, unconsciousness and software, in *Culture and Computing*. Lecture Notes in Computer Science, vol. 6259 (Springer, 2010), pp. 127–136

N. Tosa, S. Matsuoka, ZENetic Computer: exploring Japanese culture. Leonardo **39**(3), 205–211 (2006)

N. Tosa, Expression of emotion, unconsciousness with art and technology, in *Affective Minds*, ed. by G. Hatano, N. Okada, N. Tanabe (Elsevier, 2000) pp. 183–201

N. Tosa, S. Matsuoka, K. Miyazaki, Interactive storytelling system using behavior-based non--verbal information: ZENetic Computer, in *Proceedings of the Eleventh ACM International Conference on Multimedia* (ACM, Berkeley, 2003), pp. 466–467

N. Tosa, S. Matsuoka, B. Ellis, H. Ueda, R. Nakatsu, *Cultural Computing with Context-Aware Application: ZENetic Computer. Entertainment Computing – ICEC2005*. Lecture Notes in Computer Science, vol. 3711 (Springer, 2005), pp. 13–23

N. Tosa, H. Obara, M. Minoh, Hitch haiku: an interactive supporting system for composing haiku poem, in *Proceedings of the 7th International Conference on Entertainment Computing (ICEC 2008)* (Springer, Berlin, 2008), pp. 209–216

N. Tosa, P. Jong-Il, R. Nakatsu, Korea Expo 2012 and its digital gallery work focusing Asian four gods, in *Culture and Computing* (IEEE, New Jersey, 2013), pp. 45–50

M. Tribe, R. Jana, U. Grosenick, *New Media Art* (Taschen, London, 1995)

C.C. Tsai, Z. Cai, B. Bruya, *Confucius Speaks: Words to Live By* (Anchor Books, New York, 1996)

D. Uriu, N. Okude, M. Inami, T. Taketomi, C. Sato, Where Buddhism encounters entertainment computing, in *ACE'12 Proceedings of the 9th International Conference on Advances in Computer Entertainment*. Lecture Notes in Computer Science, vol. 7624 (Springer, Berlin/Heidelberg, 2012), pp. 589–592

R. Wallace, *The Elements of AIML Style* (Alice AI Foundation, San Francisco, 2003)

X. Wang, E.T. Khoo, C-R. Fu, A.D. Cheok, R. Nakatsu, Confucius chat: promoting traditional Chinese culture and enhancing intergenerational communication through a chat system, in *Culture and Computing* (IEEE, New Jersey, 2013), pp. 123–128

J. Whiting, *Auguste & Louis Lumiere: Pioneers in Cinema Film* (Mitchell Lane, 2005)

X. Wu, N. Tosa, R. Nakatsu, New Hitch Haiku: an interactive Renku poem composition supporting tool applied for sightseeing, in *Proceedings of the 8th International Conference on Entertainment Computing (ICEC 2009)* (Springer, Berlin/Heidelberg, 2009), pp. 191–196

S.P. Wyche, K.E. Caine, B.K. Davison, S.N. Patel, M. Arteaga, R.E. Grinter, Sacred imagery in techno-spiritual design, in *Proceedings of the 27th International Conference on Human Factors in Computing Systems – CHI 09* (ACM, New York, 2009), pp. 55–58

S. Xu, H. Jiang, T. Jin, F.C.M. Lau, Y. Pan, Automatic generation of Chinese calligraphic writings with style imitation. IEEE Intell. Syst. **24**(2), 44–53 (2009)

C. Yao, J. Xiao, T.Q. Guo, L. Xia, Real-time digitized shadow play performance method based on multi-point interactive controlling method. Int. J. Comput. Appl. Technol. **38**(1/2/3), 86–92 (2010)

K. Zhu, N. Ranashinghe, C. Edirisinghe, O.N.N. Fernando, A.D. Ceok, Poetry mix-up. Comput. Entertain. **9**(2) (2011), ACM, New York

Games of Chance: Explorations into Our Animal Selves

31

Siddharth Ramakrishnan and Victoria Vesna

Contents

Abstract

The authors come together from very different disciplines – (media) art and (neuro)science – in order to create an interactive work that engages the audience in a way that takes them out of the anthropocentric point of view. With this collaborative project, the authors wish to break this “human” barrier and allow an exploration and identification of the diverse world of the animals around us. *The Hox Zodiac* allows the human audience to experience the shared history and potential of genetic diversity among animals. Here, the idea of the Hox gene as a binding element is introduced, and the Chinese animal zodiac and dinner

S. Ramakrishnan (✉)
Neuroscience Program, Department of Biology, University of Puget Sound, Tacoma, WA, USA
e-mail: sramakrishnan@pugetsound.edu

V. Vesna
Department of Design Media Arts, Art | Sci Center, University of California Los Angeles, Los Angeles, CA, USA

Program in Empowerment Informatics, School of Integrative and Global Majors, University of Tsukuba, Tsukuba, Japan
e-mail: vesna@arts.ucla.edu

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_17

table as the structure/space for discussion is employed, allowing the format to build based on the audience interaction. In neuroscience this is a principle known as the emergent property of network connections, where a simple array of neurons can give rise to complex behaviors through interactions and adaptations.

Responding to the emergent nature of the game-like environment of the ancient *Book of Changes*, the *I Ching*, and the related work of John Cage serves as base for the conceptual framework of the project. Similar to Cage's ideas of chance and indeterminacy, this work looks to the ancient Eastern philosophies along with the scientific research, seeking the balance between rational and irrational, conscious and unconscious – in relation to our interconnectivity with the animal kingdom. This chapter describes the research process and variations that emerged with audience participation and interaction.

Keywords
Hox gene • Chinese zodiac • *I Ching* • Neuroscience and art • Animal-human relationship • Biotech and art

Introduction

> The eight trigrams are symbols standing for changing transitional states; they are images that are constantly undergoing change. Attention stands not on things in their state of being – as is chiefly the case in the Occident – but upon their movements in change. The eight trigrams therefore are not representations of things as such but of their tendencies in movement.
> Richard Wilhelm, Introduction, in Wilhelm, trans. I Ching (Wilhelm et al. 1967)

Our perception and experience of the world around us is limited by an anthropocentric viewpoint – in part a physical limitation of our senses – but largely because we forget that as animals we share a common inheritance both genetically and ecologically with the myriad of species around us. This chapter will detail a few explorations that the authors have embarked on to in order to break down this filter that limits our world view.

While we humans tend to place ourselves on a higher plane than the other animals around us, the fact that we share both evolutionary and genetic history with other organisms and currently coexist (albeit tenuously) within shared ecosystems and biomes cannot be escaped. Animals feature prominently in our stories and myths (from origin of the world stories to incarnations to tales of banishments), are commonly used as comparisons in linguistic usage (sly as a fox, strong as an ox), and within some societies have an indelible link with the human experience. Domestication of animals has also led to the success of the human species, transforming us from a hunter-gatherer to an agrarian society. The role of animals as pets and sources of emotional comfort has also emerged, leading not only to an increased awareness of animal needs but also to a co-evolution between humans and animals. Further, quality of human life has improved tremendously over the last two centuries primarily because of the role animals play in laboratory science, medical research, and drug testing. This is about to explode further in the near future as technology advances, and we are on the cusp of creating hybrid beings (species enhanced with capabilities of others) and xeno-harvesting (harvesting human organs grown within "donor species" Cooper and Lanza 2000).

With all the entwining of lives of humans and animals, in terms of shared spaces, genetic material, as a society, we do not often pause to consider how connected we are to all organisms around us. Our education strives to make the distinction ensuring that humans are considered different and not part of the animal "whole." Primal instincts such as sexual lust and fear are dismissed as animalistic and beneath our more perceived higher evolved state of being. Even life scientists, who in the past began their training broadly as "naturalists" and "zoologists," are now immediately thrown into studying humanistic models of diseases, with no foundation of placing humans in the context of the animal kingdom.

The Hox Gene and the Book of Change Come Together in a Dinner Table

> The matter of interest seems to be the configuration formed by chance events in the moment of observation, and not at all the hypothetical reasons that seemingly account for the coincidence.
> CG Jung, Foreward, in Wilhelm trans. I Ching. Pg. xxiii (Wilhelm et al. 1967)

The authors come from two different worlds – (media) arts and (neuro)science – and do not identify themselves with the gaming world. Despite hardly participating or playing games, allowing the audience to take the helm in this project, it took on a game mode very quickly. At one point in the development of this topic and the parallel search for best interfaces, a game designer was hired, but the established rules and expectations of the established gaming world did not apply well to the explorations in this work. Indeed, after experimenting with various formats and strategies, a conscious decision was made to not include any technology but focus on creating the context, the circumstance that makes the table into one giant petri dish. The "play" that was felt much more appropriate is based on the *I Ching*, also known as the *Classic of Changes* or *Book of Changes* in English, an ancient divination text and the oldest of the Chinese classics. Using the classic translation by Richard Wilhelm and Cary Baynes, it is seen not as a mysterious source of oracles but as a source of the Taoist and Confucius philosophies that tap into the collective unconscious.

Further, the approach to this work is closer to that of conceptual artists who applied these "rules" such as John Cage (1912–1992) (Jensen 2009; Larson 2013). Cage utilized the *I Ching* in his composition of music, writing, and visual art throughout his career, and one could safely make the assumption that *chance operations* are based on his study of the *Book of Changes*. It is known that the *Music of Changes* was composed entirely with the use of the *I Ching* and his fascination with the 64 hexagrams. It should be noted that Cage's work with the idea of chance is frequently misunderstood – he, in fact, insisted that there is no such thing and that we experience different forms of order in relation to our perception based on our societal barriers. The idea of chance was used more as a framework to point out the various different possibilities of interpretation.

With this collaborative project, the authors wish to break this "human" barrier and allow an exploration into the diverse world of the animals around us. *The Hox Zodiac* project allows the human audience to experience the shared history and potential of genetic diversity among animals that very much includes us, the human animal. It is an exploration of the rich differences among the animals around us while iterating the shared themes and common constructs that underlie the genetic basis of *all* body plans, humans included. These experiments are ongoing for a number of years and (2008 to present) and are modified based on play testing and audience feedback from exhibitions around the world.

The Hox Zodiac: Genetic Games of Chance

> It necessarily follows that chance alone is at the source of every innovation, and of all creation in the biosphere. Pure chance, absolutely free but blind, at the very root of the stupendous edifice of evolution: this central concept of modern biology is no longer one among many other possible or even conceivable hypotheses. It is today the sole conceivable hypothesis, the only one that squares with observed and tested fact. And nothing warrants the supposition – or the hope – that on this score, our position is ever likely to be revised. There is no scientific concept, in any of the sciences, more destructive of anthropocentrism than this one.
>
> Jacques Monod, *Chance and Necessity* (Monod and Wainhouse 1972)

Genetic information is the foundation of all living creatures seen around us. The DNA sequences are used to codify numerous functions – from keeping the cells alive to helping them survive. The basic genetic code comprises of four nucleotides A, T, G, and C which serve as a four-letter code (Watson and Baker 2013). Three nucleotides in sequence (e.g., ATG, UUU, GGG) are the code for an amino acid, thereby allowing the translation of the genetic sequences into functional proteins. Thus, 64 nucleotide triplets (4^3) are used to translate into the 26-letter alphabet of amino acids (Pierce 2013). It is of note that the *I Ching* also has 64 different hexagrams within its coda (Huang 2010). A few philosophers and mathematicians have compared the *I Ching* to the DNA language and researched how the 64-bit code of the DNA and *I Ching* match up (Shoenberger 1992). Further, the *I Ching* is based on the principle of the yin and yang – the combination of male and female energies, the epitome of sexual reproduction, which forms the basis of heredity and variation in genetics (Wilhelm et al. 1967). Lao Tzu wrote in the *Tao Te Ching* that the Tao gave birth to one, the one to two, the two to three, and the three gave birth to all things (Le Guin and Tzu 1998). This could be just as easily applied as a description of the 64 codons of genetics.

Genetic information, despite the basic simplicity of its nucleotide bases, has the ability to code for basic cellular and organismal functions – such as breathing, using sugars to create energy, reproduction, etc. But aside from such "housekeeping" functions that keep an organism alive, the DNA also codes for form and function – somewhat surprising if you look at the varieties of shapes and sizes of organisms around us (Carroll et al. 2004; Watson and Baker 2013). Though the world is composed of creatures of myriad varieties, there is an underlying commonality among all of them that defines certain body structures that determine the timing of their development and where they are located in the growing embryo. This specific set of information is coded in a set of genes called the Hox genes (Carroll et al. 2004).

The Homeobox (hox) genes essentially define body regions in all animals including humans – responsible for determining two arms, two legs, one nose, and so on. This gene is shared by all living beings – from the snail to the elephant to humans. Despite all the differences, our human form is as similar to that of a goat or a tiger – the coded entities making our foot and also codes for a pig hoof or a chicken leg – and a mouse eye is similar to the snakes and to us (Carroll et al. 2004). The Hox genes are arranged in a specific order in animal chromosomes – with the order of arrangement corresponding to the body alignment from head to the torso and limbs, abdomen, and hind parts. This parallel arrangement of the genes to the body structures is termed as collinearity and is also conserved across evolution (Pourquie 2009). While in mammals there are multiple copies of Hox genes, to create a cushion of redundancy, the underlying set of eight Hox genes and their collinear arrangement with the body remain the same – highlighting the fact that this common principle underlying body plans has been conserved across species through evolution (Carroll et al. 2004).

> Chance is the only true source of novelty.
> *Life Itself*. Crick, F. The Origin and Nature, pg 58 (1982)

With the current trends in technology, we are able to modify and target genes to alter how the information is coded, fast moving toward engineered organisms and humans. Laboratory manipulation of Hox genes is common in animals such as flies and tadpoles. It is commonly used as a means to understand developmental programming and the timing of development of different body structures. When does the brain develop? What structures of the brain develop first? What are the origins of the limbs? Why does the snake have a long torso and not a turtle? Why do octopuses and snails look so different even when they are mollusks? What is the timing of development of the heart? How are changes in the uterine wall during menses related to genes? How do genes play a role in transformation during puberty? These and many more are questions that can be directly answered by studying Hox genes and their mutations (Fig. 1).

Now it is fair game to modify a fly to have legs on its head or grow ears on mice – thereby creating creature blends where the demarcation between human and animal will soon be hazy. What is even more surprising is that hybrid creatures have been a

Fig. 1 Hox gene mutated fruit fly with legs growing on its head

constant feature in our myths – from the winged Pegasus to the snake-haired Medusa to the animalistic Egyptian deities to the many-armed Indian gods (Evslin 1988). Iconography and detailed description of these mythic creatures have prevailed centuries of human civilization. Hybrid creatures have also long served as themes in speculative fiction – from insect-headed humanoids to vampiric predators to superhumans with animal DNA in them (Brem and Anijar 2003). These are interesting times that we live in where myth and speculation are fast becoming reality – which raises the question of how much myth is based on history and how much speculation is based on foresight.

Genetic information is the ultimate game of chance – a success of evolution that is based on heritability, mutations, and variations that promise of survival, species propagation, and enhancement. With the advent of technologies, more accessible and readily available to a wider variety of people, we are now at the cusp of playing and toying with this game of chance (Greely 2003). The Hox Zodiac participatory project has been created to bring forth to the consciousness of the audience the commonality that we share with the animals around us by expanding the idea of their zodiac animal sign.

Hox Zodiac 1.0: Origins, Shadow Hox (2009)

When the artist (VV) and the scientist (SR) first met, the concept of Hox genes as the common blueprint of all animal body plans came up and they agreed that this underlying commonality is important and that – a nuance raising awareness about this would hopefully allow people to relate to the world of animals around them in a more empathetic way. This formed the crux of their collaboration. The intersection of art-science partnerships is not easy. There is a constant struggle between the

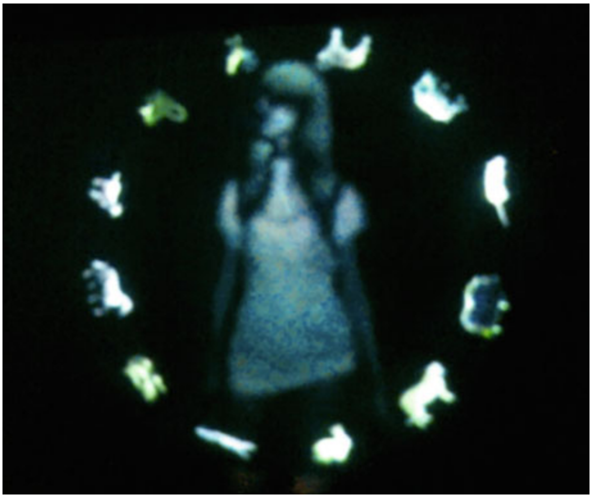


Fig. 2 Shadow Hox, 2009

scientific principle and the artistic interpretation. While the design and the representation by the artist provide the aesthetics and accessibility to the work, without consideration of the underlying scientific principle, the work misses an important dimension. The collaborative process between the artist and the scientist is one with emergent properties, with the final result (at least in this case) being an ongoing series of changes to the initial concept, in order to achieve the best audience immersion. The evolution of the audience-driven games that emerged from this collaboration is a direct evidence of their continued evolution as an artist and a scientist and the amalgamation of their partnership.

The first iteration of the Hox project involved capturing shadows of the audience in front of a screen and then projecting different limbs of animals onto them (Fig. 2). Depending on how near or close the participant was, they acquired the limbs of a horse, wings of a cock, or the body of an elephant. Body parts could be interchanged while preserving the basic body plan. However, the electronic nature of the medium felt remote, and the message of common genetics was not conveyed effectively. Further, it was felt that the interface was not allowing a longer engagement or interaction with others, which was necessary for the layered meanings to be absorbed and appreciated. This led to reimagining the piece to not involve a computer interface but be purposely "unplugged." Having considered using the roulette as an interface and noticing that our discussions often took place around the table led to the reimagining of the table more as a place for discourse about the issues the project was trying to bring to the forefront (Fig. 3).

Hox Zodiac 2.0 at Microwave Hong Kong (2011)

When invited to present the Hox Zodiac project in Hong Kong for the *Microwave New Media Festival*, it was decided to revisit the project from a different angle. One

Fig. 3 Roulette, circa 1920s used as a model

of the issues is that the Hox gene is shared by all living creatures and it was unclear how to narrow it down to a number that made sense. Thinking about presenting this project in China gave VV the idea to use the Chinese zodiac as a framework – it immediately offered the 12 different animals representing the morphological diversity that we encounter in the animal kingdom, and the zodiac offered a rich template to work on. Further, it was intriguing that people already had a default identification with the assigned animal based on their birth and that it came from the Chinese culture that is currently becoming more and more influential globally. Through work with the interface, it was revealed that six of these zodiac animals are commonly used in scientific labs (rat, pig, sheep, dog, rabbit, and monkey), while five others are wild (tiger, bull, snake, horse, rooster), with the dragon being the mythical creature from the past or the mutant hybrid of the future. This pointed to yin-yang symbology that is based on contemporary human-animal relationships and gives another dimension to the idea of the zodiac.

The zodiac allows people to relate to each other on various levels – the fact of having a birthdate grouping you with certain people and the fact of being human with form also grouping you with all forms of life. For instance, zodiac signs are interpreted in relation to career, relationships, health, and wealth. Animal personalities are assigned to humans based on these signs as well – making the zodiac an excellent anthropocentric link for us to work with when talking about common genetic blueprints encoded within all living creatures. With the Hox Zodiac the authors wished to bring forth to the consciousness of the audience the commonality that we share with animals by expanding the idea of their Chinese zodiac sign. Westerners not familiar with the zodiac discovered a new animal they had a relationship with, and those parts of or familiar with Chinese or Asian culture were confronted with a new point of view.

Thus, Hox Zodiac took form as a dinner table for 12 that would seat one person belonging to each zodiac sign. At the center was a lazy Susan (that could be rotated by the audience), with a clear mannequin representing the body as a vessel,

Fig. 4 Hox Zodiac table at Microwave Festival, Hong Kong, 2011

containing an aqueous solution (Fig. 4). Four bottles filled with different colored liquids dripped into the mannequin, representing the four nucleotides that make up our genetic codes A, T, G, and C. Audience manipulation of the lazy Susan would alter the way the color dripped into the "body" – an analogy to the genetic manipulation that is currently happening in labs around the world.

Even as the installation was progressing, the place began to reshape some of our ideas. While SR was on location in Hong Kong, he was looking for materials while jet-lagged and encountered a shop where an herbalist highlighted the connection between traditional Chinese herbs and the zodiac. Much of this information was gleaned by pointing, interpretation, and shaking of heads on both sides. For example, the tiger rules the lung, while the monkey is linked to the bladder (Table 1). It was also discovered that Chinese herbs were ingested that could cure ailments associated with each of these body parts – the snake governs the spleen which is treated with radish seeds, and the ox rules the liver treated with peppermint. Given the fact that the exhibit was to create a dinner table, with guests who were linked to the zodiac signs, it was decided to incorporate these edible herbs into the exhibit (Fig. 5).

The Hox Zodiac dinner was set, and people from the audience who belonged to a specific zodiac sign were seated at the table. Twelve audience members assumed

Table 1 Herbs associated with each body part governed by an animal of the Chinese zodiac

Animal	Body part	Herbs
Tiger	Lung	White mustard seed, genkwa flower, *Agastache*, ephedra, hyacinth *Bletilla*, plantain seed, lily bulb, mulberry leaf, ginseng, *Glehnia* root, magnolia bark, dried ginger, *Lepidium* seed, Chinese yam, tangerine peel, cinnamon twig, chrysanthemum, radish seed, astragalus root, scrophularia root, peach kernel, honeysuckle flower, polygala root, macrostem onion, forsythia fruit, dang shen, *Schizonepeta*, black plum, *Ophiopogon* root, gypsum, wolfberry bark, anemarrhena rhizome, wolfberry fruit, *Isatis* leaf, Cape jasmine fruit, dahurian angelica root, scutellaria root, inula flower, perilla seed, platycodon root, *Pinellia* tuber, *Trichosanthes* fruit, motherwort, kansui root, licorice root
Rabbit	Large intestine	Kansui root, rhubarb, magnolia bark, macrostem onion, areca seed, pumpkin seed, genkwa flower, inula flower, immature bitter orange, hemp seed, *Coptis* root, peach kernel, *Phellodendron* bark, arborvitae seed, honeysuckle flower, *Cistanche*, pulsatilla root, black plum, *Trichosanthes* fruit, perilla seed, bush-cherry seed
Dragon	Stomach	Scutellaria root, inula flower, honeysuckle flower, *Pinellia* tuber, hemp seed, dahurian angelica root, atractylodes rhizome, *Isatis* leaf, scrophularia root, gypsum, white atractylodes rhizome, Cape jasmine fruit, dried ginger, *Agastache*, amomum fruit, oriental wormwood, evodia fruit, *Ophiopogon* root, immature bitter orange, magnolia bark, germinated barley, radish seed, *Quisqualis* fruit, areca seed, pumpkin seed, *Glehnia* root, hyacinth *Bletilla*, notoginseng, grass-leaved sweet flag rhizome, licorice, dendrobium stem
Snake	Spleen	*Pinellia* tuber, radish seed, ledebouriella root, atractylodes rhizome, amomum fruit, poria, *Pueraria* root, *Psoralea* fruit, germinated barley, corydalis tuber, curcuma root, ginseng, dang shen, astragalus root, white atractylodes rhizome, inula flower, *Agastache*, oriental wormwood, *Quisqualis* fruit, tangerine peel, immature bitter orange, evodia fruit, hawthorn, cinnamon bark, Chinese yam, black plum, licorice, Chinese angelica root, white peony root, bitter cardamom, lotus seed
Horse	Heart	*Isatis* leaf, *Isatis* root, prepared *Rehmannia* root, moutan bark, red sage root, forsythia fruit, Cape jasmine fruit, *Coptis* root, poria, dried ginger, cinnamon bark, field thistle, corydalis tuber, *Curcuma* root, safflower, peach kernel, wild jujube seed, rhubarb, arborvitae seed, bush-cherry seed, polygala root, grass-leaved sweet flag rhizome, ginseng, licorice, Chinese angelica root, *Ophiopogon* root, lily bulb, light wheat, lotus seed
Sheep	Small intestine	Bush-cherry seed
Monkey	Bladder	Ephedra, *Trichosanthes* fruit, pubescent angelica root, ledebouriella root, notopterygium root, *Lepidium* seed, tetrandra root, umbellate pore fungus, *Lysimachia*, motherwort
Rooster	Kidney	*Eucommia* bark, *Psoralea* fruit, cinnamon bark, *Cyathula* root, bitter cardamom, arborvitae seed, Chinese yam, moutan bark, dendrobium stem, kansui root, *Epimedium*, hawthorn fruit, fleeceflower root, anemarrhena rhizome, *Cistanche*, lotus seed, pubescent angelica

(*continued*)

Table 1 (continued)

Animal	Body part	Herbs
		root, dogwood fruit, *Lysimachia*, *Coptis* root, mulberry mistletoe, notopterygium root, *Rehmannia* root, wolfberry bark, wolfberry fruit, scrophularia root, sweet wormwood, plantain seed, anemarrhena rhizome, tetrandra root, *Phellodendron* bark, poria, umbellate pore fungus
Dog	Pericardium	Chuanxiong rhizome, cattail pollen, red sage root, *Uncaria* stem with hooks
Pig	San Jiao	Cape jasmine fruit, *Cyperus* tuber
Rat	Gallbladder	*Bupleurum* root, scutellaria root, forsythia fruit, sweet wormwood, *Lysimachia*, oriental wormwood, chuanxiong rhizome
Ox	Liver	Moutan bark, red peony root, black plum, dogwood fruit, *Schizonepeta*, fleeceflower root, mulberry mistletoe, dried *Rehmannia* root, peppermint, mulberry leaf, chrysanthemum, rhubarb, *Bupleurum* root, gentian root, white peony root, sweet wormwood, *Lysimachia*, *Coptis* root, plantain seed, oriental wormwood, cinnamon bark, evodia fruit, motherwort, corydalis tuber, *Cyperus* tuber, germinated barley, field thistle, hyacinth *Bletilla*, notoginseng, cattail pollen, chuanxiong rhizome, red sage root, safflower, peach kernel, *Cyathula* root, wild jujube seed, *Uncaria* stem with hooks, gastrodia tuber, *Epimedium*, *Eucommia* bark, Chinese angelica root, wolfberry fruit

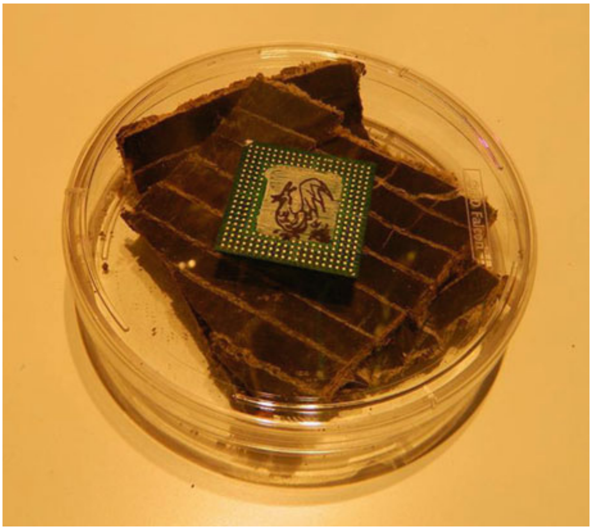


Fig. 5 Herb associated with the rooster

the role of their given animal signs and began a conversation that discussed the industrialization of food, laboratory testing, sacrifice, and animal/human relationships in our contemporary society as well as tapping into the ancient philosophies around this subject. Fear of mutation and the laboratory manipulation of genetics

Fig. 6 Hox Zodiac 2.0 installation and interaction

were being raised as well. They were encouraged to partake and taste from the herbal specimens at their plates, with frequent cross-pollination occurring between participants. All the while, the body at the center was rotated and moved, creating a dizzying alteration in the colored "nucleotides" that dripped in the center (Fig. 6).

"Hox Zodiac" participants related to each other on various levels as humans sharing similar body designs, assuming animal persona and shapes, and as creators of mutant creatures, thereby playing the role of scientists in labs. Experts emerged from the audience themselves, and within moments, a dinner where information was being shared became one of learning from the expert at the table.

While the audience engaged in the Hox Zodiac and the reception was tremendous, one main drawback was that the concept of modified Hox genes resulting in mutated body plans was not emphasized. Audience participants who were more interested in the scientific basis of the project could have explored further via the website, but there was concern that some of the takeaway message was being lost in the installation. Participants also wanted to have food at the dinner table leading to the reimagination of the next version.

Hox Zodiac 3.0: Dinner at MOCA Taipei

The success of the Hox Zodiac dinner at Microwave Festival 2011 led to an invitation to showcase the piece at the *Post-humanist Desire* Exhibition at MOCA Taipei in 2013. The authors continued to push their boundaries in making the Hox Zodiac more interactive and create a dinner table game that would evoke different conversations (Fig. 7).

The *I Ching* reemerged as a tool to play the genetic game of chance. The representation of the lines in the *I Ching* cards resembles the cartoons of chromosomes in cell biology textbooks – with the broken lines in the *I Ching* appearing

Fig. 7 Hox Zodiac at MOCA Taipei, 2013

similar to the process of crossing over and genetic information transfer during meiosis in our cells. The process of crossing over in the chromosomes where genetic material is exchanged is how we get variation in our genetics and how we appear as individuals and not identical clones of our parents. This basis of heredity and variation is once again a universal principle among all creatures (Watson and Baker 2013).

Further, certain *I Ching* cards represent specific signs of the Chinese zodiac (Fig. 8) and have also been linked to the coding sequences underlying amino acids (Fig. 9, Castro-Chavez 2012).

Other patterns began to emerge – the Hox genes are activated in a time-dependent manner; if things go wrong, the head would develop later than the foot, leading to many developmental defects (Graba and Rezsohazy 2013). The Chinese zodiac itself is represented as a clock, with the rat beginning at 11 pm–1 am moving forward with the ox (1–3 pm) and so on all the way to the pig (9–11 pm). The goal was to provide a game framework that could appear highly simplistic but has layers of details underneath for the interested participant (Wu 2010).

The Hox Zodiac presented at MOCA Taipei was set up as a circular table with a large translucent egg in the middle. This represented the proto-embryo, the potential of what could be once the genes play into its development. Twelve lab coats

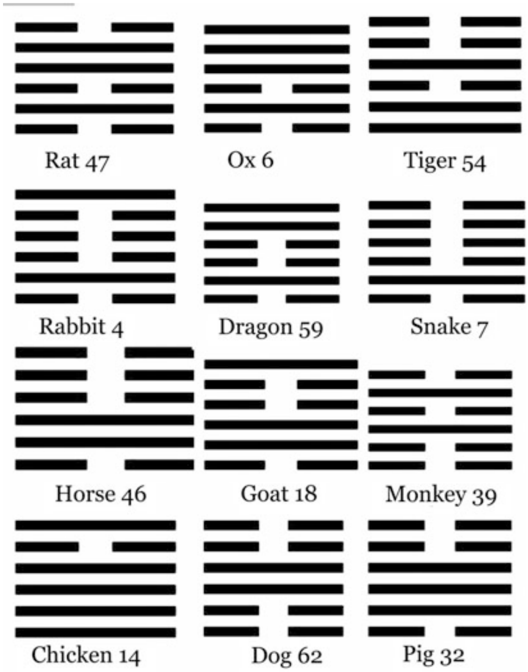


Fig. 8 *I Ching* hexagrams representing Chinese zodiac signs

with the symbols of the Chinese zodiac painted on them were lined up against the wall. The table was set for 12 – one for each animal of the zodiac. Audience participants belonging to each zodiac sign were invited to sit at the table. Once the table was full, the game began.

At each place setting was present a petri dish containing the Chinese herbs associated with the zodiac animal (Table 1), a repository for the DNA of the participant (either hair or saliva), dining tools in the form of forceps, a little repository of little known facts about the animal, and finally a drawing book. A set of cards with 12 symbols from the *I Ching* corresponding to each animal (Wu 2010, Fig. 10) and a "mutant" human card were also placed at the table.

The rules of the game were initially set but were flexible based on the audience participation and level of complexity. Every animal at the table started by placing their DNA in the repository – this was their essential commitment to the game. Then all the animals began by drawing their own head (rat with a rat, ox with an ox, and so on). The rat always began first (as the head of the clock). It started by reading a fact about the animal – these ranged from medical use of animals, animal use in cooking, or in herbs or any form of biological role. Then the rat drew from the deck of cards. This new animal now became part of the rat mutant and was drawn into the books. The play then shifted to the animal on the right.

Within a few iterations, a whole new combination of mutant creatures emerged – monkeys with horse limbs, snakes with rabbit ears, and a mishmash of different

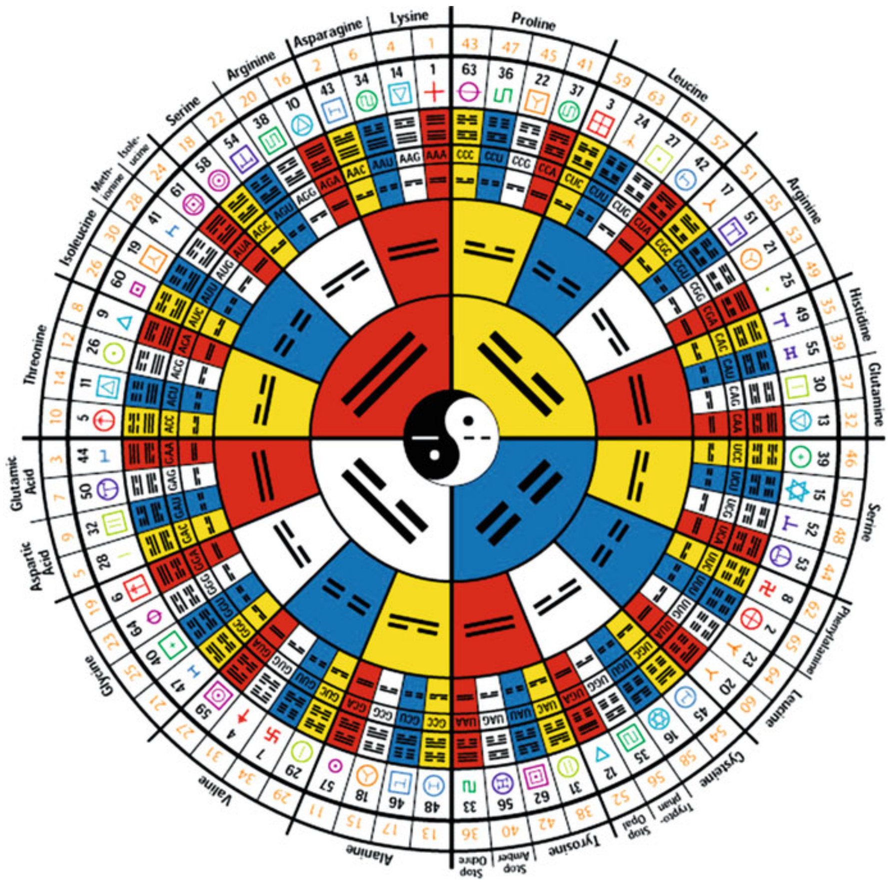


Fig. 9 *I Ching* representation of amino acids

creatures. Artists began to emerge among the participants, many of whom informed us that they had not engaged in such creative exercises in decades. They ranged from young children to older experienced scientists, who all explored the role of the mutants, learnt from the factoids and the herbs, and created hybrid creatures in the process (Fig. 11).

HOX 4.0: Zodiac Dinners

> what is it that's not art & not science? I asked, & I finally got the answer, discourse, which stands in both, and behind both. It is our discussed world, so brought into being.
> Robert Kelly 1974

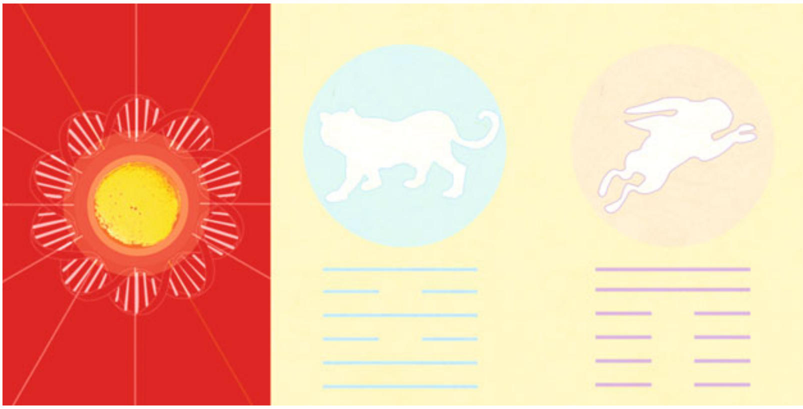


Fig. 10 The *I Ching* zodiac cards

Fig. 11 Audience draws mutant hybrid animals

Since hosting the Hox Zodiac in Taipei, the project has undergone yet another revision. Paring down the installation models suitable for museum set pieces, the new avatar of the Hox Zodiac resembles more closely a simple dinner table set for 12, with the conversations and the food becoming more of the focus.

Dinners have been organized at the UCLA Art | Sci Center (December 2014), at a private artist residence (February 2015), as a part of the College Art Association (February 2015) and for high school students taking part in the Sci | Art Nanolab (July 2015). The participants at these dinners have ranged from groups of 12, specifically invited for the zodiac dinner party game, to large crowds of up to

Fig. 12 Hox dinner for 12 at UCLA Art | Sci Center, 2014

80 people, depending on the audience, level of knowledge, and time the game is scaled from intimate gatherings to event specials.

The focus is brought back to the idea of animals, their different body plans, and how they can be manipulated in labs. Participants enter the dinner game relating to the animal portrayed in the Chinese zodiac. Food is served at the table based on food groups relating to each animal of the zodiac, and whatever one ingests symbolically brings into the genetic mutation. Research into the Chinese signs revealed that each animal is associated with specific food groups that are recommended for curing ailments relating to that sign, as well as for continued health and well-being. Based on these sources, a menu was created and food served at the first dinner hosted at the UCLA Art | Sci gallery in December 2014. Participants were seated at a table for 12 at their specific zodiac signs. Each "animal" was served an appetizer, entrée, and dessert specific for its sign. When food was shared, it was noted as sharing of traits between animals, creating hybrids and mutants (Fig. 12).

Patterns emerged – even though the seating was for 12, frequently there were excess audience in the room, leading to packs of tigers, huddles of sheep, and a horde of monkeys at certain parts of the table. Some participants were unwilling to eat food recommended for their own animal, having preference for others. Some specifically wanted traits of other animals, so willingly bartered for food across the table. All were provided with notebooks for drawing and making notes and comments about the experience, and chimeras were duly formed. The audience loved the interaction, the edible nature of the evening, and the sharing of traits and ideas. The bare-bones simplicity of the dinner table helped aid the process.

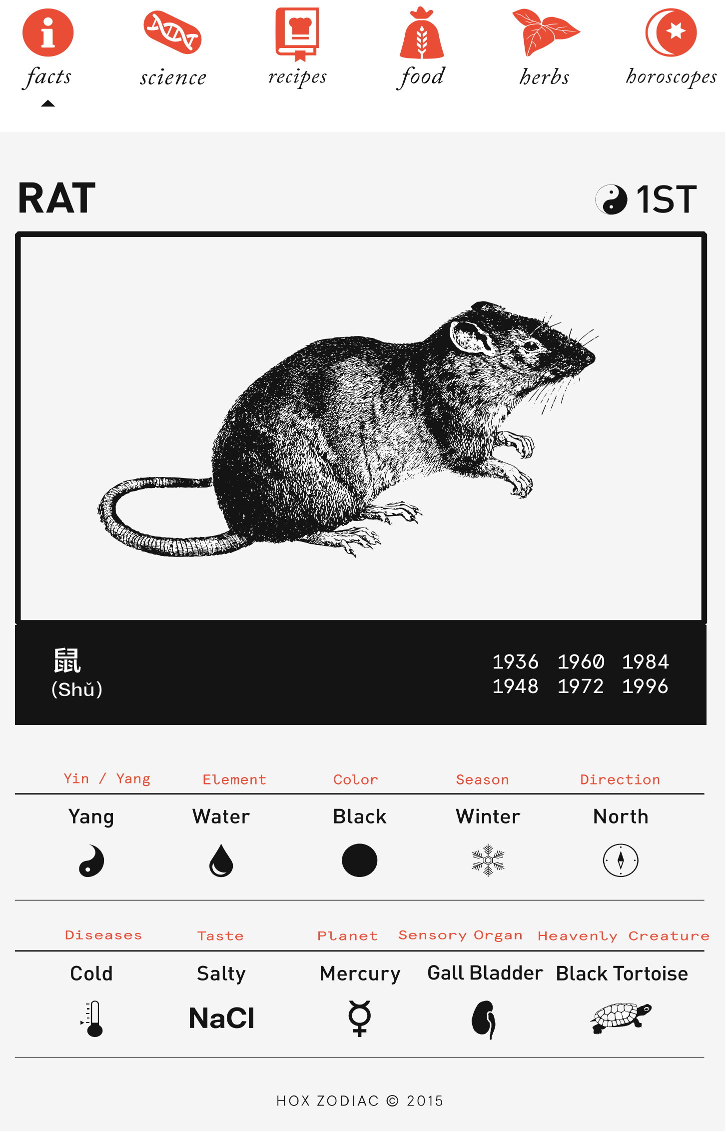


Fig. 13 Hox Zodiac information playing cards

In large installations, audience members were invited and seated in groups of ten around a number of round tables. After a brief introduction to the concept of the Hox genes and their importance in body plans, food was served pertaining to each of the animal signs of the zodiac, starting with the rat. As food was served, the uses of each animal in scientific research were highlighted, and the Chinese zodiac qualities of the animal were also expounded. Participants would imbibe foods of

Fig. 14 Hox Zodiac private dinner at Linda Weintraub's homestead, February 14th, 2015

animals whose qualities they wanted to incorporate or if it was a delicacy that they enjoyed. It was interesting to note that when some of the scientific uses of animals were elaborated, it altered perceptions of some participants, and they would then either accept or refuse the food selections. As they ate, participants were encouraged to draw their chimeras on their plates, with the qualities that they had gained from creating such hybrids.

The setup for a larger participatory public was pretty much in place, and the next challenge was to figure out a more formal way to play the Hox Zodiac in a smaller private setting. This opportunity came in the form of an invitation by author, environmentalist, and curator Linda Weintraub who was interested to test this at her homestead in upstate New York. Having a farm with animals, she contributed to the evolution of the project by suggestions such as asking each of the invited guests to bring their offering based on their sign. They were sent a set of ingredients along with information about their sign well in advance, and all came with their food prepared. Each of the 12 guests presented their food offering with a narrative, offered to the rest, and cleaned up. Thus, a 12-course dinner emerged and gave a new meaning not only to the zodiac but also to the idea of slow food and creation of an environment for meaningful dinner table discussions (Figs. 14 and 15). In a smaller setting, it became clear that it would not be so easy to have 12 people with all signs represented, and this led to the decision to highlight the absent animals.

Fig. 15 Hox Dinner setting at private dinner hosted by Linda Weintraub

Thus, it was decided that the hosts prepare dishes and present information about the animals not present.

While there will be initial suggestions of food groups and menus associated with the dinners and zodiac signs, over time audiences who indulge in the dinners will add to the menu repertoire, creating novel dishes suited for pertinent signs. The future iterations of the Hox Zodiac dinner game will become more minimalistic, with the setting and the basic ingredients provided but with the details created by the participants themselves. Hox Zodiac cookbooks are available to use and add recipes to online.

The idea is to create the Hox Zodiac in a box that will have all the elements for anyone to set up a "dinner table". The contents will be table cloth with zodiac signs, plates, petri dishes, and a set of books that give information about individual signs: general zodiac meanings, evolutionary, how it is used in science laboratories, the benefits and abuse, and as food (Fig. 13). In public presentations, we supply lab coats with the zodiac signs and empty lab notebooks that audience fills in with their findings. The basic setup is the following:

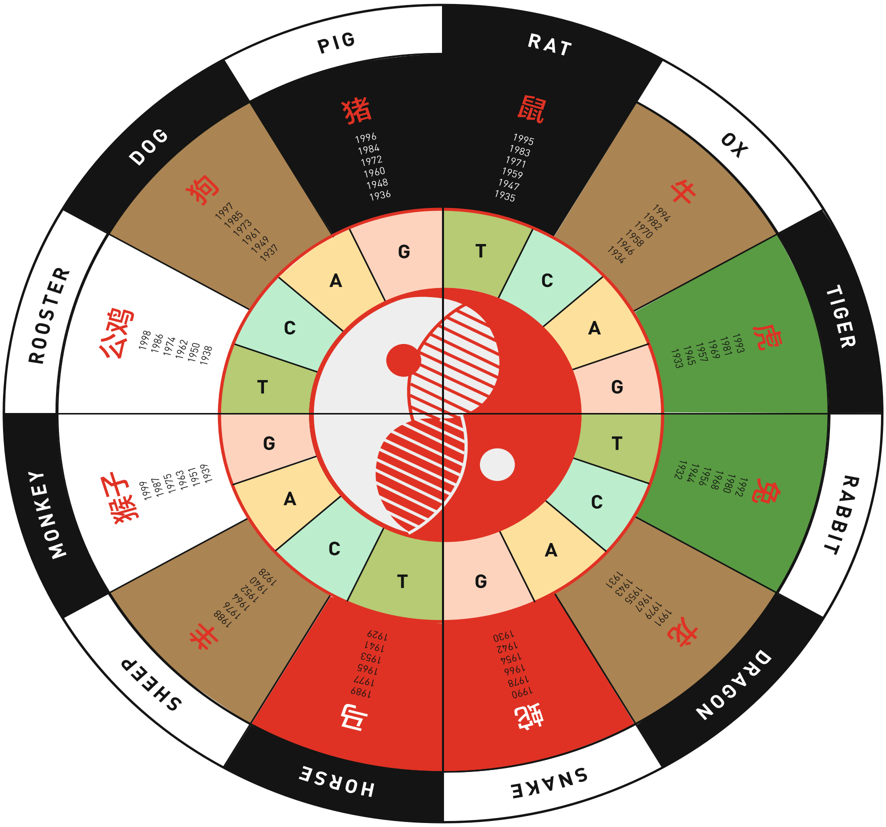


Fig. 16 Hox Zodiac interface

People sit down at their animal sign as assigned by the Chinese zodiac. The first 30 min is spent on learning about themselves and making notes on thoughts, responses, and questions. For appetizer, basic grain dishes and drinks are offered, ideally from local origins, and a menu with the various animals is passed around. Herbs associated with their signs are the plates already in the petri dish and description of the organ.

Appetizer: The first question asked is – would you eat yourself, i.e., the animal that you are represented by? If not, which animal would you eat? You have to survive so a choice is necessary, and by picking one, you are starting to evolve your being to a new variation. Either remain within yourself or start a dialogue with the animal you need to continue on. You write down your recipe and/or draw what you are creating by your choices.

Main dish: Choose something related to scientific use of your animal if one of the lab creatures, and/or find out which animal is used in relation to your organ. Write out your dish based on this.

Fig. 17 The HOX Zodiac Logo

Desert: If you wanted to create a hybrid of yourself, which of the animals would you use? What mythical/fantastic powers will you gain by this? What would this mythical beast taste like, smell like, feel like, and see like? Write your recipe for this.

Drink: By drinking the essence of the animal, you change the organs of your body. What aspect of your body do you want altered? What do you feel about this? Given such a miracle essence, would you take it (Figs. 16 and 17, Tables 2 and 3)?

Conclusion

The authors have spent a few years developing the Hox Zodiac project with the initial goal of introducing audiences to the Hox gene and our animal nature through genetic similarities. Experimentation with different approaches was deliberately open ended and allowed the audience to shape it. Over a period of 5 years, with a number of iterations, public and private, the project has evolved into a dinner table game. The participants bring to the table ingredients and share the knowledge as they ingest the food and consider the inherent deep connection to the animal world. The idea of the zodiac is expanded to contemporary scientific use of animals and discussion of food and while creating hybrids and chimeras also considering the influence of Eastern, in this case specifically the ancient Chinese culture.

Table 2 Laboratory use or evolutionary research of animals in the zodiac

Animal	Use in research
Rat	Rats and mice serve as primary genetic models of diseases Rats are predominantly used in understanding mechanisms of learning and memory tests Rats are also used extensively in addiction studies for alcohol and drugs Cancer models are also created in transgenic rats and mice
Ox	Used to extract fetal bovine serum used in all cell culture experiments Cow eye dissection is used to teach anatomy Milk used as blocking agent in immunochemistry
Tiger	Cats are used to study locomotion and spinal cord control Tiger vocalizations and vocal chords are studied The craniomandibular ontogeny (development of the jaw) is a prime field of study from both a current and evolutionary perspective
Rabbit	Many antibodies used in staining protocols are made in rabbits The rabbit retina is an excellent model to study vision Used as a model in hepatitis E infection and vaccination Eyeblink conditioning reflex in rabbits is a standard model of study
Dragon	Dragon is a neuronal adhesion protein Komodo dragon feeding is an established area of study
Snake	Snake venom has new uses in cancer therapy The snake venom was an important tool in understanding neuromuscular control using acetylcholine
Horse	Horses were historically used as model anatomical organisms in medical schools Horses are now being studied as model organisms for depression Muscle control and metabolism are studied using horses
Sheep	Sheep are used in sleep research Hormonal regulation is studied using sheep model systems Rams are used to study homosexuality in the animal kingdom Goat serum is extensively used in antibody staining techniques
Monkey	Monkeys are studied extensively in relation to their visual circuitry Motor control of movement and mirror neurons are studied in monkeys Monkeys serve as models for depression and neurodegenerative diseases Brain-computer interfaces are extensively tested on primates
Rooster	Embryogenesis is studied using chick models Chickens were the model system to study cholera The first oncogenes were discovered using chickens We can now create transgenic chickens as their genome has been sequenced
Dog	Pavlov's conditional learning is a famous example of dogs in experiments Dog genomics are powerful to understand evolution Dogs are models for studying hepatic chemotherapy Dogs may also help us understand autoimmune diseases
Pig	Pigs are used as models to understand cystic fibrosis Pigs are also serving as host models for culturing artificial organs Pig neonatal models are more comparable to humans than mice Pigs are also used to study inflammatory bowel disease and other gut diseases

Table 3 Factoids by animal

Rat
The rat rules the gallbladder Winter Psychopathy The rat is quick-witted, clever, charming, sharp, and funny You should be happy in sales or as a writer, critic, or publicist Stewed cane rat: Skin and eviscerate the rat and split it lengthwise. Fry until brown in a mixture of butter and peanut oil. Cover with water and add tomatoes or tomato purée, hot red peppers, and salt. Simmer the rat until tender and serve with rice A rat can eat a third of its body weight each day. They are true omnivorous scavengers but mostly prefer grain, livestock feed, and meat
Ox
The ox rules the liver Winter Schizophrenia The ox is steadfast, solid, goal-oriented leader, detail oriented, hardworking, stubborn, serious, and introverted but can feel lonely and insecure The buffalo would be successful as a skilled surgeon, general, or hairdresser Oxtail stew: Season oxtails with salt and pepper. Add to pot; cook, turning once, until golden brown. Add tomato paste, ginger, garlic, onions, carrot, and celery; cook until soft. Return oxtails to pot with stock, allspice, thyme, and chilies; cook and cover until oxtails are tender Oxen can eat forage, hay, grass, and grains They can eat 30 lb per day Mad cow disease is transmitted from bovines to humans Bovine tuberculosis has the ability to affect us
Tiger
The tiger rules the lung Spring Hysteria Tigers are authoritative and self-possessed, have strong leadership qualities, and are charming, ambitious, courageous, warmhearted, highly seductive, moody, and intense; and they're ready to pounce at any time You would be excellent as a boss, explorer, race car driver, or matador Tiger sauce: Worcestershire sauce, distilled white vinegar, tamarind nectar, hot cherry peppers, sugar, water, Marsala, sea salt, and pectin. All ingredients should be mixed, boiled for 3 min, simmered for 20 min, and then mixed in a blender They mainly eat large mammals, such as pigs, deer, antelope, and buffalo A tiger can consume up to 40 kg of meat at one time, but individuals in zoos are given 5–6 kg per day Tiger conservation is an important topic as most tigers across the world are endangered species
Rabbit
The rabbit rules the large intestine Spring Schizophrenia Rabbit enjoy being surrounded by family and friends. They're popular, compassionate, and sincere, and they like to avoid conflict You are successful in business but would also make a good lawyer, diplomat, or actor Place the browned rabbit pieces into the Dutch oven, and pour in the beer and chicken stock.

(*continued*)

Table 3 (continued)

Bring to a boil over high heat, then reduce heat to medium low, cover, and simmer until the rabbit is very tender, 25–30 min Wild rabbits spend much of their time foraging in meadows and fields eating wild grasses, herbs, flowers, leaves, and twigs. They also consume some fruits, vegetables, and grains A rabbit should be given about 2 lbs of veggies per 6 lb of body weight
Dragon
Dragon rules the stomach Spring Psychopathy Dragons are energetic and warmhearted, charismatic, lucky at love, and egotistic. They're natural-born leaders You would be well suited to be an artist, priest, or politician Blue dragon**:** Mix three parts vodka and one part blue curacao with crushed ice, shake or stir, and strain into a martini glass. Garnish with orange slice and cherries Dragons are known to eat any living thing that has flesh. These include bears, killer whales, giant squid, birds, walrus, leopard seals, and other animals Dragon boat racing has a positive effect on breast cancer survivors The venom in the saliva of Komodo dragons and its role in predation
Snake
Snake is in charge of the spleen Summer Paranoia Snakes are seductive, gregarious, introverted, generous, charming, good with money, analytical, insecure, jealous, slightly dangerous, smart, hardworking, and intelligent, and they rely on gut feelings Snake would be most content as a teacher, philosopher, writer, psychiatrist, and fortune teller Recipe: Kill a viper, skin it, and take out the entrails; cut the flesh into small pieces and put it into the broth, with the heart and liver cut across, two blades of mace, and a small bit of cinnamon; cover it up; and let it boil till it is reduced to a pint; by this time the flesh of the viper will be consumed; strain it off and press it very hard Snakes will eat birds, fish, and eggs. Some snakes eat insects and spiders. There is one type of snakes that will eat plants They usually only eat once every 7–10 days, and they get older and larger only about once a month Snakes use infrared "vision" to navigate the world Snakes modulate constriction in response to prey heartbeat
Horse
Horse governs the heart Summer Anxiety Horses love to roam free. They're energetic, self-reliant, and money-wise, and they enjoy traveling, love, and intimacy You should be an adventurer, scientist, poet, or politician Recipe: In a mix of 90 % olive oil and 10 % truffle oil, fry sliced onions and mushrooms for 5 min. Add green pepper slices, and fry 2 min more. Throw in a 60/40 mix of shredded roast beef/smoked horse, and stir for a few minutes till the roast beef is browned. Add some black pepper. Add a few slices of cheese on top and put a lid on the pan Horses eat eggs, chaff, apples and carrots, bran, barley, linseed, maize, horse nuts or mixes, cod liver oil, root vegetables, oats, salt, seaweed, dried sugar beet pulp, and mollichop

(*continued*)

Table 3 (continued)

An adult horse should eat between 1.5 and 3.5 % of its body weight a day. That comes out to roughly 15–35 lbs for an average 1,000 lb horse Individual recognition among domestic horses extends to humans
Sheep
Sheep govern the small intestine Summer Anxiety Goats enjoy being alone in their thoughts. They're creative, thinkers, wanderers, unorganized, high strung, and insecure and can be anxiety ridden You would be best as an actor, gardener, or beachcomber Cut the lamb into eight pieces and flatten with your fist, and then toss with salt, pepper, and the garam masala. Put into the large frying pan with one tablespoon of oil, turning when golden. Toast the cumin crunch mix in the medium frying pan until lightly golden, and then pound in a pestle and mortar Mostly sheep eat grass, clover, forbs, and other pasture plants Sheep eat about 2–3 % of their body weight
Monkey
Monkey rules the urinary bladder Autumn Personality disorder Monkeys thrive on having fun. They're energetic, upbeat, and good at listening but lack self-control. They like being active and stimulated and enjoy pleasing self before pleasing others Your sign promises success in any field you try Three-legged monkey**:** 1 oz whisky, 1 oz amaretto almond liqueur, 1 oz pineapple juice. Shake and strain into rocks glass Some eat only plants, and some eat both plants and animals. They are known for liking fruit, leaves, insects, and other small animals An average monkey eats about 7 lb a day
Rooster
Rooster governs the kidneys Autumn OCD (obsessive-compulsive disorder) Roosters are practical, resourceful, observant, analytical, straightforward, trusting, honest, perfectionists, neat, and conservative Recipe: With fingertips, separate skin from meat on each breast half. Rub equal amounts of seasoning mixture under skin of each breast. Sprinkle chicken with salt and pepper. Place chicken on grill over medium heat, and cook about 25 min or until juices run clear when thickest part of breast is pierced with tip of knife, turning over once Chickens are omnivores and will eat grain, seeds, fruit, other vegetation, corns, worms, and other insects A laying hen will require approximately one fourth to one third of a pound of feed every day or roughly 1.5 lb of feed per week

(*continued*)

Table 3 (continued)

Dog
The dog rules the pericardium
Autumn
Dependence disorder
Dogs are loyal, faithful, honest, distrustful, often guilty of telling white lies, temperamental, prone to mood swings, dogmatic, and sensitive. Dogs excel in business but have trouble finding mates
You would make an excellent businessman, activist, teacher, or secret agent
Suyuk (a boiled dog meat): Put dog meat, gravy, and ingredients into pan, steam it with a weak fire, and eat with sauce. One can eat with boiled leek. Perilla powder, perilla oil, mustard, and vinegar are to be added to the sauce
The amount of food that dogs eat will all depend on the weight of the dog. Some dogs will easily eat around 25 lb of dog food a week, and some will only eat 1 lb a week
They do eat meat, but they also eat vegetable and grain matter in small amounts
Pig
Pig rules San Jiao
Winter
OCD (obsessive-compulsive disorder)
Pigs are extremely nice, good mannered, and tasteful. They're perfectionists who enjoy finer things but are not perceived as snobs
The pig would be best in the arts as an entertainer or possibly a lawyer
Heat oil in a large nonstick skillet over medium-high heat. Flatten each pork piece to 1/2 in. thickness using your fingertips. Add pork to pan; spoon soy sauce mixture evenly over pork slices. Cook 3 min or until browned. Turn pork over; cook 3 min or until done
Farm grains such as corn, barley, oats, and wheat make up a balanced diet for a pig
Approximately 3–5 lb per day

Recommended Reading

S.K. Brem, K. Anijar, The bioethics of fiction: the chimera in film and print. Am. J. Bioeth. **3**(3), 22 (2003)

S.B. Carroll, J.B. Grenier, S.D. Weatherbee, *From DNA to Diversity* (Wiley-Blackwell, Malden, 2004)

F. Castro-Chavez, Defragged binary *I Ching* genetic code chromosomes compared to Nirenberg's and transformed into rotating 2D circles and squares and into a 3D 100% symmetrical tetrahedron coupled to a functional one to discern start from non-start methionines through a Stella Octangula. J. Proteome. Sci. Comput. Biol. **1**, 3 (2012)

D. Cooper, R.P. Lanza, *Xeno: The Promise of Transplanting Animal Organs into Humans* (Oxford University Press, Oxford, 2000)

F. Crick, *Life Itself: Its Origin and Nature* (Simon and Schuster, New York, 1982)

B. Evslin, *The Chimaera (Monsters of Mythology)* (Chelsea House, NYC, 1988)

Y. Graba, R. Rezsohazy, *Hox Genes: Methods and Protocols (Methods in Molecular Biology)* (Humana Press, New York, 2013)

H.T. Greely, Defining chimeras and chimeric concerns. Am. J. Bioeth. **3**(3), 17–20 (2003)

A. Huang, *The Complete I Ching*, 2nd edn. (Inner Traditions, Rochester, 2010)

M. Jensen, John cage, chance operations, and chaos game: cage and the "I Ching". Musical Times **150**(1907 (Summer)), 97–102 (2009)

R. Kelly, Biopoesis, Io Magazine, V 20, (ed.), Bialy, H., (Vermont Press, 1974)

K. Larson, *Where the Heart Beats: John Cage, Zen Buddhism, and the Inner Life of Artists* (Penguin, New York, 2013)
U.K. Le Guin, L. Tzu, *Lao Tzu: Tao Te Ching : A Book About the Way and the Power of the Way* (Shambala, Boston, 1998)
J. Monod, A. Wainhouse, *Chance and Necessity: An Essay on the Natural Philosophy of Modern Biology* (Vintage Books, New York, 1972)
B. Pierce, *Genetics: A Conceptual Approach* (WH Freeman Publishing, New York, 2013)
O. Pourquie, (ed.), *Hox Genes: Current Topics in Developmental Biology*, vol 85 (Academic Press, Massachussetts, USA, 2009)
M. Schönberger, *The I Ching & the Genetic Code* (Aurora, Santa Fe, 1992)
J.D. Watson, T.A. Baker, *Molecular Biology of the Gene*, 7th edn. (Benjamin Cummings, Menlo Park, 2013)
R. Wilhelm, (Translator), C.F. Baynes, (Translator), H. Wilhelm, (Preface), C.G. Jung, *The I Ching, or, Book of Changes (Bollingen Series XIX)* (Princeton University Press, New Jersey 1967)
S. Wu, *The Definitive Book of Chinese Astrology* (New Page Books, Franklin Lakes, 2010)

Interface-Centric Art Games

32

Christa Sommerer, Ulrich Brandstätter, and Laurent Mignonneau

Contents

Abstract

The design of game interfaces is a decisive factor for game perception and play experience; in many cases, interfaces are tightly interwoven with the gameplay. According to Jesper Juul "it is rare to find a clear-cut border between interface and gameplay and that the fluidity of this border characterizes games in general." Currently, input devices are regarded as "the most neglected aspect of video games."

The central argument of this article is that alterations to the interface can affect both how the game is perceived and the play experienced; this is backed by some well-known examples from the field of game arts, for example, the *Giant Joystick* of Mary Flennigan and the *PainStation* of Tilman Reiff and

C. Sommerer (✉) • U. Brandstätter • L. Mignonneau
Interface Culture, University of Art and Design Linz, Linz, Austria
e-mail: Christa.Sommerer@ufg.at; Ulrich.Brandstaetter@ufg.at; Laurent.Mignonneau@ufg.at

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_14

Volker Morawe. In both of these, a transformation of the gameplay is realized solely by altering the interface.

Specifically created user interfaces have several key advantages, as they are able to transform available games into more intimate and rewarding experiences, to improve game accessibility, to put players in a ludic state of mind, and to communicate both overall game goal and some of the relevant game rules even before the game commences. These attributes are demonstrated on the basis of relevant art game installations that have been created by students and graduates of the Interface Cultures department at the University of Art and Design in Linz, Austria.

Keywords
Videogames • Media art • Interaction • Interface • Gameplay

Introduction

In the past years a great deal has been written about how games are related to art. In two special issues of *Kunstforum* magazine, scholars and artists contemplate on the long-lasting relationship between art and play (Buchhart and Fuchs 2005a, b).

In the late seventeenth century German poet Friedrich von Schiller pointed out that artists transform realities of nature (*Naturwirklichkeit*) into art realities (*Kunstwahrheit*). He was trying to attain a union between the instinctive feeling for eternal unchanging moral and other truths (*Formtrieb*) and the constantly changing feelings, perceptions, and urges (*Sinntrieb*). He thought that the urge to play (*Spieltrieb*) is able to mediate between these two instincts, eventually giving rise to artistic beauty and freedom (Schiller 1879).

Marcel Duchamp is also often quoted as one of the first artists to elevate the mechanics of play (in his case a chess game) into an art form in its own right. Doing so, he challenged the traditional notion of art making. Gameplay was thereby used as a strategy to unmask the common notions of art. Many artists such as Yoko Ono, Valie Export, Julian Wasser, or Takako Saito in the 1960s and 1970s developed this idea further (Fuchs 2005–2006).

It is generally agreed now that art and play have several common properties: they both are not bound to a specific utilitarian use, and they both have a certain frameset within which they operate. In his philosophical hermeneutics German philosopher Hans-Georg Gadamer describes this frameset as a double-sided system: while the player becomes part of the game, he or she is becoming played with and incorporated into the game (Gadamer 1990).

Games have often been considered as a lower form of art, and Ernst Strouhal notes that, according to aesthetic theory, games are often becoming popular in times of crisis (Strouhal 2008). Whatever the context, more and more examples of artistic games in recent years deal with the very question of what it means to play and which concepts and values are hidden behind them.

Mary Flanagan et al. write extensively about the values of games. To them all games express and embody human values (Flanagan and Niessenbaum 2014). They also observe that games not only carry certain values that designers embed into these systems (such as trust, fairness, etc.). They see games as powerful activation tools for new belief systems. In the philosophy of technology value systems are studied in detail as well, as political and societal consequences are shaped by technological innovations and products.

The question of value systems is also interesting in respect to games that are related to art. They allow artists to create settings where traditional notions of art and culture can be challenged and literally "played with."

In the 1990s some scholars such as Florian Rötzer even went as far as to call computer games the new art form of our times (Rötzer 2005). The philosopher Vilem Flusser also called game theory the fundamental theory of the future, by pointing out that it is exactly the balance between rules and randomness that makes games so fascinating (Flusser 1987, 1993). The connection between limitations and openness has been a key topic for artists since the 1950s. Notions of the "Open Artwork" as defined by Umberto Eco (1989) or participatory concepts as exercised in Fluxus, Gutai, and Happenings of the 1960s had a big impact on how art and games can be connected now. Dieter Buchhart gives a good overview on the dialectics between game rules and the open frameworks of avant-garde art up until the present (Buchhart 2005).

In the 1990s the notion of the open artwork was extended by interactive installations that started to use algorithmic principles, generative design, evolutionary algorithms, and interactive technologies. An example of this time is the "A-Volve" system by Christa Sommerer and Laurent Mignonneau from 1994 (Sommerer and Mignonneau 1997). Participants could create artificial life forms by drawing any shape on a touchscreen. These forms then became alive and "swam" in a pool filled with water. Their behavior and survival was directly linked to their design. Design represented their artificial genetic code. Participants could create artificial creatures, protect them, help them propagate, or even kill them. While this installation might look like a game at first sight and does require a certain amount of playful spirit, the main intention of the artists was however not to create a game.

Rather it was to show creation of life in a metaphoric, symbolic, and artistic way. Another motivation was to show how user interaction and participation can lead to open-ended complex adaptive systems that emerge from intricate interactions between the internal algorithms and the participant's interaction parameters. In a "pool of life" setting the art piece asks questions on what life is; how it evolves, propagates, and survives; and how creation per se can be influenced and perceived.

Media artists and theorist Peter Weibel points out that often a confusion happened with interactive art pieces, when they were only called to be playful. While certainly several interactive installations from the 1990s do have playful aspects and encourage active involvement, it is the algorithmic quality and the metaphor of the open process that is more important and which can also be linked to earlier participatory art forms of the 1960s and 1970s (Weibel 2005).

Carl Goodman writes about the creative freedom in game art that has recently been discovered. To him there is now an "expanding ecology around the robust commercial center of the field that also includes fine artists, independent producers, academics, researchers, and game designers who have earned enough clout, or money, to work with complete creative freedom" (Goodman 2007).

When asked about the current state of the video game medium, Jesper Juul explains the implications of the newness of video gaming. He states,

> First, I should like to describe where we are, so to speak. The computer game is somewhere between 39 and 42 years old. Movies, by comparison, are 100 years old. So computer games are younger than movies, but not that much younger. So the computer game has a fair amount of history. While I do not think the computer game has had it's first Shakespeare - in itself a bad comparison since computer games are mostly built by groups - I do think it has had it's first, say, Dashiel Hammett. (. . .). (Juul 2000)

Henry Jenkins shares a similar viewpoint but indicates the flawed comparison of video games with other media:

> If the question is, 'Will video games become a serious art form in its own right?' I think the answer is inevitably yes. Whether the analogy is to literature or to dance or to cinema or to theater or any number of other media, it's hard to know what the right approximation is. In a way, to ask the question that way is like asking 'Will cinema become theater?' (Deen 2011)

Over the course of this article, the proximity of interfaces and video games will be examined more closely; current developments, including game accessibility, controller misuse, and video game HCI research, will be highlighted. After a brief elucidation on the exerted classification methods, the main argument is made on top of several paradigmatic interface-centric art games. The article closes with a brief summary.

Interfaces and Games

According to a popular saying among designers, designing the spoon means designing a particular way of eating; figuratively, designing a game interface affects and alters the play experience. Game controller design involves implications upon the economic success of whole game console infrastructures, as innovative interaction device design is increasingly becoming unique selling points for successful video game systems (e.g., the Nintendo Wii). Also, current video game titles tend to include gaming aspects beyond the primary player screen by using specific interaction devices: often, second screens are deployed to feature more personal game information (Nintendo Wii U, Microsoft Xbox smart glass, Sony remote play), and game controllers are endowed with multisensory player feedback mechanisms (vibrotactile force feedback, integrated speakers, camera integration, etc.); of particular interest is the current tendency of major game development companies to focus on full-body player interaction, which is usually realized with the help of depth camera technology (e.g., MS Kinect, Leap Motion, stereoscopic cameras).

For the sake of usability, many commercial titles cut through the fourth wall by incorporating in-game references to controller layouts, button key bindings, key sequences, etc. Even without in-game references, the hardware designs of game controllers implicitly explain some of the rules to the players, and this is intensified as the players accumulate increasing experience with the relevant devices.

Current research on gaming devices makes their effects on the play experience apparent. First, as indicated by Lenz (2008), the game interface has an effect on the game performance of the player, a circumstance that has an enormous effect on the evolution of commercially available game controllers toward ergonomically designed multipurpose game pads to a certain degree (Lu 2003; Ybud 2014). However, optimizing the controller of gaming systems for the main audience brings about a disadvantage for other audiences, as was pointed out by Fron et al. (2007). It is hardly surprising that the interface also has an effect on the difficulty of the video game (Johnson 2010). However, in many cases authors of relevant game studies (Barr 2007) do not deal with custom-tailored interfaces for several reasons: they have a relatively low impact, they are very heterogeneous, and in many cases they are well suited for specific tasks but are difficult to employ as general-purpose gaming devices.

Controversially, *Dogma 2001* (Adams 2001) by Ernest Adams, a voluntary set of rules for game designers, limits game interfaces to standard mouse/keyboard/joystick/pad. He states this in rule three:

> Only the following input devices are allowed: on a console machine, the controller which normally ships with it. On a computer, a 2-axis joystick with two buttons, or a D-pad with two buttons; a standard 101-key PC keyboard; a 2-button mouse. Justification: Most games that depend on gimmicky input devices are crummy games. You must not waste your time trying to design for them.

We the authors of this paper disagree, for several reasons:

– Input devices are an integral part of the game experience, and therefore should always be considered.
– There are many examples of great games that use special interfaces, ranging from independent productions to well-received and commercially successful AAA games (e.g., Guitar Hero or Nintendo WII Sports, which is considered the most successful single-platform video game ever, with more than 82 million copies sold worldwide (Nintendo Co. 2014)).
– The main idea behind Dogma 2001 is to provide boundaries to enable game developers to improve upon the status quo, or, put another way, to create less *crummy* games; game designers should never be discouraged from following innovative approaches to create novel or better gameplay.

In the following, three particular aspects of specially tailored interfaces will be discussed in more detail: game accessibility, change of use, and video game HCI research.

Game Accessibility

As video games become increasingly popular, making them more accessible becomes more important. Bei Yuan, Eelke Folmer, and Frederick C. Harris Jr. realized a classification of player impairments (Yuan et al. 2011), in which they identified four relevant types: visual, hearing, motor, and cognitive impairment.

Especially the third type, motor impairment, can be addressed by custom-tailored game interfaces. In the following, two particular examples, QuadControl, which aims at allowing quadriplegic persons to play video games, and the Equiliberator, a therapeutic video game interface for persons who have mobility problems, are featured as examples for video game accessibility that is attained by alterations in the gaming interface.

QuadControl

QuadControl (Yankelevitz 2010) by Ken Yankelevitz is a series of video game controllers for quadriplegic persons. From a technical point of view, current QuadControl devices combine a sip-and-puff interface, lip-controlled buttons, and, optionally, head-operated joysticks. Ken Yankelevitz recently received a great deal of media attention for this project, which has been ongoing since 1981, and until the present he has produced more than 800 game controllers, which are sold at cost (Dockery 2011). Again, one major challenge which QuadControl has had to surmount is the mapping problem: as the supported video game systems, including Sony Playstation and Microsoft XBox, involve games with five or more degrees of freedom that can be played with standard game controllers, this also has to apply to the device being considered here. The device maps control sequences that can be applied as single operation with mundane hardware into certain interaction sequences. When that is done they take longer to execute, and therefore QuadControl users might be at a disadvantage. Nevertheless, some QuadControl users play in progamer leagues.

Equiliberator

Interactive game systems can be very beneficial for physiotherapeutic sessions: they can track the patient's progress from session to session; collect, analyze, and edit otherwise unavailable session data for the therapists; make autonomous training sessions possible; and motivate users by providing specific rewards, such as a higher score on the game.

There are surprisingly few research examples of video games that can be employed for physiotherapy. One of them is the Nintendo Wii device, a commercially available video game system: it has been used for many purposes, including balance impairment reduction for elderly patients and therapy for patients with prosthetic limbs. It is generally considered more effective for individual physiotherapy exercise, because patients tend to exercise longer when they use it, feel more comfortable while doing so, and are motivated by sporting challenges.

The Equiliberator (Oden 2011) is a video game controller that is mainly built from five commercially available Wii balance board devices. Its hardware setup resembles a treadmill, whereby the Wii balance boards on the ground are provided with additional pressure-sensitive handrails to make training sessions possible for patients who cannot walk without support. The combination of a balance-sensing walkway and force-sensing handrails enables accurate balance assessment and improvement rating. Exercising patients are motivated and rewarded by video game mechanics that react to training progress by increasing the difficulty levels of the games. The Equiliberator is a comparatively inexpensive therapeutic device that can also be used for entertainment purposes.

Controller Misuse

Video game controllers are subject to creative examinations by artists and players alike. A current player approach to controller misuse is referred to as *stupid control method* (Gwin 2014a). It designates instances in which players use inappropriate video game controllers in conjunction with very challenging or competitive video games. For example, Benjamin Gwin managed to complete Dark Souls, a popular action role-playing game well known for its high level of difficulty, with a Rock Band guitar controller. From a technological point of view, he used the hardware without altering it, but he did provide a custom input mapping. That resulted in severely limited game control capabilities, as only a small number of the keys can be bound to the few available buttons on the guitar interface. In a similar fashion, he also beat the game by using a specific drum controller (Gwin 2014b). Benjamin Gwin summarizes this as self-challenge.

Another approach of the GameMuscleVideos group is to participate in highly competitive first-person-shooter matches with inappropriate controller hardware; concretely, they produced a gameplay video of the highly disputed game Counter Strike: Global Offense (CS:GO) with the use of a wheel controller. Besides the additional challenge, the humiliation of the opponents who have been defeated is a base motivation which is employed in this wild-goose chase (GamerMuscleVideo 2014).

Hardware hacking is a different approach to controller misuse. There are a multitude of applications for it, including

- Cheating by preprogramming certain key sequences
- Better accessibility, which is attained by alternate key triggering mechanisms
- Using available hardware as cheap foundation for building custom game controllers
- Using device-specific controllers on different hardware (and thus making it possible to use preferred devices on other platforms)
- Using cheap off-the-shelf interfaces for nongaming applications

A considerable percentage of the projects that are presented in subsequent chapters are also based on hardware hacking which is done for more artistic purposes.

On Video Game HCI Research

According to Pippin Barr, "the majority of current video game HCI does acknowledge play as a distinct form of interaction" (Barr 2007). This makes it difficult to apply standard HCI methods to the field of gaming, because these techniques are usually designed for productivity application. He describes the main differences as follows:

> First, the motivations for playing video games differ from productivity application use. Specifically, players play games for their own sake, while they generally use productivity applications to achieve some other task. Second, video game interfaces are not neutral, presenting carefully designed narratives and complex graphics to the player. Third, video games frequently dictate goals to players, while productivity applications generally facilitate user goals. Finally, video game designs purpose-fully involve conflict and constraints on the player, while productivity applications are designed to minimize them.

Yet Another Definition of Digital Games?

The interface-centric approach that is presented here makes it necessary to outline central game concepts that have been successfully deployed in a series of artistic digital game installations; this framework of concepts is presented in the following section. At the time that this paper is being written, it is already becoming tedious to keep track of current digital game definitions because there are so many of them; commonly used definitions (also including definitions for game and play) are based on the works of Huizinga, Caillois and Barash, Suits, Avedon & Sutton-Smith, Sutton-Smith, Tavinor, Crawford, Kelley, Salen and Zimmerman, Juul, Malaby, Bjork, Hughes, Ermi, Myr, Consalvo, Calleja, Sicart, Bergstrm, Esposito, and many more.

The majority of the currently cited definitions concur in fundamental respects, including the following ones:

- They regard the players as a central component (Björk and Juul 2012).
- They are interactive (Ermi and Mayrü 2005; Crawford 1984).
- They are based on rules (Juul 2000; Kelley 1998; Salen and Zimmerman 2004; Suits and Newfeld 1978; Kücklich 2003).
- They require an audiovisual apparatus (Esposito 2005).
- The players struggle toward a goal (Suits and Newfeld 1978).

Malliet summarizes the ongoing debate on the terminology as follows:

> In recent years a lot of debate has been held on the formal characteristics that mark out the boundaries of video games as a distinct medium, and on the terminology that is the most appropriate for analyzing video games as cybertexts (. . .), as rule-based systems (. . .), as simulations (. . .) or as second-order cybernetic systems (. . .). (Malliet 2007)

It is still being debated whether (video) games can be defined, should be defined, and how/in which context the definitions should be used. Therefore the authors use the Wittgensteinian concept of *family resemblance* to provide a notion of digital art games with nonstandard game interfaces, as was proposed by Jonne Arjoranta (2014). This is very similar to the suggested modus operandi suggested by Brian Sutton-Smith (1997), who argues that "it is nearly impracticable to describe play and games in positive, non-paradoxical terms" (Walther 2003). Therefore, our goal is to provide a useful set of attributes that apply to the projects that are presented here. Within our framework they constitute a set of nominal characteristics.

For the notion of video games itself, Grant Tavinor points out that "Videogames are variously referred to as 'computer games,' 'electronic games,' and even 'digital entertainments.' These terms cannot be taken to be strictly synonymous" (Tavinor 2009). As all of the projects that are presented in the following chapter make use computer-aided enforcement of game rules, in the context of this article video games are defined as games that make use of screen-based HCI and electronic games as ones in which standard displays are not used.

The primary argument is made on the basis of the game definition of Bernard Suits, as it allows us to demonstrate the holistic advantages of the custom interface approach. To sum it up, games must have three integral game properties according to his definition: a *prelusory goal* that is understandable outside of the game context, the *constitutive rules*, and a required *lusory attitude*: three particular elements concur to the definition of games by Bernard Suits, featured in *The Grasshopper: Games*, *Life and Utopia*. Briefly outlined, a game consists of the *prelusory goal*, which can be understood and achieved apart from the game (every game has such a goal), the *constitutive rules*, where the most efficient means to the prelusory goal are forbidden, and the *lusory attitude*, a person's willingness to accept the constitutive rules.

Interface-Centric Art Games

In the following chapters interface-centered examples of art games that were created by students and graduates of the Interface Cultures Department of the University of Art in Linz, Austria, will be shown and analyzed. These art games are examined in detail, with focus on specifically designed interfaces that help to transport specific values constituting integral parts of the art and game experiences. The examples that are presented here make use of novel, custom-made game interfaces which affect the underlying rules and shape the overall play experiences.

All of the game installations that are considered here display the following characteristics: first, all of them are interactive and require participation; zero-player games (Björk and Juul 2012), although artistically promising, are beyond the scope of this article. Second, most of them achieve interaction with the help of custom-made, nonstandard, physical game interfaces, which are central to the play experience. Third, all works involve specific rules, which are enforced automatically by digital technology and also encompass the game interfaces. Fourth, being video games, they make use of multiple media.

A Selection of Relevant Game Installations

In the following, brief overviews of following the interface-centric art games are given: Massage Me (2007) by Mika Satomi and Hannah Perner-Wilson, *Space Trash* (2008) by Christoph Anthes, Mika Satomi, Alexander Wilhelm, et al.; Punch Out (2009) by Mar Canet, Jayme Cochrane, and Travis Kirton; Headbanghero (2009) by Tiago Martins, Ricardo Nascimento, and Andreas Zingerle; Shopping in one minute (2010) by Varvara Guljajeva and Mar Canet; GearBox (2011) by Ulrich Brandstätter and Oliver Buchtala; and Puzzle Facade (2013) by Javier Lloret.

Massage Me

Massage Me (2007) by Mika Satomi and Hannah Perner-Wilson, shown in Fig. 1, is based upon a wearable universal game controller for the Sony Playstation 2 system and aims at transforming kinetic player interactions into enjoyable back massages for nonplayers. The authors summarize it as a "wearable massage interface that turns a video game player's excess energy into a back massage for an innocent bystander." They also include a set of instructions for nonplayers: "All you need to do is to sit or lay down in front of a video game player and you will be able to enjoy a back massage while the game lasts. Otherwise wasted button-pushing energy is transformed into a massage and the addicted game player becomes an inexhaustible masseur" (Satomi and Perner-Wilson 2007). Therefore, the player generates

Fig. 1 Massage Me by Mika Satomi and Hannah Perner-Wilson (Image reproduced with permission from Mika Satomi)

Fig. 2 Space Trash by Christoph Anthes, Mika Satomi, Alexander Wilhelm, et al. (Image reproduced with permission from Christoph Anthes)

positive effects that extend beyond the reality of the game. Players can choose among a variety of games. Their choice influences the style of game interaction, and therefore also the perceived massage quality.

Massage Me presents an alternate game interface for existing game systems, using flexible buttons made of conductive fabric that are incorporated into visually appealing vests. In addition to the vests, an installation requires a suitable game console, as well as displays and, depending on which games are involved, a speaker setup. A single-player game session requires at least two persons, whereas a total of four participants are necessary for two-player games.

Space Trash

Space Trash (2008) by Christoph Anthes, Mika Satomi, Alexander Wilhelm, et al., shown in Fig. 2, is a multiplayer VR game with custom interfaces. It is the result of a collaboration between the Institute of Graphics and Parallel Processing of the Johannes Kepler University (JKU) and the Interface Culture Department of the University of Art and Industrial Design Linz. It involves a two-site installation, whereby teams of three people engage in competitive gameplay. Space Trash is "an immersive Virtual Reality (VR) installation that combines aspects from VR software engineering, game design, physical interface and visual design. The application deals with the problem of particles floating in the earth's orbit, which causes damage to satellites. The players have to collect these particles by maneuvering a spaceship and controlling a collector arm in a coordinated manner" (Anthes et al. 2008).

Space Trash comprises a three-tier design process: first, a narration was created around the unusual topic of waste in outer space and coordinated with gameplay concepts suitable for public two-site exhibition. Both gameplay and narration were used as bases for the graphics design, which involves 3D models, textures, shader effects, GUI, and an overall play atmosphere. Finally, unusual user interfaces which enforce intrateam cooperation were created in conformity with the overall design framing. On each team there have to be three players, each of whom operates individual controllers with specific functionalities:

> The interface design for the Space Trash installation follows the approach to develop interfaces which are already well known to the user and produce a direct association between the way how to use them and the reaction of the application. (. . .) A really novel aspect is the collaboration inside the teams which is needed to successfully beat the other team. This collaborative aspect is furthermore supported by the design of the physical interfaces. (. . .) The competition is focused on the speed and timing of the game play rather than strategy or skill. Since most of the participants are playing only once during the presentation, and playing time is restricted to 10 minutes, the player's learning curve in the aspect of skill is limited. (Anthes et al. 2009)

Punch Out

Punch Out (2009) by Mar Canet, Jayme Cochrane, and Travis Kirton, shown in Fig. 3, enhances the classic 1987 version of the Nintendo game which has the same title with a highly immersive two-way interface. It aims at the creation of a full-body play experience by augmenting an available game with new interaction approaches. Of particular interest is the ability of the installation to fight back:

> (. . .) the installation "fights back" with a player. . . When the character on-screen is hit during a boxing match, the installation briefly flashes bright light in attempt to create a sense of disorientation for the player. The physicality of punching the soft screen, hopping back and forth in all directions, and the consequences of being "hit" amount to a physically embodied experience of a game that was originally playable only with one's fingertips. (Canet et al. 2009)

As the work involves a fairly complicated hardware construction, the team behind Punch Out provides thorough step-by-step construction instructions (*PunchOut instructables*).

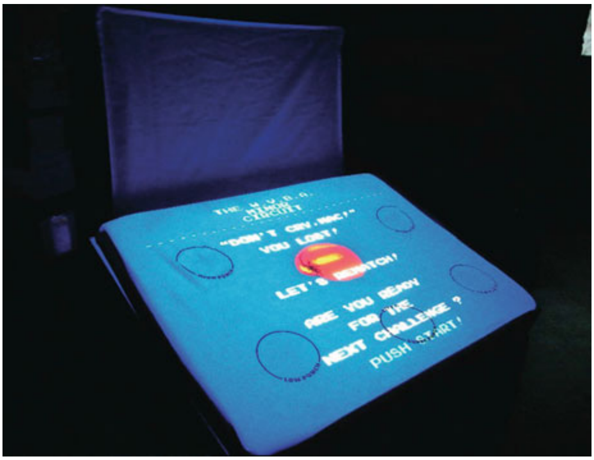


Fig. 3 Punch Out by Mar Canet and Jayme Cochrane and Travis Kirton (Image reproduced with permission from Mar Canet)

Headbanghero

Headbanghero (2009) by Tiago Martins, Ricardo Nascimento, and Andreas Zingerle, shown in Fig. 4, features a novel interface: players must test their *head-banging prowess*, and this is monitored by sensors inside a custom-made wireless wig which the player wears on his head. The game itself is classified as a music/dance game and features game mechanics similar to those found in the commercial AAA title *Guitar Hero*: players must bang their heads in accordance with a music piece that they are listening to, whereby deviations result in a lower score. In contrast to other dance games, playing Headbanghero cannot be regarded as a healthy exercise. Headbanghero warns about the health risks and even awards players with a printed health analysis based on their actual performance, which includes advice on how they can avoid major injuries (Martins et al. 2009).

Headbanghero employs a highly unconventional user interface in a game designed to seduce fans of the heavy metal music genre to participate in a unique gameplay experience. Another interesting aspect is the transformation of the play space into a stage. While they are playing the game in an exhibition space, players are transformed into performers, since spectators in the audience are watching them. In addition to the wig interface, an installation requires a PC, a video projector, a speaker setup, and a printer to provide awards for the players.

Shopping in One Minute

Shopping in one minute (2010) by Varvara Guljajeva and Mar Canet, shown in Fig. 5, challenges players to expose themselves to the stressful everyday life of a supermarket salesperson. It as an “art-game came from the craziness of people during the sales period: long lines in front of supermarkets, stone-age-style behavior, lots of useless products, etc. Thus, we are inviting audience to train for the next

Fig. 4 Headbanghero by Tiago Martins and Ricardo Nascimento and Andreas Zingerle (Image reproduced with permission from Andreas Zingerle)

Fig. 5 Shopping In One Minute by Varvara Guljajeva and Mar Canet (Image reproduced with permission from Varvara Guljajeva)

sales period. A player has to scan as many products as possible during a minute. As a reward each player will receive a ticket with bonus points according to his score" (Canet 2010).

Shopping in one minute uses a bar-code scanner as input and a simple LED display as well as a speaker that is blaring out stressful tunes as output devices. It therefore mimics and convincingly recreates the daily routine of the real-world role-model scenario. In contrast to mundane video games, it does not require a standard display device or any particular in-game narration but provides players with awards in the form of score printouts at the end of each game.

GearBox

GearBox (2011) by Ulrich Brandstätter and Oliver Buchtala, shown in Fig. 6, provides an audience with a playful interface which enables its members to cooperatively compose music. Users are invited to playfully explore and exploit the possibilities the system offers. The installation focuses on creational gaming aspects and challenges the nonproductive limitation of many game definitions: participants are seduced to engage in a novel method of music creation that demands their full player attention, since the approach is truly unique.

Fig. 6 GearBox by Ulrich Brandsttter and Oliver Buchtala

The complexity of the system is not readily apparent, since it is disguised as a sandbox game (Brandstätter and Buchtala 2011).

GearBox utilizes a common multitouch screen as a physical user interface. The virtual interface solely consists of equitable graphical gears, so that it enables multiuser participation of both novices and electronic music experts. The process of music creation is transformed into one that involves playing with extraordinary gears (setting/changing onset events, altering rotation speed/gear size) as well as interconnecting gear trees. All of the mechanisms are based on real-world gears and their machinery, which are very familiar to and readily understood by most people; therefore, the participants start with a fundamental comprehension of the system mechanics. In a short period of time this playfully evolves, so that they soon display impressive skills in respect to music composition. Thanks to the fundamental cyclic sequencing paradigm, players can incrementally change, improve, and create new variations of rhythmic and melodic patterns. The combination of a table-mounted touchscreen with reduced user interface(s) (virtual gears) makes it possible for many participants to compose and perform at the same time.

Puzzle Facade

Puzzle Facade (2013) by Javier Lloret, shown in Fig. 7, projects the process of solving a Rubik's cube on an installation in an urban space: a specifically created interface cube tracks the alterations of the cube and displays the actual state of the game on the media facade of the Ars Electronica building via wireless transmission. This installation alters the Rubik's cube play experience in several ways: first, the cube itself no longer features the current configuration, so that the player can only see two of its sides at any one time; in order to view other sides, he must rotate and/or flip the interface. He can also do so by moving around the Ars Electronica Building. Second, as implied by the projection, the game has become more difficult.

Fig. 7 Puzzle Facade by Javier Lloret (Image reproduced with permission from Javier Lloret)

Third, the game space is shifted from private to public space, enabling passive observers to become kibitzers and thereby to influence the course of the game.

Puzzle Facade augments the 40-year-old Rubik's cube mechanics with a sensor system that is able to track game states along with the position and orientation of the cube and a wireless data transmission system based on the Bluetooth protocol (Lloret 2013); the system uses the sensor data to visualize the current configuration on the Ars Electronica facade, an interactive public space display. Although the interface itself is highly portable, the game can only be played at the Ars Electronic building, due to its suitable form and its media facade technology.

Nominal Characteristics

While the dominant approach of the gaming industry toward gaming controllers involves multipurpose and standard hardware (joypad, mouse/keyboard; standardized gaming hardware generally is optimized for the main product audience, severely affecting gameplay experience of secondary target groups, as shown by Fron et al. (2007); however, it must be pointed out that companies manufacturing game peripherals try to avoid gaming-specific injuries, such as repetitive strain injuries or carpal tunnel syndrome, with designs focused on ergonomy), there are also a multitude of more specialized gaming interfaces on the consumer market, ranging from steering wheels over light guns to balance boards or toy guitars. The examples given here illustrate how custom-tailored game interfaces can benefit the overall play experience of several game genres, including arcade, dancing, and puzzle games. For the realization of most of these projects an even broader range of skills is necessary than is required for the creation of mundane video games; these include installation construction, hardware interface design, interaction design for nonstandard physical interfaces, *hacking* available off-the-shelf game controllers, and relevant software libraries. Since all of the works are presented in the form of exhibition installations, they for obvious reasons generally involve gameplay with comparably easy-to-grasp rules and conditions for winning or losing.

Table 1 features an overview of the relevant works.

The works generally succeed in establishing intelligibility by deploying four specific strategies:

- First, many examples are based on well-known games that are widely available.
- Second, the game installations communicate prelusory goals.
- Third, the interfaces communicate the constitutive rules (to some extent).
- Fourth, the devices excite curiosity and seduce spectators into the magic circle.

A classification of the strategies which are deployed in the projects that have been presented is shown in Fig. 8; the fourth strategy has, however, been omitted, since all of the projects are capable of seducing the audience into a lusory attitude.

Based on Well-Known Games

Of the seven game installations presented here, four are based on existing games:

Massage Me is built upon the Sony Playstation 2 console, and therefore relies on ready-made games. Participants are allowed to select known or promising games and are encouraged to switch game titles regularly, so that different massage styles can be carried out.

In a similar manner Punch Out uses the popular Punch-Out title (1987) from the Nintendo Entertainment System as its basis. This has several implications: as the original game is about 27 years old, its graphics representation capabilities are no longer state of the art. Today it can be regarded as a so-called retro game. However, gameplay-wise it can be disputed whether the original Punch-Out has undergone augmentation. It is the perfect title for a gameplay-augmenting interface.

Headbanghero is loosely based on Guitar Hero (/RockBand), both titles are often classified as dancing games, and they both make use of similar game mechanics. Many Headbanghero veterans consider it a humorously exaggerated version of Guitar Hero, based upon very specific characteristics of the heavy metal culture.

Puzzle Facade is the game interface project that alters the way the most prominent game, the Rubik’s Cube, is played. In 2014 Rubik’s Cube is celebrating its 40th birthday. It is probably the world’s top-selling puzzle game. Puzzle Facade is the only example in which an analog game instead of a video game is used as gameplay foundation.

Prelusory Goal Communication

From the seven game installations presented here, five communicate their prelusory goals:

Table 1 Properties of relevant works

Property	Massage me	Space Trash	Punch Out	Head Bang Hero	Shopping in one minute	GearBox	Puzzle Facade
Year	2007	2008	2009	2009	2010	2011	2013
Type of game	Game controller	3D game	Video game	Video game	Electronic game	Software toy	Video game
Genre	Multiple genres	Arcade	Fighting	Dancing	Arcade	Sandbox	Puzzle
Competitive / cooperative	Both	Both	Competitive	Competitive	Competitive	Cooperative	Competitive
Average game session duration	About 20 min	About 10 min	About 10 min	About 4 min	1 min	About 30 min	About 10 min
Based on known game	PlayStation games	–	Punch Out (NES)	Guitar Hero	–	–	Rubik's cube
More difficult than original	Yes	–	No	No	–	–	Yes
Prelusory goal known beforehand	No	Yes	Yes	Yes	Yes	No	Yes
Rules are known beforehand	Generally yes	No	Yes	Yes	Yes	Partly	Yes
Audience seduction into playing	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Game ending	Yes	Yes	Yes	Yes	Yes	No	Yes
Play score	Generally yes	Yes	Yes	Yes	Yes	No	No
Players	1/2	2 teams 3	1	1	1	Any	1
Story	No	Yes	No	No	Yes	No	No
Trophies	No	No	No	Yes	Yes	No	No
Involve audience other than players	Yes	No	No	No	No	No	Yes
Instructions required	No	Yes	No	No	No	Yes	No

Display	Standard screen	Curved screen/ portable 3D	Projection on textile	Standard projector	LCD matrix display	Table-mounted display	Ars Electronica facade
Online/networked	No	Yes	No	No	No	No	No
Sound	Yes	Yes	Yes	Yes	Yes	Yes	No
Sound necessary for gameplay	No	No	No	Yes	No	Yes	No
Playing means workout	Yes	Yes	Yes	Yes	No	No	No
Development team size	2	29	3	3	2	2	1
Setup time for new players	About 1 min	–	–	About 1 min	–	Introduction required	–
Gameplay requires provided controller	No	Yes	No	Yes	Yes	No	Yes
Based on hacked controller	Yes	No	Yes	No	Yes	No	No
Controller as output device	No	No	Yes	No	No	No	No
Controller suitable for other games	Yes	No	Limited	Limited	Limited	Limited	No
Interaction limbs	Hands	Hands/feet	Hands/feet	Full-body (neck)	Hands	Hands	Hands
Interaction in projection space	No	No	Yes	No	No	Yes	No
Portable app meaningful	No	No	No	No	Yes, available	Yes	No
Shown at public exhibition(s)	Yes	Yes	Yes	Yes	Yes	Yes	Yes

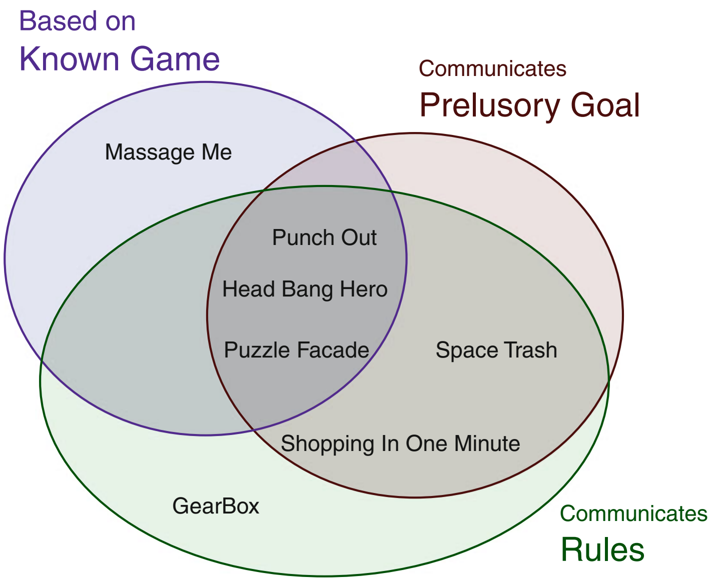


Fig. 8 Interface-based strategies

Puzzle Facade achieves this because of the popularity of the Rubik's Cube. The audience immediately understands what the game is about because they recognize the physical interface as soon as they see it.

To a certain degree, the same applies to Headbanghero: the overall game design clearly communicates the fact that a convincing game performance yields a high score, which is also true for Guitar Hero (/RockBand). Especially the graphic design of the project displays many references to its predecessor. Space Trash uses its title and its unusual game theme to give players a hint of what it is about before it even starts, and therefore it employs a more narration-based approach. Put simply, "Space Trash is a problem, do something about it."

The same is true for Punch Out, where the title directly states the prelusory goal but also hints at the game genre.

Shopping In One Minute also has a laconic title, yet it adopts a different strategy to communicate its prelusory goal: its installation format involves the game elements, consisting of a bar-code scanner as an interface and packages of various products, an annoyingly minimalistic music score, and flashy advertising; taken together, these elements clearly tell potential participants what the game is about.

Rules Communication

As pointed out previously, rules are a central facet to games, since they are reflected in many prevalent game definitions: Juul defines game as "a pastime with formal and predefined set of rules for the progression of a game session, with built-in and quantitative definitions of success and failure" (Juul 2000). Salen and Zimmermann consider it as "a system in which players engage in an artificial conflict, defined by rules, that results in a quantifiable outcome" (Salen and Zimmerman 2004). According to Calleja, "a game becomes a game when it is played; until then it is only a set of rules and game props awaiting human engagement" (Calleja 2011). Fron points out the capability of computer games to "dictate and enforce rules automatically through software" (Fron et al. 2007). Kücklich, however, points out that in video games, not all rules may be known beforehand: "(. . .) a player does not necessarily gain access to the implicit rules of the game through playing, but that he or she will find a way to interact meaningfully with the game, no matter what the actual rules encoded by its designers are. In fact, the player might even find ways to interact with the game that its creators did not think of" (Kücklich 2003). In video games, the exact rule specification and enforcement are generally unknown to players, as they are determined and set up at an inaccessible programming code level. Fortunately, players are provided with other means of gaining an understanding of the *relevant* underlying set of rules: player handbooks, in-game instructions, tutorials, read-me files, visual/audible hints, etc. It is very important for the player to have a prior knowledge of genre-specific mechanics; this is often taken for granted by game designers.

When games are played within the framework of an exhibition, the authors consider an implicit disclosure of game rules to be advantageous; it can be achieved by stating simple, easy-to-grasp game rules, assuming a certain level of prior knowledge, and communication via the installation and interface design.

Six of the seven game installations presented here communicate a large proportion of those constituent rules which are not immediately comprehensible – not even for persons who are familiar with common game controllers (such as joysticks or pads): Space Trash is featured as a two-site installation, whereby both locations showcase three distinct input interfaces, along with 3D projection technology. As soon as they see the arrangement of the device, potential players are able to deduce most of the rules. Three players are required at each site to assume different roles: "The three main controlling tasks of the navigator (navigation with pitch and heading), the accelerator (speed control including back and forward), and the grabber (manipulation of arm with pitch and heading) are assigned separately each to an individual interface" (Anthes et al. 2009). Punch Out adopts a similar strategy; the input interface features specific punch regions at different locations, allowing players to deduce the effects interactions will have.

Puzzle Facade, on the other hand, communicates its rules via the interface of Rubik's Cube, which, as a familiar toy, has rules that are generally understood and also apply to the project. Deviations from the original rules resulting from the specific installation framing are evident but add an interesting twist to the game.

GearBox features virtual gear interfaces that adopt the mechanical behavior of its real-world counterpart, thereby enabling its audience to utilize its basic understanding of mechanics to participate in a rewarding game experience. Shopping In One Minute enables the audience to deduce a large portion of the rules from the manner in which it is presented: the project name reveals the duration and time limit of the game, and the bar-code scanner, which is the central interface, makes the main game activity readily apparent.

Headbanghero combines these strategies: gamers are likely to know what the game is about from the project name, as it clearly refers to Hero game series. This makes its nature as a dance game and its central game concepts readily apparent. As soon as they see the nature of the physical user interface, the audience immediately understands that skill at headbanging is being tested.

Seduction of the Audience

In interactive art, the staging of interfaces and objects is an important component, which serves to entice the audience to participate. When media art pioneer Jeffrey Shaw presented his interactive bicycle in front of a large projection screen in "Legible City" in 1990 (Shaw 1990), many viewers immediately felt that they had to get on the bicycle and use it in order to fully experience that work of art. Active involvement and bodily participation has become a key factor in many interactive art works and art games. Artists rely on careful staging and detailed planning of the arrangements within the installation space. Every aesthetic component such as the space layout, the dimensions, the lighting, the choice of objects and various other elements can be important. Katja Kwastek analyzes these aesthetic qualities in interactive art (Kwastek 2013). Artists also often reply on *the implicit meaning of the objects*, since every one of them bears some deeper cultural connotation. Water in a pool will, for example, bring forth associations with life, a gun interface will immediately make the audience think of violence and defense. Artists are skillful at creating the desired emotional and conceptual effects by using and abusing selected or custom-designed objects. This is also linked to the term *affordance*, as coined by Don Norman from the HCI community. To him "affordances are relationships between the properties of an object and the capabilities of the agent that determine just how the object could possibly be used" (Norman 2002). Experience Design is another area where these new types of interface designs and their effects on the users are researched (Dewey 2005). John Dewey's aesthetic approach to the *expressive object* is also relevant here, as he highlights the important effect it has on the audience's experience (Dewey 2005).

All of the projects that are presented here are assumed to be capable of seducing the audience into a lusory attitude. This can be attributed to several factors: first, the novel interfaces alone excite curiosity, the audience often wants to find out how the new devices work and how the interfaces behave; where applicable, the new devices are compared to the original ones, and the differences are evaluated. Curiosity is also provoked by the uniqueness of the interface concepts, the display technology, and the prominent game sounds and music. Of course, the exhibition/ public space format itself contributes to the curiosity the pieces are regarded with. Their presentation within the framework of an exhibition makes it necessary to employ more casual gameplay concepts and focus on easy-to-grasp game ideas, short sessions, and extraordinary play experiences. Some of the games give the players printed-out scores that they can keep as souvenirs, enabling them to boast about their achievements and compare themselves to other players. However, competition-based games constitute the exception; most of the projects are primarily focused on the play experience. Another concept which is utilized is the prospect of creating something, for example, a cooperative audiovisual composition.

Additional Nominal Characteristics of the Works Presented Here

The works presented can be primarily classified as video games, but there is also an example of an electronic game, a software toy, a networked 3D game, and a universal game controller. The game genres that have been covered here include arcade, dancing, sandbox, puzzle, and fighting games.

They mostly involve competitive play and often make use of scoring mechanisms; on the other hand, two of the installations largely rely on player cooperation. One example requires at least six participants for a meaningful game session, but the others can be used by single players. An additional benefit which is conferred by the controller-centric approach toward gaming is that it provides play experiences with increased immersion: to play Headbang Hero one has to perform Headbang Hero, and when one plays Shopping In One Minute, one undergoes the stressful experience of supermarket sales personnel at peak shopping hours. However, many works also generate an additional, out-of-game benefit, that is to say an added value: playing Headbang Hero or Punch Out can be considered a sports activity which is equivalent to a moderate workout, while the jacket-bearer participating in a Massage Me session receives an enjoyable massage that relaxes her/his muscles; in playing or performing the GearBox one automatically contributes to the collaborative creation of a new piece of music.

Due to the fact that the games are played within the context of an exhibition, there is an overall lack of in-game narration elements, the game durations are relatively short, and the game objectives and victory conditions must also be adapted to the exhibition space. Besides, in many cases it is not necessary to provide (in-)game instructions.

The exhibition framing is particularly detrimental for installations where sound or music are necessary for gameplay; the difficulties are induced by sonic interferences of other audible works in close proximity, remedial actions are the deployment of headphones (often acting as a deterrent or inconvenience), low sound volumes, or selective activation of the speakers of individual works. This problem is common to media art exhibitions and game festivals. With the exception of one project, all presented games involve audible information in the form of sound effects and game music; for two pieces meaningful play requires enabled sound/music.

In respect to the device properties, most of the controllers operate only as input devices. They are often based on hacked devices that are commercially available. In one case, the controller is not suitable for other games. Some of the installations could potentially be played with mundane game controllers, involving an appropriate loss of gameplay experience, especially for the examples which require other modalities than bare hands for interactions.

Most of the projects use nonstandard displays as their primary output device; only three of them rely on standard screens. The technology which is deployed ranges from on-textile projections over simple LED screens to 3D displays and interactive building facades.

The two pieces enabling multiplayer gaming adopt different strategies: one facilitates concurrent player interaction by using an appropriately sized multitouch screen in conjunction with a rotation-invariant equitable software interfaces, so that participants can arrange around the table-mounted installation. The other example makes use of network transmission protocols in order to interconnect remote play sites.

Conclusion

The examples that are presented here by the Interface Cultures Department range between interactive art, digital games, and experience designs. They highlight the importance of the interface as an aesthetic and functional component, since it conveys specific cultural values. Carefully designed and reflected interfaces of digital art games have the power to seduce audiences and make them actively participate, and at the same time to convey deeper cultural meanings. As the authors hope to have shown, game interfaces are tightly interwoven with gameplay experiences and can completely alter and transform them. Custom-made devices are able to draw potential players into the magic circle; they succeed in explaining rules and goals outside of the game reality and prior to its initiation. These three characteristics holistically embrace the definition of games which was proposed by Bernard Suits.

Moreover, as shown on the basis of the seven selected game installations, specifically created interfaces are the key for the realization of extraordinary artistic game concepts. They feature transformational approaches common to modern media art, such as device hacking, decontextualization, and participation, and can be regarded as activation tools. In many cases they are able to provide added value, for example, by stimulating cooperative creativity, providing occasions for physical workouts, or creating pleasant situations for bystanders. Regarding the notion of the video game, new definitions have not been formulated, and the explicit adoption of any of the existing ones for our specific framework has been avoided; as was said before, the main argument was aligned with the concept of Bernard Suits, as its individual components are in their entirety conducive to interface-centric approaches. Instead, as proposed by Jonne Arjoranta (2014), a set of nominal attributes was listed to narrow down the range of works that are regarded as interface-centric art games. This was done on the basis of specifically selected projects.

Ultimately, custom-made game interfaces do not constitute a surefire success formula for particular game titles. Game designers and artists are admonished to very carefully consider the deployment of specifically created devices, particularly with regard to business applications. None of the examples presented here has yet been made available to the consumer market. On the other hand, every one of the featured projects has within the context of media art been successfully presented at several exhibitions.

Recommended Reading

E. Adams, *Dogma 2001: A Challenge to Game Designers* (2001), http://www.gamasutra.com/view/feature/131513/dogma_2001_a_challenge_to_game_.php. Last accessed 24 Feb 2015

C. Anthes, M. Satomi, A. Wilhelm et al., *Space Trash website* (2008), http://www.spacetrash.org/index.php. Last accessed 24 Feb 2015

C. Anthes et al., *Space Trash – Development of a Networked Immersive Virtual Reality Installation*. Technical report TR-GUP-01-09. Institut für Graphische und Parallele Datenverarbeitung, Johannes Kepler University Linz, 2009

J. Arjoranta, Game definitions: a Wittgensteinian approach, in *Game Studies* 14.1 (2014), URL: http://gamestudies.org/1401/articles/arjoranta. Visited on 25 Aug 2014

P. Barr, *Video Game Values Play as Human-Computer Interaction*, PhD thesis, Victoria University of Wellington, 2007

S. Björk, J. Juul, *Zero-Player Games. Or: What We Talk about When We Talk about Players*, Madrid (2012), URL: http://www.jesperjuul.net/text/zeroplayergames/

U. Brandstätter, O. Buchtala, *GearBox Video* (2011), https://vimeo.com/28270316. Last accessed 24 Feb 2015

D. Buchhart, Über die Dialektik von Spielregeln und offenem Hanlungsfeld, in *Kunstforum International*, vol. 176 (June–August 2005), pp. 39–55

D. Buchhart, M. Fuchs (eds.), *Kunst und Spiel 1*. Kunst-forum International (2005a), pp. 38–143

D. Buchhart, M. Fuchs (eds.), *Kunst und Spiel 2*. Kunstforum International. (2005b), pp. 38–145

G. Calleja, *In-game: From Immersion to Incorporation* (MIT Press, Cambridge, MA, 2011). ISBN 9780262294546

M. Canet, *Shopping In One Minute – Design Analysis* (2010), http://www.mcanet.info/papers/DesignAnalysis_Shopping1Minute.pdf.last. Accessed 24 Feb 2015

M. Canet, J. Cochrane, T. Kirton, *Punch Out website* (2009), http://www.mcanet.info/blog/?p=16. Last accessed 24 Feb 2015

C. Crawford, *The Art of Computer Game Design* (Osborne/McGraw-Hill, Berkeley, 1984). ISBN 0881341177

P. D. Deen, Interactivity, inhabitation and pragmatist aesthetics, Int. J. Comput. Game Res. **11** (2011)

J. Dewey, *Art as Experience* (Perigee Books, New York, 2005). ISBN: 9780399531972

S. Dockery, *Man Builds 30 Years of Quadriplegic Gaming* (2011), http://www.nbcnews.com/id/43365521/ns/technology_and_science-games.last. Accessed 24 Feb 2015

U. Eco, *The Open Artwork* (Harvard University Press, Cambridge, MA, 1989)

L. Ermi, F. Mayrü, Fundamental components of the game-play experience: analysing immersion, in *DIGRA* (DIGRA, 2005)

N. Esposito, A short and simple definition of what a videogame is, in *DiGRA* (2005)

M. Flanagan, H. Niessenbaum, *Values at Play in Digital Games* (The MIT Press, Cambridge, MA/London, 2014)

V. Flusser, *Ins Universum der technischen Bilder* (European Photography, Gttingen, 1987)

V. Flusser, Gesellschaftsspiele, in Künstlichespiele ed. by F. Rötzer, G. Hartwagner, S. Iglhaut (Boer Verlag, Grafrath, 1993)

J. Fron et al., The hegemony of play, in *Situated Play* (The University of Tokyo, 2007), URL: http://www.digra.org/wp-content/uploads/digital-library/07312.31224.pdf

R. Fuchs, Kunst-Sprache-Spiele: Dreiecksbeziehungen in der Kunst der 60er und 70er Jahre, in *Kunst-forum International* (November 2005–January 2006), pp. 38–47

H.-G. Gadamer, *Wahrheit und Methode, Grundzüge einer philosophischen Hermeneutik* (1990)
GamerMuscleVideo, *Killed You With A Wheel – Gamer Muscle Challenge* (2014) https://www.youtube.com/watch?v=Fl35QCbMi-c.last. Accessed 24 Feb 2015
C. Goodman, Welcome to Gameworld: games on the edge of art, technology and culture, in *Gameworld Catalogue* (LABoral Centre for Arts and Creative Industries, Gijn, 2007), pp. 44–45
B. Gwin, *Dark Souls One Handed Challenge* (2014a), http://www.twitch.tv/bearzly/c/4092128. Last accessed 24 Feb 2015
B. Gwin, *Drum Souls* (2014b), http://www.twitch.tv/bearzlyZc/5428798. Last accessed 24 Feb 2015
L. Johnson, *Game Controllers: A Critical Discussion of Input Devices in Game Design* (2010), URL: http://www.lgrace.com/documents/Critical_Game_Controllers.pdf
J. Juul, What computer games can and can't do, in *Digital Arts and Culture* (Bergen, 2000)
D. Kelley, *The Art of Reasoning* (W.W. Norton, New York, 1998). ISBN: 9780393972139
T. Kirton, *PunchOut Instructables*. http://www.instructables.com/id/PunchOut-The-Real-Installation/. Last accessed 24 Feb 2015
J. Kücklich, Perspectives of computer game philology. Int. J. Comput. Game Res. **3** (2003)
K. Kwastek, *Aesthetics of Interaction in Digital Art* (MIT Press, Cambridge, MA, 2013). ISBN 9780262317221
K. Lenz, The effect of input device on video game performance, in *Human Factors and Ergonomics Societys 52nd Annual Meeting*, SAGE Publications, New York, 2008
J. Lloret, *Puzzle Facade website* (2013), http://interactive.javierlloret.com/puzzlefacade/. Last accessed 24 Feb 2015
W. Lu, *Evolution of Video Game Controllers: How Simple Switches Lead to the Development of the Joystick and the Directional Pad* (2003), URL: http://www.stanford.edu/group/htgg/sts145papers/wlu_2003_1.pdf
S. Malliet, Adapting the principles of ludology to the method of video game content analysis, in *Game Studies* 7.1 (2007), URL: http://gamestudies.org/0701/articles/malliet
T. Martins, R. Nascimento, A. Zingerle, *Head Bang Hero website* (2009), http://www.headbanghero.com/index.html. Last accessed 24 Feb 2015
Nintendo Co., Ltd. (2014), http://www.nintendo.co.jp/ir/en/sales/software/wii.html. Last accessed 02 Feb 2015
D.A. Norman, *The Design of Everyday Things* (Basic Books, 2002). ISBN: 9780465067107
M. Oden, *The Equiliberator* (2011), http://oedk.rice.edu/Sys/PublicProfile/4836222. Last accessed 24 Feb 2015
F. Rötzer, Die Begegnung von Computerspiel und Wirklichkeit, in *Kunstforum International*, vol. 176 (Cologne, 2005), pp. 102–115
K. Salen, E. Zimmerman, *Rules of Play: Game Design Fundamentals* (MIT Press, Cambridge, MA, 2004). ISBN 9780262240451
M. Satomi, H. Perner-Wilson, *Massage Me Website* (2007), http://www.massage-me.at/index.php. Last accessed 24 Feb 2015
F. Schiller, *Über die ästhetische Erziehung des Menschen, in einer Reihe von Briefen, 14. Brief, Schillers Sämmtliche Werke, vierter Band* (J. G. Cottasche Buchhandlung, 1879)
J. Shaw, Legible City, in *Der Prix Ars Electronica*, ed. by H. Leopoldseder (Veritas Verlag, Linz, 1990)
C. Sommerer, L. Mignonneau, A-Volve an evolutionary artificial life environment, in *Artificial Life*, ed. by C. Langton, K. Shimohara (MIT Press, Boston, 1997), pp. 167–175
E. Strouhal, *Kunst und Politik der Spiele* (di:Angewandte, Son-derzahl Verlag, 2008)
B.H. Suits, F. Newfeld, *The Grasshopper: Games*, *Life and Utopia*. *Nonpareil book* (D.R. Godine, Boston, MA, 1978). ISBN: 9780879238407
G. Tavinor, *The Definition of Videogames Revisited* (Department of Philosophy, Classics, History of Art and Ideas, University of Oslo, Oslo, 2009). ISBN: 82-91670-57-9. URL: http://www.gamephilosophy.org/. Visited on 27 Jan 2010

B.K. Walther, Playing and gaming reflections and classifications. Int. J. Comput. Game Res. **3** (2003)

P. Weibel, Ein Spiel ist im Grunde ein Algorithmus, in *Kunstforum International*, vol. 176 (Cologne, 2005), pp. 127–133

K. Yankelevitz, *Quad Control - Video Game Controllers for Quadriplegics* (2010), http://quadcontrol.com/. Last accessed 24 Feb 2015

F. Ybud, *Evolution of Video Game Controllers* (2014), URL: http://thysecondgroup.blogspot.co.at/2014/03/evolution-of-video-game-controllers.html

B. Yuan, E. Folmer, F.C. Harris Jr., Game accessibility: a survey. Univers. Access Inf. Soc. **10**(1), 81–100 (2011). doi: 10.1007/s10209-010-0189-5. ISSN: 1615-5289

Love and Sex with Robots

33

Adrian David Cheok, David Levy, Kasun Karunanayaka,
and Yukihiro Morisawa

Contents

A.D. Cheok (✉)
Imagineering Institute, Nusajaya, Malaysia

City University London, London, UK
e-mail: adrian@imagineeringinstitute.org

D. Levy
Imagineering Institute, Nusajaya, Malaysia

Retro Computers Ltd, Luton, UK
e-mail: david@retro-computers.co.uk

K. Karunanayaka
Imagineering Institute, Nusajaya, Malaysia
e-mail: kasun@imagineeringinstitute.org

Y. Morisawa
Saitama Institute of Technology, Saitama, Japan
e-mail: morisawayukihiro@gmail.com

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_15

Abstract

The publication of the book *Love and Sex with Robots*, late in 2007 by David Levy, heralded a new era in this somewhat controversial field. Human-robot intimate relationships were no longer pure science fiction but had entered the hallowed halls of serious academic research. Since then, researchers have come up with many implementations of robot companions like sex robots, emotional robots, humanoid robots, and artificial intelligent systems that can simulate human emotions. This book chapter presents a summary of significant activity in this field during the seven years since that publication and predicts how the field is likely to develop.

Keywords

Love and Sex with Robots • Kissenger • Haptics • Humanoid robots • Human-robot intimate relationships • Lovotics

Introduction

Intimate relationships, such as love and sex, between human and machines, especially robots, has been one of the main topics in science fiction. However, this topic has never been treated in academic areas until recently. The topic was raised and discussed by David Levy in his book titled *Love and Sex with Robots* published in 2007 (Levy 2007b). The book found an eager public in North America who wanted to know more. During the period immediately prior to publication of the book and for a few months afterward, the topic caught the imagination of the media, not just in the USA and Canada but on a worldwide scale. During those months David Levy gave around 120 interviews, by telephone, email, and in person; to newspapers, magazines, radio, and TV stations; and to electronic media. Television interviews included an appearance on The Colbert Report (2008), as well as visits to his home by TV crews from Russia, Canada, Austria, France, Germany, Switzerland, and other countries. There was also, not surprisingly, a flurry of interest from women's magazines, including Elle and Marie Claire. And the coverage in general science publications included articles in *IEEE Technology and Society Magazine*, *MIT Technology Review*, *Scientific American*, and *Wired*.

In the academic world there has already been sufficient coverage of the topic to demonstrate rather convincingly that it is of interest not only for mainstream media.

An academically rewritten version of the book titled *Intimate Relationships with Artificial Partners* (Levy 2007a) also attracted a lot of media publicity. Conferences on robotics, AI, and other computer science-related subjects began to accept and even invite papers on the subject, and there have thus far been two conferences devoted specifically to Human-Robot Personal Relationships. In 2014 the First International Congress of Love and Sex with Robots was held in Madeira. The academic journals that have since chosen to publish papers on the topic have included *Accountability in Research*, *AI & Society*, *Artificial Intelligence*, *Current Sociology*, *Ethics and Information Technology*, *Futures*, *Industrial Robots*, *International Journal of Advanced Robotic Systems*, *International Journal of Social Development*, *International Journal of Social Robotics*, *International Journal of Technoethics*, *New Media and Society*, *Phenomenology and the Cognitive Sciences*, *Philosophy Technology*, *Social Robotics*, *Technological Forecasting and Social Change*, and various publications from the IEEE, Springer, and other highly respected technology stables. One paper, from Victoria University of Wellington, New Zealand, achieved a high profile in the general media when it appeared in 2012 for its entertaining depiction of a future scenario in the red light district of Amsterdam – a life, in 2050, revolving around android prostitutes "who are clean of sexually transmitted infections (STIs), not smuggled in from Eastern Europe and forced into slavery, the city council will have direct control over android sex workers controlling prices, hours of operations and sexual services." (Yeoman and Mars 2011).

Since the initial burst of media interest late in 2007, there have also been TV documentaries and feature movies in which sex with robots, with virtual characters, or with life-sized sex dolls was the dominant theme: Lars and the Real Girl, Meaning of Robots (which had its premiere at the 2012 Sundance Festival), My Sex Robot, Her (2013), and the BBC TV documentary Guys and Dolls as well as the 2004 remake of The Stepford Wives. This points out that it is the sexual nature of the subject matter which is responsible. Sex Sells.

Following the storm of publicity by the launch of the David Levy's book in 2007 (Levy 2007b), the subject of human-robot romantic and intimate relationships rapidly developed into an academic research discipline in its own right. The subject was named "Lovotics," and first mentioned in the literature in 2009 (Nomura et al. 2009). In his PhD thesis in 2011, Adrian David Cheok's student Hooman Samani explored certain aspects of Lovotics and describes the design and development of a hardware platform – a robot – which was capable of experiencing complex and humanlike biological and emotional states that were governed by artificial hormones within its system (Samani 2011). Samani's robot was a novel advanced artificial intelligence system and is described in a little more detail in sections below.

The interest in this field from the academic community resulted, in 2013, in the founding of a journal and e-journal devoted entirely to the subject, whose editor-in-chief is Adrian David Cheok. Lovotics (*Lovotics Journal*) defines its own domain as "Academic Studies of Love and Friendship with Robots" (Fig. 1).

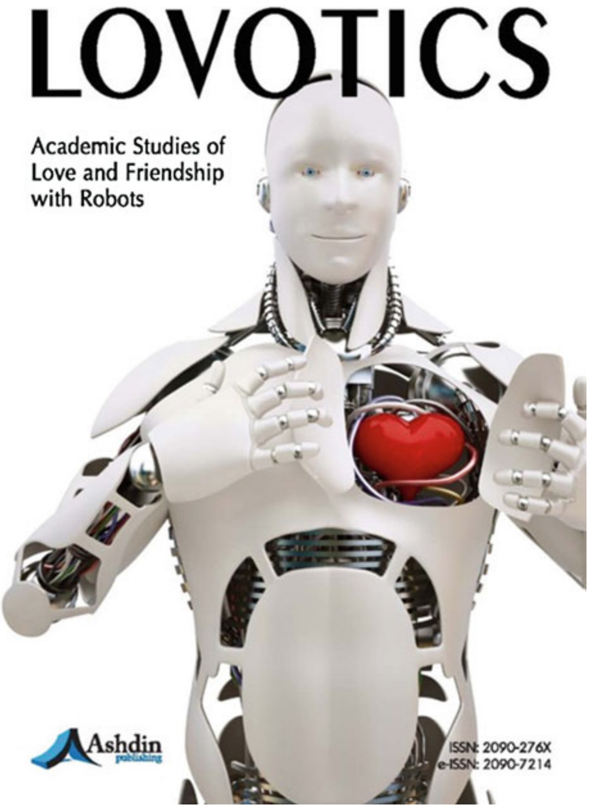


Fig. 1 Journal *Lovotics, Academic Studies of Love and Friendship with Robots*

The First Crude Sex Robot

One of the most often asked questions in media interviews with the David Levy in 2007–2008 was this: "How soon do you think the first sex robots will be on the market?" His consistent response was that the technologies necessary to create a crude sex robot were already available, and therefore it would probably not be more than 2–3 years before some enterprising entrepreneur(s) put these technologies together. For example, a sex doll with certain parts vibrating and some sexy synthetic speech would be a significant step up for those customers who have hitherto purchased just a static sex doll. This applies equally to a malebot as to a fembot – the worldwide commercial success of female vibrators indicates that a male sex doll endowed with a well-designed vibrating penis would be a good start in that direction.

Late in 2009 publicity began to appear in the media about a "sex robot" developed by a New Jersey entrepreneur, Douglas Hines. His Web site www.truecompanion.com proudly proclaimed that:

> We have been designing "Roxxxy TrueCompanion", your TrueCompanion.com sex robot, for many years, making sure that she: knows your name, your likes and dislikes, can carry on a discussion and expresses her love to you and be your loving friend. She can talk to you, listen to you and feel your touch. She can even have an orgasm!

Other amazing claims on the truecompanion.com site included:

> She also has a personality which is matched exactly as much as possible to your personality. So she likes what you like, dislikes what you dislike, etc. She also has moods during the day just like real people! She can be sleepy, conversational or she can "be in the mood!"

and

> Roxxxy also has a heartbeat and a circulatory system! The circulatory system helps heat the inside of her body.

and

> She can talk to you about soccer, about your stocks in the stock market, etc.

For millions of men eagerly awaiting the next major technological development that would enhance their sex lives, the announcements about Roxxxy probably seemed too good to be true. And they were! The press launch of Roxxxy took place at the Adult Entertainment Expo in Las Vegas on January 9th 2010, but it posed more questions than it answered. It appeared, for example, that touching Roxxxy's hand caused it to exclaim that "I like holding hands with you," but what does that prove? Only that an electronic sensor was linked to some sort of recorded sound output. It was not a demonstration of the speech technology that would be needed in a talking conversational robot. And furthermore, Hines's behavior during the demonstration prompted the question – how much of the technology was inside Roxxxy and how much in the computer or whatever electronics were located behind the prototype?

The media hype surrounding Hines's launch in Las Vegas seems to have attracted the attention of many prospective customers for Roxxxy's supposedly seductive charms. At the beginning of February 2010, Hines's Web site started to take orders for Roxxxy, advertising the product at a "sale price" of $6,495, which it claimed represented a reduction of $500. Accompanying the invitation to place an order, the site also presented a "master agreement" that extended to 15 clauses of legalese, covering the purchase of Roxxxy and subscriptions to associated services, but the "returns, refunds and cancellation policy" of that agreement (clause 12.1) made it clear that once production of a customer's Roxxxy commenced, the purchaser could not get any of their money refunded. This begs the question why would any prospective customer be willing to part with their money without any possibility of recovery, when there had been no public demonstration or independent product review of a fully working Roxxxy that could perform as advertised?

Shortly after truecompanion.com started taking orders for Roxxxy, various news sites posted comments such as:

> Roxxxy won't be available for delivery for several months, but Hines is taking pre-orders through his Web site, TrueCompanion.com, where thousands of men have signed up.

Doubts about Roxxxy persist to this day (July 2015). David Levy wrote an exposé entitled "Roxxxy the "Sex Robot" – Real or Fake?" and posted it on www.fembotcentral.com. And the Wikipedia entry for Roxxxy (2012) includes the following:

> According to Douglas Hines, Roxxxy garnered about 4,000 pre-orders shortly after its AEE (*A*dult *E*ntertainment *E*xpo) reveal in 2010. However, to date, no actual customers have ever surfaced with a Roxxxy doll, and the public has remained skeptical that any commercial Roxxxy dolls have ever been produced.

If it is true that Hines received 4,000 preorders, then he would have raked in something over $20 million for those orders, since his Web site demands payment in advance. But as the above extract from the Wikipedia entry indicates, neither Hines himself or any of his customers has demonstrated, in public or to reputable media, the advertised features of Roxxxy actually working. Three years after its "launch," there still appears to be absolutely no sign of a demonstrable product that can talk about Manchester United (as Hines claimed Roxxxy could do) or perform in the other ways that Hines's advertising blurb claimed for Roxxxy.

Despite all the negative aspects of Hines's operation and of the product itself, the launch of Roxxxy at the January 2010 Adult Entertainment Expo can be viewed as some sort of milestone – a vindication of the forecast for a 2–3 year time span from late 2007 to the launch of the world's first commercially available sex robot. Hines has proved that there is indeed a significant level of interest in sex robots from the buying public.

Lovotics

Samani describes the design and development of a robot aimed at imitating the human affection process so as to engender attraction, affection, and attachment from human users toward the robot (Samani 2011). Then Samani summarizes the design of the robot thus:

> The artificial intelligence of the robot employs probabilistic mathematical models for the formulation of love. An artificial endocrine system is implemented in the robot by imitating human endocrine functionalities. Thus, the robot has the capability of experiencing complex and human-like biological and emotional states as governed by the artificial hormones within its system. The robot goes through various affective states during the interaction with the user. It also builds a database of interacting users and keeps the record of the previous interactions and degree of love.

The artificial intelligence of the Lovotics robot includes three modules: the artificial endocrine system, which is based on the physiology of love; the probabilistic love assembly, which is based on the psychology of falling in love; and the affective state transition, which is based on human emotions. These three modules collaborate to generate realistic emotion-driven behaviors by the robot.

The next four subsections summarize the formulation of love that underpins much of Samani's work, as well as the three software modules of the system mentioned above. The combined effect of these modules is to provide an artificially intelligent model that can display a range of emotions, adjusting its affective state according to the nature and intensity of its interactions with humans. The goal is to develop a robotic system that can exude affection for the user and react appropriately to affection from the user.

The Formulation of Love

The robot's intimacy software employs parameters derived and quantified from five of the most important reasons for falling in love (Levy 2007b): proximity, repeated exposure, attachment, similarity, and attraction. Intimacy in the robot is thereby related to those same factors that cause humans to fall in love. The robot utilizes audio and haptic channels in order to provide these different types of input which communicate the user's emotional state to the robot (Samani et al. 2010). The audio channel carries data for five audio parameters that characterize emotional cues within a human voice. The haptic channel carries data relating to the user touching the robot – the area of contact between robot and human and the force of that touch.

The Lovotics robot includes mathematical models for those five causal factors of love, creating a mathematical formula to represent each factor as well as a single "overall intimacy" formula which combines these five individual formulae into one. As an example of the five models, the proximity formula incorporates various distances between robot and human that indicate, inter alia, how closely the robot and human are to touching each other and how close they are emotionally.

The Probability of Love

The robot algorithm has taken account of the various factors that can engender human love, in order to develop a systematic method for assessing the level of love between a robot and a human. This is achieved by formulating probabilistic mathematical models for these factors, which in turn enable the robot to determine the level of intimacy between humans and robots. These models can be represented in a Bayesian network that depicts the relationship between love and its causal factors. The factors involved in this model include proximity (the physical distance between human and robot), propinquity (spending time with each other), repeated exposure (this can increase familiarity and liking in the other individual), similarity (this is directly related to the feeling of love), etc.

The probabilistic nature of these parameters allows a Bayesian network to be employed to link the parameters to relevant audio, haptic, and location data, leading to an estimate of the probability of love existing between robot and human. For example, audio proximity is employed in the calculations to emulate the effects of physical distance.

From the various causal parameters, the system calculates the probabilistic parameters of love, resulting in an appraisal of the level of love between human and robot (Samani and Cheok 2010).

The Artificial Endocrine System

The human endocrine system is a system of glands that secretes different types of hormones into the bloodstream. The purpose of those hormones is to maintain homeostasis, i.e., to regulate the internal environment of the body in order to keep certain functions stable, such as body temperature, metabolism, and reproductive functions.

The Lovotics artificial endocrine system is based on the human endocrine system, employing artificial hormones to create a simulation of the human system. The artificial hormones are software simulations of those human hormones which are related to the emotions – dopamine, serotonin, endorphin, and oxytocin, inter alia. The levels of these artificial hormones change dynamically due to the robot's interactions with users and according to its awareness of its emotional and physical circumstances.

The Affective State Transmission System

The affective state of the Lovotics robot depends largely on the various inputs it receives that are caused by its interactions with humans. Every interaction provides input data that is mapped onto a combination of six basic emotional parameters: happiness, sadness, disgust, surprise, anger, and fear. These six emotions are widely employed and described in the emotion literature.

The manner in which the robot's emotional state changes with the various inputs it receives is controlled by a model of emotion referred to as affective state transition (Samani and Cheok 2010). The Lovotics robot has a novel transition system which governs the immediate emotional changes in the robot. Their transition system functions in collaboration with the "probabilistic love assembly" module in order to control the overall emotional state of the robot. The short-term affective state of the robot is thereby transformed repeatedly into other affective states which are determined by the robot's previous affective states, its current mood, and the influences of the various input data it received during its interactions with humans, including audio and touch, and with its environment. For example, temperature could be one environmental input that might be programmed to influence the robot's affective state, if it "dislikes" being cold.

The Kissenger

In order for robots, such as the Lovotics robot, to have realistic physical interactions with humans, technology needs to be developed for human – machine kissing. In order to address this issue, Adrian David Cheok and his research team in Mixed Reality Lab has developed a kissing robot messenger called "Kissenger" (Fig. 2) (Samani et al. 2012).

We live in a global era, and more and more couples and families are apart due to work and business. New technologies are often employed to help us feel connected with those who we care about, through an increasing interest in touch and feeling communication between humans in the human-computer interaction community. Research like "Hugvie" (Kuwamura et al. 2013) and the "Hug over a Distance" project (Mueller et al. 2005) tested the feasibilities of telepresence and intimacy technology. However, these are big, bulky, and impractical.

There is some commercial work like "The HugShirt" (2002) and "Huggy Pajama" (Teh et al. 2008), which explore hugging in remote with love ones using wearable fashion technology. But these still lack a proper interface for "abstracted presence." Thus, Kissenger propose a new system to feel real presence using communication over internet for humans or robots.

Kissing is one of the most important modes of human communication as it conveys intimacy and many deeply feel positive emotions such as respect, greeting, farewell, good luck, romantic affection, and/or sexual desire through the physical joining or touching of lips by one individual on another individual's cheek, forehead, etc. (Millstein et al. 1993).

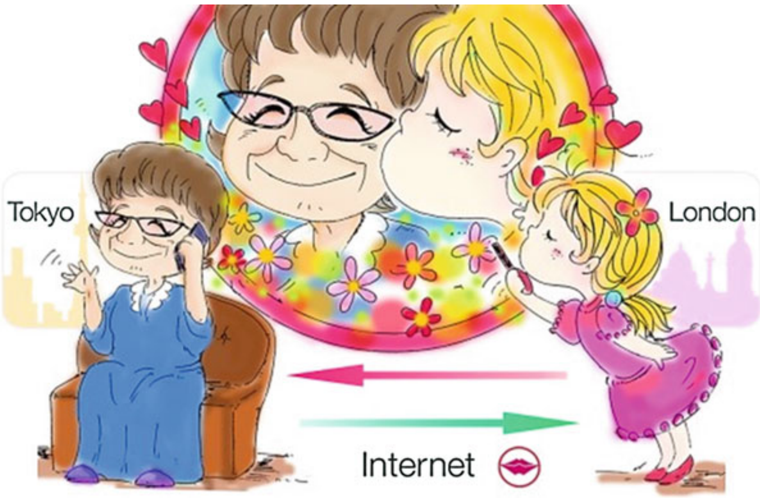


Fig. 2 The concept of kiss communication

The first Kissenger device to be developed by the Lovotics community was unveiled at the Designing Interactive Systems Conference in Newcastle in June 2012 (Samani et al. 2012). Then its development was started in the Mixed Reality Lab under the supervision of Adrian David Cheok. It would be possible to integrate the Kissenger technology into a sex robot but initially its use will be in teledildonic products for enabling lovers to kiss each other via the Internet.

The Kissenger employed soft, pressure-sensitive, vibrating silicone lips which, in the early prototypes, stood out from the surface of a smooth plastic casing shaped somewhat like a human head. Those early prototypes have since been replaced by a version for mobile phones.

When a user kisses the device on its lips, the changes in shape of the lips are detected by sensors and the resulting data is transmitted over the Internet to a receiving Kissenger, which converts the data back to lip shapes. This reproduces the changes in the kisser's lip shape, changes which are felt by the kisser's partner.

The Kissenger technology could perhaps be enhanced with an idea from a rather more ambitious haptic device of the same ilk which has been developed in Tokyo at the Kajimoto Laboratory in the University of Electro-Technology. Their invention is a French-kissing device (Takahashi et al. 2011), whose prototypes are not yet at a stage where they are likely to inspire erotic thoughts, being based on a straw-like tube that moves when in contact with a user's tongue. But we can expect to see an enhanced form of this idea in a future version of the Kissenger and similar inventions – enhancements under consideration at the Kajimoto Laboratory include adding taste, breath, and moistness to the experience.

During a kiss, along with its strong emotional and affectionate connections, a series of physical interactions takes place. The touch of the lips exchanges the pressure, softness, and warmth of each lip in a convincing way. The inventers of Kissenger approached this design problem carefully, given the intimate nature of the interaction and iteratively designed Kissenger which consists of two paired devices that can send and receive kisses simultaneously as shown in concept images of Figs. 3 and 4.

After studying the biological and psychological parameters of a kiss, a series of exploratory form factors were drawn to help visualize the possible interfaces. Figure 5 shows some of our initial concept designs.

At this stage, they looked for designing a system that effectively transmits the same sensation of kiss to one another. The one key issue was that the use of the device should be comfortable and not distract or obstruct the natural interaction of the kiss. Hence, they decided to integrate the initial concept design for a lip-like portable device with a minimalistic shape. However, one of the main concerns was the lip needed to be equipped with sensors and actuators. Hence, they looked into the possible technologies and sizes which could be fit into the form factor of our device. Figure 6 shows the 3D depiction of the proposed device with the new shape which can be attached to a smartphone, allowing a video call and virtual kiss simultaneously.

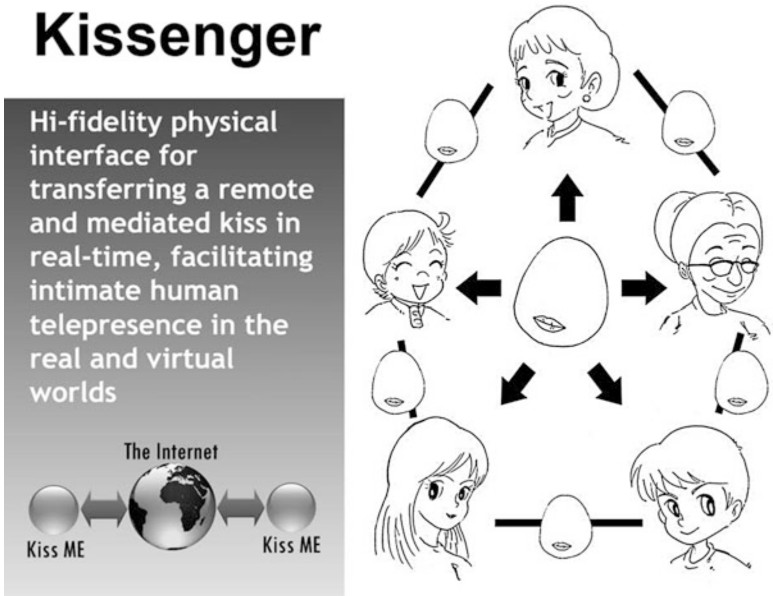


Fig. 3 Kissenger usage scenario A

Design Features

The interaction mechanism for Kissenger was devised with a number of features that make kiss communication between two users more meaningful. The system consists of following key features:

- Lip sensor push and pull reverse feedback for kiss behavior
- Lip rotation force feedback
- Sending scents
- Feeling LED light color communication (red, orange, green, and blue)
- Apps for kiss communication with video chat (Facetime, Google + Hangout, Skype, Facebook, Zoom, etc.)
- Changing the user characters and voices (face images)
- One-to-one pair and one-to-many user connections
- Recording the behavior of the partner's lips
- Scent tank that changes the scent to suit the partners
- Soft silicon cover made with gel for kiss communication

Design Flow

The hardware design of Kissenger with all the features listed above specifies the use of a light feeling sensor, pressure sensors, actuators, a vibration motor, a scent tank, and a smartphone connector in the Kissenger design flow. Their design role is as follows:

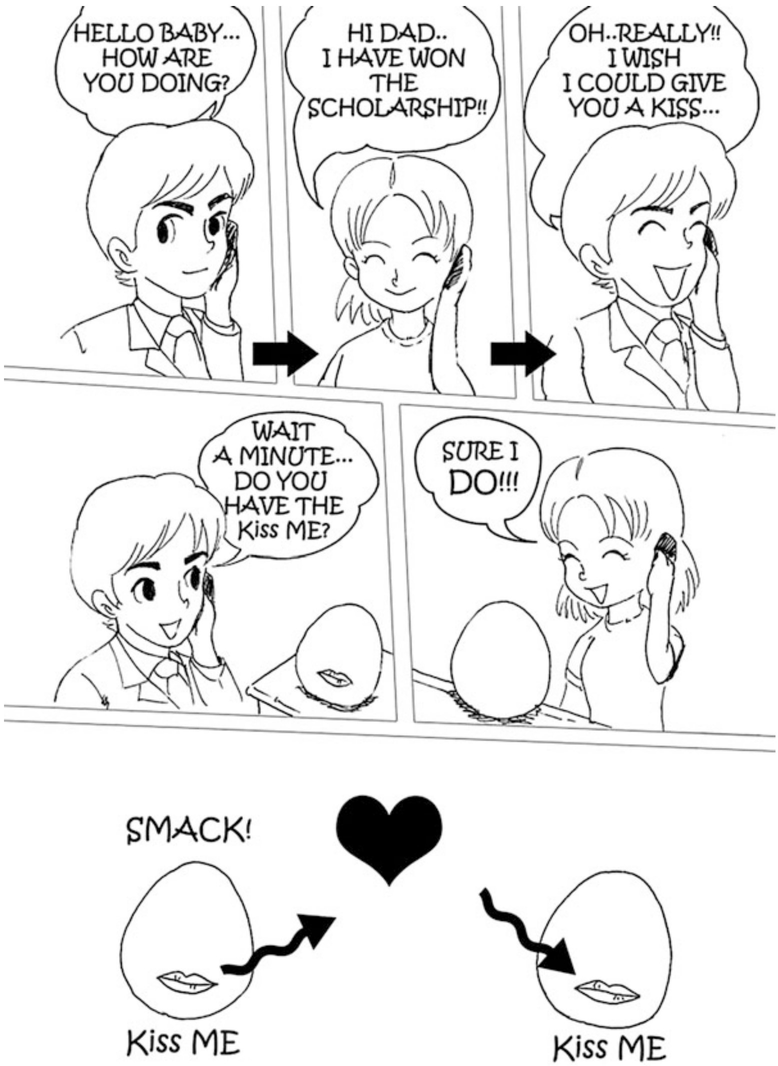


Fig. 4 Kissenger usage scenario B

Input Kiss Sensing

The front of the lip has pressure sensors placed just below the outer surface to initiate the Kissenger for the transmitter (the kissing person) and the receiver (the kissed person) and also to sense varying levels of soft touches. The features for the lip sensor push and pull reverse feedback for kiss behavior as shown in Fig 7.

Upon initialization, the front end of Kissenger can be tilted to a maximum of 18 degrees to replicate different styles of kissing. Thus, this design simplifies the

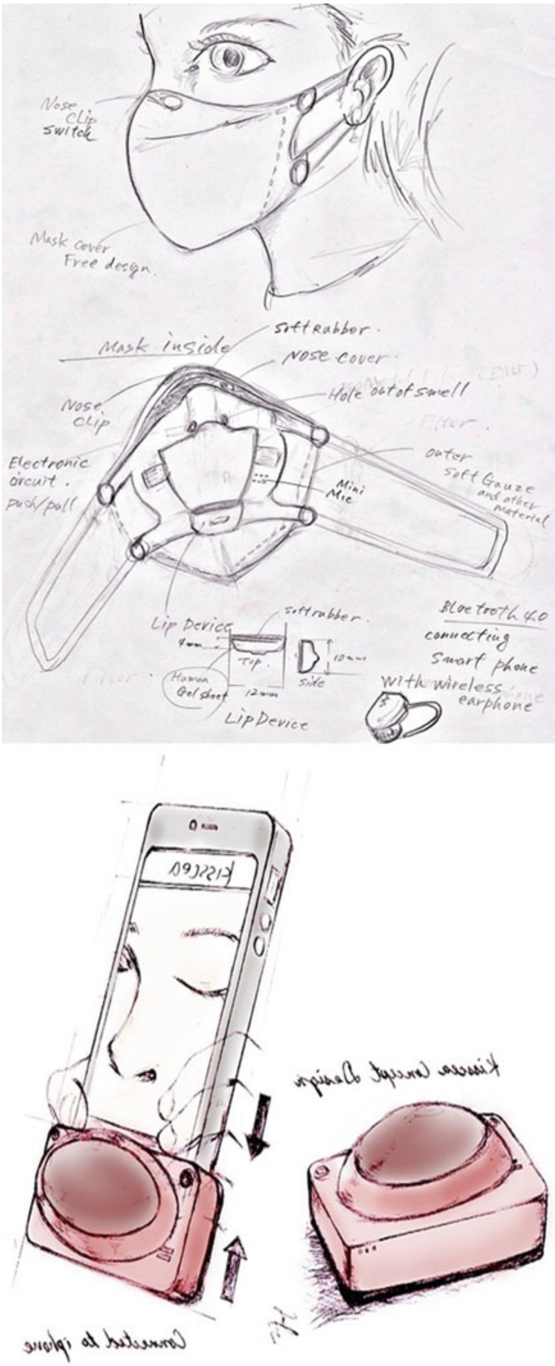


Fig. 5 Preliminary concept designs of Kissenger

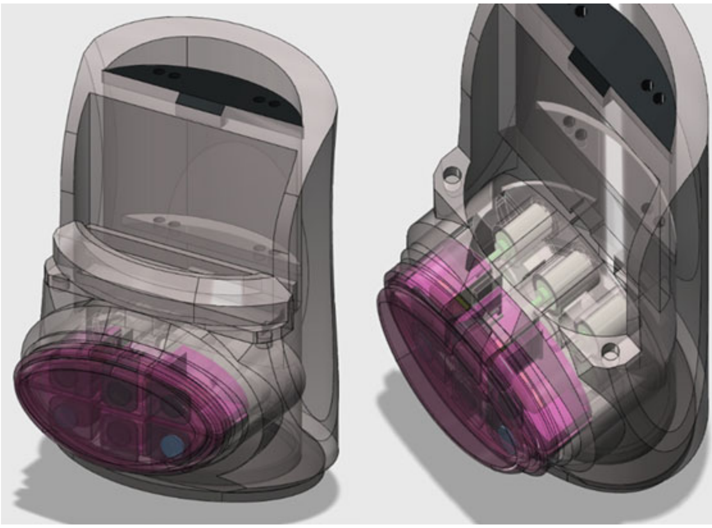


Fig. 6 The new design of Kissenger which can be attached to a mobile phone

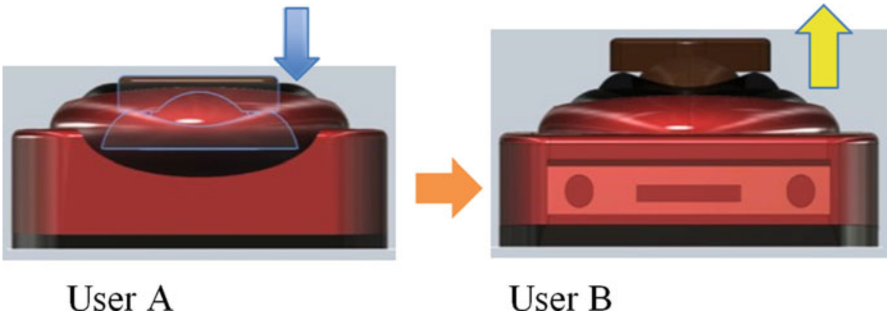


Fig. 7 Lip sensor push and pull reverse feedback for kiss behavior

interface and enables users to form a correct and semantically meaningful mental representation of the system with great feasibility for real kissing. The system also can be used for kissing a robot or a virtual 3D character.

Control and Wireless

Each Kissenger device is equipped with lip sensors (pressure sensor + heat sensor), a scent tank, a smartphone connector, and voice speaker (Fig. 8) connected to an embedded circuit that controls the sensors and actuators with your phone and

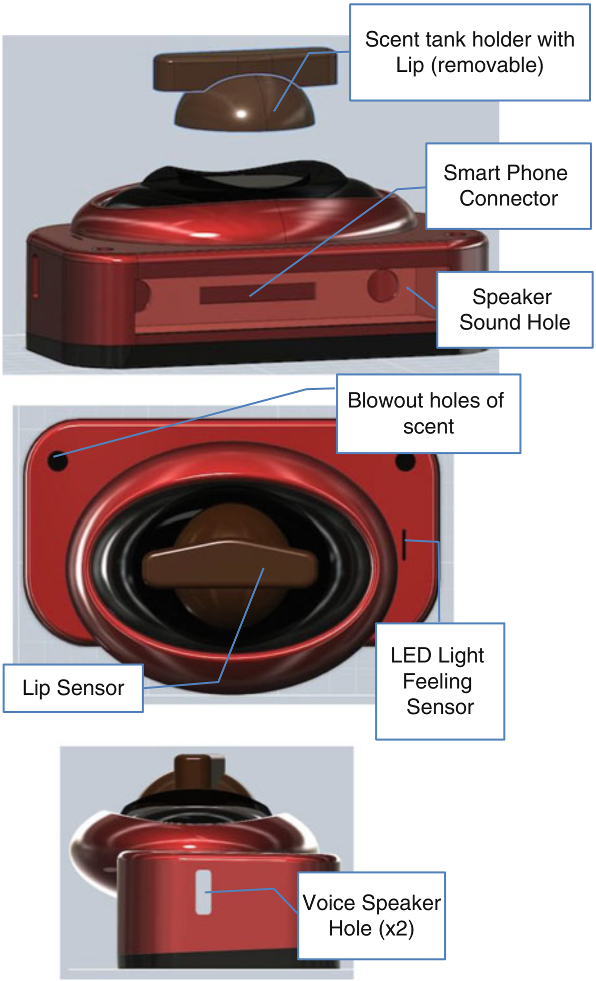


Fig. 8 Key design features of Kissenger

thereon with other Kissenger devices through the Internet. Data from the pressure sensors is read continuously until a change is detected. If there is a substantial change, the resulting increase is transmitted wirelessly to a receiver circuit that then actuates a servomotor array to produce similar motion of the lips.

Output Kiss Actuation

The kiss sensation on receiver (the kissed person) is produced through movement of servomotors that distend the surface of the lip. Simultaneously, the scent, LED light feeling sensor, and voice speaker are actuated for pheromone scents, colors to depict different moods, and sounds, respectively (Fig. 9). Pheromones are the scents

Fig. 9 LED light feeling sensor color depiction

LED Color	User's Feeling Image
	Blue: (Chat start) Feeling: Normal Chat Time: 0s~120s Number of Kiss 0 Number of Sent 0
	Green: Feeling: Peace Chat Time: 120s~240s Number of Kiss 1~3 Number of scent 1
	Orange: Feeling: Empathy Chat Time: 240s~300s Number of Kiss 3~5 Number of scent 2
	Red: Feeling: Love (sincerely) Chat Time: 300s~ Number of Kiss 5~ Number of scent 3 more (1/s)

used in Kissenger that are capable of acting outside the body of the secreting individual to impact the behavior of the receiving individual giving the feel of real presence of the partner. The shape and size of the lip covers hide the inner electronics that go into the sensing, control, and actuation of the device. Thus all these features make the user more amicable to this device and help evoke emotional responses and feelings for kiss communication.

Communication

Two or more Kissenger devices are wirelessly connected to each other via the smartphone Kissenger app, which are internally connected to their respective

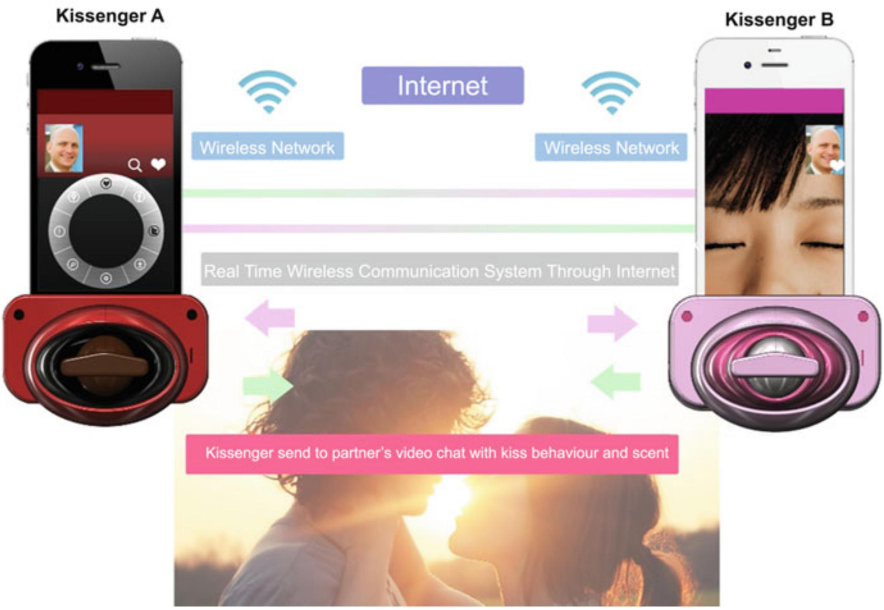


Fig. 10 Kissenger system diagram

smartphones as shown in Fig. 10. One of the unique added features of the app is that it allows one-to-many user communication along with one-to-one user communication as shown in Fig. 11. With the Kissenger app, the user can also actuate and transmit different colors to their partners to depict different moods with different scents, thus giving a real sense of kissing.

An assessment of the new proposed shape and its implementation was conducted with a wide variety of people including researchers not involved in our project, mall shoppers, and friends over a period of time with around fifty people from different cultural backgrounds, age, and sexes who participated in the evaluation process and provided feedback for the proposed shape and features. The major feedback is to integrate the size to make it more portable and user-friendly and provide the room for asynchronous kissing. There is the ability for the device to store a kiss that can be read at a later time on which the researchers will be working in the future for the social impact of this project.

The Ethical and Legal Debate

The ethics of robot sex were first aired in an academic forum at the EURON Workshop on Roboethics in 2006 (Levy 2006a, b, c). The following year this David Levy has discussed five aspects of the ethics of robot prostitution at an IEEE conference in Rome (Levy 2007c): the ethics of making robot prostitutes

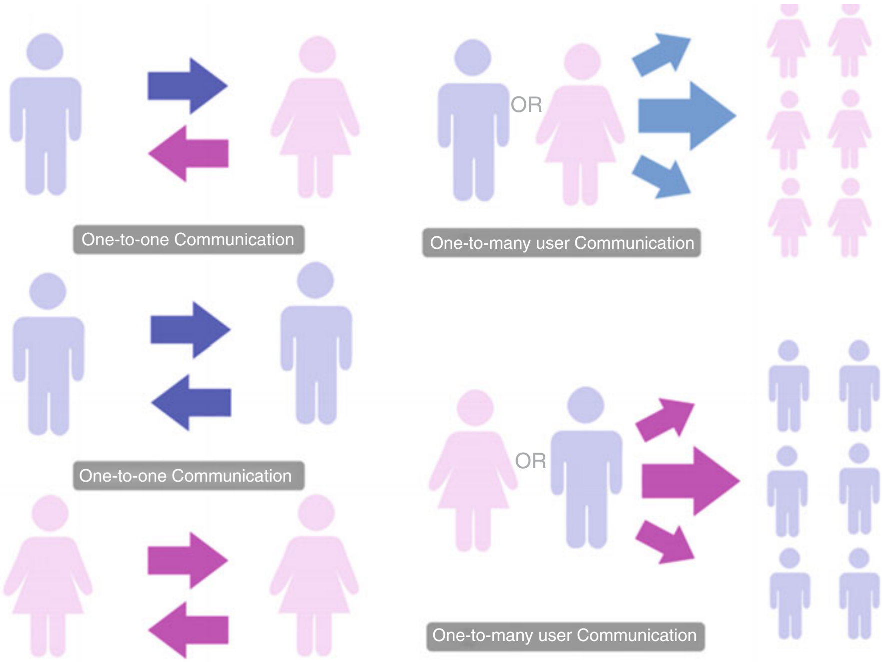


Fig. 11 User communication via Kissenger app

available for general use; the ethics vis à vis oneself and society in general, of using robot prostitutes; the ethics vis à vis one's partner or spouse, of using robot prostitutes; the ethics vis à vis human sex workers, of using robot prostitutes; and the ethics vis à vis the sexbots themselves, of using robot prostitutes. Since the last of these issues is only of significance if robots are eventually developed with (artificial) consciousness, it is also relevant when considering this particular issue to contemplate the ethical treatment in general of artificially conscious robots (Levy 2012).

A somewhat broader airing of the ethical impacts of love and sex machines was presented by John Sullins (2012). Sullins explores the subject partly on the basis that such entities are programmed to manipulate human emotions "in order to evoke loving or amorous reactions from their human users." He submits that there should be "certain ethical limits on the manipulation of human psychology when it comes to building sex robots," and accordingly he identifies three design considerations which he proposes should be applied to the development of robots designed for love: (i) robots should not fool people into ascribing more feelings to the machine than they should; (ii) robot designers should be circumspect in how their inventions exploit human psychology; and (iii) robots should not be designed to intentionally lie to their users in order to manipulate their user's behavior.

A considerably more strident attitude to the ethics of robot sex pervades a 2012 paper by Yusuff Amuda and Ismaila Tijani (Amuda and Tijani 2012), which views the subject from an Islamic perspective. These authors appear to have no doubts that "having intercourse with robot is unethical, immoral, uncultured, slap to the marriage institution and respect for human being." While many might not concur with the robustness of their position on the subject, it cannot be denied that the question of robot sex within the confines of marriage, or indeed within any existing human sexual relationship, is a serious issue. The question most often asked of the present author in media interviews has been: "Is it cheating for someone who is married or in a committed relationship to have sex with a robot?"

In this author's opinion, the answer is a resounding "no." A partner or spouse who has sex with a robot is no more guilty of cheating on their other half than are any of the tens of millions of women who use a vibrator. But not everyone agrees with this position, and in parallel with the possibility that sex with a robot should be regarded as cheating on one's spouse, there comes an interesting legal question which has been flagged by the California lawyer Sonja Ziaja (2011). Could a sex robot be legally regarded as the enticing protagonist in a law suit brought for the enticement of one's spouse? In the eight states of the USA where this type of law is still on the statute books, where they are called amatory or "heart balm" laws, Ziaja questions whether a sex robot could be held to be the cause, or a contributing cause, to the breakdown and dissolution of a marriage, and if so, who should be held legally liable to pay whatever damages a court might assess? Ziaja suggests a few obvious possible culprits for cases of enticement by a robot: the robot's inventor, its manufacturer, its owner, or even the robot itself. But the attribution of liability for a wrong wrought by a robot is an extremely complex issue, one which this author believes will not be adequately solved in the foreseeable future. Instead it has been suggested (Levy 2012) that robot wrongs could be compensated by an insurance scheme, much akin to that which works well for automobiles and other vehicles.

The only form of punishment considered by Ziaja for transgressing the American heart balm laws is to compensate the plaintiff, which is a notion that pales into insignificance when compared to the punishments discussed by Amuda and Tijani. They point out that, under Sharia law, judges are permitted to invoke lashes or even capital punishment for having sex with a robot, provided there is sufficient credible evidence of the crime (Ziaja 2011). "To this study, death penalty by hanging may not be applicable and implemented unless there are enough and credible evidences to justify the death by hanging of robot fornicator or adulterer."

Ziaja's paper largely avoids discussing punishment in relation to enticement cases in which a robot is the protagonist, preferring to prevent the problem from occurring by having robots designed in such a way as to incorporate feelings of heartbreak together with the goal of caring for those in its owner's circle of friends and relatives. "In order for robots to enter into human romantic relationships in a way that is consistent with the values underlying the heart balm torts, it may also need to experience heartache and empathy as we do." Ziaja's position thus supports that of John Sullins.

An in-depth consideration of whether or not human-humanoid sexual interactions should be legally regulated was discussed by Anna Russell in *Computer Law & Security Review* (Russell 2009). The very fact that such a discussion should appear in the pages of a respected legal journal points to the seriousness with which the legal profession is viewing the legal implications of the human-robot relationships of the future. Russell suggests that:

> Regulation of human-humanoid sexual interaction either by the state or federal government (In the USA.) will be sought when the level of interaction either (1) mimics human sexual interactions currently regulated or (2) will create a social harm if the interaction is not regulated... currently, in places where humans are using robots for pleasure in a sexual way that pleasure is either not regulated or is regulated in the way the use of any sexual device may be regulated" but that when more advanced robots – humanoids – are used for sexual pleasure, "then in many places, traditional norms and social mores will be challenged, prompting the development of state regulation. Will such regulation, then, be at odds with accepted notions of rights and freedoms?

Russell then delves further into the question of how regulation of human-humanoid sexual encounters would work and highlights some of the questions that will arise, including:

> How many rights will humans allow if humanoids clamor for sexual freedoms? How will humanoids be punished for sexual transgressions? Will humanoids need legal protection from the abuse of human sexual proclivities?

Russell's conclusion is a call for the:

> ...early discussion of the ramifications of a future species' demand for legal rights... the legal profession should develop legal arguments before attest case occurs in order to avoid the illogic and danger of arguments that stem from species bias.

In 2011 the *MIT Technology Review* conducted a poll on people's attitudes to the idea of loving a robot. 19 % of those questioned indicated that they believed they could love a robot, 45 % said "no," and 36 % responded "maybe." When it came to a question of whether or not people believed that robots could love humans, 36 % said "yes," only 23 % responded "no," and 41 % said "maybe." So already the idea of human-robot love was taking root as a serious proposition.

In a later poll about robot sex rather than robot love, which was conducted in February 2013 by The Huffington Post and YouGov among 1,000 American adults, 9 % of respondents indicated that they would have sex with a robot and 42 % opined that robot sex would constitute cheating on one's human partner (31 % said "no" to the cheating question, while 26 % said they were uncertain). This can be taken as further evidence that a significant portion of the population already regards robot sex as a serious subject. Just how serious can perhaps be judged by a news story that hit the media in March 2013 about an online auction for the virginity of a Brazilian sex doll called Valentina (Gates, 26) which was inspired by a 20-year-old Brazilian woman, Catarina Migliorini, who had auctioned her own virginity for $780,000 (sold to a Japanese buyer). True, a sex doll is only an inanimate product, lacking all the interactive capabilities of the sex robots of the future. But the level of interest demonstrated by this news story bodes well for the commercial possibilities of sex robots.

For the Brazilian sex doll auction, the online retailer Sexônico offered a complete "romantic" package for the successful bidder, which included a one-night stay with Valentina in the Presidential Suite at the Swing Motel in Sao Paulo, a candlelit champagne dinner, an aromatic bath with rose petals, and a digital camera to capture the action. If the successful bidder lived outside Sao Paulo, Sexônico also offered to provide a round-trip air ticket. Valentina's charms were not able to match the great commercial success of Ms Migliorini, but considering that most sex dolls retail at prices in the range $5,000–$10,000, the final bid of $105,000 was still a good result for Sexônico, not to mention the value of all the media exposure they attracted.

Robot Love

In parallel with the developments we have discussed in the field of robot sex and teledildonics, there is a continuing and burgeoning research interest in robot love. Among the fundamental conditions for engendering human love, physical appearance and attractiveness rank highly. The translation of these conditions to the field of robotics has a champion in Professor Hiroshi Ishiguro, whose research teams are based at the Graduate School of Engineering Science at Osaka University and at the Hiroshi Ishiguro Laboratory in the Advanced Telecommunications Research Institute International in Kyoto.

Ishiguro is famous for, inter alia, the amazingly lifelike robots he has developed in various human images (Hofilena 2013). These include one in his own image which is sometimes sent to deliver his lectures when he is too busy to do so himself. Another of his robots, called "Geminoid F" (Fig. 12), is made in the image of an attractive young woman who can blink, respond to eye contact, and recognize and respond to body language (Torres 2013). Ishiguro is encouraged in this aspect of his work by his conviction that Japanese men are more prone than are western men to develop amorous feelings toward such robots because, in Japan, with the influence of the Shinto religion, "we believe that everything has a soul and therefore we don't hesitate to create human-like robots."

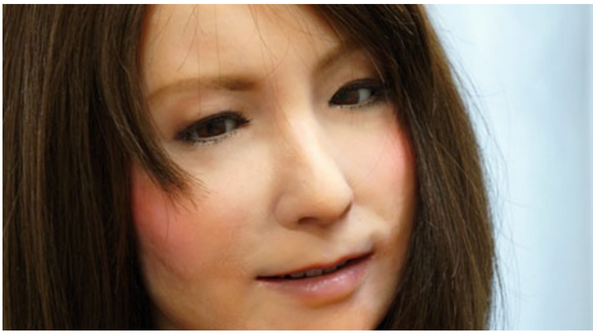


Fig. 12 "Geminoid F" robot (Countdown.org, 2015)

Another strand of Ishiguro's research into artificially engendering feelings of love in humans is concerned with promoting romantic forms of communication. The "Hugvie" (2011) is a huggable pillow, shaped in a somewhat human form, that is held by a user close to their body while they speak to their human partners via their mobile phone, located in a pocket in the Hugvie's head. (The Hugvie project grew out of an earlier Ishiguro project called "Telenoid.") The Hugvie incorporates a vibrator to simulate a heartbeat, and the vibrations emanating from it are synchronized with the sounds of the partner's voice. This allows the simulated heartbeat to be changed according to the volume of the partner's voice, with the result that the listening user feels as though they are close to their partner. The comfort felt by holding the cushion, the sense of hugging one's partner, hearing one's partner's voice close to one's ear, and the simulated heartbeat aligned with that voice, all these combine to create a sense that the partner is in some way present, which in turn intensifies the listener's feelings of emotional attraction for their partner. Ishiguro expects this intensified affinity to increase the sense of intimacy between couples who are communicating through their respective Hugvies. Ishiguro shared in a breakthrough study that the Hugvie could decrease blood cortisol levels, therefore reducing stress (Sumioka et al. 2013). Integrating the Hugvie technology into the design of an amorous robot might therefore enable a human user of such a robot to experience an enhanced feeling of a humanlike presence and a greater sense of intimacy from and for the robot.

Yet another direction of Ishiguro's research into having a robot engender emotions in humans is his investigation of the emotional effects, on a human user, of different facial expressions exhibited by a robot (Nishio et al. 2012). That research is currently in its early stages, but there is already some indication that it will be possible for robots, by their own facial expressions, to affect a user's emotional state. Emotional facial expression is also a hot topic at the MIT Media Lab, where the Nexi robot was developed (Allman 2009).

Predictions

Robot Sex

Clearly a significant sector of the public is now ready for the advent of commercially available sex robots, and the public's interest in and appetite for such products seems to be growing steadily. We have noticed a steady increase in the number of requests for media interviews on the subject during the past two years. Also growing steadily is the interest within the academic research community.

In our opinion nothing has occurred since the publication of *Love and Sex with Robots* to cast doubt on his 2007 prediction that sophisticated sex robots would be commercially available by the middle of this century. On the contrary, the increase in academic interest in this field has reinforced David Levy's conviction regarding that time frame.

What will be the next significant steps in this field? Intelligent electronic sex toys are gaining in popularity, for example, the SaSi Vibrator, which "comes pre-loaded with sensual intelligence which learns movements you like, specifically tailoring a unique experience by remembering movements that suit you," and the "Love Glider Penetration Machine" which can be purchased from Amazon.com at around $700 and which is claimed to "give you the most comfortable stimulating ride you will ever have!" The Amazon Web site also offers a very much more primitive looking sex machine at around $800, a machine of the type seen in many variations on the specialist site www.fuckingmachines.com, which "supports multiple positions and has adjustable speeds, strong power, and remote control." (The sole review on Amazon.com as of May 2013 suggests that this product is poorly made and describes it as "a piece of junk.")

Another research direction that perhaps offers even greater commercial potential comes from a combination of augmented reality with digital surrogates ("dirrogates") of porn stars. A recent (June 2013) posting by Clyde DeSouza (2013) posits that the 3D printing of human body parts will enable the downloading, from "hard-drives in Hollywood studios" of "full body digital model and "performance capture" files of actors and actresses." DeSouza continues:

> With 3D printing of human body parts now possible and blue prints coming online with full mechanical assembly instructions, the other kind of sexbot is possible. It won't be long before the 3D laser-scanned blueprint of a porn star sexbot will be available for licensing and home printing, at which point, the average person will willingly transition to transhuman status once the 'buy now' button has been clicked.
>
> If we look at Digital Surrogate Sexbot technology, which is a progression of interactive porn, we can see the technology to create such Dirrogate sexbots exists today, and better iterations will come about in the next couple of years. Augmented Reality hardware when married to wearable technology such as 'fundawear' (2013) and a photo-realistic Dirrogate driven by perf-captured libraries of porn stars under software (AI) control, can bring endless sessions of sexual pleasure to males and females.

Fundawear is a prime example of the increase in popularity of intelligent electronic sex toys and teledildonic devices. It is a wearable technology project currently under development by the condom manufacturer Durex, which allows lovers to stimulate their partner's underwear via their respective mobile phones. Such products seem likely to benefit from the increased academic interest in Lovotics, which will surely lead to at least some of the academic research in this field being spun off into commercial development and manufacturing ventures. And the more prolific such products become in the market place, the more the interest in them and in fully fledged sex robots will grow. How long will it be before we see a commercially available sexbot much more sophisticated than Roxxxy? Almost certainly within the next five years.

Robot Love

The past five years has seen a surge of interest in research projects aimed at different aspects of love with robots. One aspect is concerned with enabling humans to convey amorous feelings to artificial partners or to remotely located human partners with whom they communicate by artificial means (i.e., technology). Another aspect works in the opposite direction, enabling artificial partners to exhibit their artificial feelings, including love, to human partners. Some of this research has already demonstrated promising results, for example, the experiments conducted with Hugvie by Ishiguro and his team in Japan. They plan further research with the Hugvie to investigate how vibration can further enhance the feeling of presence experienced by a user. Additionally they plan to employ tactile sensors to monitor the emotional state of a user, which will provide feedback for the Hugvie and thereby enhance its ability to influence a user's emotions. Ishiguro's team has already found that hugging and holding such robots "is an effective way for strongly feeling the existence of a partner."

Another domain to become an important catalyst for the development of human-robot emotional relationships is what might be called girlfriend/boyfriend games. An example of this type of game is "Love Plus," which was first released in 2009 for the Nintendo DS games console and subsequently upgraded for re-release. A recent (February 2013) article describes the relationship between a 35-year-old Tokyo engineer, Osamu Kozaki, and his girlfriend Rinko Kobayakawa (Belford 2013). When she sends him a message:

> . . .his day brightens up. The relationship started more than three years ago, when Kobayakawa was a prickly 16-year-old working in her school library, a quiet girl who shut out the world with a pair of earphones that blasted punk music.

Kozaki describes his girlfriend's personality as being:

> . . .the kind of girl who starts out hostile but whose heart gradually grows warmer. And that's what has happened; over time, Kobayakawa has changed. These days, she spends much of her day sending affectionate missives to her boyfriend, inviting him on dates, or seeking his opinion when she wants to buy a new dress or try a new hairstyle.

But while Kozaki has aged, Kobayakawa has not. After three years, she's still 16. She always will be. That's because she is a simulation; Kobayakawa only exists inside a computer.

Kozaki's girlfriend has never been born. She will never die. Technically, she has never lived. She may be deleted, but Kozaki would never let that happen.

Because he's "in love."

Conclusion

In this chapter, we discussed about the possibility of human-robot intimate relationships and humanoid robot sex. We detailed Lovotics, which is a new research field that study emotions of robots with an artificial endocrine system capable of simulating love. We also presented the design and principle of Kissenger, an interactive device that provides a physical interface for transmitting a kiss between two remotely connected people. Finally we have discussed ethical and legal background and future predictions of love and sex with robots.

Recommended Reading

T. Allman, *The Nexi Robot* (Chicago, Norwood House Press, 2009)

Y.J. Amuda, I.B. Tijani, Ethical and legal implications of sex robot: an Islamic perspective. OIDA Int. J. Sustain. Dev. **3**(06), 19–28 (2012)

A. Belford, That's Not a Droid, That's My Girlfriend. The Global Mail, 21 Feb 2013. Available at www.theglobalmail.org/feature/thats-not-a-droid-thats-mygirlfriend/560/

Countdown.org (2015). [online] Available at http://www.countdown.org/media/attachments/covers/b2f5aa2c15f47659e9b2860ba2580eb9.jpg. Accessed 6 Nov 2015

C. DeSouza, Sexbots, ethics and transhumans (2013), http://lifeboat.com/blog/2013/06/sexbots-ethics-and-transhumans

Fundawear Reviews (2013), www.fundawearreviews.com

S. Gates, Brazilian sex doll's virginity: bids for Valentina's flower surpass $105,000. The Huffington Post, 7 Mar 2013

J. Hofilena, Japanese robotics scientist Hiroshi Ishiguro unveils body-double robot. Japan Daily Press, 17 June 2013. Available at http://japandailypress.com/japanese-robotics-scientist-hiroshi-ishiguro-unveils-body-double-robot-1730686/

Hugvie (2011) www.geminoid.jp/projects/CREST/hugvie.html. Accessed 8 Sept 2015

K. Kuwamura, K. Sakai, T. Minato, S. Nishio, H. Ishiguro, Hugvie: a medium that fosters love, in *RO-MAN, 2013 IEEE* (IEEE, 2013), pp. 70–75

D. Levy, A history of machines with sexual functions: past, present and robot, in *EURON Workshop on Roboethics* (Genoa, 2006a)

D. Levy, Emotional relationships with robotic companions, in *EURON Workshop on Roboethics* (Genoa, 2006b)

D. Levy, Marriage and sex with robots, in *EURON Workshop on Roboethics* (Genoa, 2006c)

D. Levy, Intimate relationships with artificial partners. PhD thesis, Universiteit Maastricht, 2007a

D. Levy, *Love + Sex with Robots* (HarperCollins, New York, 2007b)

D. Levy, Robot prostitutes as alternatives to human sex workers, in *Proceedings of the IEEE-RAS International Conference on Robotics and Automation (ICRA 2007), Workshop on Roboethics, Rome, Italy* (2007c). 10–14 Apr 2007

D. Levy, The ethical treatment of artificially conscious robots. Int. J. Soc. Robot. **1**(3), 209–216 (2009)

D. Levy, When robots do wrong. *Invited paper, Conference on Computing and Entertainment, Kathmandu*, 3–5 Nov 2012. Available at http://share.pdfonline.com/87cad18d73324e8fb2eaae1cddb60f77/Kathmandu_final_text_October31st.htm

S. Millstein, A. Petersen, E. Nightingale, *Promoting the Health of Adolescents* (Oxford University Press, New York, 1993)

F.F. Mueller, F. Vetere, M.R. Gibbs, J. Kjeldskov, S. Pedell, S. Howard, Hug over a distance, in *CHI'05 Extended Abstracts on Human Factors in Computing Systems* (ACM, 2005), pp. 1673–1676

S. Nishio, K. Taura, H. Ishiguro, Regulating emotion by facial feedback from teleoperated android robot, in *Social Robotics* (Springer, Berlin/Heidelberg, 2012), pp. 388–397

S. Nomura, J.T.K. Soon, H.A. Samani, I. Godage, M. Narangoda, A.D. Cheok, O. Katai, Feasibility of Social Interfaces based on Tactile Senses for Caring Communication, in *The 8th International Workshop on SID,* Vol. 68, No. 3 (2009)

Omicsonline.com, Lovotics Journals|Peer Reviewed|High Impact Articles List (2015), [online] Available at http://www.omicsonline.com/open-access/lovotics.php. Accessed 7 Sept 2015

A.C. Russell, Blurring the love lines: the legal implications of intimacy with machines. Comp. Law Secur. Rev. **25**(5), 455–463 (2009)

H.A. Samani, A.D. Cheok, Probability of love between robots and humans, in *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on* (IEEE, 2010). pp. 5288–5293

H.A. Samani, A.D. Cheok, F.W. Ngiap, A. Nagpal, M. Qiu, Towards a formulation of love in human-robot interaction, in *RO-MAN, 2010 IEEE* (IEEE, 2010). pp. 94–99

H.A. Samani, *Lovotics: Love + Robotics, Sentimental Robot with Affective Artificial Intelligence*. Doctoral dissertation, 2011

H.A. Samani, R. Parsani, L.T. Rodriguez, E. Saadatian, K.H. Dissanayake, A.D. Cheok, Kissenger: design of a kiss transmission device, in *Proceedings of the Designing Interactive Systems Conference* (ACM, 2012). pp. 48–57

J. Sullins, Robots, love, and sex: the ethics of building a love machine. IEEE Trans. Affective Comput. **3**(4), 398–409 (2012)

H. Sumioka, A. Nakae, R. Kanai, H. Ishiguro, Huggable communication medium decreases cortisol levels (2013). Scientific reports, 3

N. Takahashi, Y. Kuniyasu, M. Sato, S. Fukushima, M. Furukawa, Y. Hashimoto, H. Kajimoto, Kiss interface for intimate communications. J. Human Interface Soc. **13**(4), 53–62 (2011) (in Japanese)

J.K.S. Teh, A.D. Cheok, R.L. Peiris, Y. Choi, V. Thuong, S. Lai, Huggy Pajama: a mobile parent and child hugging communication system, in *Proceedings of the 7th International Conference on Interaction Design and Children* (ACM, 2008), pp. 250–257

The Colbert Report, [TV programme] CC.COM (2008)

The HugShirt (2002), [online] CUTECIRCUIT. Available at http://cutecircuit.com/collections/the-hug-shirt/. Accessed 7 Sept 2015

I. Torres, Japanese inventors create realistic female 'love bot'. Japan Daily Press, 28 Mar 2013. Available at http://japandailypress.com/japanese-inventors-create-realistic-female-love-bot-2825990/

Wikipedia, Roxxxy (2015), [online] Available at http://en.wikipedia.org/wiki/Roxxxy. Accessed 7 Sept 2015

I. Yeoman, M. Mars, Robots, men and sex tourism. Futures: The journal of policy, planning and futures studies **44**(4), 365–371 (2012)

S. Ziaja, Homewrecker 2.0: an exploration of liability for Heart balm torts involving AI Humanoid consorts, in *Social Robotics* (Springer, Berlin/Heidelberg, 2011), pp. 114–124

Media, Art, and Society: Interface of the Digital Image, Aesthetics, and Culture

34

Alistair D. Swale

Contents

Abstract
This chapter surveys the dominant trends and significant turning points in the changing perception of how media, art, and society have been discussed academically, with an emphasis on the twentieth to twenty-first centuries.

A.D. Swale (✉)
Screen and Media Studies, School of Arts, Faculty of Arts and Social Sciences, University of Waikato, Hamilton, New Zealand
e-mail: alexei@waikato.ac.nz

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_21

This incorporates an overview of salient developments in the philosophy of aesthetics during this period. It then proceeds to examine the more recent question of how digital technology has impacted on this nexus of interaction. While there is no commonly accepted consensus among academia with regard to the long-term implications of this impact, it is nonetheless possible to distinguish certain emergent changes in the spheres of narrative, persona, and spectacle within contemporary media and art, and it is possible to identify an emergent aesthetic regime which accounts for artistic experiences across a number of increasingly integrated media platforms and an expanded conception of the contiguity of aesthetic experience transcending individual perception and traversing the organic and the material. As a final section several symptomatic and suggestive case studies are raised to discuss how these play out in more concrete instances.

Keywords
Aesthetics • Art • Avatar • Challenge-based immersion • Christopher Nolan's Inception • Control theory • Digital art • Digital image • Digital Image • Final Fantasy: the spirit within • Imaginative immersion • Inception • Input/output theory • Machinic assemblage • Making use theory • Media studies • Procedural/ participatory theory • Resident Evil: Damnation • Sensory immersion • Silent Hill • The Anime Machine

Introduction

The role and nature of art in society has been fundamentally transformed through the development and implementation of new technologies of artistic expression and reproduction. Certainly technologies facilitating mass production and, concomitantly, mass consumption stand out as historically highly significant in transforming the experience of "high art" from one for the select few to something that could be more genuinely social and accessible to the entire population. This has not simply been a matter of quantitative expansion but also an intrinsic qualitative transformation as well. As the movable-type printing press ushered in greater numerical productions of texts it also entailed the ushering in of new forms of imagination, expression, and new audiences. In the last century, photographic and cinematic technologies have had a similarly profound impact on the volume of moving image works produced, the development of new modes of narrative and spectacle, as well as a new kind of spectatorship (and a new kind of expectation as concerns notions of entertainment).

Rather than covering this process exhaustively from a historical perspective, it should suffice to focus on academic and intellectual responses to the legacy of the photographic and cinematic image since the early twentieth century. Key to this discussion is a contextualizing of visual art entertainment in relation to the emergence of computer graphic imaging (CGI), augmented reality (AR), and digital cinema. In particular this entry discusses more specifically the impact of digital image production on the process of increasing integration between the audiovisual platforms most broadly dispersed in society today: the video game, social media streaming, television, and cinema.

Accordingly some discussion of key commentators on the impact of media transformations on art in society should precede the latter discussion. Figures such as Walter Benjamin (Benjamin 1968, 1978, 1999), Raymond Williams (1974), Marshall Macluhan (1962, 1964), Gilles Deleuze (1989a, b), and Lev Manovic (2001, 2006, 2013) in varied ways stand out as making particularly significant contributions to our understanding of this nexus of technology, art, and society.

The Problem of Technology and Human Agency

Walter Benjamin's commentary on the mass reproduction of artistic artifacts (1999) provides a useful springboard for highlighting a key question in our understanding of the relationship of technology to art: to what degree does technology obviate human agency in art? Benjamin famously suggested that the interpellation of a technical medium to reproduce art implied the loss of the original artistic expression's "aura," the authentic touch of the authentic artist. The response to this depends on what aspect of the process of artistic production and aesthetic experience one wishes to highlight. At the point of production, there may well be the authentic creative moment, but it can be argued that an aesthetic experience of the artifact that is produced at that moment is still possible regardless of whether it is the original or a highly faithful reproduction (this would apply not only in audiovisual art but also music and literature). It certainly remains moot, however, that in the more plastic arts, the capacity to reproduce the texture or the very sense of physical presence of a piece is problematic (although the advent of 3D printing poses an intriguing possible complication for that problem).

Benjamin's "problem" finds a further avenue for discussion in relation to the photographic image, in that it is not an image produced "by hand" so to speak but through the technological agency of the lens. Roland Barthes (2010), in contrast to Benjamin, did not apprehend this as a loss of artistic agency, although it certainly constituted a different order of image from those pertaining to the traditions of fine art. It was a new form of art, and one that had distinctive constraints and affordances.

Andre Bazin (2009) provides the cinematic correlate to Barthes' photographic commentary, in that he recognized the primacy of the camera's capacity to capture images that transcend subjective or authorial intentions. He embraced the immanence of the cinematic image in its own right and routinely lauded auteur cinema directors who gave primacy to techniques that would enhance precisely that affordance. Where Bazin goes further than Barthes is to suggest that this new medium was capable of transcending and in some sense surpassing prior art forms altogether.

Bazin's influence was profound, and yet there would be those who would re-engage with cinema as a medium that did not simply entail the recording of a personal vision as a visual event but as a medium with distinctive capacities to manipulate perception and consciousness. Gilles Deleuze's *Cinema I* and *Cinema II* (1989a, b) in particular investigate the impact of the manipulation of both camera and the reordering of cinematic footage to create new perceptions of time and space. Broadly speaking he saw the history of cinema in the twentieth century as entailing a move from the thorough apprehension of possibilities of the "movement image" (the movement of the camera to create new notions of perspective previously unattainable) to the "time image," the accentuation of the camera's capacity to "linger" without explicit narrative or spectatorial purpose, at times with simply the sole purpose of highlighting the tenuousness of chronographic representation of time in cinema itself (e.g., as seen with the "jump cut").

The Technological Focus Within "Media Studies"

In parallel to this response to this emergent commentary on the medium of cinema, there has been a broader commentary developed on media in general, largely initiated from within humanities and social science disciplines that had some degree of intellectual concern with mass communication and popular media. The appearance of Marshall McLuhan's *The Gutenberg Galaxy: The Making of Typographic Man* in 1962 paved the way for a more fully wrought analysis of the implications of mass media in *Understanding Media: The Extensions of Man* (McLuhan 1964). McLuhan's key insight that "[the] message of any medium or technology is the change of scale or pace or pattern that it introduces in to human affairs" was further developed in media-specific ways with, for example, Raymond Williams' seminal analysis in *Television: Technology and Cultural Form* of the distinctive characteristic of "flow" in television broadcasting (Williams 1974). This in turn paved the way for the emergence of "Media Studies" as a distinct discipline with a particular concern with the newly emergent experiences of the audience and the adaptation of new media technologies to making that experience more immersive and ultimately interactive.

The question of whether it was the media driving the change or the adaptive agency of the viewer driving it remained moot and found a focal theme for disputing the impact of technologies through the analysis of "media convergence." The notion of a single "black box" technological solution facilitating the integration of all media experienced in the domestic situation in one device, though promoted as an ideal by some home entertainment corporations, nonetheless has remained elusive. In *Convergence Culture: Where Old and New Media Collide*, Henry Jenkins (2006) arguably produced the most succinct and nuanced appreciations of the fact that new media technologies seldom completely obliterate "obsolete" ones, and in fact all media are in a constant state of retooling and redeployment in at times unanticipated ways (Jenkins 2006).

McLuhan's legacy has found a more recent reworking through Lev Manovich, who in *The Language of New Media* introduced a new layer of analysis rooted in the logic of computing and digital data (Manovich 2001). Most recently Manovich has engaged more specifically with the manner in which the computer was notably relatively absent from McLuhan's initial discussion of media and has himself posited "software" as the now pivotal dimension of contemporary media production and consumption (see *Software Takes Command: Extending the Language of New Media*, 2013).

Deleuze and the Cinematic Understanding of Time and Space

The thinker to perhaps most perceptively grasp the impact of the moving image as a novel media technology while also traversing some of the foregoing concerns regarding human agency vs technology is Gilles Deleuze. Though primarily noted for his writing on cinema, his commentary, as an attempt to map the interface between perception and its action within a particular medium, remains highly innovative and influential. His collaboration with Felix Guattari also spawned some of the conceptual tools (such as the rhizome) that have emerged as central to our understanding of interactive media and digital gaming.

As already outlined, Deleuze's distinctive insight was based on the observation that the employment of a camera to record a sequence of images that could then be reordered or interspersed in new ways would facilitate new perceptions of movement and time (Deleuze 1989). For instance, the recording of images where the body being filmed is stationary but nonetheless the camera moves around the object signifies a pure moment of cinematic motion, hence the "movement-image." Perhaps one of the best and most well-known instances of this is in the Matrix, where the main character, Uno, leaps into a dramatic attack pose and is then rapidly circled by the camera view creating an extraordinarily intense accentuation of both the camera as an independent source of motion and indeed the complete subsuming of the constraints of chronological ordering of images representing the actor's physical movement to a moment of utter stasis and, in effect, a continuously stretched moment of time that could not be represented any other way. Deleuze describes instances when these cinematic visions are achieved transcending convention models of perception as "singularities."

The way in which Deleuze escapes from a merely apparatus-centered account of the cinematic image is to characterize these cinematic interventions in space and time as instances of revealing and negotiating with a deeper "plane of immanence." At no time does Delueze privilege the cinematic apparatus as creating a totally independent "reality" – it must always be recognized as be rooted on a deeper plane of existence that we can never in fact totally encapsulate or articulate in any case. Cinema is nonetheless a vehicle for engaging in philosophy, and Deleuze has a distinctive way of articulating how our acts as philosophers move from the deeper realm of existence to explicitly conscious acts of perception. At base there is the realm of "affect," the sphere of sense and unorganized perception. "Concepts" are our attempts to give explicit and organized form to our perceptions, and these in turn can form the basis for testing and developing "ideas," categories of thought which transcend experience.

It is important to note that "affect" here refers to "sense" in its most broad meaning and is not limited to individual cognition but incorporates experience in a manner that acknowledges a profound inter-relatedness including even the physical body and physical matter. This is the basis for Deleuze and Guattari's notion of the "machinic assemblage" (Deleuze and Guattari 1988), a totality that is not in fact a "machine" per se (in the conventional English sense of the word) but a productive matrix wherein there is no predetermined privileging of any particular aspect of the world or human experience. In this manner, whether we accept Deleuze and Guattari's characterization of human existence or not, we see that Deleuze developed a highly novel solution to issues of human agency, the role of a particular apparatus (the camera) and its utility in exploring human experience.

The question that emerges next, of course, is what happens when we no longer use cameras to create images? This is the question that places the impact of digital imaging in clear relief; there is a point at which a computer-generated image can produce a semblance of the photoreal image that is almost next to indistinguishable. But if we jettison the assumption that the primary function of CGI is to replicate the cinematic image, then there emerges a new set of possibilities. In a sense, we are confronted with a parallel set of issues that emerge when we consider the implications of the cinematic vision for our capacity to rework space and time; what kinds of alternate modes of imagination do digital technologies afford?

Toward a "Postcinematic" Understanding of Digital Technology and Art

Overall, then, the terrain for the analysis of the relationship between media, art, and society has been located firmly in the realms of cinema, mass entertainment, and emergent digital technologies, with a pivotal dichotomy of technological determinism versus human agency. It has been perhaps unavoidable, to some extent, that cinema would be central in the commentary of mass media and art for the greater part of the twentieth century; the impact of the moving image, particularly in terms of the reworking of perceptions of space and time and the development of new forms of mass consumption of cultural tropes and narratives, has been immense. Indeed, some of the most cogent analyses of the transforming of human perception and expression have emerged out of film theory, and we are now confronted with a rapidly transformed terrain through the arrival of the Internet, digitization of popular media, and the emergence of increasingly interactive and immersive technologies. Consequently, we can be said to be entering a “postcinematic” phase of discourse around contemporary media and entertainment.

The Problem of Aesthetics Within Modern Philosophy

One major lacuna in the discussion of contemporary media and entertainment is aesthetics. The reasons for resisting a discussion of aesthetics in media studies are multiple but can be attributed to some fairly self-evident and palpable causes. The first is the globalization of world media; if aesthetics is about the development of rigorous philosophical criteria for the evaluation of art, cultural relativism immediately emerges to problematize the very prospect from the outset. The second is the inherent conceptual instability of art and a philosophical resistance to permitting that to contaminate a scholarly analysis through an engagement with it through aesthetics. Finally, there is the tendency to regard aesthetic perception itself as being grounded in individual consciousness, which posits the more fundamental problem of avoiding a subjectivist account of art and aesthetic experience (Danto 1986; Scruton 1998; Forsey 2003).

The question of the significance of aesthetics in relation to digital culture of course stems from a more fundamental consideration of the place of aesthetics in contemporary philosophy. Considering two of the more influential commentators on aesthetics in the last three decades – Danto and Scruton – there is a common argument that aesthetics has been historically marginalized within broader philosophical investigation in the Anglo-American context and they contend that this is indeed a deplorable situation. Their respective accounts of how we got there are different, Danto accentuating the rise of a particularly detrimental hegemony of logical positivism, while Scruton has more broadly identified the pervasiveness of scientific and empirical methodologies which have been brought to bear on considerations of art and thereby exterminated it. As Jane Forsey has argued, there are some prima facie circumstances and identifiable philosophical “culprits” that can be “blamed” for their role in undermining aesthetics as a valued facet of philosophy more broadly construed.

A. J. Ayer presents as a prime advocate of a radically constrained view of “legitimate” philosophical questions which excises, rather dogmatically as many might now acknowledge, any problems rooted in either ethics or aesthetics from the purview of philosophy. “Stealing is wrong,” for example, is a statement that indicates a series of premises and assertions that cannot in fact be verified empirically, and if that is the case, philosophy should not have a bar of such propositions. As Forsey indicates, this was such a constrained view of philosophy that in fact it did not hold sway quite as profoundly as might seem. Her litmus is the degree to which ethics, as a field of philosophical concern dealing with philosophical matters demanding practical guidance for life, became the key to demolishing the dogma of empiricism – a philosophy without ethics was not satisfactory in any degree. This, of course, begs the question why aesthetics was not rehabilitated at the same time.

And here Forsey offers the insight that aesthetics at this point was “twice removed” from philosophy; it was already being radically differentiated from the spheres of ethical concerns and therefore contending for a space autonomous from all other facets of philosophy.

Forsey argues persuasively that it was in fact a movement from within the community of artistic practitioners – the advocates for “art for art’s sake” – who had a pivotal role in separating aesthetic concerns from ethical ones. In fact she elucidates how this movement entailed a three-pronged refutation of association with other aspects of public concern; in sum, it resulted in a refutation of any assumption that art had a concern with “truth,” “beauty,” or the “ethical” (Forsey 2003).

These observations are salutary in counterbalancing a causal account of aesthetics being marginalized through the hegemony of a particular scientific intellectual tradition. Even so, it can be argued with equal force that perhaps this particularly “autonomous” view of art was in fact precipitated by a perception that art was under attack from positivist and empiricist views of the world and that this pleading for a “special case” for art was in fact the only recourse open to anyone hoping to defend it from Philistinism or moral conservatism.

In any case, the commonly accepted view is that aesthetics has suffered from a degree of marginalization in the “modern” period, and anyone who wishes to reintroduce a discussion of aesthetics into a consideration of contemporary culture has their “work cut out.” For Forsey’s part she identifies the institutions shoring up perceptions of “fine art” as being the culprits of excluding aesthetic concerns from broader social awareness, and by extension broader philosophical concerns. As a response to this condition she argues for a return to identifying the aesthetically beautiful in everyday life, focusing primarily on the sphere of design and the creation of artifacts for daily use.

At the same time there has gradually been a refocusing of aesthetic discussion away from artistic outputs and products to the creative process of bringing them into existence, and indeed the process by which we as spectators engage with such artistic artifacts. This has a precursor in the “spectatorial” conception of aesthetics, but the contemporary difference is that the focus on creative process expands concerns further than simply a subjective (and implicitly isolated) reception of art.

Aesthetics and the Digital Image

Jacques Ranciere has emerged in the last decade as a leading advocate for a reinvigorated engagement with aesthetics in the analysis of contemporary media (cinema in particular). In the translated collection of essays entitled *The Future of the Image* (Ranciere 2007) he works from the premise that “the image,” and how we apprehend and respond to an image, needs to be understood in a manner that transcends particular media. All images represent to some degree or another, and when it comes to the exercise of imaginative faculties to engage with any image, there remain persistent continuities. Ranciere even suggests that the imaginative dimension of aesthetic experience remains constant whether indeed we view a painting or watch a film.

This is not to say, however, that all images produced at all times are subject to the same criteria of scrutiny or aesthetic judgment. Building on Aristotle's classical framework of aesthetic theory as developed in *The Poetics*, Ranciere focuses on the changing relation between *muthos* (narrative) and *opsis* (spectacle) over time, and he describes a fundamental transition in aesthetic judgment that occurs toward the end of the eighteenth century. The late Enlightenment philosophy of aesthetics he describes as the representative regime, the primary concern being with accurate representation with a view to securing more accurate knowledge and ultimately freedom. This gives way from the early nineteenth century onward to the aesthetic regime, and he attributes the transition as being largely due to the emergence of photography (and later on cinema). There is a rupture, a contradiction that emerges in the wake of the photographic image. Spectacle in a sense "usurps" the place of narrative and while not altogether obliterating it engenders an inherently unstable relation.

The foregoing offers a plausible account of the inherent paradox of cinematic art and ultimately the even more precarious position of narrative in new media when the aspect of spectacle is augmented by interactivity. Again, it is not the case that narrative recedes into oblivion (although it may sometimes seem like that for those who are comfortable with more explicit and serialized plot structures).

It is at this point that some comment can be made on the relevance of contemporary aesthetics to "digital culture" and "digital art." There is a great deal of diversity in what is intended by terms such as "digital art." Christiane Paul has perhaps one of the most serviceable general definitions in that she highlights the heterogeneity of approach that makes a unified exceptionally problematic, but there is a consensus that "digital arts" can be gathered together on the basis of digital practices that either transcend "analog" modes of artistic expression or augment traditional creative practices to the extent that they become highly contingent on digital platforms for their articulation and dissemination.

The digital also has a significance for our understanding of the status of the creative agent and the ontology of artifacts. As will be discussed in more detail in the ensuing sections, we have on one hand the media theories of Lev Manovic which radically subsume the agency of artists within the digital tools that are employed. On the other hand we have theorists such as Thomas Elsaesser and Malte Hagener who accentuate the deep integration of the corporeal and digital technology, as they explore in their work *Film Theory: An Exploration Through the Senses* (Elsaesser and Hagener 2010). And there is also Steven Shaviro's substantial contribution to this branch of theoretical exploration which commences with (Shaviro 1993) and most recently has culminated in *The Universe of Things: On Speculative Realism* (Shaviro 2014).

Understanding the Implications of the Digital Image

It could not be said that there has been a clear and coordinated academic response to the digital revolution in contemporary media. What we find is that representatives from various disciplinary backgrounds assay into the realm of the relative unknown and attempt to make reasonable sense of where they believe the revolution is headed based on their own background and experience. We also find that any existing predilection for a particular emphasis on either the technological apparatus or human cognitive agency finds its way into the analysis proffered. Others, picking up on Deleuze's notions of affect and the "machinic assemblage," aim to explore the contiguity of digital image production and consumption with the haptic and the corporeal, whether this be in the context of the experience of 3D cinema or the immersive computer game.

In almost all cases there is a degree of acknowledgment that dominant position of cinema as it was once understood in some classical sense is no longer tenable; in this very general sense we can say that the paradigm for discussing contemporary digital media and mass entertainment has shifted to a postcinematic one (although there is as yet little agreement on what that might mean).

Lev Manovich and the "Velvet Revolution"

Lev Manovich has been one of the most articulate, prolific, and outspoken advocates for a radical reconceptualization of the nature of the relation between art and technology by incorporating his substantial background in computer science to an appreciation of the changing logic of image production. Accordingly he accentuates the emergence of the database as one of the key differentiators of where the material that ultimately constitutes the image is reposited. He also accentuates, as already discussed earlier, the profound impact of software technology in terms of facilitating modes of open-ended image construction that are carried out by a practitioner audience. The increasing accessibility of powerful open-source software packages means that the computer is increasingly structured around facilitating creative production rather providing ready-made content.

The Language of New Media (Manovich 2001) was a *tour de force* in terms of systematically outlining how digital technology was engendering new media that reflected the architecture of the computer. The products of new media were unavoidably structured according to new principles: numerical representation, modularity, automation, variability, and transcoding. The screen was being reconfigured as part of a new human-computer interface, and the operations of digital image construction would be pursued as part of a distinctive practice of compositing (rather than sequencing). This in turn ushered in a new set of imaginative possibilities, a kind of synthetic realm with new possibilities of illusion, narrative, and interactivity (this he discusses more specifically under the theme of "teleaction").

For a more detailed and focused example of how the foregoing would work out in practice, *After Effects, or Velvet Revolution* (Manovich 2006) describes the impact of a particular software package (Adobe After Effects) on image construction, dissemination of content, and new imaginative tropes of expression. It is here that he highlights the role of the technology in driving the emergence of a new set of aesthetics, diverse but nonetheless sharing a broadly consistent logic.

While few would debate Manovich's grasp of the technological dimension of contemporary digital media, there remains, arguably, the difficult question of how far human agency becomes subsumed in that technology. The prospect of a cybernetic aesthetic impulse is problematic; do machines have aesthetic capacities? And if they do, what is the more precise nature of the interface between human aesthetic impulses and binary code? There is no simple answer to this; however, Manovich's more recent discussion of the impact of software holds out the prospect of a more detailed investigation of this terrain.

The Response from Within Film Theory

In contrast to Manovich's highly technical exposition of the impact of the digital image, there have been two broad lines of response to the digital revolution in cinema from within film studies. The first is rooted in the Deleuzian tradition and focuses on developing new notions of "affect"; the other has emerged as something of a rejection of that tradition focusing instead on a broadly held belief in the ongoing relevance of approaching the filmic experience, regardless of whether it is digital or analog, from the dimension of cognition. This latter movement has been championed by some of the veteran scholars of film studies such as David Bordwell (Bordwell and Carroll 1996), who rather strikingly gave the cognitivist school its broad sobriquet of "Post-Theory." While the cognitivist approach has been characterized by increasing interdisciplinarity and a proclivity with scientific evidence to underpin conclusions, there seems to be no pronounced movement to date to problematize the digital in the manner it has been treated by either Manovich or by theorists broadly influenced by Deleuze.

The Corporeal Turn: Steven Shaviro and "Post-Cinematic Affect"

With *The Cinematic Body* (Shaviro 1993) Steven Shaviro initiated a two-pronged initiative that would have some deeper significance in the future development of film theory. On the one hand, his deeply felt ambivalence toward certain strands of postmodern theorizing, particularly those preoccupied with the broad application of psychoanalytic concepts to media analysis, was presented and aimed to refocus film theory back toward the experience of the moving image in some more fundamental sense. On the other hand, he aimed to explore and expound through a series of reflections on various case studies taken from film and television which adumbrated the physically contingent aspects of experiencing cinematic images.

Shaviro's work was in part resonant with the emergent discourse of disquiet with regard to postmodernism which was to find expression in David Bordwell and Noël Carroll's *Post Theory* (Bordwell and Carroll 1996), a work which became something of a manifesto for film theorists who wanted to distance themselves from Marxist structuralism, cultural studies, and varieties of psychoanalysis. However, Shaviro's concerns with ontology and affect would decisively mark a point of disjuncture with the "Cognitivist" school that arose out of Bordwell and Carroll's critique.

The work of developing the concept of affect in relation to film theory by Shaviro can be fruitfully contrasted with that of Brian Massumi, who in *Parables of the Virtual: Movement, Affect, Sensation* (Massumi 2002) re-engages with the legacy of Deleuze and Guattari to explore avenues of affective consciousness that cannot be contained in a closed relation to explicit signifiers.

Film Through the Senses: Elsaesser and Hagener

Thomas Elsaesser and Malte Hagener's *Film Theory: An Introduction Through the Senses* (Elsaesser and Hagener 2010) is a thorough investigation of the relationship between cinema, perception, and the human body, commencing with the more conventional terms of reference ("Cinema as Window and Frame," "Cinema as Door – Screen and Threshold") but quickly incorporating a language steeped in the physiological ("Cinema as Skin and Touch," "Cinema as Brain – Mind and Body"). Much of the exposition serves as a general overview of film theory (albeit from the rather distinctive angle as just alluded to); what makes this a rather distinctive commentary is the manner in which the authors reflect on the foregoing material with an extended conclusion which deals with the question of digital cinema, virtual reality, and media convergence.

The observation that Elsaesser and Hagener commence with is the fact that all three of the foregoing phases – digital cinema, virtual reality, and media convergence – are profoundly self-contradictory. For a film that is 100 % percent digitally produced, there is an implication that it was not "shot" with a camera and therefore lacks one of the key defining elements of the cinematic medium. Virtual reality has a more self-evident sense of contradiction: the reality that has not existed. Media convergence is perhaps more subtly problematic; cinema is now viewed on a variety of screens and handheld devices, repackaged and recirculated in multiple formats, often online. Even so the IMAX cinematic experience remains as a reminder that cinematic projection, like any "superseded" technology, can find a new lease on life in unanticipated ways.

These contradictory phrases in a sense signify an inherent rupture in media history. As Elsaesser and Hagener convincingly argue, the appearance of some of the first entirely digital films for the mainstream produced from Pixar, such as *Toy Story* (1996) and *Monsters Inc* (2001), arguably presented a "meta-commentary" on the reversal of the position of the photographic image and were emblematic of the fact that it was a tradition of graphic art preceding the photographic that was reascendant. Perhaps the most incisive comment, however, resides in their commentary on the transformation of the use of CGI away from merely simulating a photoreal world to creating completely alternative worlds. The telling metaphor is the transition from the screen as "window" to the screen/device as *portal*.

"Cinema 3.0"

Elsaesser and Hagener's account of digital cinema can be usefully augmented by the work of Kristine Daly, who in "Cinema 3.0" goes further to elucidate on the palpable impact of digital technologies on narrative, character, and spectacle. Referring to film texts such as the *Pirates of the Carribean* franchise and televisual texts such as *Heroes* and *Lost*, she persuasively argues that they reflect a narrative logic embedded in social media and the video game – they are episodic and driven by a rhythm of spectacle rather than an intrinsic imperative toward coherent plot. Within this context character also gives way to become a weaker facet of the production's rationale, to the point that a character can become more easily dispensable or, conversely, diversifiable (or even replicable).

Christopher Nolan's *Inception* (2010) arguable provides a further example of this: the settings of the action are deliberately diverse and unrelated, the protagonists do not exist necessarily in "real" time, and we are not necessarily left with any assurance that their commentary within the film is at all the "authoritative" one. Moreover, the CGI effects are not employed to portray a plausible alternative version of the world we live in but in fact a sublime version of a world that is somehow plausible but only exists beyond the bounds of physics as we know and experience them.

Thomas Lamarre and "The Anime Machine"

The Anime Machine: A Media Theory of Animation, (Lamarre 2009) has a particular significance in that it explores the internal dynamic of the emphatically 2D tradition of animated cinema pursued by Studio Ghibli and explains it within the contemporary milieu of digital design. In keeping with the implicitly Deleuzian premise of the title, Lamarre sets out to explicate how "animatic" images are constructed, how they cohere in space and time, as well as how this distinctive "machine" structures thought.

Lamarre discusses the relation between cinema and animation more directly in the earlier article "The First Time as Farce: Digital Animation and the Repetition of Cinema" (Lamarre 2006). In a fashion similar to Elsaesser and Hagener, Lamarre focuses on the film *Final Fantasy: The Spirit Within* (2001), which he suggests epitomizes what Manovich conceived as the "ultimate digital cinema": a series of thousands of images crafted "by hand" but indistinguishable from the photoreal. Lamarre is not altogether convinced that this constitutes the definitive litmus of digital cinema but argues that at the very least it provided a pivotal moment where animation was freed from its rather tragicomic fate of being a slave to "farce." Ultimately he sees animation, retooled through digitization, as having the capacity not to altogether surpass cinema but in a sense remake it (hence the reference to "repetition").

Interactivity, Immersive Media, and Telematic Performance

Insofar as digital imaging technologies are used simply to replicate the experience of cinema in some classical sense, there can be no apparent need to revise the conventional understanding of the screen experience. Yet, as was discussed above, a number of subtle transformations have occurred within the nature of media experiences and have identifiable impact on certain aspects of narrative, persona, and spectacle. In the following section we will look more particularly at three concrete consequences of the impact of digital technology as they illustrate a transformation of the experience of the moving image. Interactivity, immersion and "telemetry" are symptomatic of a fundamental break with the classical notion of cinematic spectatorship where media platforms become highly integrated and interchangeable and media experiences are increasingly "transversal." The move toward affect in its more extensive conception (including the "corporeal turn") reflects a raft of media experiences that entail to varying degrees haptic, performative, and at times deeply visceral engagement.

Interactive Digital Art

Many an academic commentary will commence by conceding that "interactivity" as a media concept is employed prolifically but remains persistently difficult to define with a great degree of consensus. Lev Manovich has gone so far as to critique the use of the term as in some contexts useless. In the first instance, he highlights the case where a person modifies images on a computer screen in real time through the manipulation of the interface; this is hardly "interactive" but rather a case of data input. Secondly, he highlights the case where the active audience member could be construed as interacting with media simply by mentally augmenting the visual stimulus with their own intuitions or information.

These objections aside, there are potentially other instances of media experience that could be construed as more or less genuinely interactive. In his article "What is Interactivity?" (Smuts 2009) Aaron Smuts undertakes to delineate five theories of interactivity, each of which he regards as coming closer to realizing the interactive ideal.

The first, "Control Theory," resonates with Manovich's first instance cited above. The theory focuses on the implications of the DVD and the audience's capacity to control access and display, quite apart from the intent of the creator. While this might indeed be frustrating for a filmmaker, it nonetheless does not rate as a particularly high-grade instance of interactivity. The second, "Making Use

Theory," responds to an account of interactivity where a concept of interactivity as a continuum from merely static information provision to more multifaceted platforms for exploratory accessing of various data sources is developed. Accordingly, the Internet is regarded as inherently "interactive." But as Smuts notes, the facility for flipping from one information source to another online does not differ intrinsically from having multiple print sources laid out in front of one. The third, "Input/Output Theory," focuses on a more concretely defined continuum based on the presence of sensors, data triggers, and output data generated in real time. Though more detailed, this too seems susceptible to some of the earlier critiques. The fourth variant, "Modifiable Structure Theory," discusses a model where differentiated criteria for "weak" and "strong" interactivity are identified – but this too is regarded as problematic due to the inherent difficulty of defining what aspects of a particular structure lend themselves to such variable intensities.

A more persuasive model seems to emerge in the form of what Smuts describes as the "Procedural/Participatory Theory" of Janet Murray. The procedural dimension refers to the current computer programming paradigm of integrating an assemblage of interrelated programming procedures to respond to specific but nonetheless not preordered stimuli. Even so, as Smuts suggests, a procedural assemblage is still an ordered process, even if it only operates under certain circumstances. By contrast the participatory dimension of Murray's account is endorsed as being closer to the heart of the matter, yet the implications can be ambiguous. The stimuli that the assemblage of procedurals respond to may well be cued from, for example, the movement of a participant in an art installation, but it is still possible to characterize this as *responsiveness* rather than interactivity per se.

Overall, Smuts acknowledges that responsiveness remains core to any interactive media platform, – the only proviso is that it should be neither completely random or preordered (Smuts 2009). This is augmented in his conclusion with a reference to a concept of "concreativity" which he adopts from the art theory of R.G. Collingwood, and it is here that perhaps a core of the essence of genuine interactivity comes into clear relief. If we were to discuss a game, for example, that was truly interactive, rather than simply rhizomatic, the environment would be modifiable, actions not originally accounted for in the game design would be possible, and the teleological schema of the "game" itself could be altered.

The Digital Video Game: Media Integration and Interchangeability

The video game has evolved in multiple ways to surpass the initial relatively rudimentary interfaces, game designs, and ludic structures of the first computer game titles. Central to this evolution has been the introduction of more streamlined technical refinements to enhance the performative aspects of the gaming experience. In particular these modifications have facilitated more detailed 3D environments (and multiple options within the same game platform), and modifications of the HCI have enabled more accurate and variable player input options. Digitization has enabled substantially expanded possibilities in terms of the sheer data processed, the sophistication of the images depicted, and the array of real-time processes that can be computed and rendered in real time in response to those inputs. This has culminated in the transformation of the ludic logic of the video game away from mere "puzzle solving" or "reward gathering" to the "rhizomatic" environment that the player or players explore with less predetermined objectives or narratives.

It is commonplace to regard the video-gaming experience as "immersive" in the sense that the player becomes integral to the unfolding of the game play and also is, in most cases, injected directly into the game action through an avatar. As will become evident in the following section of interactivity, there are some qualifications to just how immersive these experiences technically are. Ermi and Mayra (2005) have developed the following schema of different types of immersion:

(A) Challenge-based immersion
(B) Imaginative immersion
(C) Sensory immersion

Each of these indicates subtle but significant qualitative differences in what we might understand by the term. The first signifies the engagement with the game as a practical task-solving exercise with a clear flow; it requires concentration and application but remains in essence a technical exercise with explicit constraints and objectives. The second relates to the emotional investment of the participant who suspends disbelief and "enters into the spirit of the thing" so to speak. The third signifies the most exciting possibility of human-computer interaction, the possibility that sensory outputs become integrated in the experience of the computer game world.

James Bizocchi (2007) has persuasively related this to Gunning's work on the "cinema of attraction," and it is no accident that gaming, such as we experience it to date, retains deep connections with the aesthetics of cinema, both in terms of narrative and affect. What differs is the degree to which narrative is subordinated to the production and generation of affect; narrative must follow and facilitate multiple lines of development and may even seem to be disintegrated. Nonetheless, it still persists, and, as seen in the discussion of Jacques Ranciere earlier, we can find a theoretical basis for discussing this development in a manner consonant with our broader consideration of the aesthetics of mass media and popular entertainment.

One of the most intriguing media phenomena of the last two decades has been the emergence of a profound overlap in the design, production, and distribution of digital content across previously relatively exclusive platforms. We have already had passing reference to *Final Fantasy* which was notably a video game before being reworked into a film feature. But there have been a raft of other films, of particular note titles such as *Silent Hill*, *Tomb Raider*, and *Resident Evil* that started out as video game productions before being refashioned.

It is noteworthy that many of the video games that have successfully made the transition from console to film have been produced from Japanese production houses such as Konami, Capcom, and Square Enix. One plausible reason is that the Japanese game design industry has benefited from the relatively advanced animation industry with a concomitant reserve of highly trained specialists in motion graphics and character design. It is also now a well-established practice to move production projects from manga to television and thence to video game.

Telematic Performance and the Transposition of Place and Persona

While telemetry is a sphere of technology that facilitates the coordination of data collection in disparate places to be integrated and processed at some kind of central hub (with applications from the workplace to medicine) it is more recently that the possibilities of using information technologies and the real-time processing of multiple inputs across multiple spaces has come into the purview of more general users of digital technology rather than specialist performance artists. The potentialities of telematic performance as a concreative event have been discussed afresh, for example, in Bob Giges and Edward C. Warburton's discussion of a collaborative project Lubricious Transfer, which was a performance staged through the real-time collaborative performance of two dance troupes, one in California and another in New York (as discussed in Giges and Warburton (2010)). What distinguishes this project is that it does indeed fulfill the "concreative" ideal as articulated by Smuts in his analysis of interactivity – it is not simply a splicing of two performances randomly and simultaneously. More importantly, it very consciously interrogates the spaces beyond the screen and the creative space that emerges through the traversing of the two geographic locales.

Through the increasing popularity of open-source platforms such as Isadora (produced by the California-based company Troikatronix) artists are able to create spaces that fundamentally subvert the screen as the focal nexus for presenting the creative output. Participants in the installations, whether as artists themselves or members of a "public," become performers and creators of content. The addition of audial recording devices, webcams, and projectors interspersed within the space allow other artistic expressions in real time that do not, strictly speaking, entail a screen at all. Even more telling is the prospect of developing a space where the human-computer interface is so intangible as to be consciously manipulated through the devices. Indeed recent advances in electroencephalography, as with the release of the Emotiv Brainwear system, puts forth the possibility that in fact thought itself might be sufficient to provide triggers and inputs without any handheld platform whatsoever.

At the same time another significant technological advance with some relevance to telematic media experiences has been presented through the development of the Oculus Rift VR unit which provides an alternate visual reality through a head-mounted visual display. The system can also be augmented with sound inputs and position tracking which holds forth the prospect that a person may feel, for most intents and purposes, as though they are in an alternative space and yet remain within the confines of their physical space. As possibilities for modifying the virtual reality through physical actions and responses are developed further, the possibility for a more thoroughly immersive virtual media experience would seem to be that much more profound.

The most fundamental implication of such technologies, however, is the release of the creative process from the logic of a projector and a screen; it is a most emphatic fulfillment of the promise of escaping the tyranny of the frame and the tyranny of the screen, as Peter Greenaway famously argued. It also implies a release from the other two tyrannies that he enumerated, the tyranny of the actor and the tyranny of the text, although these are less obvious consequences. Certainly human agency is not negated through such installations, but the possibility of transcending the traditional individual subject either as "artist" or as "spectator" should be apparent. Also, the tenuousness of a set "text," or what might otherwise be framed in terms of narrative, "the story" that must be stuck to, is evident from the minimalism of scripting.

As with the advent of digital computer graphic imaging in relation to cinema, none of the foregoing implies the end of screens, frames, actors, or texts; yet the ecology of media, art, and society is substantially restructured, and in keeping with that, aesthetics also needs to be recalibrated to take account of the transformation. More recently, theorists such as Mark B.N. Hansen have explored how the "digital image" needs to be understood not so much as a vehicle of representation but a highly embedded product of the human body which negotiates the affective and bodily impact of images which are in turn imaginatively generated as images through this process.

Some Symptomatic Media Case Studies

Silent Hill (Konami 1999) and *Resident Evil* (Sony 1996) stand out as two video games with exceptional success in terms of creating a broad fan base and indeed spawning a franchise of successor productions which, especially in the latter's case, are almost unmatched. More intriguing is the facility with which they both engendered parallel franchises in film as well. Though critical opinion remains somewhat divided, *Silent Hill* (2006) demonstrated the viability of a video game concept traversing into cinema with arguably surprising ease. As Bernard Perron's insightful analysis in *Silent Hill: the Terror Engine* persuasively demonstrates, there was an aesthetic and a strategy of accentuating affective responses which differentiated the film from a conventional filmic impetus of narrative and character. The world of *Silent Hill* is underpinned by a profound ambivalence of place; parallel worlds with deeply conflicting "truths" collide with graphic transformation of the mundane and municipal to the ambience of a macabre dungeon of pestilence and death. Moreover, the identity of the characters is deeply ambiguous; they belong to both of these worlds in some deep sense, and the horror within the film is generated from the uncertainty of which world will claim them in the end. *Resident Evil* (2002) does not share quite the same dynamics, but its apparent success as a crossover into film is even more emphatic than is the case with *Silent Hill*. The apocalyptic demise of human society is perhaps a familiar enough scenario for a futuristic dystopic science fiction film, but the manner in which the main characters serve as almost unreconstructed game avatars that mediate the projections of the viewer who arguably comes to regard the protagonist as their own self within a game world – the zombie apocalypse – is a problem that must be "solved." Furthermore, the *Resident Evil* franchise was taken one step further in being remediated into novels and animated works with *Resident Evil: Damnation* (2012) being the epitome of the original text's "closing the loop" and returning to an all-digital imaging format, albeit in this case an animated digital feature.

The foregoing cases demonstrate an intensifying fluidity of image production across a variety of media platforms, and indeed it seems that the logic of the video game is finding expression in the cinematic format accordingly. At the same time there are films that have not been generated out of video game platform precursors but also reflect some debt to the logic of that platform. *Avatar* (2009) by its very title pursues the radical traversing of different states of being, most emphatically represented by the male protagonist, a disabled marine, who takes on the physique and attributes of a completely alien life-form, the Na'vi. The apparent pleasure that the marine takes in exploring the capacities of his new avatar form is resonant with the experience of exploring the potential of an avatar figure in a game. However, as the hero pursues two parallel states of existence, with the implication of an increasingly divergent source of identity, the dramatic tension stems from the question of which one of these states will become the definitive one.

By contrast, *Inception* (2010) presents a radical exploration of the potential multiplicity of states of being in time and space, and does so through the device of a technology that enables the exploration of multiple layers of consciousness, each with a distinctive relation of time to the preceding one, there also being a constantly implicit danger that "failure" on one level could jeopardize the capacity of the main characters to successfully reconnect with an original "reality." There is an almost ludic logic to how conquering the problem of one level relates to "success" in proceeding to a successful securing of an overall objective. At the same time, the implicit relativity of the "reality" of any given world leads to a fundamental subverting of the certainty of who a person is based on the belief that one level in particular is the "genuine" one. And the possibility of a consistent narrative line is also radically subverted by the diving from one world to another, there being no guarantee that the next level down will "pick up" from the preceding level.

Overall, there is evidence of a rather fundamental restructuring of the nature of our understanding of some of the traditional elements of artistic expression – narrative, persona, and the ontology of the visual in particular. The problem for aesthetic theory, which is still currently in a condition of evolution toward a fresh level of clarification, is now very much to take account of a reconfiguring of the relations between some of these classical elements. It would seem that Jacques Ranciere's articulation of the transition from the "representative regime" to the "aesthetic regime" is to be followed by some new regime, as yet clearly identifiable as in some sense "post-cinematic" and profoundly traversing the sensory, the corporeal, and the material (Ranciere 2006).

Conclusion

The foregoing overview has sought to contextualize the current thinking about multimedia and digital technologies in relation to the origins of contemporary theoretical preoccupations with the photographic and the cinematic image. This has been augmented with a commentary on how scholars and commentators from various backgrounds, with highly varied academic and technical specializations, have attempted to gauge the implications of digitization in the visual media. As we move toward some collective understanding of a "Post-Cinematic" vision it would seem clear that the following conclusions stand out:

The first is that an engagement with the moving image, postdigitization, requires us to return to a more fundamental understanding of visual and graphic arts (as indeed has been argued by Lev Manovich). This requires the articulation of an aesthetic theory reflecting this transition, and Jacques Ranciere has been presented as one leading example.

The second is that the emergence of new media has not led to the instantaneous obsolescence of cinema but its rebirth (much as Lamarre has argued when discussing its relation to animation). Indeed we are required to understand the elements of contemporary media texts as profoundly hybridized and integrated, often finding their structure (in terms of narrative and other fictive dimensions) profoundly modified by media platforms often regarded as unrelated or antagonistic (as we saw with Kristine Daly's analysis of "Cinema 3.0").

It can also be concluded that the realms of possible integration of digital media with human experience remain open; what we commonly refer to as "immersive" or "interactive" media is at times highly sophisticated and capable of multiple procedural responses, yet the "grail" of full human-computer interactivity will not have arrived until the fulfillment of the preconditions enumerated by Aaron Smuts above. In particular it will be the facilitation of "concreativity" that will be the litmus of any more substantial development in computer technology that will push the horizon of cultural production to new frontiers.

Finally, aesthetic philosophy is currently in a condition of transition, and its future would seem to lie in being reconfigured to accommodate the impact of the foregoing transformations.

Recommended Reading

P.W.S. Anderson (dir.) *Resident Evil* (Sony Pictures, 2002)

R. Barthes, *Camera Lucida: Reflections on Photography* (Richard Howard, trans.). (Hill and Wang, New York, 2010)

A. Bazin, *What is Cinema?* (Timothy Barnard, trans.). (Caboose, Montreal, 2009)
W. Benjamin, in *Illuminations: Essays and Reflections,* ed. by H. Arendt (Harcourt Brace Jovanovich, New York, 1968)
W. Benjamin, *Reflections: Essays, Aphorisms and Autobiographical Writings* (Harcourt Brace Jovanovich, New York, 1978)
W. Benjamin, *The Arcades Project* (Howard Eiland & Kevin McLaughlin, trans.). (Belknap Press, Cambridge Massachusetts, 1999)
J. Bizocchi, Games and narrative: ana analytical framework. Load. J. Can. Games Stud. Assoc. **1** (1), 5–10 (2007)
D. Bordwell, N. Carroll, *Post-Theory: Reconstructing Film Studies* (University of Wisconsin Press, Madison, 1996)
K. Daly, Cinema 3.0: the interactive image. Cinema J. **50**(1), 81–98 (10/2010)
A.C. Danto, *The Philosophical Disenfranchisement of Art* (Columbia University Press, New York, 1986)
G. Deleuze, *Cinema I* (Athlone Press, London, 1989a)
G. Deleuze, *Cinema II* (Athlone Press, London, 1989b)
G. Deleuze, F. Guattari, *A Thousand Plateaus*. (Athlone Press, London, 1988)
T. Elsaesser, M. Hagener, *Film Theory: An Introduction Through the Senses*. (Routledge, New York, 2010)
L. Ermi, F. Mayra, Fundamental components of the gameplay experience: analysing immersion in *Changing Views: Worlds in Play*, Digital Games Research Association Conference Proceedings, Vancouver, British Columbia, ed. by S. Castell, J. Jenson. (2005)
J. Forsey, The disenfranchisement of philosophical aesthetics. J. Hist. Ideas **64**(4), 581–597 (2003)
C. Gans (dir.) *Silent Hill* (Sony Pictures, 2006)
B. Giges, E.C. Warburton, From router to front row: lubricious transfer and the aesthetics of telematic performance. Leonardo **43**(1), 24–32 (2010)
T. Gunning, The cinema of attraction in *Early Cinema: Space, Frame, Narrative,* ed. by T. Elsaesser (Palgrave, London, 2008)
M.B.N. Hansen, *New Philosophy for New Media* (MIT Press, Cambridge, MA, 2004)
H. Jenkins, *Convergence Culture: Where Old and New Media Collide* (NYU Press, New York, 2006)
T. Lamarre, The First Time as Farce: Digital Animation and the Repetition of Cinema. see *Cinema Anime: Critical Engagements With Japanese Animation* (Palgrave, New York, 2006)
T. Lamarre, *The Animation Machine: A Media Theory of Animation* (University of Minnesota Press, Minneapolis, 2009)
L. Manovich, *The Language of New Media* (The MIT Press, Cambridge, MA, 2001)
L. Manovich, After effects: or the velvet revolution. Millenn. Film J. **45/46**, 5–19 (2006)
L. Manovich, *Software Takes Command* (Bloomsbury, London, 2013)
P. Marrati, *Gilles Deleuze: Cinema and Philosophy* (Alisa Hartz, trans.). (Johns Hopkins University Press, Baltimore, 2008)
B. Massumi, *Parables for the Virtual: Movement, Affect, Sensation* (Duke University Press, Durham, 2002)
M. McLuhan, *The Gutenberg Galaxy: The Making of Typographic Man* (University of Toronto Press, Toronto, 1962)
M. McLuhan, *Understanding Media: The Extensions of Man* (McGraw Hill, New York, 1964)
S. Mikami, et al. (dir.) *Resident Evil* (Capcom, 1996)
J. Ranciere, *Aesthetics and its Discontents* (Polity Press, Cambridge, UK, 2009)
J. Ranciere, *Film Fables* (Emiliano Battista, trans.). (Berg, Oxford, 2006)
J. Ranciere, *The Future of the Image* (Verso, London, 2007)
R.G. Scruton, *The Aesthetic Understanding: Essays in the Philosophy of Art and Culture Indiana* (St. Augustine's Press, 1998)
S. Shaviro, *The Cinematic Body* (University of Minnesota Press, Minneapolis, 1993)
S. Shaviro, *Without Criteria: Kant, Whitehead, Deleuze, and Aesthetics* (MIT Press, Cambridge, MA, 2009)

S. Shaviro, Post-cinematic affect: on grace Jones boarding gate and Southland Tales. Film Philos. **14**(1), 1–102 (2010), Open Humanities Press

S. Shaviro, *The Universe of Things: On Speculative Realism* (Minneapolis: University of Minnesota Press, 2014)

R. Sinnerbrink, *New Philosophies of Film: Thinking Images* (Continuum, New York, 2011)

A. Smuts, What is Interactivity? J. Aesthet. Educ. **43**(4), 53–73 (2009)

K. Toyama (dir.) *Silent Hill* (Konami Computer Entertainment, 1999)

R. Williams, *Television: Technology and Cultural Form of the Distinctive Characteristic of "flow" in Television Broadcasting*. (Collins, London, 1974)

Part VIII

Edutainment

Edutainment in Sport and Health

35

Josef Wiemeyer and Lars L. Tremper

Contents

Abstract

Edutainment technologies offer new and fascinating options in sport and health. In this chapter, the impact of edutainment in the fields of sport and health is reviewed. Two edutainment technologies are focused on: educational TV as a more or less receptive technology and serious games as a more activating and interactive approach, i.e., digital games serving a double mission to establish a "serious" goal (like learning and training) without compromising fun and motivation.

J. Wiemeyer (✉) • L.L. Tremper
Institute of Sport Science, Technische Universitaet Darmstadt, Darmstadt, Germany
e-mail: wiemeyer@sport.tu-darmstadt.de

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_67

In the field of sport edutainment, predominantly off-the-shelf games are used for improving skills, abilities, and knowledge. Educational TV plays a minor role. Research shows evidence that games can improve sport-related skills, abilities, and knowledge on a low and elementary level. Currently, no evidence is available that sport edutainment can improve high- or top-level outcomes.

In the field of health, educational TV and serious games are both used. Existing evidence confirms weak to strong effects on health-related knowledge, attitude, and behavior.

In general, the review shows the great potential of TV and serious games in the fields of sport and health. On the other hand, there are still numerous open questions concerning sustainability of the results, appropriate settings, personalization, individualization, and social context. Future system developments and interventions should be more based on sound theory. Furthermore, the quality of research methods and designs needs to be improved favoring randomized controlled trials. Future directions like integration of social media and cross-platform edutainment are discussed.

Keywords

Games for health • Exergames • Educational TV • Educational games • Gamification

Introduction

The term "edutainment" denotes the convergence of *edu*cation and enter*tainment*. The term "education" means teaching individuals to acquire, retain, and transfer knowledge, skills, and abilities in formal or informal contexts. Formal educational contexts mean education in institutions like schools or university, whereas informal education denotes education happening in less structured contexts like visiting a website or communicating with peers. "Entertainment in the sense of the Latin word *tenere* means to keep somebody steady, busy or amused" (Bosshart and Macconi 1998, p. 4). "Entertainment" as a multidimensional construct denotes the subjective experience of primarily, but not exclusively, positive or pleasant psychophysical or emotional states like enjoyment, fun, joy, relaxation, stimulation, curiosity, and pleasure (Vorderer 2001; Bryant and Vorderer 2013). Bosshart and Macconi (1998) state a close relationship of entertainment and play in the sense of experiencing an "as-if" world being different from the real world. Typical entertainment media are books, television (TV), movies, the Internet, and games.

Therefore, edutainment denotes the intention of integrating entertainment technologies into learning and education "to attract and hold the attention of the learners by engaging their emotions" (Okan 2003, p. 255), "as the individual is enjoying himself and learning at the same time" (Addis 2005, p. 730). However, there is no clear line of distinction between "pure" education or entertainment and edutainment. "Accidentally," education can be entertaining or entertainment can educate. In

Table 1 Comparison of "entertainment experience" and "game experience"

Entertainment experience	Game or player experience
Receptive involvement	(Inter)active involvement, flow, immersion
Presence	Presence
Fun	Fun
Positive/pleasant emotions	Positive/negative emotions
Psychological relaxation or stimulation Suspense	Tension
Change and diversion	Curiosity
Fantasy	Fantasy
–	Competence
–	Challenge
–	(Social) interactions

contrast, typical edutainment is characterized by the deliberate and intentional integration of entertainment and education.

Technologies deemed most appropriate for edutainment applications are modern interactive audiovisual or multimodal systems like augmented reality (AR) and virtual reality (VR) as well as the Internet, mobile communication, and game technologies (e.g., Addis 2005; Anikina and Yakimenko 2015). Furthermore, less interactive technologies like films, videos, and TV are also applied to education. Entertainment technology is predominantly applied in informal educational contexts (Okan 2003; Anikina and Yakimenko 2015).

Whereas the significance of TV for education and learning in general has been a matter of controversy, particularly with respect to children (e.g., Carson et al. 2015), educational TV has been proven to be beneficial for education and learning in numerous areas including language and mathematics (e.g., Fisch 2014; Jarvin 2015; Lali et al. 2015). In this regard, *Sesame Street* is probably the most investigated TV program.

A similar discussion concerning the educational relevance of digital games, i.e., games controlled by a microprocessor (video, computer, and mobile devices), has been present since their advent (e.g., Charsky 2010). On the one hand, the adverse effects of digital games, e.g., addiction or increasing aggressive or sedentary behavior, are criticized; on the other hand, "serious games," i.e., digital games serving a particular purpose beyond "pure game experience," have been controversially discussed concerning their claimed positive effects on education and learning. It is important to note that although entertainment and game experience share some common features, they denote different concepts (see Table 1). Whereas entertainment experience is a pleasant receptive experience, playing games is much more (inter)active, challenging, and empowering.

Before the significance of edutainment in sport and health is analyzed, these areas have to be defined. Unfortunately, there is no consensus on the meanings of "sport" and "health." In this chapter, we define the term "sport" as physical activities that are voluntarily performed, unproductive, and normally governed by rules. Sport includes competition, performance, and chance. Typical sport activities are swimming, track and field, (apparatus) gymnastics, volleyball, and basketball.

According to the World Health Organization (WHO), health is a "state of complete physical, mental and social well-being and not merely the absence of disease or infirmity" (WHO 2006, p. 1). Therefore, being healthy means to be in a favorable physical, mental, and social state. One of the most important factors contributing to health is regular physical activity (PA; e.g., WHO 2010; ACSM 2011). Therefore, under certain conditions, sport activities as a specific type of PA can also contribute to health, for example, aerobic and resistance exercises.

Edutainment plays an important role in sport and health. There is a wide scope of applications ranging from online and TV edutainment to gamification and educational games. Whereas technologies like the Internet and educational TV are important (e.g., Bellotti et al. 2011), particularly in health rather than sport education (e.g., Aksakal 2015), it is serious games, i.e., digital games addressing game experience and at least one serious (characterizing) goal simultaneously, which offer new options for edutainment in sport and health (e.g., Göbel et al. 2012, 2015; Göbel and Wiemeyer 2014).

In this chapter the potential benefits of edutainment in the areas of sport and health are critically discussed using educational TV and digital games as two of the most visible technologies (e.g., Iacovides et al. 2012). Particularly, games with an educational purpose can be considered a prototypical edutainment technology because they claim to integrate two goals in a double mission: education and experience of gaming. Compared to "pure" simulations or animations, serious games are claimed to add entertaining, activating, and motivating components to education (Dondlinger 2007). Games may elicit a specific psychophysiological state of the brain enhancing plasticity and thus supporting learning and training (e.g., Bavelier et al. 2010; Howard-Jones et al. 2010). However, there are also critical arguments like "shallow learning" due to superficial information processing or increased attentional distraction (e.g., Spitzer 2014) or "Shavian reversal," i.e., combining suboptimal learning activities and less than entertaining gameplay (e.g., Charsky 2010). There are several subcategories of serious games depending on the characterizing goal, games genre, or content. In the context of edutainment in sport and health, particularly educational games, persuasive games, games for health, rehabilitation games, and exergames are important subcategories, i.e., games with an educational purpose and games including health promotion, therapeutic, or physical exercises.

Furthermore, the existing evidence is analyzed to establish a detailed view on the impact of educational TV and games in sport and health.

The guiding questions are: What are the characteristics of edutainment by TV and serious games in sport and health? Which potential benefits do TV and serious games offer for sport and health? What are the requirements and challenges of serious games in sport and health? Which impact do serious games have on outcomes and attitude in sport and health?

Edutainment in Sport

Many people are engaged in formal or informal sport activities. In Germany, for example, more than 27 million people owned a membership in a sports club in 2014 (DOSB 2014), and more than 25 million were engaged in (informal) sport activities in leisure in 2014 (Statista 2015a). In addition, many people perform training or workout in fitness studios. This fact indicates that participating in sport is an entertaining or at least motivating activity in itself (e.g., Allender et al. 2006).

Furthermore, sport is an attractive activity for watching. There is hardly any nation where sport is not a matter of great public interest. In the USA 70 % of the adult population claims to follow sports with football, baseball, and basketball receiving most attention. The most widely used methods to consume sports are TV (96 %), the Internet (68 %), and attendance (52 %). Furthermore, sport consumption via social network platforms as well as mobile devices has increased significantly over the last 4 years (SportBusiness Group 2014). The big sport events receive tremendous attention worldwide. More than 4.8 billion people worldwide followed the 2012 Olympic Summer Games in London on TV or online (IOC 2012). The Soccer World Cup final in 2014 has been watched by an average of 34.65 million Germans (Statista 2015b). In 2015, 114.4 million Americans watched the Super Bowl final (Statista 2015c).

Therefore, the sport domain is an attractive subject of entertainment as well as edutainment (for a critical analysis in college sports, Benford 2007). There is much to learn about sport, e.g., the sports-specific rules and strategies, match reports and analysis, the training and coaching system, the characteristics of successful athletes, the health effects of engaging in sports, and training options offered by clubs or commercial institutions like fitness studios.

However, the broadcasting stations offer a limited range of sports they deem attractive for their target audience. Furthermore, many broadcasts include edutainment elements like match analysis by the animation of player trajectories or tactical constellations in soccer. In addition, specific TV magazines offer edutainment regularly. Sample topics address doping, sport technology, sport and aging, or the future of elite sport.

An interesting example for the application of edutainment during sport events is the “Join the Game” campaign during the 2013 UEFA Champions Festival in Stratford (UEFA 2013). In the run-up to the Champions League final between FC Bayern Munich and Borussia Dortmund, all 34,000 visitors of the festival have been invited to join the campaign by registering with their RFID cards and linking to their Facebook or Twitter accounts. The visitors then competed in multiple challenges like target shooting, penalty shooting against a camera-guided robot, dribbling through parcours, or playing virtual soccer games. Altogether, 85,000 RFID interactions of 13,000 registered visitors have been counted during the festival. Further gamification systems are available like activity trackers (see section “Edutainment in Health”) and mobile activity games (e.g., Zombies, Run! Erenli 2012).

Another option for edutainment is digital sport games. These games can be considered a particular further development of multimedia (e.g., Wiemeyer and

Mueller 2015). In the last years, digital game industry has developed to a “blockbuster business.” It has clearly overtaken music and movie industry as an economic factor, with an estimated overall revenue of more than $ 80 billion worldwide in 2012 (Marchand and Hennig-Thurau 2013). According to the Entertainment Software Association (ESA), sport games constituted about 13 % of video games sold in 2014 in the USA (ESA 2015). The sport genre comprises all digital games addressing traditional and established sports like soccer, football, or car racing. These digital games use mouse and keyboard, specific gamepads or joysticks, or sensors and interfaces requiring movements of bigger body parts as input devices. Beyond commercial off-the-shelf (OTS) games, there are also custom-made digital games that have been developed especially for sport education (e.g., Kliem et al. 2012).

On the one hand, OTS games have been successfully used for combining education and game experience. On the other hand, custom-made games have been developed and applied for learning and training.

In this section a selection of sport games for education will be introduced and critically analyzed.

Potential Benefits, Goals, and Requirements

In general, digital sport games and edutainment technology can be used to pursue various educational goals in sport (see Fig. 1): people can gain knowledge about theory and practice of sports, enhance cognitive skills and abilities, or improve their sensory-motor skills and abilities. In addition, edutainment goals often include a change in attitude as a by-product.

A particular goal at least implicitly addressed by digital games is game or player experience (see Table 1; e.g., Nacke 2009). Pasch et al. (2009) analyzed the motivation of players using sport games. By means of Grounded Theory-based analysis of four semi-structured interviews with experienced gamers, they identified two distinct motivations to play sport games – with two corresponding strategies: to achieve (strategy: game) and to relax (strategy: simulation). These motivations and strategies could be validated by a study including a boxing game. Four movement features were reconstructed to influence the players’ immersion: natural mode of control, mimicry of player movements by the avatar, proprioceptive feedback, and physical challenge.

To ensure high quality of serious games in sport, relevant educational concepts as well as appropriate game design have to be considered. This means that on the one hand, the relevant models, theories, and principles of education, learning, and training have to be integrated into the games (e.g., Kebritchi and Hirumi 2008; Adcock et al. 2011). Furthermore, the educational environment or setting plays an important role; serious games can be used in informal and formal educational settings (e.g., Iacovides 2012). Education and sport science have established a complex body of knowledge concerning these issues. For example, appropriate instruction, feedback, and guidance are required to support sensory-motor learning

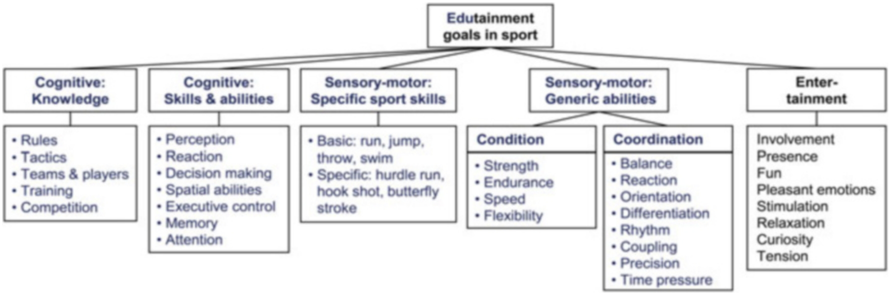


Fig. 1 Edutainment goals in sport(s): *blue color*, educational goals; *black color*, entertainment goals

(e.g., Wiemeyer and Hardy 2013). Concerning conditioning, specific training methods have been established for the training of endurance, strength and resistance, flexibility, and speed (e.g., Hardy et al. 2015). On the other hand, edugames require an appropriate game design, including gameplay, game environment, user interaction, and social interaction as well as adequate goals, choices, competition or cooperation, rules, challenges, and fantasy (e.g., Hays 2005; Bond and Beale 2009; Charsky 2010).

Evidence

Currently, there is no comprehensive review or meta-analysis known to the authors which analyzes the impact of educational TV or serious games in sport in general. Concerning educational TV, non-sport areas seem to be focused like literacy, mathematics, and science for edutaining children and young adults (Lali et al. 2015). Therefore, the edutainment goals illustrated in Fig. 1 are ignored by research on educational TV.

Concerning serious games, selected issues are addressed. An example is the impact of exergames, i.e., serious games including physical exercises to increase energy expenditure, on the cardiovascular and metabolic system (e.g., Biddiss and Irwin 2010; Peng et al. 2012). Therefore, selected studies are discussed first, followed by existing reviews, to identify generic results and basic issues of applying serious games to sport.

Learning of Sensorimotor Skills

One of the first studies to address the learning of sport skills was published by Fery and Ponserre (2001). The authors developed a putting game for golf. The players were allowed to control the swinging movements of an avatar and/or a gauge symbolizing the force applied by lateral mouse movements. The authors performed an experiment comparing two different types of instruction (enjoyment versus

Fig. 2 GUI of the sailing game developed by Hebbel-Seeger (2008) – reproduced by permission of the author – center, boot on the lake; *top left*, coach; *bottom left*, bird's-eye view; *bottom center*, display of speed, position of rudder, weight of the sailor, and sail; *bottom right*, top view of boot illustrating wind conditions (*blue* and *orange arrows*)

learning) and attentional focus (avatar versus gauge). The study revealed a superiority of the symbolic (i.e., gauge) to the analogue representation (i.e., avatar); the participants focusing on the gauge were able to perform real putting movements with significantly greater precision than the group focusing on the avatar. Secondly, the study confirmed the significance of the intention to learn for transfer of virtual gaming experience. However, the study did not address issues of game or player experience (e.g., fun, motivation, and immersion). Therefore, it is not clear whether the participants used the game just as a simulation tool or really experienced playing a game.

Hebbel-Seeger (2008) applied a sailing game (see Fig. 2) to the education of sailing novices. The sailing game included two locations: the sailing area and the teaching room. In the sailing area, the players could choose between sailing course, free sailing, and competition. In the teaching room, the players had four options: instructional videos, knots, wind and sailing courses, and model of the boat. The mission of the game is to get a sailing license by completing all the tasks within the game. A pilot study revealed that by playing the sailing game, ten of eleven members of the treatment group were able to solve the real sailing task (i.e., steering a wherry through a triangular course), whereas only two of ten members of the no-treatment control group successfully passed the sailing test. The author believes that the superiority of the game group was due to a cognitive rather than a perceptual-motor transfer. Again, it is not clear whether the participants just used the game as a simulation because game experience was not assessed.

Sohnsmeyer (2011) performed three studies to test the impact of the Wii game "Table Tennis" on cognitive and perceptual-motor functions. One dependent variable was quality and speed of anticipation in a specific perceptual test using the temporal occlusion technique, i.e., occluding visibility of the opponent's actions at the moment of contact of ball and bat. A second variable was knowledge in table tennis which was assessed by a multiple-choice test. The influence of Wii gaming on knowledge and anticipation was assessed in three contexts and target groups: university education of students (age: 20–24 years), school education of pupils (age: 16–18 years), and summer camp program for youth (age: 12–14 years). In all three studies, Sohnsmeyer found a significant increase of knowledge, whereas anticipation only improved in the leisure (summer camp) and university group, but not in the school group. Again, game experience was not assessed.

Wiemeyer and Schneider (2012) performed a study in basketball. Young elite sportsmen attended either a virtual training (VT) or a real training (RT) of basketball throws. Training consisted of 750 throws distributed over 10 training units. As a result, VT and RT groups improved in virtual and real performance, but only the RT group transferred training to the virtual condition. Furthermore, the RT group enjoyed training more than the VT group.

Vernadakis et al. (2015) compared the influence of a traditional versus game-based training (8 weeks, two times per week, and 30 min per session) on the fundamental motor skills (FMS) of children. Whereas traditional training included typical exercises addressing the FMS to control objects, i.e., throwing, catching, dribbling, striking, rolling, and kicking, the game-based learning comprised Xbox Kinect games like baseball, basketball, soccer, and bowling. The authors found that the short-term and long-term gains in FMS tests addressing object control (six tests: striking a stationary ball, stationary dribble, catch, kick, overhand throw, and underhand roll) were similar for traditional and game-based training. In addition, the game group enjoyed training slightly, but significantly more than the traditional group.

Improving Coordinative Abilities

Numerous studies addressed the improvement of balance skills by playing exergames. The samples analyzed in the studies cover a wide range of age and health conditions.

Vernadakis et al. (2012) as well as Kliem and Wiemeyer (2014) reviewed existing evidence concerning game-based balance training. The game technology predominantly used in these studies is the Wii fit system which includes a balance board, i.e., a four-sensor force platform to determine the trajectory of the center of mass. Existing evidence shows that balance can be improved by game-based learning. However, effect sizes are small to medium and learning gains do not automatically transfer to real-world settings. Furthermore, fun and game experience seem to be increased only in younger samples, i.e., children and young adults. Older adults seem to prefer traditional training interventions.

Strength and Endurance Training

The only study addressing strength training has been published by Sohnsmeyer et al. (2010). The authors compared a 6-week intervention program (20-min sessions playing Wii bowling, twice a week) to a no-intervention control condition in a sample of older adults (>60 years). The gains in isometric strength of the quadriceps muscle of the game group were strong ($\eta^2 = 0.265$) and significantly greater than in the control group. However, due to lack of a traditional group, it is not clear whether the effects can be ascribed to gaming rather than other causes like repeated standing up for playing and sitting down again.

Surprisingly, no real training experiment exists for endurance training. Instead, there are numerous studies examining the question whether exergames are able to increase energy expenditure above a critical level required to elicit adaptations in the cardiovascular, respiratory, and metabolic systems. Due to their relevance for health, these studies will be addressed in the next section.

Science and Teaching Education

Another application field of sport games is university education. Here often minigames are applied. For example, many minigames address biomechanical laws and principles like throwing mechanics or conservation of momentum. Usually, these minigames are integrated into larger e-learning systems.

For teacher education, the serious game "Virtual Sports Teacher" has been developed (e.g., Kliem et al. 2012; see Fig. 3). In this game, players have to take the role of a teacher. Their mission is to organize physical education (PE) lessons in the gym. They have to perform all the tasks that are required by a PE teacher, e.g., giving appropriate instructions and feedback or managing accidents and conflicts. Formative evaluations show that the game elicits realistic experiences and appropriate game experience.

Evidence: Summary

Taken together, the existing studies support the assumption that serious games have the potential to enhance learning sport or sport-relevant motor skills and abilities. However, evidence is sparse in many application areas (with the exception of balance), and the mechanisms seem to involve either elementary perceptual-motor processes or cognitive transfer (knowledge and strategy). Improvement of the specific coordination in motor skills seems to depend on the biomechanical or physical correspondence of game movement and sport movement. In this regard,

Fig. 3 GUI of the educational game "Virtual Sports Teacher" – scene: The teacher is gathering the pupil to start a dialogue; *top left*, info box; *top center*, display of elapsed time and awarded points; *top right*, supervisor button; *bottom left*, gauges indicating learning atmosphere and activity level; *bottom right*, help, support, and dialogue buttons

OTS sensors and interfaces as well as the game mechanics seem to lack the precision required for accurate feedback. Furthermore, effects of learning and training seem to be particularly pronounced in samples with low skill and ability levels. Game-related effects are low to moderate, are primarily specific, and do not always transfer to real-world situations and tasks. In addition, research almost exclusively focuses on OTS videogames in informal conditions of learning and training. Although these games consider selected aspects of learning and training theory (e.g., Wiemeyer and Hardy 2013), custom-made games are more likely to include the tenets of learning and training theory in much more systematic way (for an example; Hoffmann et al. 2015).

Existing studies focus almost exclusively on the outcome of the movements rather than the quality of the movements. Furthermore game or player experience is also widely neglected.

Open Questions and Perspectives

The studies reviewed above show a heterogeneous and preliminary picture for edutainment in sport. On the one hand, many edutainment goals (see Fig. 1) are not addressed at all. Concerning selected application fields, many studies find

Table 2 Structure of an ideal experimental design for the study of edutainment in sport

Group	Pretest	Intervention	Immediate posttest	Delayed posttest(s)
Edutainment	Attitudes Knowledge and skills	Edutainment	Attitudes Knowledge and skills Subjective experience Retention and transfer	Attitudes Knowledge and skills Subjective experience Retention and transfer
Traditional		Traditional exercises		
Control		–		

encouraging (preliminary) results indicating that serious games have the potential to make education in sport more engaging and motivating. However, the systematic development, application, and evaluation of edutainment in sport are still in their infancy. Many questions are still unanswered. For example, the long-term effects of edutainment in sport are rarely addressed as well as appropriate settings for edutainment in sport. An issue extensively addressed in current research and development activities is personalization and adaptation. Another issue is the systematic integration of state-of-the art models and theories of learning and training. In this regard, interdisciplinary cooperation of experts in edutainment technology and experts of sport (science) has to be improved.

On the other hand, many existing studies addressing edutainment in sport suffer from more or less severe methodological problems. An ideal experiment analyzing the effects of edutainment interventions should be a "randomized controlled trial" (RCT). This means that the participants are randomly assigned to the treatments and all potentially intervening influences are carefully controlled. An optimal design requires at least three groups (see Table 2): a game or edutainment group, a traditional intervention group, and a no-intervention group. Furthermore, the sample needs to be determined in an appropriate way; in particular, the sample size should be calculated in advance based on α error, β error (or power), and effect size. At least three measurements have to be performed: pretest, immediate posttest (interval $>$ 10 min), and delayed posttest measurements (interval $\geq$ 24 h). Particularly delayed posttests are required to analyze the long-term effects of the intervention. Posttests should equally consider edutainment and real-world contexts in order to evaluate transfer effects. Due to the double mission of serious games, outcomes concerning the characterizing goal and game experience or entertainment have to be tested.

Edutainment in Health

Compared to the sport domain, the health domain is even much more complex. Many factors influence health, for example, nutrition, lifestyle, regular physical activity, social relations, consumption of alcohol and drugs, smoking, sexual behavior, stress-relaxation balance, and working conditions. The prevention and therapy of diseases and injuries is an important health-related objective. In this regard, the prevalence of a sedentary lifestyle and the demographic changes are serious challenges to public health.

In this extremely complex area of health, edutainment can also play an important role. Due to numerous challenges, e.g., prevailing sedentary lifestyle and increased life expectancy, main goals of health interventions are either to maintain or restore health by means of prevention, rehabilitation, and therapy or to support people with chronic diseases to regain their well-being. According to the definition mentioned above, the concept of health regards at least three levels (WHO 2006): physical, psychic, and social level. Existing health models address numerous factors located at these three levels influencing health and health-related behavior in a particular way. For example, risk factors like physical inactivity or inappropriate nutrition are influences favoring the appearance of cardiovascular diseases (CVD), whereas protection factors like physical activity, social support, and personal attitudes may prevent or alleviate certain diseases like CVD or neurological disorders.

Therefore, one important goal for edutainment in health is to initiate, maintain, or resume health-enhancing behavior. In general, this often means to induce sustainable changes of human behavior. This is a challenging task because human habits are usually extremely stable.

Due to the great significance of health for society, many health campaigns have been implemented worldwide. In this regard, mass media like TV and the Internet play an important role. These mass media also offer edutainment like health quizzes, TV shows, or infotainment programs. The Henry J. Kaiser Family Foundation (2008) investigated the number of health storylines of top-rated TV shows in the USA produced during 2004 and 2006. In this time span, 59 % of the top-rated scripted TV shows had at least one health-related storyline. The content of the health information focuses mainly on symptoms and treatment of diseases rather than on health promotion, e.g., exercising, nutrition, or obesity (less than 1 %). Evidence exists that the power of drama narrative in mass media is able to trigger changes in the consumers' health behavior (e.g., The Henry J. Kaiser Foundation 2008; Movius et al. 2007). To further exploit the educational potential of prime-time TV shows, projects like "Hollywood Health & Society" or the Entertainment Education Program of the Centers for Disease Control and Prevention support entertainment industry. For example, they offer expert consultation, education, and professional resources for the development of scripts with health-related storylines and information. In addition, health insurance companies offer numerous edutainment or infotainment options, particularly for children and youth. In a recent analysis of 221 systematic reviews related to digital media use in public health, Clar et al. (2014) complain the dominance of passive reception and the lack of participatory digital media. In this regard, applications including activity trackers like FitBit, Nike FuelBand, Jawbone, or Garmin vivoactive (e.g., Meyer et al. 2015) offer new options for active participation in HEPA. However, systematic studies of the impact of activity trackers are currently not available. Another category of edutainment plays an increasing role, both in prevention and rehabilitation: serious games or – more specifically – games for health (G4H). Again, high-quality studies dealing with the impact of serious games are rarely to be found.

In this section, the benefits and requirements of applying edutainment, and serious games in particular, to the field of health are discussed, followed by the critical analysis of existing evidence concerning prevention and therapy.

Potential Benefits, Goals, and Requirements

Due to the complexity of the health domain, goals for edutainment located at different levels depend on the specific health area (see Fig. 4).

To explain the mechanisms of changing human health-related behavior, numerous models have been proposed (for a comprehensive review, see Schwarzer 2008). Furthermore, lots of behavior change techniques have been derived from these models (e.g., Michie et al. 2011; Williams and French 2011).

Human behavior can be modeled as a cycle of planning, action, and evaluation. Planning is influenced by intentions, i.e., to visit a fitness center in order to increase physical fitness. Actions and their outcomes are evaluated, e.g., concerning achievement of goals and causes of success or failure. Human behavior shows an interaction with factors like perception of health risks and benefits, self-concept and self-efficacy, attitudes toward health, health-related knowledge and skills, and communication and social support.

There are numerous techniques to influence one or more components of human behavior. For example, Michie et al. (2011) proposed a taxonomy comprising 40 behavior change techniques (BCT). Williams and French (2011) proposed a list of 39 BCT. In addition, the authors propose a ranking of BCT identifying instruction, reinforcement, planning, and social comparison as the most promising BCT for enhancing health-related self-efficacy.

For edutainment in health, these findings are important guidelines to support systematic development and evaluation of applications. The challenge is to address at least the most important factors influencing health and health-related human behavior and to systematically exploit the entertaining effects like fun, motivation,

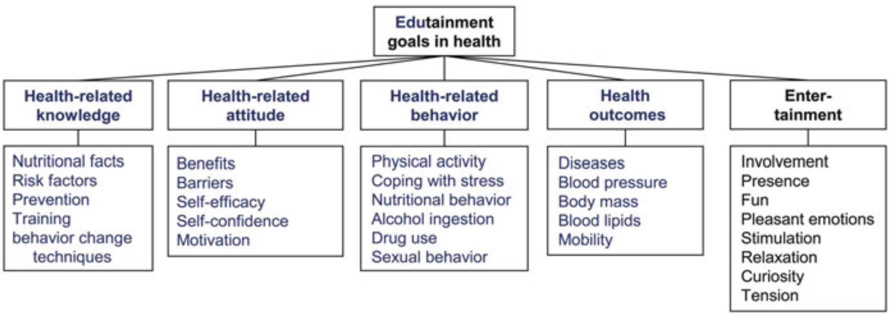


Fig. 4 Edutainment goals in the health domain: *blue color*, educational goals; *black color*, entertainment goals

and emotions. Further added values discussed in the literature are attractive narrative framing, interactivity, immersion and flow, and fantasy.

Analogous to serious games in sport, an interdisciplinary team consisting of experts of edutainment technology, psychology, medicine, and health science is required.

In the next sections, the application of serious games to the areas of prevention and therapy is analyzed separately.

Evidence: Prevention

Prevention, i.e., all interventions aiming at protecting people against (re)appearance of diseases, can pertain to different areas like physical activity, nutrition, alcohol, smoking, drug abuse, relaxation, work organization, etc. Furthermore, different target groups, e.g., depending on age, gender, health state, and risk category, can be addressed. Methods often used are narratives or interactive storytelling (Frank et al. 2015) as well as persuasive games (Knaus 2015). Existing reviews clearly confirm the great potentials of game-like preventions. However, high-quality studies are rarely available.

Shen and Han (2014) published a meta-analysis including 22 entertainment education studies communicating health information. Whereas 17 studies addressed TV, five studies used radio programs. Health topics were detection and prevention of diseases and organ donation. The authors found a small but significant overall effect (effect size $r = 0.12$) which was moderated by study design favoring field studies (versus lab experiments), time of exposure favoring multiple (versus single) episodes, and gender favoring males. Channel (TV versus radio) was not found to have a differential influence on health outcomes. Burzyńska et al. (2015) identified 47 papers published between 2010 and 2014 addressing the impact of TV on health-related knowledge, behavior, and outcomes. One of the eight categories distinguished by the authors is "television and health education/edutainment." Two studies addressing TV edutainment dealing with alcohol consumption showed positive results on knowledge, awareness, and behavior. A study published by van Leeuwen et al. (2012) applied a series of 11 single-story episodes to adolescents with seven episodes addressing alcohol consumption. First, they found a significant interaction of viewing (versus non-viewing) and social status. Particularly less educated viewers showed positive effects on attitude and behavior. Secondly, narrative realism and enjoyment showed mean scores of 5 on a seven-point scale. Thirdly, utility and narrative realism significantly predicted enjoyment ($r^2 = .54$). Frank et al. (2015) revealed the differential role of identification with characters and narrative involvement on selected components of cancer-related attitudes and behavior in a sample of 353 women. The authors produced a narrative film including a young woman suffering from cervical cancer, her mother, and her sister.

Lager and Bremberg (2005) were the first to review the health effects of digital games. They located 30 studies addressing six types of outcome: spatial abilities ($n = 8$), reaction time ($n = 3$), aggressive play ($n = 4$), aggressive thoughts/

interpretations ($n = 5$), aggressive feelings ($n = 4$), and aggressive behavior ($n = 4$). Consistently positive effects could be confirmed for spatial abilities and reaction time, whereas findings concerning aggression were inconsistent.

Baranowski et al. (2008) analyzed evidence for the effectiveness of G4H in different areas. Their review comprises the fields of nutrition and diet ($n = 3$), physical activity and energy expenditure (EE) for health and disabled people ($n = 9$), combination of diet and PA ($n = 4$), and other health-related issues like asthma, diabetes, and cancer ($n = 7$). The nutrition and diet games elicited changes in both knowledge and behavior. The activity games, i.e., exergames or active video games (AVGs), elicited low to moderate increases in EE and physical activity (PA). However, effects declined with longer study duration.

The authors concluded that G4H have much to promise to promote health behavior change, but much work has to be done to improve and systematically evaluate the quality of G4H.

Some reviews have particularly addressed the issue of EE and PA increase in healthy or diseased persons by exergames. It is important to note that "Exergaming is not only physically-engaging but is also cognitively-engaging, and this combination of physical and cognitive engagement may translate into improved cognitive function" (Best 2013, p. 72).

Papastergiou (2009) analyzed 34 articles addressing the possible impact of computer and video games on health and physical education. She concluded that digital games have the potential to enhance physical fitness, motor skills, and the motivation to exercise, but evidence is still very limited.

Biddiss and Irwin (2010) analyzed 18 studies testing the influence of AVG on EE. They confirmed a significant increase of mean heart rate (HR) by 64 % and mean EE by 222 % indicating light to moderate PA levels. Lower body movements elicited a significantly greater EE increase than movements of the upper body.

Similar results were found by Peng et al. (2011) who analyzed 18 studies testing the impact of AVG on EE. Indicators of EE were HR, EE (unit: kcal or kJ), and oxygen uptake (VO_2; unit: ml/min or l/min). A significant increase above resting level was found in all three indicators. However, compared to real physical activities, EE increase in games was slightly lower. Increase of EE can be categorized as light to moderate PA level (see also Peng et al. 2012). The EE effects were moderated by age and AVG type. Children and activities of the whole body or lower extremities showed higher increases of EE compared to adults and upper extremities, respectively.

Peng et al. (2012) analyzed 13 intervention studies aiming at increasing regular PA. Only one out of nine studies (11 %) was able to increase PA attendance in children, whereas two out of four interventions (50 %) were successful in adults.

LeBlanc et al. (2013) analyzed 51 AVG intervention studies addressing health-related parameters and behavior. The authors confirmed a positive effect of AVG on EE and enjoyment, whereas effects on adiposity, cardiometabolic indicators, and regular PA were inconclusive. Furthermore particularly adherence to AVG or PA declined gradually. As in the previous reviews, EE increases indicated light to

Table 3 Overview of the effect of serious games on indicators of prevention

Indicator	Behavior and outcome	Knowledge and attitude
Nutrition	++	+
Physical activity (PA)	+ (adults > children)	++ (gradual decline)
Energy expenditure (EE)	++ (low to moderate; children > adults)	N/A
Reaction ability	++	N/A
Spatial ability	++	N/A
Aggression	+	N/A

++ strong evidence, + weak evidence, – no evidence, N/A not applicable

moderate PA level. Weight loss is most probable in obese and overweight children rather than normal weight.

Deutsch et al. (2015) analyzed the impact of AVG on EE in patients suffering from stroke or cerebral palsy. The authors located seven studies. The results show that patients suffering from neurological diseases reached a light to moderate level of PA which is lower than in healthy adults.

The evidence found in the literature is summarized in Table 3. According to the GRADE system (Guyatt et al. 2011), evidence is graded (weak versus strong). Strong evidence can be found for nutrition behavior, EE, reaction, and spatial abilities as well as attitude to PA.

Evidence: Therapy

In the therapy of many diseases, it is a “sine qua non” that the patients adhere to the therapy. However, particularly in long-term interventions including the repetitive application of equal or similar methods, motivation and compliance of the patients tend to gradually decline. Edutainment and particularly serious games (specifically rehab games) have the potential to maintain or increase motivation.

One of the most frequently cited and most successful examples for the application of digital games to the therapy of diseases is the game “Re-Mission” issued by HopeLab (see Fig. 5). The narrative of the game features Roxxi, a nanorobot, who has to fight different types of cancer. Roxxi is supported by her holographic advisor, a sympathetic and friendly teacher. Kato et al. (2008) could confirm that playing the game improved cancer-related knowledge, self-efficacy, and adherence to therapy. In addition, Cole et al. (2012) showed that during active gameplay, certain brain areas controlling motivation and emotion are activated.

Primack et al. (2012) analyzed 38 RCTs addressing different health outcomes. The authors distinguish psychological therapy ($n = 9$), physical therapy ($n = 6$), physical activity ($n = 4$), clinician skills ($n = 3$), health education ($n = 7$), pain distraction ($n = 5$), and disease management ($n = 4$). Physical therapy included balance training and exercises of the upper extremity, whereas psychological therapy addressed a wider variety of perceptual, cognitive, and emotional functions. Health education aimed at improving knowledge concerning nutrition, diet, asthma, or

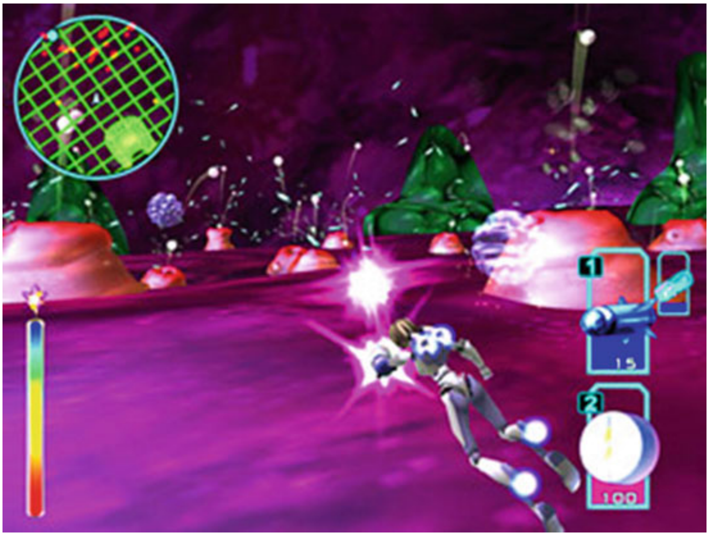


Fig. 5 GUI of the game "Re-Mission" issued by HopeLab (reproduced by permission of HopeLab)

diabetes. Positive effects were most pronounced in psychological therapy (18 out of 26 outcomes) and physical therapy (19 out of 32 outcomes). In the other areas, the portion of positive outcomes was 50 % and below. However, concerning primary outcomes, all studies had a proportion of positive effects equal to and above 50 %.

Wiemeyer (2014) analyzed evidence concerning the application of rehab games in neurorehabilitation. He identified 34 studies including 4 RCT. The studies addressed stroke ($n = 22$), cerebral palsy ($n = 2$), acquired brain injuries ($n = 3$), and various other neurological disorders ($n = 7$). Whereas all four RCT showed a significantly greater improvement of clinical scores by playing digital games compared to conventional treatments, the remaining studies were almost exclusively pilot, feasibility, or case studies lacking minimum quality. Positive effects concerning performance were reported in 13 studies, whereas 10 studies found in increase of attitudes.

Staiano and Flynn (2014) also performed a systematic review of AVG in therapy. The authors located 64 studies addressing cerebral palsy ($n = 11$), stroke ($n = 11$), extremity limitations and amputation ($n = 5$), Parkinson's disease ($n = 4$), distraction during burn treatment ($n = 3$), spinal injuries ($n = 2$), or cancer ($n = 1$). The authors found that in the majority of studies, AVG demonstrated equal or better outcomes as compared to usual care. Furthermore, acceptance and enjoyment were high in almost every population.

Lohse et al. (2014) performed a meta-analysis concerning the impact of virtual therapy interventions on stroke rehabilitation. The authors identified three studies

Table 4 Overview of the effect of serious games on outcome indicators of therapy

Disease/therapy	Behavior and diagnostics	Knowledge and attitude
Stroke	++	++
Cerebral palsy	++	+
Multiple sclerosis	+	
Parkinson	+	N/A
Spinal cord lesions	+	+
Acquired/traumatic brain injury	+	+
Extremity limitations/amputations	+	+
Cancer	+	+
Burns	+	++

++ strong evidence, + weak evidence, – no evidence, N/A not applicable

applying commercial video games to stroke therapy yielding a strong effect on body functions (e.g., visual perception, attention, or muscle functions). Four studies addressing activity outcomes (e.g., balance, walking, or lifting objects) showed a strong, but nonsignificant effect size.

Pietrzak et al. (2014) reviewed the effects of virtual therapy interventions on traumatic brain injury (TBI) rehabilitation. They identified six studies using games for TBI rehabilitation. Three case studies applying commercial games to balance training found an increase of balance abilities. One case study addressing mental functions showed a better performance in visuospatial and verbal memory as well as executive functioning after playing a commercial game. Two quasi-experimental pre-post studies revealed positive effects of a custom-made game requiring reaching movements of the upper extremities on reaching movements and balance.

Table 4 illustrates the overall effects. Strong performance effects were found in the therapy of stroke and cerebral palsy. Strong effects on attitude were found in stroke rehabilitation and distraction from burn therapy.

Existing studies either used OTS digital games or custom-made rehab games. Custom-made rehab games either used OTS game interfaces like the Kinect camera, the Wii Balance Board, or game controllers or specific therapeutic devices (often rehabilitation robots). One advantage of OTS games and interfaces is the low costs. However, low-cost systems normally are accompanied by a loss of precision and flexible application.

Furthermore, as in the field of sport, RCT is rarely to be found. Compared to sport, much more custom-made systems have been developed.

Some interesting trends in the area of rehabilitation can be observed: First, low-cost solutions are developed that may also be used at home (reha@home). Secondly, games become increasingly personalized and adaptive (e.g., Hardy et al. 2015). Thirdly, social settings of games for rehab games fostering either cooperation or competition are increasingly recognized (for a review, see Marker and Staiano 2015).

Open Questions and Perspectives

Although a considerable amount of research and development in the area of games for health has been published, there are still a lot of unanswered questions and unsolved issues.

One of the most important issues is the development of game solutions that are specifically tailored to the target group. This requires on the one hand considering numerous characteristics like age, or grade or stage of disease that will change very slowly or not at all. On the other hand, there are numerous dynamic characteristics requiring instantaneous changes in the game like an inappropriate rate of success or failure that may corrupt game flow. According to the flow model, too much success may lead to boredom, whereas too much failure may lead to anxiety and disappointment. Another important option is the inclusion of the target group into the development process, i.e., player-centered design (Wiemeyer et al. 2015).

Furthermore, specific models of health and disease (intervention) as well as human health-related behavior have to be considered in a more systematic way to ensure that games for health are built on a sound theoretical fundament (for an example, see Lieberman 2001).

Finally, quality of the solutions should be systematically tested by appropriate research designs including formative and summative evaluation (e.g., Baranowski et al. 2010; López-Arcos et al. 2014; Jarvin 2015; Wiemeyer et al. 2015). In this regard, there is still a great heterogeneity of studies.

Furthermore, current evidence for the success of edutainment applications in the field of sport and health lacks long-term studies. Currently, reasonable doubts on edutainment's ability to adhere consumers over a longer period of time exist (e.g., Dixon et al. 2010; Peng et al. 2012). The general criticism bases on the assumption that the evidence for short-term effectiveness results mainly from the novelty factor of these applications. However, this novelty factor is wearing off quickly. Because sustainable health benefits can only be achieved by interventions of longer periods, the restricted motivating nature of edutainment applications thus limits their long-term effectiveness. From this point of view, addressing the application of edutainment in sports and health from a holistic perspective considering the complex interactions of the numerous influencing factors gains in importance.

To enhance long-term attractiveness, multiple recommendations exist (e.g., Sinclair et al. 2007; Mueller et al. 2015). One promising possibility to meet these recommendations could be to enhance the narrative power of edutainment applications in sports and health (Tremper 2015). What established story models have in common is that narrative tension is continuously raised up to a point near the end (the climax), where it is suddenly released. Provided that consumers are interested in a presented story, increased narrative tension seems to be suitable to enhance suspense and curiosity of edutainment applications. If these two traits can be maintained over a longer period, long-term attractiveness may result. To increase the narrative power of edutainment applications in sports and health, the storyline, the interaction possibilities, as well as the game mechanisms should be mapped onto each other in order to maximize coherence (Tremper 2015).

Conclusion and Outlook

The application of edutainment to the fields of sport and health offers a variety of fascinating options. Due to the complexity of both fields, there are a plethora of promising perspectives. On the one hand, mass media like the Internet and TV offer a huge amount of edutainment. However, these offers are received in a comparatively passive way by the target groups. In this regard, digital games and in particular serious games have much more to offer, e.g., immediate feedback, challenging tasks, and motivating environments.

In sport, OTS games dominate. The integration of these games into educational contexts shows positive effects which are limited to low skill level and elementary functions. To reach more ambitious educational goals concerning cognition and sensory-motor coordination, much more custom-made games are required. Particularly for motor skill learning, the precision of the sensors and interfaces as well as the algorithms for activity recognition needs to be improved.

In the broad and heterogeneous field of health, educational TV and serious games are applied to change health-related behavior and to improve adherence and compliance to rehabilitation and therapy. In the health area, much more custom-made products exist, often using low-cost technology, e.g., sensors. Respective studies confirm significant effects. However, sustainability seems to be a problem. Furthermore, many studies suffer from poor quality.

In general, existing evidence shows that in many subfields of sport and health, edutainment has been successfully applied. However, theoretic substantiation of these applications is often weak and evaluation studies show heterogeneous quality. Therefore, the full potential of edutainment has not yet been exploited in these fields. Particularly tailored solutions including adaptation and personalization are needed. Furthermore, interdisciplinary teamwork of experts for content, education, and technology is required to enhance quality of edutainment in health and sport.

What will the future bring for edutainment in sport and health? On the one hand, social and mobile information and communication technologies will gain importance (e.g., Kim et al. 2015). These technologies seem to satisfy additional needs of humans to interact in social networks, for example, the "quantified self" network or other web 2.0 platforms addressing sport and health. Therefore, social learning in edutainment environments addresses competence, autonomy, and social relatedness (e.g., Ryan and Deci 2000). Furthermore, the well-known benefits of collaborative learning (e.g., Dunlosky et al. 2013; Hattie 2013) and the attractiveness of collaborative and cooperative gaming can be combined (e.g., Staiano et al. 2012; Novak et al. 2014).

A new emerging field is cross-platform edutainment. Considering the problem that every medium or platform has its specific strengths and limitations, this approach advocates to combine platforms to take the best of every single platform. However, first results show that there is no simple addition of platform-specific effects (Fisch 2013, 2014). Rather an appropriate combination of media has to be established fitting the capabilities of the respective media and the target audience in relation to the educational content.

Recommended Reading

A.B. Adcock, G.S. Watson, G.R. Morrison, L.A. Belfore, Effective knowledge development in game-based learning environments: considering research in cognitive processes and simulation design, in *Gaming and Simulations: Concepts, Methodologies, Tools, and Applications*, ed. by Information Resources Management Association, vol. 3 (IGI Global, Hershey, PA, 2011), pp. 409–423

M. Addis, New technologies and cultural consumption – edutainment is born! Eur. J. Mark. **39**(7/8), 729–736 (2005)

N. Aksakal, Theoretical view to the approach of the edutainment. Proc. Soc. Behav. Sci. **186**, 1232–1239 (2015)

S. Allender, G. Cowburn, C. Foster, Understanding participation in sport and physical activity among children and adults: a review of qualitative studies. Health Educ. Res. **6**, 826–835 (2006)

American College of Sports Medicine [ACSM], Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. Med. Sci. Sports Exerc. **43**(7), 1334–1359 (2011)

O.V. Anikina, E.V. Yakimenko, Edutainment as a modern technology of education. Proc. Soc. Behav. Sci. **166**, 475–479 (2015)

T. Baranowski, R. Buday, D.I. Thompson, J. Baranowski, Playing for real video games and stories for health-related behavior change. Am. J. Prev. Med. **34**, 74–82 (2008)

T. Baranowski, D. Thompson, R. Buday, A.S. Lu, J. Baranowski, Design of video games for children's diet and physical activity behaviour change. Int. J. Comput. Sci. Sport **9**, 3–17 (2010)

D. Bavelier, D.M. Levi, R.W. Li, Y. Dan, D.K. Hensch, Removing brakes on adult brain plasticity: from molecular to behavioral interventions. J. Neurosci. **30**, 14964–14971 (2010)

F. Bellotti, R. Berta, A. De Gloria, A. Ozolina, Investigating the added value of interactivity and serious gaming for educational TV. Comput. Educ. **57**, 1137–1148 (2011)

R.D. Benford, The college sports reform movement: reframing the "edutainment" industry. Sociol. Q. **48**, 1–28 (2007)

J.R. Best, Exergaming in youth. Zeitschrift für Psychologie **221**, 72–78 (2013)

E. Biddiss, J. Irwin, Active video games to promote physical activity in children and youth: a systematic review. Arch. Pediatr. Adolesc. Med. **164**, 664–672 (2010)

M. Bond, R. Beale, What makes a good game? Using reviews to inform design, in *Proceedings of the 23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology*, (British Computer Society, Swinton, 2009), pp. 418–422

L. Bosshart, I. Macconi, Media entertainment. Commun Res. Trend. **18**(3), 3–6 (1998)

J. Bryant, P. Vorderer (eds.), *Psychology of Entertainment* (Routledge, New York, London, 2013)

J. Burzyńska, M. Binkowska-Bury, P. Januszewicz, Television as a source of information on health and illness–review of benefits and problems. Progr. Health Sci. **6**, 174–184 (2015)

V. Carson, N. Kuzik, S. Hunter, S.A. Wiebe, J.C. Spence, A. Friedman, M.S. Tremblay, L.G. Slater, T. Hinkley, Systematic review of sedentary behavior and cognitive development in early childhood. Prev. Med. **78**, 115–122 (2015)

D. Charsky, From edutainment to serious games: a change in the use of game characteristics. Games Cult. **5**(2), 177–198 (2010)

C. Clar, M. Dyakova, K. Curtis, C. Dawson, P. Donnelly, L. Knifton, A. Clarke, Just telling and selling: current limitations in the use of digital media in public health: a scoping review. Public Health **128**, 1066–1075 (2014)

S.W. Cole, D.J. Yoo, B. Knutson, Interactivity and reward-related neural activation during a serious videogame [Online]. PLoS One **7**(3), e33909 (2012). http://journals.plos.org/plosone/article?id=10.1371/journal.pone.003390 (19 Mar 2012)

J.E. Deutsch, P. Guarrera-Bowlby, M.J. Myslinski, M. Kafri, Is there evidence that active videogames increase energy expenditure and exercise intensity for people poststroke and with cerebral palsy? Games Health J. **4**, 31–36 (2015)

Deutscher Olympischer Sportbund [DOSB]. Bestandserhebung 2014 [Online]. http://www.dosb.de/fileadmin/sharepoint/Materialien%20%7B82A97D74-2687-4A29-9C16-4232BAC7DC73%7D/Bestandserhebung_2014.pdf

R. Dixon, R. Maddison, C. Ni Mhurchu, A. Jull, P. Meagher-Lundberg, D. Widdowson, Parents' and children's perceptions of active video games: a focus group study. J. Child Health Care **14**, 189–199 (2010)

M.J. Dondlinger, Educational video game design: a review of the literature. J. Appl. Educ. Technol. **4**, 21–31 (2007)

J. Dunlosky, K.A. Rawson, E.J. Marsh, M.J. Nathan, D.T. Willingham, Improving students' learning with effective learning techniques: promising directions from cognitive and educational psychology. Psychol. Sci. Public Int. **14**(1), 4–58 (2013)

Entertainment Software Association [ESA]. Essential facts about the computer and video game industry [Online] (2015). http://www.theesa.com/wp-content/uploads/2015/04/ESA-Essential-Facts-2015.pdf

K. Erenli, The impact of gamification: a recommendation of scenarios for education, in *Interactive Collaborative Learning (ICL)* (IEEE, New York, 2012), pp. 1–8

Y.A. Fery, S. Ponserre, Enhancing the control of force in putting by video game training. Ergonomics **44**, 1025–1037 (2001)

S.M. Fisch, Cross-platform learning: on the nature of children's learning from multiple media platforms. New Dir. Child Adolesc. Dev. **139**, 59–70 (2013)

S.M. Fisch, *Children's Learning from Educational Television: Sesame Street and Beyond* (Routledge, New York and London, 2014)

L.B. Frank, S.T. Murphy, J.S. Chatterjee, M.B. Moran, L. Baezconde-Garbanati, Telling stories, saving lives: creating narrative health messages. Health Commun. **30**(2), 154–163 (2015)

S. Göbel, J. Wiemeyer (eds.), *Games for Training, Education, Health, and Sports*. Springer LNCS 8395 (Springer, Cham, Heidelberg, New York, Dordrecht, London, 2014)

S. Göbel, W. Müller, B. Urban, J. Wiemeyer (eds.), *E-Learning and Games for Training, Education, Health and Sports* (Springer, Berlin, Heidelberg, 2012)

S. Göbel, M. Ma, J. Baalsrud Hauge, M. Fradinho Oliveira, J. Wiemeyer, V. Wendel (eds.), *Serious Games* (Springer, Cham, Heidelberg, New York, Dordrecht, London, 2015)

G. Guyatt, A.D. Oxman, E.A. Akl, R. Kunz, G. Vist, J. Brozek, S. Norris, Y. Falck-Ytter, P. Glasziou, H. DeBeer, R. Jaeschke, D. Rind, J. Meerpohl, P. Dahm, H.J. Schünemann, GRADE guidelines: 1 Introduction—GRADE evidence profiles and summary of findings tables. J. Clin. Epidemiol. **64**, 383–394 (2011)

S. Hardy, T. Dutz, J. Wiemeyer, S. Göbel, R. Steinmetz, Framework for personalized and adaptive game-based training programs in health sport. Multmed. Tools Appl. **74**, 5289–5311 (2015)

J. Hattie, *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement* (Routledge, New York and London, 2013)

R.T. Hays, *The Effectiveness of Instructional Games: A Literature Review and Discussion. Naval Air Warfare Center Training System Division (No. 2005–004)* (Naval Air Warfare Center, Training Systems Division, Orlando, FL, 2005)

A. Hebbel-Seeger, Videospiel und sportpraxis - (k)ein Widerspruch. Zeitschrift für e-learning **3**, 9–20 (2008)

K. Hoffmann, S. Hardy, J. Wiemeyer, S. Göbel, Personalized adaptive control of training load in exergames from a sport-scientific perspective: an algorithm for individualized training. Games for Health Journal **4**(6), 470–479 (2015)

P. Howard-Jones, M. Ott, T. van Leeuwen, B. de Smedt. Neuroscience and technology enhanced learning [Online] (2010). http://www.futurelab.org.uk/projects/games-in-education

I. Iacovides, *Digital Games: Motivation, Engagement and Informal Learning* (Ph.D. thesis). The Open University, Milton Keynes, UK (2012)

International Olympic Committee [IOC]. Olympic Marketing Fact File [Online]. (IOC, Lausanne, 2012). http://www.olympic.org/Documents/IOC_Marketing/OLYMPIC-MARKETING-FACT-FILE-2012.pdf

L. Jarvin, Edutainment, games, and the future of education in a digital world. New Dir. Child Adolesc. Dev. **147**, 33–40 (2015)
P.M. Kato, S.W. Cole, A.S. Bradlyn, B.H. Pollock, A video game improves behavioral outcomes in adolescents and young adults with cancer: a randomized trial. Pediatrics **122**, e305–e317 (2008). doi:10.1542/peds.2007-3134
M. Kebritchi, A. Hirumi, Examining the pedagogical foundations of modern educational computer games. Comput. Educ. **51**, 1729–1743 (2008)
S.Y. Kim, H.J. So, S. Kwon, S. Oh, K. Park, M. Ko, J. Yoo, G. Oh, Towards designing a mobile social learning application with meaningful gamification strategies, in *Advanced Learning Technologies (ICALT), 2015 I.E. 15th International Conference on.* (IEEE, New York, 2015), pp. 170–174
A. Kliem, J. Wiemeyer, Gleichgewichtstraining mit serious games. Neurol. Rehab. **20**, 195–206 (2014)
A. Kliem, V. Wendel, C. Winter, J. Wiemeyer, S. Göbel, Virtual sports teacher – a serious game in higher education. Int. J. Comput. Sci. Sport **10**, 100–110 (2012)
M. Knaus, Persuasive technologies and applications in health and fitness. Pers. Technol. Appl. **3**(2), 5–10 (2015)
L.E. Nacke. *Affective Ludology: Scientific Measurement of User Experience in Interactive Entertainment* (Ph.D. thesis). Blekinge Institute of Technology, Doctoral Dissertation Series No. 2009:04, 2009
A. Lager, S. Bremberg, Health effects of video and computer game playing. A systematic review [Online] (2005). (Swedish National Institute of Public Health, Stockholm). http://www.researchgate.net/profile/Anton_Lager/publication/233903988_Health_effects_of_video_and_computer_game_playing_A_systematic_review/links/02bfe50cb22f22f3eb000000.pdf?inViewer=true&disableCoverPage=true&origin=publication_detail
T.A. Lali, A.A. Gill, N.U. Hassan, M.S. Juni, An empirical review of television as potentially beneficial medium for children; exploring some realities. J. Stud. Manage. Plan. **1**(3), 98–124 (2015)
A.G. LeBlanc, J.P. Chaput, A. McFarlane, R.C. Colley, D. Thivel, S.J. Biddle, R. Maddison, S.T. Leatherdale, M.S. Tremblay, Active video games and health indicators in children and youth: a systematic review. PLoS One **8**, e65351 (2013). doi:10.1371/journal.pone.0065351
D.A. Lieberman, Management of chronic pediatric diseases with interactive health games: theory and research findings. J. Ambul. Care Manage. **24**, 26–38 (2001)
K.R. Lohse, C.G. Hilderman, K.L. Cheung, S. Tatla, H.M. van der Loos, Virtual reality therapy for adults post-stroke: a systematic review and meta-analysis exploring virtual environments and commercial games in therapy. PLoS One **9**, e93318 (2014). doi:10.1371/journal.pone.0093318
J.R. López-Arcos, N. Padilla-Zea, P. Paderewski, F.L. Gutiérrez, A. Abad-Arranz. Designing stories for educational video games: a player-centered approach, in *Proceedings of the 2014 Workshop on Interaction Design in Educational Environments*, (ACM, New York, 2014), pp. 33–40
A. Marchand, T. Hennig-Thurau, Value creation in the video game industry: industry economics, consumer benefits, and research opportunities. J. Int. Mark. **27**, 141–157 (2013)
A.M. Marker, A.E. Staiano, Better together: outcomes of cooperation versus competition in social exergaming. Games Health J. **4**, 25–30 (2015)
J. Meyer, J. Fortmann, M. Wasmann, W. Heuten, Making lifelogging usable: design guidelines for activity trackers, in *MultiMedia Modeling*, ed. by X. He, S. Luo, D. Tao, C. Xu, J. Yang, M.A. Hasan (Springer, Cham, 2015), pp. 323–334
S. Michie, S. Ashford, F.F. Sniehotta, S.U. Dombrowski, A. Bishop, D.P. French, A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the CALO-RE taxonomy. Psychol. Health **26**, 1479–1498 (2011)
L. Movius, M. Cody, G. Huang, M. Berkowitz, S. Morgan. Motivating television viewers to become organ donors [Online] (2007). *Cases in Public Health Communication and Marketing*. http://www.casesjournal.org/volume1/peer-reviewed/cases_1_08.cfm 6 Aug 2015
F.F. Mueller, D. Altimira, R.A. Khot, Reflections on the design of exertion games. Games Health J. **4**, 3–7 (2015)

D. Novak, A. Nagle, U. Keller, R. Riener, Increasing motivation in robot-aided arm rehabilitation with competitive and cooperative gameplay. J. Neuroeng. Rehabil. **11**(1), 64 (2014)
H. Okan, Edutainment: is learning at risk? British J. Educ. Technol. **34**(3), 255–264 (2003)
M. Papastergiou, Exploring the potential of computer and video games for health and physical education: A literature review. Comput. Educ. **53**, 603–622 (2009)
M. Pasch, N. Bianchi-Berthouze, B. van Dijk, A. Nijholt, Movement-based sports video games: investigating motivation and gaming experience. Entertainment Comput. **1**, 49–61 (2009)
W. Peng, J.-H. Lin, J.C. Crouse, Is playing exergames really exercising? a meta-analysis of energy expenditure in active video games. Cyberpsychol. Behav. Soc. Network. **14**, 681–688 (2011)
W. Peng, J.C. Crouse, J.-H. Lin, Using active video games for physical activity promotion: a systematic review of the current state of research. Health Educ. Behav. **40**, 171–192 (2012)
E. Pietrzak, S. Pullman, A. McGuire, Using virtual reality and videogames for traumatic brain injury rehabilitation: a structured literature review. Games Health Jl **3**, 202–214 (2014)
B.A. Primack, M.V. Carroll, M. McNamara, M.L. Klem, B. King, M. Rich, C.W. Chan, S. Nayak, Role of video games in improving health-related outcomes: a systematic review. Am. J. Prev. Med. **42**, 630–638 (2012)
R.M. Ryan, E.L. Deci, Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am. Psychol. **55**(1), 68–78 (2000)
R. Schwarzer, Modeling health behavior change: how to predict and modify the adoption and maintenance of health behaviors. Appl. Psychol. **57**, 1–29 (2008)
F. Shen, J. Han, Effectiveness of entertainment education in communicating health information: a systematic review. Asian J. Commun. **24**(6), 605–616 (2014)
J. Sinclair, P. Hingston, M. Masek, Considerations for the design of exergames, in *Proceedings of the 5th international conference on Computer graphics and interactive techniques in Australia and Southeast Asia*. GRAPHITE. (ACM, New York, 2007), pp. 289–295
J. Sohnsmeyer, *Virtuelles Spiel und realer Sport – Über Transferpotenziale digitaler Sportspiele am Beispiel von Tischtennis* (Hamburg, Czwalina, 2011)
J. Sohnsmeyer, H. Gilbrich, B. Weisser, Effect of a six-week-intervention with an activity-promoting video game on isometric muscle strength in elderly subjects. Int J Comput Sci Sport **9**, 75–79 (2010)
M. Spitzer, Information technology in education: risks and side effects. Trend. Neurosci. Educ. **3**, 81–85 (2014)
SportBusiness Group. *Know the Fan: The Global Sports Media Consumption Report 2014 – US Overview* [Online]. (2015). http://sportsvideo.org/main/files/2014/06/2014-Know-the-Fan-Study_US.pdf
A.E. Staiano, R. Flynn, Therapeutic uses of active videogames: a systematic review. Games Health J. **3**, 351–365 (2014)
A.E. Staiano, A.A. Abraham, S.L. Calvert, Motivating effects of cooperative exergame play for overweight and obese adolescents. J. Diabetes Sci. Technol. **6**(4), 812–819 (2012)
Statista, Bevölkerung in Deutschland nach Häufigkeit des Sporttreibens in der Freizeit von 2010 bis 2014 [Online]. (2015a). http://de.statista.com/statistik/daten/studie/171911/umfrage/haeufigkeit-sport-treiben-in-der-freizeit/
Statista, Fernsehzuschauer der Spiele der deutschen Nationalmannschaft bei der Fußball-Weltmeisterschaft 2014 in Brasilien (in Millionen) [Online]. (2015b). http://de.statista.com/statistik/daten/studie/305427/umfrage/tv-quoten-der-spiele-der-deutschen-nationalmannschaft-bei-der-wm/
Statista, Anzahl der TV-Zuschauer bei den Super-Bowl-Finales in den USA in den Jahren 2001 bis 2015 (in Millionen) [Online]. (2015c). http://de.statista.com/statistik/daten/studie/286519/umfrage/tv-zuschauer-beim-super-bowl-finale-in-den-usa/
The Henry J. Kaiser Family Foundation, *How Healthy Is Prime Time? An Analysis of Health Content in Popular Prime Time Television Programs* (Menlo Park, Henry J. Kaiser Family Foundation, 2008)
L.L. Tremper, *Narrative Exergames* (MSc thesis), (Technische Universität, Darmstadt, 2015)

UEFA, Wembley final proves global pulling power [online]. http://www.uefa.com/uefachampionsleague/news/newsid=1957523.html 28 May 2013
L. Van Leeuwen, R.J. Rene, C. Leeuwis, Televised entertainment-education to prevent adolescent alcohol use perceived realism, enjoyment, and impact. Health Educ. Behav. **40**(2), 193–205 (2012)
N. Vernadakis, A. Gioftsidou, P. Antoniou, D. Ioannidis, M. Giannousi, The impact of Nintendo Wii to physical education students' balance compared to the traditional approaches. Comput. Educ. **59**, 196–205 (2012)
N. Vernadakis, M. Papastergiou, E. Zetou, P. Antoniou, The impact of an exergame-based intervention on children's fundamental motor skills. Comput. Educ. **83**, 90–102 (2015)
P. Vorderer, It's all entertainment—sure. But what exactly is entertainment? Communication research, media psychology, and the explanation of entertainment experiences. Poetics **29**(4), 247–261 (2001)
WHO, *Constitution of the World Health Organization [Online]* (WHO, Geneva, 2006)
WHO, *Global Recommendations on Physical Activity for Health* (WHO, Genf, 2010)
J. Wiemeyer, S. Hardy, Serious games and motor learning – concepts, evidence, technology, in *Serious Games and Virtual Worlds in Education, Professional Development, and Healthcare*, ed. by K. Bredl, W. Bösche (IGI Global, Heshey, 2013), pp. 197–220
J. Wiemeyer, F.F. Mueller, ICT-enhanced learning and training, in *Computer Science in Sport – Research and Practice*, ed. by A. Baca (Routledge, London, 2015), pp. 187–213
J. Wiemeyer, P. Schneider, Applying serious games to motor learning in sport. Int. J. Game-Based Learn. **2**, 61–73 (2012)
J. Wiemeyer, Serious games in neurorehabilitation: a systematic review of recent evidence, in *Proceedings of the 2014 ACM International Workshop on Serious Games (*ACM, New York, 2014), pp. 33–38.
J. Wiemeyer, J. Deutsch, L.A. Malone, J.L. Rowland, M.C. Swartz, J. Xiong, F.F. Zhang, Recommendations for the optimal design of exergame interventions for persons with disabilities: Challenges, best practices, and future research. Games Health J. **4**, 58–62 (2015)
S.L. Williams, D.P. French, What are the most effective intervention techniques for changing physical activity self-efficacy and physical activity behaviour — and are they the same? Health Educ. Res. **26**, 308–322 (2011)
G.N. Yannakakis, J. Hallam, Modeling and augmenting game entertainment through challenge and curiosity. Int. J. Artif. Intell. Tools **16**, 981–999 (2007)

Gamification

36

Alke Martens and Wolfgang Müller

Contents

Abstract

This article looks at the concepts, the historical roots, and the scope of gamification. The term *Gamification* has its roots in terms like *game* and *game-based learning*. Moreover, there exists a strong relation to the psychological background of playing, of using game elements and game mechanics. Also, the term *game thinking* is part of interpreting the term *gamification*. But what is gamification? Is it old wine in new skins? Or is it a brand new and bright idea? This plethora of things and interpretations related to the term *gamification* led to a broad set of application domains and pseudoknowledge. The term and the related fields are analyzed, and similarities and differences to games, edutainment,

A. Martens (✉)
Institute for Computer Science and Electrical Engineering, University of Rostock, Rostock, Germany
e-mail: alke.martens@uni-rostock.de

W. Müller
Media Education and Visualization Group (MEVIS), University of Education Weingarten, Weingarten, Germany
e-mail: muellerw@ph-weingarten.de

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_66

game-based learning, and serious games are shown in this chapter. Additionally, an overview of elements typically applied in gamification approaches is given. The intersection of gamification with edutainment and learning is focused, and empirical and other results on the effectiveness of gamification are analyzed and discussed. Finally, a summary and brief discussion of open questions and challenges in this field are presented.

Keywords
Gamification • Definition • Empirical and other forms • From games to game Elements • Game elements and game mechanics • Historical roots and relations • Instructional design

Introduction

Game-based learning and gamification gained a lot of interest in the last years not only on academia but also in the fields of edutainment and learning in general. Many voices discount them as mere temporary fashions. However, both approaches are not new. They have deep roots in instructional design, in education, and psychology. After all, the human being can be called a homo ludens (Huizinga 2006).

Game-based learning and gamification are both trends, which occurred several years ago in the context of digital teaching and training (aka. eLearning). Both trends have an inherent promise, which made them buzzwords and nice-to-have in quite a short time: gaming has the association to "having fun," and what might be better than teaching and training content that can be learned while having fun? For example, an empirical study showed that players of World of Warcraft are quite sure that playing this game offers them the opportunity to train for leadership skills required in their business activities (Rausch et al. 2012). Moreover, the mysterious state of "flow" has been observed to be reached by persons while playing games (Csikszentmihalyi 1990). Additionally, what can be better to have school kids playing games and meanwhile learning? And it is obvious that, given the fact that quite a lot of people use computers for playing and gaming, computers are the right medium to offer game-based learning and gamification. The comparably long history of the known relation between learning and gaming (e.g., described in Fuchs et al. 2015) is contrasted with what currently happens in the scientific field and in the field of commercial applications. Looking closer at the development of game-based learning and gamification in computer science, it has to be admitted that game-based learning seems to have ended in an impasse point where the instructionally smooth integration of learning and gaming has not really been realizable yet – at least not without enormous effort by the learning and gaming content author (Maciuszek and Martens 2012).

Gamification, the "smaller" version of game-based learning, initially had the inherent promise to be easier to realize, and to have similar effects as game-based learning. As a result, gamification has made it into various application domains. For instance, technology consultancy Gartner has projected that 50 % of corporate innovation will be "gamified" already by 2015 and that this trend will continue (Burke 2012). But the concept of Gamification has also been criticized. Critique depicts fundamental deficits in the leveraging of game thinking and games, fundamental weaknesses in learning in gamified contexts, and address that the focus on motivating by gamifying is problematic. Gamification targets to increase motivation for completing (learning) activities providing extrinsic recognition and reward. However, this is at risk of demotivating learners with an existing high intrinsic motivation. Robertson claimed existing gamification approaches to regularly limit the adaption of game elements to points and badges, thus "taking the thing that is least essential to games and representing it as the core of the experience" (2010). Bogost (2011) claims gamification to represent mostly a marketing-driven movement, with no real interest in tapping the greater values of games, which would need more thoughtful approaches. On the other hand, Mekler et al. (2013) concluded from own studies that points, levels and leaderboards by themselves would neither make nor break users' intrinsic motivation in non-game contexts, but that they may be applied as effective progress indicators, guiding and enhancing user performance. However, they also admitted that more research would be required in this area.

The divergent receptions lead also to quite mixed predictions on the future of gamification. In a survey (Anderson and Rainie 2012) tech stakeholders and analysts forecasted that the use of game mechanics, feedback loops, and rewards will become more embedded in daily life by 2020, but they were split about the implications, between a mostly positive forecast expecting aiding education, health, business, and training, and quite negative, warning of invisible, insidious behavioral manipulation.

Many definitions have been tried for the term *gamification*, often explaining it from the perspective of fun or motivation. In contrast to this, and ignoring the fact that the term might be coined by chance, as many buzzwords are, the term comprises many more facets, which are interesting to investigate. In this chapter, some of these aspects are brought together. As a first step, the relation between instructional design and gamification in learning is investigated. Afterward, an analysis from the perspective of game to gamification is given. Finally the term *edutainment* is compared to *gamification*. A short glance is thrown into the field of evaluation. The chapter is closed with summing up the results and putting the big picture together.

Historical Roots and Relations

The term *gamification* was apparently coined by Nick Pelling around 2002 to describe services of his consultancy, but it did not get accepted and popular in the business and academic domain until 2008, when this term appeared in an article in *Loyality Management* (Loyality360 2009), and a number of scientific articles followed. The basic idea in these early approaches has been to have companies use gamification for attracting customers, for marketing, and advertisements. Many approaches in gamification have targeted toward customers in service marketing (see, e.g., Huotari and Hamari 2012). Huotari and Hamari (2012) also argued that the usage of the term *gamification* is not restricted to marketing or computer science.

They reasoned in their definition that gamification is used with the goal to support "user's overall value creation." This value creation might be related to feelings and emotions (feeling happy, having fun), but the added value might also come in a completely different shape, e.g., extend the cognitive activity in rethinking things and broadening the horizon of possibilities. In the game, this aspect can also be found: it relates to the rule "to act as if." Thus, looking at the deep meaning of the term, gamification can do even more: it can lead toward new ways of thinking. Gamification can, for example, mean to use the conservative park bench for climbing and balancing (ad hoc usage) or to develop a bus station with a swing instead of a bench to sit on (design-based usage) (see Swing Seat Bus Stops in Montreal (http://www.goodnewsnetwork.org/swings-placed-near-bus-stop-make-music-2/, Accessed 13 July 2015)).

Even if the term *gamification* might have been coined by chance, in the years of usage a set of relations, explanations, and definitions have been tried, with the goal to make clear what is part of gamification and what is not. Most prominent related term is *game-based learning and game playing*.

Taking all these aspects together: marketing, computer science and everyday usage, value creation, creativity, and motivation, it becomes obvious why it is so difficult to come to a precise definition of the term *gamification*. The historical analysis of this quite new trend does not help. It is necessary to have a closer look at the term *gamification* itself and on related terms. This will be done in the following sections.

Relation to Instructional Design

The term *gamification* appeared at a time when many researchers and practitioners have been frustrated about game-based learning, as the goal-orientated and fruitful combination of gaming and learning content comes with a large amount of development costs – both on the software design (Maciuszek and Martens 2011) and on the instructional design side (Maciuszek and Martens 2012; Martens et al. 2008). Moreover, results of using game-based approaches are very difficult to investigate with empirical methods, as usually long-term studies are required to get rid of the innovation hype. This new approach, gamification, seemed to allow to smoothly overcome the difficulties experienced with game-based learning (related literature is analyzed, e.g., in Caponetto et al. 2014). The basic idea reduces to just integrate some game elements and you have a motivated and completely engaged learner. As a consequence, this is exactly how results from this line of development look like: traditional teaching and training approaches with some game-like elements. Looking a bit closer, it is only correct to ask whether the integration of game elements, like levels or badges, really offers anything new in teaching and training. Levels and badges can also come in the shape of grades, classes, and experience levels, and almost all first-graders (at least in Germany) know the praise the teacher gives by making a small bee/smiley/sun stamp under the first written words, meaning: "well done." Is this gamification? Or is this just a (well-known) form of instructional design, which is used in a goal-oriented way to support learning, at least since the late 1960s? Remembering school and University time, the best lectures have been the interesting ones. From the perspective of a student, the term *interesting* is more important than the term *game-like*. From the perspective of a teacher, the term *interesting* is related to very hard work, including a careful instructional design, which consists of at least the following steps (Jank and Meyer 2002):

- Define learning goals or targets to be reached in the educational unit (it is not necessarily a lecture, can also be something smaller).
- Reduce the material to a level appropriate for the target group and the learning goal.
- Redesign the learning content in a way that it is “interesting” – including aspects like closeness to the everyday experiences (the world) of the target group, allow for mental knowledge construction, growing levels of complexity (find the right level of challenge).

These steps always require a very good knowledge of the target group!

Next to these steps, which are also prerequisites of all game-based and gamified approaches, which target at learning, everything around games requires even more! Research in game design offers different and sometimes even more demanding things. A good game design, and also a good game-based learning design, requires a lot of knowledge about the constituents of a game (e.g., game elements and game mechanics), more aspects regarding the target group (e.g., their culture, their traditions, their generation-dependent language and interests) and the gaming situation.

From Games to Game Elements

Looking from the ad hoc usage perspective, everybody has experienced gamification: from the perspective of playing, a stick can become a broadsword or a horse. But this is not quite true, as a mixture of perspectives is taking place. Before going deeper into these thoughts, a question has to be answered: is there a difference between gaming and playing? Kampan Walther (2003) went back to the epistemological roots of the terms and came out with the following insight: play and game are different, as they appear to be on different levels of complexity. As Kampan Walther explains, there is a first transgression between the modes nonplay and play (e.g., the child has to explain the other child that “this stick is currently my sword!” which means a permanent rearticulation to keep the order between the play world and the nonplay world). The game mode does not necessarily require play, but playing can enrich the game experience. The permanent negotiation which takes place during play (I am playing versus I am not playing) does not play a role in the game world, as the game world has its own spatial and temporal rules, which are accepted when taking part in the game.

Coming from this perspective, the term *gamification* seems to be a bit blurry: in most cases, gamification is related to having fun while doing something serious.

A quite better perspective on the term would be to use aspects of games (often called game design, game techniques, game mechanics, game elements, etc.) for structuring and communicating things. This unburdens us from the claim that gamification brings fun into everyday activities and helps to judge gamification approaches. And playing is also not necessarily fun! It is not funny to be attacked by a broadsword if you play knights, even if this broadsword is represented by a stick and you wear a costume.

It is clear that the term *gamification* is related to the English term *game* (and not to play!). Deterding et al. (2011) define that gamification is "the use of game design elements in non-game contexts." Huotari and Hamari investigated this and came to the result that there are no game elements which are unique to games (Huotari and Hamari 2012). They relate the term *game element* to game mechanics and game design patterns. The authors in Wikipedia Germany go one step beyond: there, gamification is related to game design principles, game mechanics, and game thinking (Wikipedia 2015). Interestingly, these levels of detail are not explicated in the English version of the Wikipedia entry. The aspect game thinking is also addressed by Kapp: "Gamification is using game-based mechanics, aesthetics, and game thinking to engage people, motivate action, promote learning, and solve problems" (Kapp 2012). Also the terminology in this domain to distinguish (certain types of) game elements is not consistent. Terms such as *ludemes* (Parlett 2015), *game mechanics* (Sicart 2008), *game design patterns* (Bjork and Holopainen 2004; Ecker et al. 2011), *game atoms* (Brathwaite and Schreiber 2009), and also *game elements* are being used, often without a very clear distinction of the individual concepts.

Looking closer at the terms, game design elements are only one aspect of game design principles, and can thus be subsumed. The same can be said about game mechanics. Game design shows a way how to design games – and for this purpose one can use game elements, and game mechanics. Only the game thinking adds something new. So the next step will be to analyze these terms in more detail. For this chapter, the perspective is chosen that gamification design has its roots in game design principles and game thinking. However, to get a precise definition of these terms is also not an easy task.

Generally, **game design** principles relate to the theoretical background of game development (Martens et al. 2008). Design principles stem from psychological mechanisms, game traditions (e.g., character design, story design), experiences, cultural dependencies, marketing aspects, and the like. The term *game design* is the heading for a list of parts, which can be used for designing a game. These are, for example,

- Game elements: From our perspective, an element is (similar to the latin basic form elementum) the base (in the sense of "parts") for designing a game. In contrast to (Huotari and Hamari 2012), the position in this chapter is that there exist certain game elements, as there they are expected to be part of a game, have a certain role or behavior in this game, and the culture of game play has a basic set of rules how to treat these elements. From the structural perspective, types of game elements can be distinguished:

  - Interaction elements: things the game player can interact with. In real life, these things can come in the shape of 6-sided dice, 12-sided dice, figures, the board, etc. In the digital world nonplayer characters, the virtual world, and other things extend the list of simple virtual versions of real-life things (like virtual dice).
  - Rating elements, like levels, badges, scoring systems, and the like.
- Game mechanics: Sometimes the term *game mechanics* is also subsumed under game design. At the core, game mechanics are the rules and modes of interacting with the game (aka gameplay) (see, e.g., Fullerton 2012). Structurally, a distinction of the following game mechanics can be found:
  - Rules – for interaction, for communication, for role behavior, the strategies, and the like.
  - Roles – of the game elements and of the game players. There are, for example, quite a lot of role patterns for player and for nonplayer characters in role playing games. Roots of these patterns are often found in literature, but they are reused and sometimes extended for game design (e.g., Maciuszek and Martens 2014).
  - Stories – structuring the game and the interaction with elements and with coplayers. Similar to role patterns, there exist several story patterns, which partially also stem from literature science (see, e.g., Maciuszek and Martens 2010).

From the computer science perspective, the game mechanics reduce to states in the game and the rules used for state transitions. These rules relate to human-human interaction (e.g., in role play) or human-element interaction (e.g., throwing a dice). Every interaction (aka action) leads to a state transition.

Terms like *game structures* and *game interaction* would be certain aspects of game mechanics.

- Additional terms (e.g., Fullerton 2012):
  - Game interface: it might be worth to distinguish the outer appearance of the game from the rest of it, but it might also be a way to embed this in a graphical design of game elements. Worth to notice might be that game elements have a graphical design, whereas game mechanics have not. However, game interfaces are not further investigated in this chapter.
  - Game culture: traditions of game play and cultural dependencies exist, but are also not investigated in this chapter. They might influence the game elements (and their outer appearance) as well as the game mechanics.
  - There might also be a game psychology and game testing as additional categories.

One way to describe different aspects of game design principles is to apply **game design patterns**. Design Patterns (Alexander 1977) represent a general method to describe recurring things and best practices in different domains, and they have been successfully applied to the game elements and game mechanics (Björk and

Holopainen 2004). Game patterns for game mechanics, as distinguished above, can come in the shape of story or character patterns (e.g., Maciuszek and Martens 2014), they can be realized as programming patterns for computer-based games (e.g., Maciuszek and Martens 2011), and additionally, as all patterns, they can be used as the basis for communication in interdisciplinary teams – and thus, patterns can also be used to describe game mechanics and game elements.

Another way to describe certain types of games are game models. A model can then be used similar to a pattern as a draft for designing a certain type of pattern. Moreover, a model of a game can be used for first testing and experimental playing of a game. From the perspective of game design, a game model would be constructed by game elements and game mechanics and the potentially existing other aspects of game design. And it is also possible to develop the game model based on game patterns.

First result: There are game design principles, which can at least be game mechanics and game elements. A method to describe them can be game patterns, which can occur in every aspect of game design.

The term *game thinking* (Werbach and Hunter 2012) stems mainly from a combination of design thinking and game design and brings both of these trends together. Design thinking is a general (not game related) method for problem solving, like other methods, for example, hybrid thinking. Design thinking focuses on the end user's (or customer's) added value and is best used in interdisciplinary teams (see, e.g., Brown 2006). Bringing game design and design thinking together shall lead to new designs, which now occur under the term *gamification* (e.g., Marczewski 2015). In the combination of game and design thinking, it looks like the goal of game thinking might be comprised by all what is associated with games, e.g., having fun, being motivated. But looking at the deep meaning of design thinking, game thinking would lead to an interdisciplinary perspective on end users, which are perceived to be game players.

Second result: Game thinking brings together design thinking and game design, which leads to a way to practice problem analysis and problem solving taking aspects of game design principles into account, and keeping a user-centered design in mind.

From this perspective, some of the basic rules for games (e.g., Caillois 2001; Huizinga 2006) should also be remembered, to make clear what game players (culturally dependent) expect as a game. McGonigal (2011) sees four defining traits that all games share: a goal, rules, a feedback system, and voluntary participation. This is quite close to the basic set, which leads to the basic rules for games, which are

- Games are fictitious and symbolic.
- Games are voluntary.
- Games have rules.

Caillois (2001) defined that all games have at least one of the following four principles: Agon (contest), Alea (chance), Mimikry (to mask), and Ilinx (emotional state, often also called flow, but this does not necessarily relate!). There are many more aspects in games, and in play, but at least these should be kept in mind for this chapter. A close relation to gamification can be found in the principles agon, as gamification also addresses the aspect contest, and alea, as gamification always addresses chance. Mimikry and Ilinx can not directly be found, but might be interpreted, as gamification masks the original situation (be it customer service or University lecture) and Ilinx is the target of using gamification, and also a prerequisite, as persons which don't like to be gamified might very easily be repelled.

The Relation to Edutainment

Definitions of the term *Edutainment* typically stress the instructional basis of corresponding approaches. Buckingham and Scanlon (2000) define it as "… a hybrid genre that relies heavily on visual material, on narrative or game-like formats, and on more informal, less didactic styles of address. The purpose of edutainment is to attract and hold the attention of the learners by engaging their emotions through a computer monitor full of vividly coloured animations. It involves an interactive pedagogy and, … totally depends on an obsessive insistence that learning is inevitably 'fun'" (Buckingham and Scanlon 2000, after Okan 2003).

Edutainment suffered from a critical reflection in the last years, on the one hand based on sometimes too simplistic and too excessive approaches to enrich instructional material with entertaining elements to make learning fun (e.g., Okan 2003), on the other for not considering the state-of-the-art in gaming, e.g., with respect to the integration of learning and play and gameplay in general (e.g., Egenfeldt-Nielsen 2005).

Gamification, when applied to the domain of learning, clearly depicts intersections with Edutainment. In this context, both share the motivation to make learning and instruction more fun. As Kapp (2012) puts it in his definition of gamification in the context of learning, "Gamification is … a careful and considered application of game thinking to solving problems and encouraging learning using all the elements of games that are appropriate."

Figure 1 depicts the relationships of the different areas.

Gamification in learning also relates to the field of Serious Games. Similar to the fields of Gamification and Edutainment, there is no generally accepted definition of the term *Serious Games*, and there are different understandings of the objectives, approaches, and methodologies in this field. Serious games may be defined in short as "computer games with embedded pedagogy" (followinging e.g., Zyda 2005), or more extended as "games that do not have entertainment, enjoyment or fun as their primary purpose" (Michael and Chen 2005). Serious games describe the design of full-fledged games for nonentertainment purposes, while "gamified" applications merely incorporate elements of games (Deterding et al. 2011). However, in reality it may be difficult to draw a clear line between games and "artifacts with game elements" (see section "From Games to Game Elements"), and the boundary may be fuzzy. Deterding et al. also point out that such a classification might be a question of perspective: while a gamification designer might intend to create a system that just includes elements of games, and not a full game, a user might perceive it as game experience, a gameful experience, play, or something completely different. Deterding et al. see this openness as central aspects of gamification approaches.

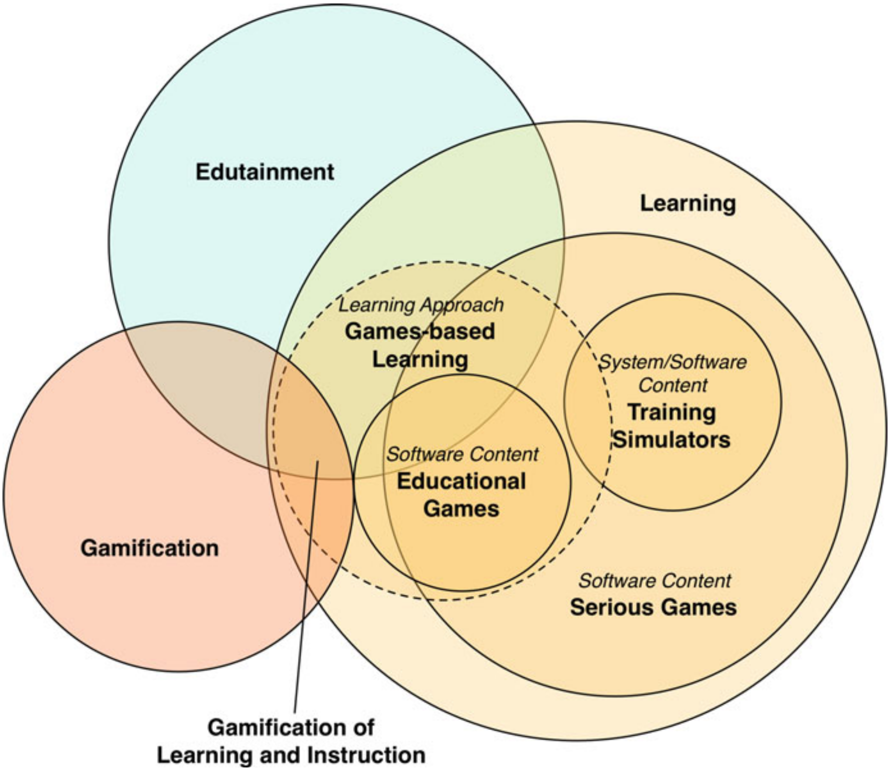


Fig. 1 Relation between edutainment, gamification, and learning (After Tang et al. 2009, with own adaptions)

Thus, in the context of edutainment, the design of gamification in learning would typically not lead to an integrated concept of an instructional game. However, the differences here go further, since gamification typically differs from instructional and learning games in that it is not explicitly based on instructional principles, theories, and assumptions.

Definitions of *Gamification* stated above quite often relate to game elements to be applied in a nongaming context. This raises the questions, which game elements are referred to, and what kinds of game elements are available? These aspects will be discussed in more detail in the following paragraphs.

Game Elements and Game Mechanics

There exist a number of approaches to collect and list typical game elements and game mechanics (not only the typical game design books, like (Fullerton 2012). Examples include

Table 1 Levels of game design elements

	Part of gamification	Description
Game interface elements	Profile and, settings dialogues, information displays	GUI components representing the game world, interactive game elements providing the affordances to interact with the gameplay, and means to represent game information and to control and personalize game behaviour
Game elements	Points, states, levels, badges, leaderboards	Components connected to a media representation directly accessible by the user that are part of the game and represent a certain role or behaviour in the game play
Game mechanics	Rules, roles, time constraint, limited resources, turns	Components defining gameplay: rules and modes of interacting with the game, elements, which cannot be "seen", have no visual representative

– Squidi.net – 300 Mechanics (Howard 2014), a list of 210 game mechanics entries
– Gamification.org (2015) with 24 game mechanics entries, classified into behavioral, feedback, and progression
– The Octalysis gamification framework (Chou 2013), listing some 77 gamification elements classified with respect to 8 core drives of gamification: epic meaning and calling, development and accomplishment, empowerment of creativity and feedback, ownership and possession, social influence and relatedness, scarcity and impatience, unpredictability and curiosity, and loss and avoidance

However, all these listings and classifications appear arbitrary and not necessarily scientifically grounded.

Deterding et al. (2011) distinguish five levels of game design elements in the context of gamification, which represent different levels of abstraction. While the general approach is to clarify the terminology in the field and to provide a classification of terms in this area, the presented approach depicts several deficits. First of all, Deterding et al. (2011) do not clearly distinguish between principles, models, and methods. Second, they do not separate the concept of design patterns as a semiformal documentation of best practices as an approach independent from the general terminology in the field. The following table shows the extraction of all relevant literature used in this chapter, and leads to another categorization (Table 1):

In the following we would like to discuss a selection of most often applied game elements in some more detail: points, badges, levels, leaderboards, and challenges.

Points represent one of the most important elements applied in games and also in Gamification. Points represent a numerical value awarded for achievements or (combinations of) activities in a game, and are thus a means to measure a player's performance. Reaching a certain amount of points may lead to additional rewards, such as more power, access to the next level, a badge, a virtual item, and so on (Gamification Enterprise Wiki 2014). Points can also be understood as a feedback component, allowing players to progress. They also provide players with fundamental means to compare their performance with their own performance in the past and with that of other players. Points may also serve as a motivating element on its own, when players formulate own goals based on them (e.g., "I want to reach 10.000 points!").

Leaderboards are based on points and extend the granular feedback functionality of these into an aggregated feedback (Enterprise Gamification Wiki 2014). They simplify the comparison of players with past performances (personal bests) and other players. Here, it also supports competition, and it may also serve the display of status.

Badges represent one of the most important elements applied in gamification approaches on a regular basis. They represent virtual awards rewarded for specific accomplishments or achievements. Hamari and Eranti (2011) define (achievement) badges as "an optional challenge provided by a meta-game that is independent of a single game session and yields possible reward(s)." Different to points, badges clearly relate to a specific achievement reached and reward gained. As such, they carry the potential to be used as an evidence of achievement and proof of performance, also beyond the gamified activities.

Badges have a long history in social media and are often understood to be an effective instrument to motivate to engage and motivate users.

Hamari and Eranti (2011) distinguish three components of achievement badges: (1) a signifier, (2) a completion logic, and (3) a reward. The signifier resembles the visible part, conveying information on the achievement, such as the name, a description, and a visual representation. The completion logic defines what is required from the player and with respect to the game state for the achievement in terms of a set of foundational rules. The completion logic itself may be decomposed into further components: (1) a precondition on the game setting, (2) a trigger, and (3) a rule or conditional requirement that determines whether the triggered action will be counted toward the achievement. This is a first pattern for badges. Hamari and Eranti (2011) further list a multiplier for cases an action has to be successfully performed several times to fulfill the task and achievement. Rewards refer to a form of price or "payback" that a user receives after completion. Hamari and Eranti (2011) categorize such rewards into meta-game, in-game, or extra-game, where in-game rewards are awarded that can apply in the current game or gamification context, meta-game rewards refer to awards to be applied in a bigger meta-game context that the achievement system belongs to, and extra-game rewards are something the user receives outside the game and outside the achievements meta-game.

Antin and Churchill (2011) analyze badges from a social psychology perspective, resulting in an analysis, which is from our perspective more close to the game mechanics. They differentiate between five primary functions of achievements:

– Goal setting
– Instruction
– Reputation
– Status/affirmation
– Group identification

Goal setting refers to the motivating function of badges as a challenge. They may represent a clear mark set up for a user to be reached, and fun and interest of goal seeking is often the primary reward itself. Most motivating goals seem to be those ones that are just out of comfortable reach. Such goals do not necessarily have to be explicitly set up. They might just be indicated, or the necessary activities may be subjective or imprecisely defined. Goal setting seems to be most effective when users can clearly see and retrace their progress toward the goals. In specific, people often seem to increase their efforts when they know they are near their goal. Feedback is therefore an indispensable element to support goal setting.

The Antin and Churchill function “instruction” leads to a close relation between game mechanics and game thinking: a badge can be even more than just a realization of a game mechanic, it can be another form of instruction about how to interact with the gamified environment (which actions are possible). Badges can be used as visual representatives for social norms or can be used as trigger for group dynamics. In this form, they implicitly show what is allowed and what is highly valued by the developers of the system (or by the group) (e.g., reputation, status, and affirmation, see (Antin and Churchill 2011)). Thus, badges influence gaming on another level, even if certain users never earn a badge or – even more – are not motivated in earning badges. Even these users react on the meta level on the badges as a game element.

As mentioned above, badges are frequently applied in gamification approaches. Well-known examples in the domain of learning include Khan Academy, where badges are awarded for achievements in learning (https://www.khanacademy.org/badges, Kamens 2010), and the Peer 2 Peer University (P2PU, https://courses.p2pu.org), where badges are used as a form of feedback that learners give to each other in order to improve their projects. Further examples from the field of Edutainment include the New York Museum of Modern Art (MoMA), which awarded badges for its Catalyst course Credly.com (2013), and the Smithonian, rewarding badges for a large number of learning activities on its Smithonian Quest platform, for learning at school, at home, and everywhere else (www.smithonianquests.org). Several initiatives targeted to extend the recognition of badges beyond individual organizations awarding such badges. At the time being, the Mozilla Open Badges (http://openbadges.org) represents the accepted technological standard in this field.

Levels, the next game elements, may be defined as “. . . a system, or ‘ramp’, by which players are rewarded an increasing value for a cumulation of points” (Gamification.org 2015). Subsequent levels are typically connected with increasing challenge and demand in terms of performance and skills. Simultaneously, features or abilities are often unlocked as players progress to higher levels, extending the in-game skillset. Leveling is being considered one of the highest components of motivation for players in gaming.

Challenge refers to an ideal amount of difficulty and improbability of obtaining goals presented to a player (Wilson et al. 2009). This includes multiple, clearly defined goals, which have to be meaningful for the player (Garris et al. 2002). Furthermore, progressive difficulty levels should be employed, and a certain amount of informational ambiguity to ensure an uncertain outcome is required. Challenge may also add fun and competition by creating barriers between current state and goal state (Wilson et al. 2009). Performance feedback and score keeping are indispensable to allow a user to assess her progress toward the desired goals (Garris et al. 2002), which represents a connection to points and levels.

Similar to the game element badge, the game element level can have the five game mechanics, which have been differentiated by Antin and Churchill (2011). Levels have a certain instructional effect; levels are a form of reputation, status and affiliation, and group orientation (e.g., you are inside the club or masters or not, you are part of a guild or not, etc.).

As levels and badges are typical game elements, quests represent a typical example for game mechanics: quests are a pattern for story design, they contain role models and usually also a set of rules, which are often also defined by the quest pattern (see, e.g., (Maciuszek and Martens 2014). In the context of gaming, challenge is closely connected to the concept of quests. Quests can be understood as "goals within games with rewards associated to their completion" (Björk 2015). Quests add a narrative aspect and provide story structures. This can be used to communicate intentions of the tasks presented to the user and the effects of their completion (Björk 2015). In the optimal case, the story is meaningful to the user and provides a connection to the user's own experience (Tosca 2003). The motivating and fascinating aspects of quests may be enhanced when they are enriched with an epic meaning: Players will be highly motivated if they believe they are working to achieve something great, something awe inspiring, something bigger than themselves (McGonigal 2011). Challenge and quests have been applied as a guiding principle in several Gamification approaches. Examples include the Quest to Learn (www.q2l.org, Salen et al. 2011), an approach to gamify a public 6–12 school in New York, and the Smithonian Quests, learning opportunities provided by the Smithonian institution and the related museums (http://smithsonianquests.org).

Empirical and Other Forms of Results

Lately, quite some work has been published on the application of gamification in the context of edutainment and learning. However, the reported results are mixed. Barata et al. (2013) were able to increase motivation and participation in a university class on Information Systems and Computer Engineering with a gamification approach. They also observed mild positive effects on learning outcomes, which were, however, not statistically significant. On the other hand, Kruse et al. (2014) failed to enhance motivation for participation and to reduce dropout during the semester for a voluntary Math class in higher education applying a gamification approach. Snow et al. (2015) were able to increase motivation of students based on the introduction of game currency to be earned in learning activities and which could be spent for game-based features, such as integrated minigames and personalizable features. However, they observed a clear relation and negative effects of students' propensity to use this system currency to unlock game-based features on in-system performance, immediate learning outcomes, and skill transfer outcomes (Snow et al. 2015). These three examples may be considered typical for the current situation on the evaluation of gamification in learning.

In general, the number of scientifically valid approaches to evaluate the effectiveness of Gamification is still comparably low. In general, studies may be classified in approaches to prove benefits with respect to the followings aspects:

– Increase of motivation
– Increase of participation
– In the context of edutainment and learning, increase of learning

Few studies exist, trying to analyze and summarize the general effects of gamification in learning and on the effectiveness of certain game elements in these approaches.

Hamari et al. (2014) performed a literature review on empirical studies on Gamification. It therefore does not target to represent a meta-analysis, trying to integrate results from individual studies with statistical methods. Hamari et al. (2014) also identify a number of methodological typical problems of existing studies, including small sample sizes, a lack of control groups, a lack of validated measures, and a lack of consideration of long-term effects. In general, they point out in their review that many gamification studies are descriptive in nature and more research based on rigorous methodologies is required to understand the effects and to conclude on potentials and benefits.

Falkner and Falkner (2014) describe a metastudy on the effect of gamification in Computer Science Education, focusing on badges and virtual awards. Their study also does not comply completely to the requirements of a metastudy, foregoing to conduct a statistical meta-analysis and remaining descriptive in the analysis. In their study, they consider 11 studies and analyze them in some detail, however apparently with motivation and participation as sole aspects, not learning. The conclusions they draw from this analysis is quite negative. They summarize that there is not enough evidence to clearly indicate that badges, by themselves, make any contribution to student engagement or effectiveness in forming a learning community. Even more, they point out that virtual rewards can have strongly negative effects on behavior without providing any substantial list of benefits. Nevertheless, they conclude that there might be situations where badges may represent a valuable component of gamification, but more research and evidence is needed to support this.

Given the lack of expressive studies in the field of Gamification, it makes sense to take a look on the general field of games, trying to identify those studies and results which might be relevant for the field of Gamification. There exist quite a few empirical studies on the effects on games and on the effectiveness of learning with (computer) games, but it is not completely clear in how far the results of such studies may be mapped to the more general domain of Gamification.

With respect to learning, various studies provided evidence that game-based features may lead to increased students' engagement and motivation in learning tasks (Rai and Beck 2012; Jackson and McNamara 2013).

Cognitive gains are more difficult to prove. There is quite a number of work on the evaluation of the educational effectiveness of game-based learning and serious games for learning. However, these studies quite often depict methodological weaknesses (Mayer 2014). More thoroughly performed studies provide in fact mixed results (e.g., Randel et al. 1992; Hays 2005; Vogel et al. 2006; Tobias et al. 2011; Young et al. 2012). Mayer (2014) carefully concluded from a meta-analysis of a substantial number of media-comparison studies that games provided better learning outcomes than conventional learning in second-language learning and science, with medium-to-large effect sizes, and that effects were strongest for college students and adults, weakest for elementary school children. Also, adventure games seemed to be more effective than simulation games, which again appeared more effective than quizzes and puzzle games. However, Mayer himself questioned the approach of media-comparison studies in general due to the large number of sometimes intimately connected variables involved. In specific, it is often claimed the instructional methods cause learning, not instructional media, and it is therefore difficult to attribute learning outcome aspects to learning games as media (Mayer 2014).

It may be concluded that the application of games in instruction and training may facilitate transfer and learning sometimes as well and sometimes also better than traditional methods (Tobias et al. 2011). However, this raises the questions, which game elements and mechanics do foster learning, and should therefore be considered of value or even indispensable when applying gamification in learning.

Wilson et al. (2009) performed a literature review and analyzed a number of publications with respect to results on game attributes and their effects on learning. The study focused on the attributes fantasy, representation, sensory stimuli, challenge, mystery, assessment, and control. While it proved difficult to break down effects on single attributes, Wilson et al. nevertheless proposed possibly positive effects of elements such as conflict presented within a game to cognitive learning, challenge to declarative knowledge and learner's retention, adaptation features to cognitive strategies, and interaction and control to skill-based learning. Further effects were identified with respect to affective outcomes, in specific on motivation. The study does not provide a statistical validation of the proposed effects.

Mayer (2014) proposes five promising features that have shown to improve learning when added to games based on a metastudy on value-added research approaches: modality, personalization, pretraining, coaching, and self-explanation. These features do not directly correspond to game elements and game mechanics, but they do correspond to best practices in game design (cf., Koster 2014). In addition, Mayer lists additional six features that may be effective to foster learning, but where evidence on their effectiveness is insufficient so far. These features are competition, segmenting, image, narrative theme, choice, and learner control (Mayer 2014). Here, competition relates to the introduction of a goal connected to a prize that can be gained, narrative theme to the introduction of an overall theme and story, choice to the ability of players/learners to choose aspects of the game format, and learner control to the ability of players/learners to determine the order of levels in a game. Clearly, these four aspects may be related to game mechanics, though these have to be considered quite abstract. Segmenting refers to segmenting the screen into parts, and, thus, a concept more related to effective multimedia presentation in general (Mayer 2009).

In summary, the question whether points, levels, leaderboards, or badges can be considered value tools to motivate students and to enhance learning remains difficult to answer, also from the perspective of game-based learning and edutainment.

Especially the assessment on long-term effects is difficult in the absence of expressive studies in this direction.

Gee (2011) states a quite general and valid critique to such existing studies on the effectiveness of games and game elements to promote learning, which in parts also extends to the field of Gamification.

The quite diverging results of empirical studies may be due to very different types of gamification approaches and game ideas applied in the individual cases, comparable to different types of games and game genres. In games, the differences between such genres are big, and it is therefore obvious that it should be considered in empirical studies as an additional free variable when assessing their effectiveness, for instance, to promote learning. Results from such studies should therefore clearly be related to the applied Gamification methodology. However, a widely agreed classification of such methodologies, which would allow for a comparison of results in one category, is so far missing.

Another crucial point is that gamification can be done good or bad, and learners may like it and devote much time, attention, and energy to the gamified system, or not. This is very similar to field of games (Gee 2011). That is, game elements and game mechanics have been successfully integrated with task-related elements in a motivating manner, and goals are clearly communicated to provide the feeling of a challenge, quest, and epic meaning, or not. It is obvious that a bad gamification approach will fail to enhance learning, and may even hinder learning. So, when evaluating a gamification approach to get empirical evidence on its ability to enhance learning, it has to be shown that the game idea and implementation itself are convincing and motivating. However, so far there are no good theories about what makes a good game and a good gamification, and there is a lack of instruments to assess the perceived quality of a gamification approach. It should be stressed at this point that assessment of motivation and participation are not sufficient to the perceived quality of a game by a user.

Furthermore, game quality is rated differently by different users (Hays 2005). Evaluation studies therefore – whether positive or negative in results – need to clearly reference the category of gamers they are making claims about (Gee 2011). This is often not the case, neither in studies in Gamification nor in game-based learning.

While the introduction and application of game concepts may not represent the driving factor for enhanced motivation, participation, and learning, the question whether gamification may be accompanied by positive effects deserves further attention. In fact, experiences from the introduction of e-learning in teaching and learning also showed that the introduction of new instructional media may also provide the opportunity and trigger rethinking of established teaching strategies, leading to the introduction of novel teaching concepts and strategies. It is therefore sometimes difficult to say whether improvements in learning may be related to the introduction of new media, or whether the same improvements could have been accomplished based on the novel teaching strategies alone.

The implementation of gamification concepts in learning may therefore be accompanied by similar effects. Gamification is often attributed to create situated learning experiences with incentives and clear objectives (Miller 2012). That is, introducing a gamification approach, the development of an overall game theme and story may lead to a clearer presentation of learning goals and structure of the learning process. It may also introduce a framework for learning in context and possibly foster situated learning. The introduction of elements such as levels and quests helps to implement an incrementally staged process in the spirit of cognitive apprenticeship approach. Points and badges provide opportunities for immediate, meaningful feedback to learners. Especially the aspect of appropriate feedback seems to be a central element of successful gamification approaches. Feedback has been identified as one of the most important aspects in learning (Hattie 2013), and this makes one wonder whether it is not the much more important aspect than the introduction of game concepts. Or, as Herger put it in the title of an ISAGA keynote, "… Gamification is not about games, but feedback and information." (Herger 2014). Still, similar to the field of e-learning, though, comprehensive studies to analyze such secondary effects of gamification are missing.

In summary, it can be stated that there is a lack of evidence on the benefits of gamification in the context of learning. This is well in line to critical assessments such as Perrotta et al. (2013) and Falkner and Falkner (2014). It can be concluded that more research is needed to prove the impact of such techniques and to understand possible mechanisms and effects on learning outcomes Perrotta et al. (2013). Regarding game elements and game mechanics, the question remains whether these elements could apply isolated without losing their meaning or potency, and whether they could be used effectively to add value to traditional teaching and learning (Perrotta et al. 2013).

Conclusion

This article provided an overview on the field of Gamification and its historical roots. Furthermore, definitions of the term *Gamification* were presented and discussed, and a comparison to approaches in the fields of instruction, gaming, and edutainment were given. This discussion revealed that the boundaries between gamification and other discussed fields are blurry, and the terminology is sometimes inconsistent with the one in other subject areas, e.g., gaming. This is especially true with respect to the discussion and classification of concepts such as game elements, game mechanics, and game design patterns. For this reason, a classification of levels of game design elements was provided, closer grounded to the accepted terminology in the field of computer games.

The chapter also presented an introduction and discussion of the most important game elements and game mechanics applied in the field of Gamification to date. Furthermore, a detailed discussion on approaches on empirical assessments of the effects of Gamification is given. There exists evidence that Gamification may foster motivation and participation of users. Regarding the application in the field of education with additional objectives targeted to instruction, evidence that Gamification may foster learning has to be considered weak. Valid empirical studies on the effects on Gamification approaches and specific game elements applied are rare. It can be concluded that further research is required in this direction. However, it may be questioned whether learning effects can be related to individual game elements and game mechanics, since a successful gamified experience will always be based on game thing that carefully selects of a variety of these and integrates them to an exciting game. Furthermore, it seems problematic to assess the effects of gamification to motivation, participation, and learning, without a good understanding whether the gamified scenario or game is really good from the point of the players, and without taking into consideration the differences of user groups with respect to their stance toward games. It seems that appropriate methodologies and instruments for the assessment of both, quality of games and gamifications and the assessment of divergent attitudes with respect to games, are missing.

In summary, Gamification depicts potential to inject new life into the field of Edutainment, leveraging and extending best practices from the field of gaming and mapping it to more general domains and forms of application. It clearly succeeded to spread ideas from game design and game thinking, and made game concepts acceptable in various domains. However, an approved methodology on how to get to successful gamification is missing so far. Instead, many approaches seem to rely on an ad hoc adaptation of selected game elements and game mechanics, failing to fully exploit the potentials of game thinking and to generate really convincing gameful experiences. It seems that a more thorough understanding of game design principles is required to tap the existing potential. Similarly, when applied in the field of learning and instruction, existing results from the few valid empirical studies in the field seem to support that positive effects on learning effects cannot be reached by focusing on the introduction of game elements, but that a corresponding gamification process requires a simultaneous instructional design. Here, it makes sense to connect to the closely related field of Edutainment, where instructional approaches have always played a more important role and where experiences exist on how to integrate learning and fun successfully. Still, Gamification will have its limits, and there will be areas where the application will not be appropriate or may not bring the expected results. For instance, in the context of commemorative culture first experiments introducing an edutainment and gamification approach appeared problematic, leading to distraction and a level of trivialization, which was not considered necessarily positive (Haake and Müller 2015). More experiences are required to clearly understand the potentials and limits of Gamification.

Recommended Reading

C. Alexander, *A Pattern Language. Towns, Buildings, Construction* (Oxford University Press, New York, 1977)

J. Anderson, L. Rainie, *The Future of Gamification* (Pew Research Center, Washington, DC, 2012). Retrieved from http://www.pewinternet.org/2012/05/18/the-future-of-gamification/

J. Antin, E.F. Churchill, Badges in social media: a social psychological perspective, in *Proceedings of the CHI 2011* (Vancouver, 2011)
G. Barata, S. Gama, J. Jorge, D. Gonçalvez, Improving participation and learning with gamification, in *Gamification'13* (Stratford, 2013)
S. Björk, Gameplay design patterns collection: quest (2015), http://129.16.157.67:1337/mediawiki-1.22.0/index.php/Quests. Retrieved 25 Oct 2015
S. Björk, J. Holopainen, *Patterns in game design. Game development series* (Charles River Media, Hingham Massachusetts, 2004)
I. Bogost, Gamification is bullshit (2011), http://bogost.com/blog/gamification_is_bullshit/. Retrieved 1 Jan 2011
B. Brathwaite, I. Schreiber, *Challenges for Game Designers – Non-digital exercises for video game designers* (Charles River Media, 2009). Retrieved from http://ircg.ir/Lib/challenges-for-game-designers.pdf
T. Brown, *Innovation through design thinking, MIT video*. Video of a talk given at Massachusetts Institute of Technology, Boston (2006). Retrieved from http://video.mit.edu/watch/innovation-through-design-thinking-9138/
D. Buckingham, M. Scanlon, *That Is Edutainment: Media, Pedagogy and the Market Place* (International Forum of Researchers on Young People and the Media, Sydney, 2000)
Burke, B. (2012). Gamification 2020: What is the future of gamification? Retrieved from https://www.gartner.com/doc/2226015?refval=&pcp=mpe#dv_2_source_google
R. Caillois, *Man, Play and Games* (Original in French from 1958, English translation from 1961) (University of Illinois Press, Urbana/Chicago, 2001)
I. Caponetto, J. Earp, M. Ott, Gamification and Education: A literature Review. In Proc. ECGBL 2014: Eighth European Conference on Games Based Learning, (Berlin, 2014), pp. 50–57
Y.-K. Chou, Octalysis: complete gamification framework (2013), http://www.yukaichou.com/gamification-examples/octalysis-complete-gamification-framework/#.VWV912BY99R. Retrieved 27 May 2015
Credly.com, *Credly to Issue Digital Badges for The Museum of Modern Art's Online Courses* (Credly.com, 2013), http://blog.credly.com/moma-open-badges/. Retrieved 25 Oct 2015
M. Csikszentmihalyi, *Flow: The Psychology of Optimal Experience* (Harper & Row, New York, 1990)
S. Deterding, D. Dixon, R. Khaled, L. Nacke, From game design elements to gamefulness: defining "gamification", in *Envisioning Future Media Environments* (MindTrek, 2011), pp. 9–15. doi:10.1145/2181037.2181040
M. Ecker, W. Müller, J. Zylka, Game-based learning design patterns – an approach to support the development of "better" educational games, in *Handbook of Research on Improving Learning and Motivation Through Educational Games: Multidisciplinary Approaches*, ed. by P. Felicia (IGI Global, Hershey, 2011), pp. 137–152
S. Egenfeldt-Nielsen, *Beyond Edutainment: Exploring the Educational Potential of Computer Games* (Lulu.com, 2005), Retrieved from https://books.google.com/books?hl=de&lr=&id=snupBAAAQBAJ&pgis=1
N.J. Falkner, K.E. Falkner, Whither, badges? or wither, badges!: A metastudy of badges in computer science education to clarify effects, significance and influence. In Proceedings of the 14th Koli Calling International Conference on Computing Education Research, Koli, Finland, 2014, pp. 127–135
M. Fuchs, S. Fizek, P. Ruffino, N. Schrape, *Rethinking Gamification* (Menson Press, 2015), http://meson.press/read/rethinking-gamification. Accessed 10 Aug 2015
T. Fullerton, *Game-design workshop*, Kindle Edition (CRC Press, Boca Raton, Florida, 2012)
Wikipedia entry Gamification (2015), https://de.wikipedia.org/wiki/Gamification. Accessed 14 Oct 2015
Gamification Enterprise Wiki, *Enterprise Gamification Wiki: Point* (Enterprise Gamification Wiki, 2014), http://www.enterprise-gamification.com/mediawiki/. Retrieved 23 Oct 2015

Gamification.org, *Gamification Wiki: Game Mechanics* (Gamification.org, 2015), https://badgeville.com/wiki/Game_Mechanics. Retrieved 23 Oct 2015

Garris, R., Ahlers, R., & Driskell, J. E., Games, motivation, and learning: A research and practice model. Simulation & Gaming, 33(4), 441–467 (2002)

J.P. Gee, Reflections on empirical evidence on games and learning, in *Computer Games and Instruction*, ed. by S. Tobias, J.D. Fletcher (Information Age Publishing, Charlotte, 2011), pp. 223–232

S. Haake, W. Müller, New memory spaces for cultural history. Novel approaches for mobile applications in the field of commemorative culture, in *Proceedings of the Digital Heritage 2015* (IEEE Computer Society, Granada, 2015)

J. Hamari, V. Eranti, Framework for designing and evaluating game achievements, in *Proceedings of the DiGRA 2011: Think Design Play* (Digital Games Research Association DiGRA, 2011), pp. 122–134

J. Hamari, J. Koivisto, H. Sarsa, Does gamification work? A literature review of empirical studies on gamification, in *System Sciences (HICSS), 2014 47th Hawaii International Conference on System Science* (Hawaii, USA, 2014), pp. 3025–3034

J. Hattie, *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement* (Routledge, New York, 2013)

R.T. Hays, The effectiveness of instructional games: A literature review and discussion. Naval Air Warfare Center, Training Systems Division. Orlando, Florida. October. Retrieved from: http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA441935

Herger, M, Why gamication is not about games, but feedback and information. ISAGA 2014 Keynote lecture (2014). http://www.isaga2014.com/

S. Howard, Squidi.net – three hundred mechanics (2014), http://www.squidi.net/three. Retrieved 23 Oct 2015

J. Huizinga, *Homo ludens: Vom ursprung der kultur im spiel* (Rowohlt, Hamburg, 2006)

K. Huotari, J. Hamari, Defining gamification – a service marketing perspective, in *Proceeding of the 16th International Academic MindTrek Conference* (ACM, 2012). Retrieved from http://www.rolandhubscher.org/courses/hf765/readings/p17-huotari.pdf

G.T. Jackson, D.S. McNamara, Motivation and performance in a game-based intelligent tutoring system. J. Educ. Psychol. **105**(4), 1036 (2013)

W. Jank, H. Meyer, *Didaktische Modelle* (Cornelsen Verlag, Berlin, 2002). (In German, book is about instructional design)

B. Kamens, Khan Academy now has badges (2010), http://bjk5.com/post/2426884194/khan-academy-now-has-badges. Retrieved 25 Oct 2015

B. Kampman Walther, Playing and gaming – reflections and classifications, game studies. Int. J. Comput. Game Res. **3**(1), (2003). http://www.gamestudies.org/0301/walther/

K.M. Kapp, *The Gamification of Learning and Instruction* (Pfeiffer/Wiley, San Francisco, 2012)

R. Koster, *A Theory of Fun for Game Design*, 2nd edn. (O'Reilly, Sebastopol, 2014)

V. Kruse, C. Plicht, J. Spannagel, M. Wehrle, C. Spannagel, Creatures of the Night: Konzeption und Evaluation einer Gamification-Plattform im Rahmen einer Mathematikvorlesung, in *Proceedings of the Pre-Conference Workshops 12. e-Learning Fachtagung Informatik DeLFI 2014*, ed. by C. Rensing, S. Trahasch (2014), pp. 246–253. In German. Online

Loyality360, *Gamification of Loyality: Driving Deeper Customer Engagement Through the Power of Play* (Loyality Management, 2009), p. 64, http://de.scribd.com/doc/17718638/Loyalty-Expo-2009-in-Review

D. Maciuszek, A. Martens, Chapter 9: Patterns for the design of educational games, in *Educational Games: Design, Learning, and Applications*, ed. by F. Edvardsen, H. Kulle (Nova Publishers, 2010), pp. 263–280, ISBN 978-1-60876-692-5

D. Maciuszek, A. Martens, Patterns for designing digital educational game content. Int. J. Game-Based Learn. **1**(3):65–82 (2011)

D. Maciuszek, A. Martens, Integrating cognitive tasks in game activities, in *Proceedings of the 12th IEEE International Conference on Advanced Learning Technologies* (Rome, 2012), pp. 376—378. doi:10.1109/ICALT.2012.205, ISBN 9780769547022

D. Maciuszek, A. Martens, Construction spaces for the design of learning tasks in virtual worlds, in *Proceedings of the 8th International Conference on Games Based Learning (ECGBL 2014)* (Academic, Reading, 2014)

A. Marczewski, Game thinking – differences between gamification and games, http://www.gamified.uk/gamification-framework/differences-between-gamification-and-games/. Accessed 14 Oct 2015

A. Martens, H. Diener, S. Malo, Game-based learning with computers – learning, simulations, and games. J. Trans. Edutainment **5080**, 172–190 (2008)

R.E. Mayer, *Multimedia Learning* (Cambridge University Press, Cambridge, MA, 2009)

R.E. Mayer, *Computer Games for Learning – An Evidence-Based Approach* (MIT Press, Cambridge, MA, 2014)

J. McGonigal, *Reality Is Broken: Why Games Make Us Better and How They Can Change the World* (Vintage Digital, London, 2011)

E.D. Mekler, F. Brühlmann, K. Opwis, A.N. Tuch, Do points, levels and leaderboards harm intrinsic motivation? in *Proceedings of the First International Conference on Gameful Design, Research, and Applications – Gamification'13*, (2013), pp. 66–73. doi:10.1145/2583008.2583017

D. Michael, S. Chen, Serious games - Games that educate, train, and inform. Thomson (2006)

A. Miller, K-12 gamification of education Q&A – Gamification Co (2012), http://www.gamification.co/2012/09/27/gamification-of-k12-education/. Retrieved 13 Oct 2015

Z. Okan, Edutainment: is learning at risk? Br. J. Educ. Technol. **34**(3), 255–264 (2003). Retrieved from http://www.savie.ca/SAGE/Articles/1084_453-Okan-2003.pdf

D. Parlett, What's a ludeme? And who really invented it (2015), http://www.davidparlett.co.uk/gamester/ludemes.html. Retrieved 13 Oct 2015

C. Perrotta, G. Featherstone, H. Aston, E. Houghton, *Game-Based Learning: Latest Evidence and Future Directions* (Slough, 2013). Retrieved from https://www.nfer.ac.uk/nfer/publications/GAME01/GAME01.pdf

D. Rai, J.E. Beck, Math learning environment with game-like elements. Int. J. Game-Based Learn. **2**(2), 90–110 (2012). doi:10.4018/ijgbl.2012040106

J.M. Randel, B.A. Morris, C.D. Wetzel, B.V. Whitehill, The effectiveness of games for educational purposes: a review of recent research. Simul. Gam. **23**(3), 261–276 (1992)

S. Rausch, U. Faßhauer,, A. Martens, Evaluation of competence development in WoW, in *Proceedings of the International Conference on Edutainment and Game Days 2012 No 7516 LNCS,, Springer Lecture Notes in Computer* (2012), pp. 78–88. doi:10.1007/978-3-642-33466-5_9, ISBN 9783642334658

M. Robertson, Can't play, won't play (2010), http://www.hideandseek.net/2010/10/06/cant-play-wont-play/. Retrieved 13 Oct 2015

K. Salen, R. Torres, L. Wolozin, R. Rufo-Tepper, A. Shapiro, *Quest to Learn: Developing the School for Digital Kids* (MIT Press, Cambridge, MA, 2011)

M. Sicart, Defining game mechanics. Game Stud. **8**(2), 1–14 (2008). Retrieved from http://gamestudies.org/0802/articles/sicart?viewType=Print&viewClass=Print

E.L. Snow, L.K. Allen, G.T. Jackson, D.S. McNamara, Spendency: students' propensity to use system currency. Int. J. Artif. Intell. Educ. **25**(3), 1–21 (2015)

S. Tang, M. Hanneghan, A.E. Rhalibi, Introduction to games-based learning. In T. Connolly, M. Stansfield, & L. Boyle (Eds.), Games-based learning - advancements for multi-sensory human computer interfaces (New York: IGI Global, 2009)

S. Tobias, J.D. Fletcher, D.Y. Dai, A.P. Wind, Review of research on computer games, in *Computer Games and Instruction*, ed. by S. Tobias, J.D. Fletcher (Information Age Publishing, Charlotte, 2011), pp. 127–221

S. Tosca, The quest problem in computer games. Technologies for interactive digital storytelling and entertainment (TIDSE) conference. Darmstadt (2003)

J.J. Vogel, D.S. Vogel, J. Cannon-Bowers, C.A. Bowers, K. Muse, M. Wright, Computer gaming and interactive simulations for learning: a meta-analysis. J. Educ. Comput. Res. **34**(3), 229–243 (2006)

K. Werbach, D. Hunter, *For the Win. How Game Thinking Can Revolutionize Your Business* (Wharton Digital Press, Philadelphia, 2012)
K.A. Wilson, W.L. Bedwell, E.H. Lazzara, E. Salas, C.S. Burke, J.L. Estock, K.L. Orvis, C. Conkey, Relationships between game attributes and learning outcomes. Review and research proposals. Simul Gam. **40**(2), 217–266 (2009)
M.F. Young, S. Slota, A.B. Cutter, G. Jalette, G. Mullin, B. Lai, . . . M. Yukhymenko, Our princess is in another castle: A review of trends in serious gaming for education. Rev. Educ. Res. **82**(1), 61–89 (2012)
Zyda, M, From visual simulation to virtual reality to games. Computer, 38(9), 25–32 (2005)

Part IX

Entertainment Robots

Challenges for Robots Acting on a Stage 37

Creating Sequential Structures for Interaction and the Interaction Process with the Audience

Mayumi Bono, Perla Maiolino, Augustin Lefebvre, Fulvio Mastrogiovanni, and Hiroshi Ishiguro

Contents

M. Bono (✉)
National Institute of Informatics, National Center of Sciences, Tokyo, Japan
e-mail: bono@nii.ac.jp

P. Maiolino
Goldsmiths University of London, London, UK
e-mail: p.maiolino@gold.ac.uk

A. Lefebvre
Sorbonne Nouvelle University Paris 3, Paris, France
e-mail: nitsuguata@hotmail.fr

F. Mastrogiovanni
Department of Informatics, Bioengineering, Robotics, and Systems Engineering, University of Genoa, Genoa, Italy
e-mail: fulvio.mastrogiovanni@unige.it

H. Ishiguro
Osaka University, Osaka, Japan
e-mail: ishiguro@sys.es.osaka-u.ac.jp

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_62

Abstract

In this chapter, we describe issues and challenges in deploying robots on stage, specifically in the context of theatrical plays. We are interested in understanding how an audience may perceive the robots. To this aim, we first survey available literature about cognitive human–robot interaction, with an emphasis on direct perception of robots and the observation of other humans interacting with robots. Then we discuss state-of-the-art approaches to select and sequence robot behaviors for human–robot interaction. Finally, we report about long-term interaction experiences involving robots and human actors on stage.

Keywords

Human–robot interaction • Robot actors • Motor resonance • Robot behaviors • Robot theater project

Introduction

During the past decades, robots have gained a prominent role in industrial and manufacturing settings. Industrial robots are usually designed to perform heavy and repetitive tasks within well-defined and structured environments. Even the most common robot-based shop floor operations do not require any form of interaction with the environment, not even considering humans.

However, the role of robots in our society is rapidly changing. The possibility of employing robots in tasks where interaction and cooperation with humans is necessary is under active debate both in academia and the industry. Possible application scenarios involve direct interaction with humans, for instance, when robots are used in assistive tasks (e.g., housekeeping, healthcare support, or entertainment, just to name a few).

The interaction between humans and robots, as reported by Goodrich and Schultz (2007), *is the process of working together to accomplish a goal*. This, by definition, requires the unfolding of a communication process between humans and robots. Communication should be as much *natural* and *appropriate* as possible and should convey in the observer a sense of *believability* (Breazeal 2002; Fong et al. 2003).

This new class of robots (which we refer to as *social robots*, according to the definition by Matarić 2006) should be qualified not only in terms of mechanical performance (e.g., speed or accuracy) but also for their capabilities in communicating with humans. The study of human–robot interaction (HRI) considers the evaluation of the interaction processes between humans and robots, as well as the study of design methodologies to provide robots with advanced sensorimotor, cognitive, emotional, and social skills for achieving desirable interactions (Goodrich and Schultz 2007; Dautenhahn 2007). If we consider robots for entertainment purposes, this is particularly evident (Bruce et al. 1999; Burke et al. 2001; Arkin et al. 2003).

In this chapter, we investigate two intertwined issues related to how humans perceive robots (when they interact with them) and how they perceive robots interacting with other humans. The two questions are:

Q1. Do humans perceive a robot they are interacting with as if it was a human as well?
Q2. Do humans perceive a robot interacting with other humans as *natural*?

The two questions may look similar; however, they are characterized by a subtle difference. On the one hand, Q1 must be answered at a perceptual level, in order to avoid subjective judgments that may be originated by subjective beliefs (Oztop et al. 2005a; Oberman et al. 2007; Stanley et al. 2010; Shen et al. 2011; Wykowska et al. 2012). One is interested in understanding whether a robot has the same effects on human perception as humans do and under which conditions. On the other hand, Q2 requires considering those subjective judgments that must be avoided in Q1, since the credibility and *naturalness* of a human–robot interaction, as seen in "a third-person perspective," involves emotions and subjective beliefs (Gallese and Goldman 1998; Carr et al. 2003).

We will discuss how robots acting on a stage are perceived by other actors and, possibly, the audience. It is noteworthy that in such scenario, different aspects are involved: the audience's expectation, the fact that a script must be followed by actors very precisely (in that respect, robots may be expected to act in order to reach a given goal and therefore assigned with a *sense of agency* as pointed out by Sciutti et al. 2013), and that close-proximity cues, such as gaze and subtle facial expressions (Minato et al. 2006; Sciutti et al. 2012b; Andrist et al. 2015; Huang et al. 2015), can be hardly appreciated by the audience, just to name a few.

Furthermore, we aim at providing a new methodology to design long-term human–robot interactions, and we investigate the audience's perception and reactions related to performances on the stage. We call this methodology "interaction science" (IntSci). Indeed, the audience is expected to focus on nonverbal cues in actor performance, such as gestures (Chaminade et al. 2010; Georgescu et al. 2014), movements on stage, and timing, i.e., the ability to timely coordinate with other actors (Kose-Bagci et al. 2010; Huber et al. 2013), whereas verbal cues follow the script of the theatrical piece on stage. As it has been mentioned by Mastrogiovanni and Sgorbissa (2013), professional actors intervene at very precise moments during a scripted dialogue, which is defined at rehearsal time by the director, synchronize with other actors and their movements, move precisely in specific stage locations, and do not leave much space for improvisation.

As discussed above, in a broader HRI perspective, robots are expected to exhibit a *natural* and *appropriate* interaction with humans in order to convey in them a sense of *believability*. What are we talking about when we talk about *natural interaction*? What are the main characteristics of a natural interaction? We argue, and we are supported by the available literature, that *natural* means that the robot's behavior (if a human is interacting with it) or an observed interaction process involving robots and humans (if a human acts as an external observer) must be easily recognizable and understandable. As Rizzolatti et al. (2001) put it, "an action is understood when its observation causes the motor system to *resonate*."

Therefore, recognition is related to a phenomenon usually referred to as *direct-matching hypothesis* or *motor resonance*. When an observer witnesses an action performed by another individual, the neurons representing the observed action are activated in the motor cortex. This leads to a motor representation of the action corresponding to the one generated when the observer performs the same action (Sadeghipour et al. 2011; Chaminade et al. 2010). Such an "internal simulation" mechanism allows understanding the meaning of the observed action, because it links the action itself with its outcome (Csibra 2004); furthermore, it allows naturally and interactively coordinating one's behavior with others across different communication channels (such as speech, gesture, posture, and gaze), as posited by Andrist et al. (2015). The effect seems to disappear when stimuli look "artificial" (Rizzolatti and Craighero 2004; Chaminade et al. 2007; Stanley et al. 2010). This suggests that this effect arises when an observer has recognized the observed individual as a conspecific (Sciutti et al. 2012a).

Understanding in which way and under which conditions motor resonance emerges during an interaction process is decisive for designing robots that can interact in a natural way with humans and to understand if the occurring interaction is perceived as natural (Oztop et al. 2005a). It has been demonstrated that observing someone, while they move, has a measurable interference on qualitatively different simultaneous movements executed by the observer. Kilner et al. (2007) and Cook et al. (2014) posit that observed movements are processed in a different way according to whether they are performed by a human or a robot. This is defined *interference effect* (IE) and it is grounded in the motor resonance effect. It is necessary to understand whether such interference effect is specific to the observation of a biologically plausible motion trajectory, to the observation of a human shape, or to a combination of both.

A movement performed by a human has a characteristic nonlinear velocity profile and follows a minimum jerk trajectory, which is quite different from a constant velocity profile typically characterizing robot motions. Initial studies by Kilner et al. (2007) demonstrate that observed human movements interfere with executed robot arm movements only when these are biologically plausible. On the contrary, observed, nonbiologically plausible motions have no interference effect on executed movements. These two facts allow Kilner et al. (2007) to conclude that interference is due to the biological nature of human movements, rather than to the shape of the human body. Kupferberg et al. (2012) emphasize that the degree of motor interference is determined by various features of the observed agent and investigate how different aspects of appearance and motility influence it. They conducted a study where participants were asked to perform horizontal and vertical arm motions while observing videos showing a human, a humanoid robot, and an industrial robot arm, respectively, with either artificial (i.e., industrial) or humanlike joint configurations, performing similar motions. They show that, given a humanlike joint configuration, motor interference is elicited by observing arm motions of both humanoid and industrial robots. However, if the joint configuration of the robot does not resemble a human arm, motor interference does not occur. These results suggest an evidence for the importance of humanlike joint configurations rather than other humanlike features when observing inanimate agents. In non-humanlike robots (but characterized by a human-inspired face), such traits as skin color and masculine or feminine aspects play a significant role in acceptability, as pointed out by Eyssel and Kuchenbrandt (2012) and Eyssel and Hegel (2012).

Besides motor interference, *proactive gaze behavior* (PGB) is a related phenomenon occurring when observing actions, specifically when the observing human's gaze anticipates the forthcoming goal of the observed movement rather than tracking the corresponding motion (Sciutti et al. 2012a). Such behavior appears when the observed object is clearly visible and disappears when the object moves by itself. This suggests that it can be exploited to understand how a human observer perceives the interaction. Sciutti et al. (2012a) compare gaze behavior when participants observe a human and the iCub humanoid robot performing a simple object-transportation task. They conclude that participants exhibit a very similar anticipatory gazing pattern when observing the human and the robot and argue that the robot evokes gaze proactivity as well.

In a more general perspective, gaze behaviors are employed in interaction processes to communicate intentions, while interpreting another's gaze behaviors allows one to predict theirs. As pointed out by Huang et al. (2015), this dual process allows people to adapt their behaviors to the intents of others and vice versa. Timing plays an important role in communication, whereas delays, gaps, and mismatches in responses, during an interaction process, have a negative impact on humans (Kose-Bagci et al. 2010; Huber et al. 2013). Minato et al. (2006) analyze gaze behaviors in human–android interaction and compare the results with those obtained in similar human–human interaction processes. Their goal is to evaluate the human likeness of an android robot investigating how humans break eye contact. They posit that a human may break eye contact more often if the interlocutor (i.e., the android robot) was not perceived as a human.

Beside behavioral analyses related to interference, gaze, and timing, a further question arises, which is related to the specific beliefs of human observers and their influence on how humans perceive robots. As pointed out by Weil and Rosen (1995) and Kanda et al. (2001), age, gender, and cultural background must be considered. Studies reported by Stanley et al. (2010) suggest that the interference effect is strongly dependent on beliefs or instructions about the agency of the observed action, especially in situations where the visual stimulus is ambiguous. Different brain areas are involved when human observers are informed that an identical visual stimulus is either of a human or computer origin: in the first case brain regions involved in emotional intelligence are involved, whereas in the second case it is believed that areas related to spatial and temporal analysis are activated.

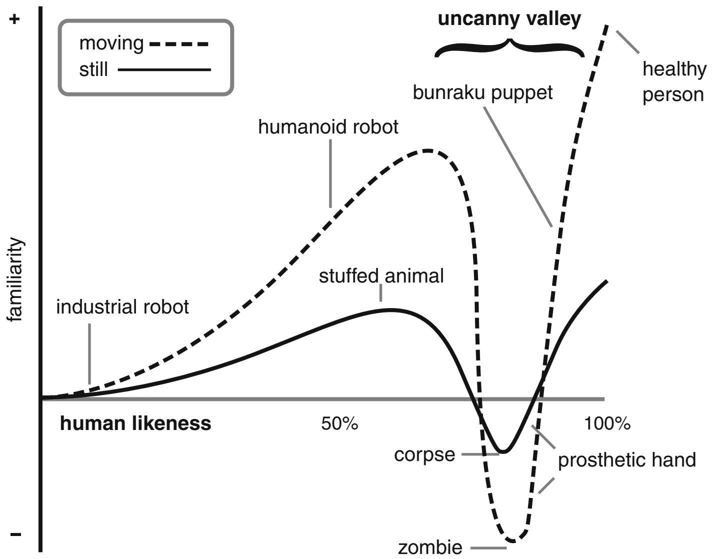


Fig. 1 The familiarity of robots does not increase with their humanness, but drops even for small imperfections when the robot is designed to look like humans (Picture from Wikipedia)

The relationship between the robot humanness and the perception of familiarity felt by a human has been introduced by Mori (1970). It is stated that familiarity of a robot increases with its similarity to humans, until a certain point where even small imperfections cause the robot to appear repulsive (Fig. 1). This sudden drop is defined as the uncanny valley.

A possible explanation may be related to the framing theory as reported by Bartneck et al. (2007). When humans face a new situation, they select from their memory a frame, i.e., a data structure representing stereotyped situations and giving expectations, anticipating what will happen next, and guiding them in case expectations are not fulfilled. According to this framework, Bartneck and colleagues posit that when humans encounter a machinelike robot, they select a "machine frame." Robot's humanness attracts their attention because they deviate from their expectations, and such deviation is positive since humans tend to like other humans. On the contrary, when an android is encountered, a "human frame" is selected and its robot-like features lead to a negative deviation. Bartneck and colleagues highlight also that it is difficult to find examples of entities in the uncanny valley, which are located in between the deepest dip and the human level. The question arises about the existence of the uncanny valley in the first place, leading to the possibility that the uncanny valley be considered more of a cliff than a valley, where robots strongly resembling humans could either fall from the cliff or they could be perceived as being human.

According to the previous discussion, we are interested in assessing how natural long-term human–robot interaction processes can be, and we would like to focus on a particular scenario. Robot actors on a stage can be very useful in assessing how natural HRI processes are. The scenario gives the possibility to work in a constrained context where the director can define with a high precision how the interaction takes place, not only from the point of view of verbal clues but also from the point of view of nonverbal communication and synchronization between actors.

The contribution of this chapter is twofold: (i) a discussion about general principles and requirements for robot design, which allow achieving a sense of believability and naturalness in the perception and observation of human–robot and robot–robot interaction processes, and (ii) an analysis of current robot architectures targeted at enforcing such requirements and presenting a real-world example about using robots on a stage.

This chapter is organized as follows. Section "What Makes a Robot Believable When Interacting with a Human? Or Do We Perceive Human–Robot Interactions as Natural?" discusses a number of issues related to robot believability. Section "Composition and Sequencing of Robot Behaviors for Human–Robot Interaction" introduces techniques for behavior composition that are suited to human–robot interaction processes. Section "Experiences with Robots on Stage" reports about experiences with an android robot on stage, whereas section "A Case Study: Creating Conversations on a Stage" discussed a number of trials. Conclusions follow.

What Makes a Robot Believable When Interacting with a Human? *Or* Do We Perceive Human–Robot Interactions as Natural?

Human Expectations About Social Robots

In the past few years, we witnessed a rapid growing of public interest and research activities in robotics, also fostered by huge investments by public bodies, funding agencies, and companies worldwide. Starting from initiatives to literally "provide each home with a robot" (e.g., refer to the Robot Companions for Citizens initiative in Europe (web: http://www.robotcompanions.eu/) or the Japanese ERATO Symbiotic Human–Robot Interaction project (web: http://www.jst.go.jp/erato/ishiguro/en/index.html)), much effort is currently devoted to provide robots with skills to easily interact with people using a variety of once-disregarded media, such as speech (Randelli et al. 2012; Bastianelli et al. 2014) as well as physical contact (Schmitz et al. 2011; Denei et al. 2015).

A few visionaries argue that robots are reaching a sort of *technological singularity* in many respects, and this may allow them to play a significant, long-term role in our everyday life, especially through the development of advanced techniques for learning, reasoning, and interacting with humans. Each communication process between humans and robots is subject to an *interaction gap*, which arises at the cognitive levels whenever information must be mutually transferred. As pointed out by Pratama and colleagues (2014, 2015), the interaction gap can be overcome by fostering a process of *mutual understanding* at various levels of the communication spectrum (Stoytchev 2009).

In the literature, a number of studies investigate the expectations that people have when interacting with social robots. Such expectations are characterized by a complex variability depending on age, sex, robot's shape (i.e., how the robot walks or gazes at the observing human), as well as on specific conditions (e.g., normally developed children versus children affected by autism) and personal beliefs of people and, eventually, culture (e.g., European versus Japanese culture). Such studies are characterized by different aims: this first category aims at assessing generic qualities a robot must have in order to be classified as *socially aware agents*, whereas the second category defines psychological scales which may be used by humans when evaluating robots. However, both categories provide a number of insights for the design of social robots. In this chapter, we focus only on the most relevant contributions for each category.

Robots as socially aware agents. This class aims at identifying general principles that humans deem necessary to provide social robots with believability and intentionality traits.

Ishiguro et al. (2012) describe their experience with the Telenoid robot. Telenoid is a tele-presence robot designed to resemble a *minimal human being* and expected to elicit intimacy. Such design choice is motivated by studies by Kanda and colleagues (2008), where an initial principled study about the robot's appearance in human–robot interaction processes is discussed. Telenoid can move its small arms, face, and eyes, and it is endowed with a speaker and a microphone to allow a remote operator to interact with the person in physical contact with it. In the study, more than 100 people without any prior experience with humanoid robots interact with Telenoid, both in free-form and task-targeted scenarios. The free-form scenario is aimed at creating natural conditions for human–robot interaction. After the sessions with Telenoid, subjects are required to fill in a questionnaire aimed at assessing two social dimensions of the interaction, namely, how they *perceive* Telenoid and its *believability* as an interaction medium. The main finding of the study is that humans report to have a clear feeling about them *sharing the space* with the robot, therefore assigning it with intentionality and presence and, as a consequence, believability.

Shen et al. (2009, 2011) conduct a study to assess the interference effect in humans when interacting with the KASPAR2 humanoid robot. As firstly argued by Oztop et al. (2005a) and better described in section "The Role of Anthropomorphism in Robot Design," motor resonance has been proposed to study human–robot interaction (Marin et al. 2009). Motor resonance can take the form of two contradicting effects: on the one hand, *motor coordination* between any two agents can be regarded as a behavioral manifestation of a social relationship, i.e., a desirable feature for a successful interaction (Richardson et al. 2005; Schmidt and Richardson 2008); on the other hand, the *interference effect* (i.e., the effect of being *distracted* when executing a motor task while observing a contrasting task executed by another agent) has been used in various studies to validate subconscious-level perception of robot appearance (Oztop et al. 2005a). The hypothesis by Shen and colleagues is that the attribution of social awareness to robots depends on both bottom-up (i.e., robot appearance and bioinspired robot motions) and top-down (i.e., a priori human beliefs about robot's agency and engagement) aspects of human perception, a concept synthetized as *overall perception*. Only robots characterized by a higher level of overall perception are regarded as social entities.

As discussed by Sciutti et al. (2014), normally developed children's opinions about robots are very peculiar and change as a result of age and the possibility to interact with robots themselves. Starting from a number of previous studies, which relate human age and acceptability of robots as companions and social agents (Broadbent el al. 2009; Beran et al. 2011; Wu et al. 2012; Diehl et al. 2012), Sciutti and colleagues exposed children of different ages to a number of tasks carried out by the iCub robot (Tsagarakis et al. 2007) during the Genoa Science Festival (web: http://www.festivalscienza.eu) and asked them to fill in open questionnaires *before* and *after* the interaction. Examples of interaction tasks include witnessing the robot in whole-body and fine hand movements, experiencing different robot control schemes by physically interacting with it through large-scale robot skin (Maiolino et al. 2013; Anghinolfi et al. 2013), seeing how the robot detects and tracks their faces and showing the robot a red ball that is tracked by gaze and head motions. The main findings of this study are that: (i) *appearance* is less important for children than it is for adults, (ii) robot *skills* and overall *capabilities* are progressively more important from children to adults, and (iii) after children can interact with a robot, they are more aware of the importance of skills and capabilities, including dexterous motions (e.g., fine manipulation).

Psychological scales to classify robots. This class is aimed not only at identifying principles governing human qualitative perception of robots in social settings but at identifying principled psychological scales to classify robots according to a number of parameters defined by humans.

Kamide et al. (2012) focus on how ordinary people of various ages perceive 11 different humanoid robots. After having seen videos where humanoid robots perform everyday tasks and interact with people, subjects are required to fill in open questionnaires, which are then evaluated by a team of psychologists. Employed robots belong to three classes, namely, wheeled humanoids (i.e., Robovie, Wakamaru, and Enon), walking humanoids (i.e., ASIMO, HRP-2, HRP-4C, and Neony), and androids (i.e., CB2, Repliee Q2, Geminoid, Geminoid F). Kamide and colleagues build up from previous work on the development of psychological scales, e.g., the *robot anxiety* scale (i.e., the anxiety preventing an individual from interacting with a robot) by Nomura et al. (2006), and the studies employing the so-called *semantic differential* method to evaluate perceptions among humans by Lorr (1965) and Rosemberg et al. (1968). The team of psychologists identified nine fundamental traits, namely, familiarity (i.e., general feeling about robot design), repulsion, utility (including a utility-cost trade-off), performance (with respect to interaction and intelligence), motion, sound (including the robot's voice), humanness, emotion (i.e., facial expressions, gaze, and lip-syncing), and robotness. From these traits, three macro-groups are derived, namely, *familiarity*, *utility,* and *humanness*.

Kamide et al. (2012) extend their previous work (2012) in order to analyze the classification obtained using the developed psychological scale. The authors focus on the effects of age, sex, and robot types. The macro-groups identified by Kamide et al. (2012) are considered. Key findings are that (i) older females attribute to (both wheeled and biped) humanoids a higher familiarity score with respect to other population groups, (ii) adolescents feel a higher level of familiarity with wheeled humanoids than with bipeds or androids, (iii) middle-aged females and old-aged males consider robot utility more important than it is for younger populations, and (iv) independently of age, androids have high scores on humanness, but a low utility, and they are considered as less familiar.

From this general analysis about human expectations from social robots, it is possible to derive a few general comments, which will guide us through the remaining of this chapter.

Both anthropomorphism and human beliefs are important factors. While the initial work by Oztop et al. (2005a) stresses the importance of humanlike shape and biologically plausible motions, the later work by Shen and colleagues (2009, 2011) integrates high-level human beliefs. As a consequence, both low-level (i.e., perceptual) mechanisms and high-level (i.e., cognitive) processes seem to be involved.

Robot engagement elicits believability. In all experiments and in real life, naïve subjects and ordinary people, as well as people acquainted with robots, are subject to cognitive biases (Kanda et al. 2001). When interacting with robots, a priori assumptions about their engagement may elicit an instance of what Dennett (1987) calls *intentional stance*, i.e., the attribution of mental states to (possibly complex) objects. As discussed by Shen et al. (2009, 2011), believability of KASPAR2 is elicited by letting the interacting human assume that the robot is engaged in the play scenario (i.e., by *human instructions*). Furthermore, it has been discussed how humans tend to attribute believability (through a positive reaction to interference tests) also to moving dots on a screen, if they are told that their motions originate from human movements (Stanley et al. 2007; Bouquet et al. 2007; Kilner et al. 2007).

Human age plays a significant role. Different population groups tend to attribute believability to robots according to their own scales. The study by Sciutti et al. (2014) tells us that the importance of perceptual (e.g., related to robot morphology and biologically plausible behaviors) and cognitive (e.g., engagement, function) aspects is modulated by age. Although this is not surprising per se, given that today's children are immersed in technology, this factor has profound implications on robots' acceptability and believability.

Not all robots are equal. Different population groups (in terms of age, sex, and cultural background) tend to perceive robots in different ways. Kamide and colleagues (2012) argue that humans attribute different scores (as far as familiarity, utility, and humanness are concerned) to different classes of humanoid robots. Again, according to the target human group a robot may interact with, its shape may prove to be fundamental to elicit believability.

The Role of Anthropomorphism in Robot Design

In the past few years, robotics researchers adopted tests from cognitive psychology (Sciutti et al. 2012b) to formally define measures and benchmarks to quantify the degree of the intrinsic perception of a robot as a human. As posited by Oztop et al. (2005a), it is necessary to understand “whether a humanoid robot can be treated as a human by the perceptual system of a human observer.” The focus on the perceptual system by Oztop and colleagues plays a significant role for our analysis. It allows isolating specific parts of the perception chain, but – we anticipate – is no longer valid to address Q2, where high-level and social aspects are important.

Tests are based (in a way or another) on motor resonance and the companion concept of motor interference. Both these concepts are tightly related to the well-known discovery of mirror neurons (MNs) and the existence of sensorimotor loops (Rizzolatti et al. 2001; Gutemberg and Yiannis 2007; Wykowska and Schubö 2012). They imply that observing and doing share many similarities and even parts of the brain. As reported by Oztop and colleagues, in macaque monkeys, MNs have been discovered in the ventral premotor and the parietal cortices, which are tightly connected. Similar connections have been found in the human brain (Wilson and Knoblich 2005). Motor resonance plays a central role in adaptation and survival, for it allows imitating an observed behavior (e.g., a gesture), either consciously or unconsciously, therefore enabling the ability to understand others (Hamilton and Grafton 2007).

As argued by Demiris and Hayes (2002), an observed behavior can be imitated in two ways. If the behavior is known, *active imitation* is enabled. In this case, the motor system is directly activated during perception (in a forward fashion). If the behavior is unknown, it cannot be imitated directly: novel motor knowledge and the associated representation must be created first and then passive imitation (which involves goal-oriented inverse models) is possible.

In the literature, first results by Kilner et al. (2003) are contradictory. They employ a robot manipulator to verify whether interference occurs in arm-waving motions. On the one hand, an interference effect is found in case of incongruent motions (which are detected as an increase of variance in trajectories). On the other hand, the opposite is *not* detected, i.e., no facilitation effect is observed for congruent motions.

Oztop et al. (2005a) perform the same set of experiments, but using a humanoid robot whose motions reproduce trajectories generated by humans. They show that incongruent motions are characterized by a higher variability than congruent motions, in a way similar to what happens with humans. They conclude that both humanoid form and biologically plausible motions play a fundamental role in eliciting motor interference.

Shen et al. (2009) perform the same class of experiments, but in a more unconstrained scenario and using the KASPAR2 robot. Both children and adults must synchronize with the robot’s motions. KASPAR2 executes both horizontal and vertical arm-waving gestures, and it follows a well-known nursery rhyme. Differently from the humanoid robot used in the experiments by Oztop and colleagues, which is still “mechanically looking,” KASPAR2 is very close to fall into Mori’s uncanny valley (1970). Results do not show interference effects. However, most of the participants are affected by the robot’s behavioral rhythms, although children and adults behave differently: the children’s rhythm associated with arm waving is strongly synchronized with robot’s rhythms, whereas adult rhythms are synchronized only in part. They explain such a difference by recurring to the playful nature of the interaction, which is expected to engage children. This seems consistent with the findings by Robins et al. (2004) and Woods et al. (2004), which report studies suggesting evidence about the fact that children prefer "mechanically looking" robots with respect to humanlike robots. Overall, a behavioral adaptation of people's motions to the robot can be observed.

Wykowska et al. (2012) report about similar experiments, but with the aim of investigating *congruence effects*. These effects are related to better performance in perceptual tasks when a concurrent motion is congruent with the dimension of the target. The main findings are that congruency is observed in the first place and it is independent from the perceptual cue.

Considered altogether, these studies reveal that at least two elements contribute to Q1: the first is related to the robot's shape, which must be at the *proper* level of human resemblance; the second has to do with appropriate motions, which must be biologically plausible.

A number of approaches in the literature consider the study of a robot's shape in a relevant perspective for our purposes.

Chaminade et al. (2007) posit that it is not human resemblance per se, but its connection with internal simulations in the observer, which elicits the attribution of agency to a robot. They discuss a study where participants are asked to label as "biological" or "artificial" the movements of a series of computer-generated characters. Motions have been obtained in two ways: the first involves human actors, whose gestures have been captured by a motion capture system, and therefore effectively biological; the second is entirely generated by a computer using a traditional key-frame animation algorithm. The actual shape of computer-generated characters spans different levels of anthropomorphism. The main finding of this study is that the correct attribution of biological motions negatively correlates with an increase in the anthropomorphic level of characters, thereby confirming the role played by the uncanny valley effect.

Press et al. (2007) aim at linking motor interference with sensorimotor experience. They set up an experiment where human imitation of both robot and human (either congruent or noncongruent) actions is assessed. This is done *before* and *after* a training phase. Training involves reinforcement of motor resonance with congruent and noncongruent movements. They show that training increases the likelihood of correctly imitating robot-based stimuli, in terms of a reduced delay. They conclude that the mirror neuron system is coupled with sensorimotor experience and postulate that the bias toward the attribution of agency to agents with humanoid shape is biased given the abundance of human examples in everyday life.

In the comparative study about different humanoid robots (including androids) carried out by Kamide et al. (2012), the role of robot shape as pointed out by Chaminade and colleagues is not fully confirmed. In particular, the study reports analyses about the attribution of humanness depending on the sex and the age of participants. Surprisingly, androids are ranked the best in this scale (although they are very close to the uncanny valley), independently of age and sex variations, except for older individuals, who prefer mechanical-looking robots.

From these studies, and recalling the work by Shen et al. (2011), we can devise a few principles contributing to Q1.

Attribution of agency based on shape has limits imposed by the uncanny valley. All the studies agree that the proper level of anthropomorphism is required to elicit motor resonance and mental imitation. When we say "proper," two factors must be considered. The first is related to the fact that *an archetype model for agency exists*, *which is constituted by the human itself*. This probably originates a bias as pointed out by Press et al. (2007). The second is that the correct *attribution of agency correlates with age*, which is also in agreement with the studies by Sciutti et al. (2014).

A few studies investigate the role of biologically plausible motions in human–robot interaction. The most relevant for our purposes are briefly reported in the following paragraphs.

Kilner et al. (2003) asked subjects to perform arm-waving motions while observing a video of either a human performing an arm movement or a ball moving across the screen. Both human and ball videos contained either biological (i.e., minimum jerk) or nonbiological (i.e., constant velocity) motions. The executed and observed arm motions were either congruent (i.e., same direction) or noncongruent (i.e., tangential direction) with each other. Results show that observed motions are processed differently according to whether they are made by the human or the ball. For ball videos, both biological and nonbiological noncongruent motions interfere with executed arm motions. In contrast, for human videos, the motion velocity profile is the critical factor: only noncongruent, biologically plausible human motions interfere with executed arm motions. According to these results, the authors propose that the interference effect be due either to the information the brain has about different types of motion stimuli or to the impact of a priori experience with different types of form and motion.

Human movements comply with characteristic kinematic laws of motion. According to previous studies, low-level motion perception is biased toward stimuli complying with these laws and specifically with those modulating the responses of neuronal circuits also involved in action recognition predominantly located in the left frontal lobe (Casile et al. 2010).

Georgescu et al. (2014) highlight how kinematic properties of perceived motions play a key role in communicating meaningful information about social goals of other people. Two neural networks circuitry appear to be involved in nonverbal behavior understanding: the action–observation network (AON) and the social neural network (SNN). Georgescu and colleagues present also a study determining the different contributions of the two networks for understanding dyadic social interactions through the display of virtual characters. These are parameterized with respect to motion fluency and gestures. Experimental results show that both fluency and contingency significantly influence the experience of the animations and that AON is preferentially engaged when processing contingent motion patterns. Results do not discriminate between different degrees of motion fluency. In contrast, regions in SNN are more strongly engaged when observing dyads with disturbed movement fluency.

Noy et al. (2009) highlight the importance of kinematics with respect to similarity in visual appearance. They discuss how the brain is responsive to the motions of an abstract visual stimulus moving in a biologically plausible way. However, kinematic features seem to be a sufficient, but not necessary, condition for the automatic mapping of perceptions to actions.

According to the discussed studies, we conclude that both motion velocity and kinematics play an important role with respect to visual appearance in eliciting motor resonance. However, other factors seem to be involved as well.

Observing a Human and a Robot, or Two Robots, Interacting with Each Other

From the analysis carried out in the previous section, it is evident how the attribution of agency to robots, and in general a sense of believability in the direct interaction with robots, transcends the pure robot physical shape and behavior, although these aspects are fundamental. As pointed out by Shen et al. (2011), an overall perception of the robot's shape, biologically plausible motion, observer's beliefs and culture, and social behavior is of the utmost importance to generate believability. An appropriate match between a robot's appearance and its social function may facilitate human acceptance and cooperation in the interaction (Goetz et al. 2003).

This observation is not new. In Braitenberg's *Vehicles* (1984), the author describes a set of simple robots which, using hard-coded sensorimotor loops, are able to achieve credible behaviors such that an external observer may consider them *intelligent*. The discrepancy between simple design and exhibited intelligence is considered a prototypical example of the fact that "intelligence resides in the eyes of the observer." The overall perception of the vehicles cannot but convey a sense of complexity that is only *imagined*.

When interacting with a robot, a number of cues (e.g., shape, biological plausibility of motion, beliefs about the robot's engagement) contribute to the attribution of a sense of agency to the robot. Subject to certain limitations, we have seen that it is possible to provide a positive answer to Q1. However, what can be said about Q2? Is it possible to adopt the same parameters used for Q1?

We believe that different aspects are involved when judging the believability of characters interacting with each other, or with other humans. In this case, there is no direct interaction with the observer. Additional social aspects play a fundamental role.

Observing actors on a stage defines a social scenario where the audience cannot but assume the actors behave as goal-oriented agents. Actors have well-defined goals (e.g., concluding the scene, executing a given gesture, or moving to a certain location on the stage), which are set by the director very precisely, as discussed above and described also in section "A Case Study: Creating Conversations on a Stage."

Interaction patterns between humans and robots have been employed to evaluate how humans attribute social behaviors to robots, as discussed by Kahn et al. (2010a, b). Surprisingly, with the sole partial exception of Mastrogiovanni and Sgorbissa (2013), we did not find any specific study aimed at investigating Q2, i.e., determining how an interaction between a human and a robot (or between two robots) is perceived by another human watching the scene. However, we try here to trace a possible pathway to that direction.

Our proposal is to extend cognitive psychology tests to encompass the social sphere of interaction. Two studies in the literature seem to confirm the feasibility of this direction, the first related to the possible use of MNs to predict intensions through internal simulation, the second linking action representation to empathic aspects.

Gallese and Goldman (1998) assume that MNs may enable an organism to estimate (otherwise inaccessible) mental states of observed conspecific organisms, as well as to attribute them a propositional attitude. The link between the perception of a movement and the attribution of intentionality is the presence of a goal. If a goal is present (i.e., the agent shows a certain level of engagement, or the observer believes the agent is engaged, as posited by Shen et al. 2011), then one can interpret a movement as a consequence of a mental state. Gallese and Goldman assume that this process underlies a simple form of *mind reading*.

In the literature, two main approaches to mind reading have been put forward, namely, the "theory theory" (TT) and the "simulation theory" (ST) (Davies and Stone 1995a, b; Carruthers and Smith 1996). According to TT, mind reading can be achieved by developing a *common sense* theory of how the mind works, i.e., by constructing a *folk psychology corpus*. This is done through a set of causal/explanatory rules relating external stimuli, unobservable internal states, and actual observable behavior. On the other hand, ST posits that internal simulation mechanisms are activated to estimate and predict other's mental states. The use of internal simulation implies that no rule must be learned. In other terms, ST assumes that an attempt to mimic the mental states of the observed individual is done through simulation (Goldman 2002).

Interestingly enough, once internal simulation is activated, it can be used not only to mimic the current mental states of the observed agent but also to predict and estimate future and past mental states, i.e., to determine what mental states may occur or have occurred already. When an action is observed, internal simulation allows one to ask what is the goal to be achieved, which the executed action contributes to. This may lead to assume a specific goal and then mentally simulate an action path to achieve the goal involving the observed action. At this point, the attribution of agency to the robot is assessed.

This goal-oriented internal simulation process can be grounded to neurophysiological studies (Carr et al. 2003). In the brain, it has been found that the superior temporal cortex may be involved in an early representation of an observed action (Perrett et al. 1989; Perrett and Emery 1994). This representation is shared with MNs in the posterior parietal cortex, which codes all the kinesthetic aspects of the observed motion (Lacquaniti et al. 1995). The representation of kinesthetic aspects is transmitted to MNs in the inferior frontal cortex. The inferior frontal cortex is believed to encode the goal of the action (Kohler et al. 2002). Such motor plans are transmitted back from parietal and frontal areas to the superior temporal cortex

(Iacoboni et al. 2001). Here, a match is attempted between the early description of the action and predicted sensory consequences of the planned imitative action, which are based on the goal. If such a match exists, internal simulation (i.e., an imitation) occurs.

The work by Carr et al. (2003) tries to link such a simulation mechanism to the onset of empathy in the observer. They want to show the relationships between the internal representation of (observed) behaviors and their role in modulating inner experiences, as well as the attribution of goals. They argue that MNs in the superior temporal and inferior frontal cortices may link motor representation to the limbic system, which processes emotional content (Augustine 1996). Furthermore, they hypothesize that we understand what other people feel through action representation, an effect also known as *chameleon effect*, through the unconscious mimicry of others (Gallese 2001). In the study by Carr and colleagues, participants are asked to simply observe or imitate emotional facial expressions, while being subject to functional MRI imaging. Their hypothesis, which is confirmed, is that observation and imitation of facial expressions should yield to comparable patterns of activated brain areas, with an increased activity during imitation, in premotor areas, the inferior frontal cortex, the superior temporal cortex, the insula, and limbic areas.

Summarizing, we believe that Q2 cannot be satisfactorily answered at this point. We are confident that extending previous studies can lead to a definitive assessment of such a question. To this aim, we need to extend the notion of motor resonance to that of *social resonance*.

Composition and Sequencing of Robot Behaviors for Human–Robot Interaction

In section "What Makes a Robot Believable When Interacting with a Human? Or Do We Perceive Human–Robot Interactions as Natural?," it has been observed that believability and the attribution of agency to robots depend on many factors, both at the perception and social levels. In this section, we focus on techniques dealing with the composition of basic robot behaviors, with the aim of devising guidelines for robots having to interact with humans in a credible manner, as well as to act in front of humans.

It is noteworthy that this requires to address a number of diverse and hard intertwined subproblems. On the one hand, it is necessary to choose the most appropriate behaviors and coordinate them in real time (Mastrogiovanni and Sgorbissa 2013). The selection may depend on the "robot mental state" (in terms of interaction goals or communication message conveyed through gestures) and on unexpected events (including the interaction with humans and other robots). On the other hand, proper mechanisms to integrate deliberative and reactive behaviors in a timely manner must be devised.

In order to ground the aspects of human–robot interaction outlined in section "What Makes a Robot Believable When Interacting with a Human? Or Do We Perceive Human–Robot Interactions as Natural?," the need arises to design an effective SW architecture for social robots able of: (i) sequencing behaviors, (ii) composing sequences of behaviors, and(iii) allowing for simple forms of variations or improvisation. Different relevant examples of SW architectures explicitly dealing with behavior composition are discussed in the literature. In the following paragraphs, we focus on methods which best resemble the possibilities that a director has when setting up a script with human actors, i.e., logic-based methods.

Among them, a set of approaches aims at precisely arranging behaviors in graph-like structures, which implicitly define all the possible sequences in a fairly rigid way.

Hart and Grupen (2011) propose a control architecture in which behaviors are based on low-level "potential" functions. Behaviors are activated using a four-value grammar: functions (i.e., Skolem symbols) are used to represent predicates that, if verified, correspond to successful behavior executions. Unfortunately, explicit methods to manage temporal constraints or behavior sequencing strategies are not properly addressed. Jenkins and Matarić (2004) introduce the so-called performance-derived behavior vocabulary (PDBV) architecture, in which different behaviors are acquired from human demonstrations, in order to enforce the believability of robot motions. Therefore, the architecture is a de facto library of behaviors that can be either primitive or composed on top of simpler primitive ones. Composed behaviors can be represented as a graph where nodes are primitive behaviors, and edges (i.e., sequences) are learnt as probabilistic models (first-order Markov models are used). Behavior sequences are defined a priori. The work by Knoop et al. (2008) introduces the concept of flexible programs (FPs), which are task compositions based on hierarchical task networks. The basic representation unit is the elementary operator (EO). EOs can be grouped to form macro-operators (MOs), which are represented as predefined sequences of EOs. A task precedence graph (TPG) is used to encode all the precedence relationships between MOs and EOs. The TPG is made up of a number of precedence relationships. Unfortunately, it is not possible to form complex compositions of behaviors (e.g., including alternative sequences or temporal constraints). Lim et al. (2008) describe an SW architecture explicitly focused on behavior composition for teams of small humanoid robots. A number of predefined motion primitives are available: these can be either atomic (e.g., "walk," "pull," or "push") or complex (e.g., "pick an object" or "close a door"). Behavior sequencing is specified using finite state diagrams. Eskridge and Hougen (2010) propose a system, called multiple composite levels (MCL), which allows for the modulated composition of fuzzy behaviors structured in hierarchies. Behavior sequencing is fixed; however behavior modulation is learnt using reinforcement learning techniques. Since the focus is on behavior learning, issues related to complex behavior composition are not investigated. Sadeghipour and Kopp (2014) propose a system able to learn a motor grammar based on gestures. Their system is capable of learning both basic motion primitives and their hierarchical organization.

Interestingly enough, neurobiological studies seem to suggest that the human brain exploits a graph-like structure to decompose complex motions in simpler ones. Each motion can be regarded as a motion primitive. During execution, the brain generates complex motions both parallelizing and sequencing basic motions

(Flash and Hochner 2005; Gutemberg and Yiannis 2007). Furthermore, each internal motor representation seems to be modeled separately in the motor cortex.

Other approaches aim at pushing robot's autonomy up to the point in which robots can get *out of control*.

The work by Fujita and Kitano (1998), which introduced the term "entertainment robot," presents a pet robot named MUTANT (later commercialized by Sony as the Aibo robot). The robot is a fully actuated quadruped, provided with a microcamera, a stereo microphone, as well as touch sensors. It is controlled using a three-layered control architecture for the generation and execution of highly complex behaviors by merging social skills with high-level cognition capabilities. Fujita et al. (2003) extend this approach and describe a similar architecture to control the small humanoid robot SDR-3X (commercialized by Sony as the Qrio robot). The robot is equipped with a number of sensors to interact with the environment. Behavior generation is obtained through an internal homeostatic regulation mechanism, which selects and composes behaviors to keep internal parameters within proper ranges in the presence of external and internal stimuli. The approach proposed by Petterson and Wahde (2005) relies on utility functions. At any given time, the system activates the (set of) behaviors with the largest utility. Utility functions are evolved artificially by controlling the robot for a certain period of time and then by assigning a fitness value based on the robot's overall performance.

The approaches described by Fujita and Kitano (1998), Fujita et al. (2003), and Petterson and Wahde (2005) deal with fundamental issues in human–robot interaction and autonomous behavior organization. However, they are only partially related to our study. We do not aim at studying robots that behave spontaneously in everyday situations. Our focus is on believable characters like actors performing on a stage, where everything must be under control, a direct interaction with the audience is almost null, and – in general – unforeseen events are reduced to a minimum.

Finally, a few SW architectures explicitly *mimicking the role of motor resonance* have been proposed as well. We focus on those architectures stressing the role of prediction and internal simulations (Oztop et al. 2005a).

MOSAIC is an SW architecture composed of multiple forward–inverse model pairs, which can be used for control and behavior recognition (Wolpert and Kawato 1998; Haruno et al. 2001; Wolpert et al. 2003). At runtime, pairs compete for the overall control of the robot. Pairs characterized by better predicting forward models become more influential in control using the corresponding inverse models. When an external behavior is observed, forward models contribute in this way to select the best controller. MOSAIC allows a robot to implement a simple form of motor interference. The observation of an action biases the control, since the influence of pairs controlling congruent motions is higher, whereas the influence of pairs controlling noncongruent motions is lower.

Sadeghipour and Kopp (2009) describe a system based on a probabilistic model for gesture detection and motor resonance. Forward and inverse models correspond to active and passive imitations. The interesting assumption is the multilayer organization, where motor commands, programs, and schemata can resonate.

The so-called mental state inference model models MNs as an exclusive forward prediction system (Oztop et al. 2005b). During behavior execution, the efficient forward model deletes delays possibly originating from a planning phase, whereas when observing behaviors, the forward model acts as an internal simulation loop used to determine the goal associated with the observed action (in accordance with Carr et al. 2003 and Gallese 2001). Observing an action activates an internal simulation mechanism, which causes a conflict between the perceived and the currently executed behavior, if they are not congruent.

Kahl and Kopp (2015) describe a system integrating mirroring and internal simulation processes. The authors focus on how it is possible to attribute mental states during a situated communication, such as a handshake coupled with a social gaze. Unfortunately, preliminary results are only available in simulation.

A part from a few approaches, current SW architectures for robots do not take explicitly into account results from cognitive psychology studies related to believability. As a consequence, much research activities are expected to this long-term objective.

Experiences with Robots on Stage

In this section, we report on one of the first long-term studies concerning human–robot interaction in a setting that can be considered both experimental and natural: a theatrical play. During a theatrical rehearsal, the director and actors reconstruct a social interaction by focusing on small behaviors that are repeated and incrementally transformed. This is a natural environment, as the rehearsal would occur if the researcher were not present (see the difference between hypothetico-deductive and empirico-inductive approaches below).

Theater is relevant to HRI studies because participants deal with the problem of integrating a robot in interactions with humans. That is, behaviors are created in which a robot must coordinate with humans.

Interaction Science: Beyond Android Science

Ishiguro (2006) describes "android science" (AndSci) as a field that contributes to our understanding of human beings by using androids in test beds to validate concepts developed by cognitive science and psychology. Several projects based on AndSci have attempted to create a humanlike appearance and behavior in androids. However, the approach used can only be employed for the analysis of short-term human–android interactions, such as a first encounter with a guide robot in a museum setting (Kuno et al. 2007).

As illustrated by the examples of Searle's Chinese room and ELIZA in classical artificial intelligence (AI), it is evident that computers can successfully keep up ongoing interaction with humans for some time. Studies in AI have demonstrated that humans tend to believe that the partner behind the computer is a human being, without questioning this fact. As discussed in section "Composition and Sequencing of Robot Behaviors for Human-Robot Interaction," to build a face-to-face environment in which a social robot can maintain interaction with humans, with their own bodies over a long period of time, robot engineers need to focus not only on appearance and individual behaviors but also on determining the appropriate sequential structure of interaction.

Prolonged sequences of actions between humans and robots have been investigated in a robot-theater project. During a theatrical performance, it is essential that human actors and androids can maintain a long-term interaction. Participants (engineers, the director, and actors) face the problem of coordinating long sequences of interaction between humans and robots.

Our approach is grounded on a field of research focusing on naturally occurring interaction (Atkinson and Heritage 1984). Ethnomethodology (Garfinkel 1967), conversation analysis (Sacks et al. 1974; Schegloff 2007), and multimodal interaction studies (Goodwin 2000a, b; Mondada 2006) focus on methods by which members of society continually produce and coordinate intelligible actions. These studies show that phenomena usually described as cognitive processes can be redefined as interactional accomplishments. In any activity, what an individual understands and perceives is not something located in the brain. It is, rather, the result of a sequential coordination of actions with another human during social interaction (Goodwin 2013).

Ethnomethodology (EM) considers that the members of society organize rational and intelligible actions by selecting relevant social expectancies at each moment of their interaction. Most of the time, these expectancies are "seen but unnoticed." However, when problems occur, members can describe these expectancies and change them if needed. The goal of EM is then to describe how participants to any activity produce mutual understanding by activating relevant social knowledge. Conversation analysis (CA) extends EM investigations by video-recording naturally occurring interactions. CA focuses on the temporal and sequential unfolding of social interaction and shows that mutual understanding is a social and sequential phenomenon. It is important to understand that mutual understanding in naturally occurring interaction is the focus of sequence organization. For instance, any speaker in a *turn-at-talk* manifests his actual understanding of the previous turn-at-talk of the previous speaker (Schegloff 2007). In their reply, the first speaker can confirm or invalidate the way the second speaker has understood their original comment. This process occurs in many social interactions. But the analysis of sequence organization does not only focus on speech: any embodied behavior can be examined following the same method as that shown by numerous studies about the multimodality of social interaction (Kendon 2004; Goodwin 2000a, b; Mondada 2006).

To summarize, the goal of EMCA research is to understand any kind of human activity by analyzing the methods that participants use to accomplish their tasks. Analysis is done at the sequential level of interaction by examining how participants coordinate multimodal resources. These studies show that humans can coordinate complex courses of action with great accuracy by combining speech, gaze, gesture, body movements, and object manipulation. Recent studies on multi-activity show, furthermore, that humans can accomplish more than one activity at a time

(Haddington et al. 2014). Taking the complexity of human action into account can only benefit AndSci if researchers can conceive and produce social robots that can interact with humans. The term "interaction science" refers then to the application of EMCA to AndSci.

While AndSci focuses on creating the robot's appearance and behavior, "interaction science" (IntSci, as defined in section "Introduction") focuses on the creation of an appropriate sequential organization of interaction between humans and robots. While AndSci describes how humans perceive robots in the fields of cognitive science and psychology, IntSci describes how human perception of the robot behavior is constructed during social interaction and becomes a resource for maintaining the interaction over an extended period of time. In the next section, we specify an important difference between AndSci and IntSci.

Hypothetico-deductive Approach Versus Empirico-inductive Approach

Two approaches are available to researchers investigating human behavior: the hypothetico-deductive approach and the empirico-inductive approach. Hypothetico-deductive analysis establishes hypotheses to explain particular behaviors that are confirmed or invalidated through experiments. This approach is used by AndSci. In empirico-inductive analysis, researchers observe behavior without preconceived hypotheses, looking for repeat occurrences in social interaction. Through comparison of systematic recurrent behaviors, researchers describe what is systematic in the organization of participant behaviors in a naturally occurring interaction. This is the approach used in EMCA and IntSci.

According to EMCA, the problem with the hypothetico-deductive approach is that the scope of observation is reduced:

> Experimental procedures are generally successful to the extent that, through experimental manipulation, behavioural variation is limited to those aspects selected for investigation under controlled conditions. (. . .) Without previous exposure to a range of naturally occurring interactional data, the experimenter is unlikely to anticipate the range, scope, and variety of behavioural variation that might be responsive to experimental manipulation, nor will he or she be in position to extrapolate from experimental findings to real situations of conduct. (. . .) The most economical procedure, therefore, has been to work on naturally occurring materials from the outset. Naturally occurring interaction presents an immense range of circumstances – effectively amounting to a "natural laboratory" – for the pursuit of hunches and the investigation of the limits of particular formulations by systematic comparison. (Heritage and Atkinson 1984)

EMCA studies adopt the empirico-inductive approach to describe human behavior:

> The use of recorded data [of naturally occurring interaction] (. . .) exposes the observer to a wide range of interactional materials and circumstances and also provides some guarantee that analytic conclusions will not arise as artifacts of intuitive idiosyncracy, (. . .) or

> experimental design. (. . .) Each recording necessarily preserves a very substantial range of interactional phenomena, and an initial noticing of any one of these can motivate a search through other data for similar occurrences which were previously overlooked or the significance of which had remained unrecognized. Such a research process is possible only by virtue of the fact that the data-collection procedure is not constrained by a specific research design, with reference to some particular hypothesis. (Heritage and Atkinson 1984, p. 4)

The empirico-inductive approach allows the researcher to identify the practical solutions that participants use to solve their problems. In the case of our study, we investigate how participants that interact with a robot on stage find practical solutions to problems.

Current HRI studies adopting the hypothetico-deductive approach focus on creating believable interactions between humans and social robots. Real-world issues are frequently faced because robots do not always operate according to the hypotheses of their engineers and designers. However, if these problems are considered from an empirico-inductive perspective, the naturally occurring interactions in the Robot Theater Project provide an opportunity to discuss how humans solve problems to facilitate ongoing interaction with robots, e.g., by considering the effect of a robot's unexpected reaction (Gehle et al. 2015).

To collect beneficial insights into the design of social robots, we carried out fieldwork and collected video recordings of naturally occurring interactions in a rehearsal setting to study the behavior and response of a robot acting on stage with human actors in the context of an android–human theater project (The Robot Theater Project, Osaka University et al. 2010). Research in social robotics has mostly been conducted in carefully designed experimental settings. Recently, robot engineers have considered settings in which robots and humans have real interactions. Collaborative fieldwork with sociologists and anthropologists in the real world, e.g., museums (Kuzuoka et al. 2010), has provided valuable information about the capabilities of social robots developed in research labs worldwide.

Figure 2 summarizes the two methods for approaching a real-world human–robot interaction scenario.

In the hypothetico-deductive approach, researchers design and conduct an experiment, and individuals participate and contribute to the experiment. The results are predictable within the framework of the experiment. The researcher controls the tasks, the methods, and the results. Studies based on hypothesis testing are typically performed by engineers, robot designers, and participants in the experiment.

One major difference between the two approaches for the field of HRI is whether targeted practices to be implemented in the robot would have occurred in the real world without the researchers. As shown in Fig. 2, the human–robot interaction of the empirico-inductive approach would have occurred also in the absence of the researcher. Researchers observe the naturally occurring construction of a human–robot interaction; they accept the presented scenario and do not contribute to, nor influence, the field conditions.

The interactional practices observed in the Robot Theater Project occur on stage where unexpected events, such as a robot's unexpected actions, digressions from lines in the script, or addition of long pauses for an interlocutor's line, are repaired by

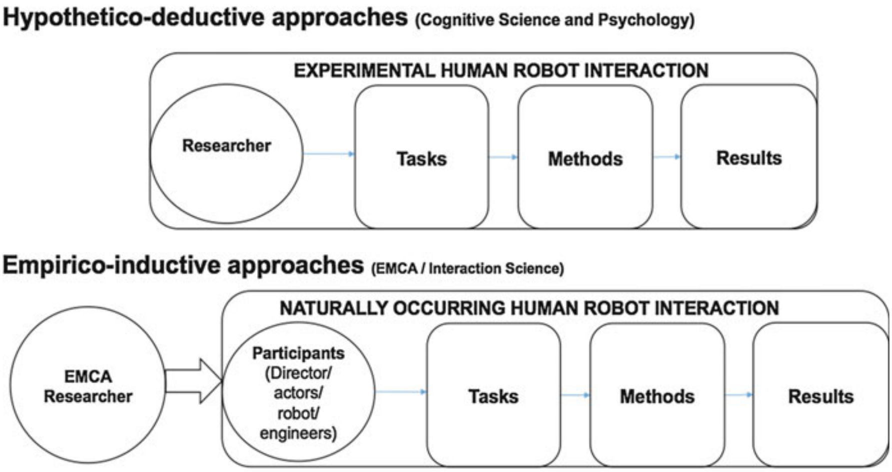


Fig. 2 Two methods to approach real-world scenarios: hypothetico-deductive and empirico-inductive approaches

participants (engineers, playwrights, and human actors) because they would be, otherwise, unintelligible to an audience. To avoid such situations, several rehearsals may be used to fine-tune human–robot interactions, to make the audience believe that the android is working autonomously. In our fieldwork with the Robot Theater Project, we investigated how both director and human actors create a "real" conversation with robots through rehearsals.

Contemporary Colloquial Theater Theory and the Robot Theater Project

Oriza Hirata is a Japanese playwright and director, a key figure in the contemporary theater scene in Japan. He proposed an original theory to create performances, which is referred to as the contemporary colloquial theater theory (CCTT) (Hirata 1995) or, as well-known theater critics called it, "quiet drama" (Rimer et al. 2014). CCTT represents a paradigm shift for Japanese theatrical plays, away from the then widespread and fundamental method rooted in modern theatrical performance in Europe and the United States. To avoid using unnatural Japanese directly translated from Western languages and using deformed embodied actions, Prof. Hirata strongly advocates for the use of ordinary, everyday Japanese conversations on stage.

In 2008, Oriza Hirata and Hiroshi Ishiguro began a collaborative research project, called the Robot Theater Project. The former started to use robots and androids developed at the Ishiguro Labs as actors in a number of theatrical performances, namely, "I, Worker," using two Wakamaru robots, developed by Mitsubishi Heavy Industries, Ltd., with a premiere in 2008; "In the Heart of a Forest," again using two Wakamaru robots, with a premiere in 2008; "Three Sisters," using one Robovie robot, developed by ATR and Vstone Co., Ltd., and one android, Repliee Q1, developed by ATR, with a premiere in 2012; “Night of the Galaxy Express,” using one Robovie robot, with a premiere in 2013; and “Metamorphosis,” using Repliee S1, developed by ATR and Osaka University, with a premiere in 2014. These theatrical productions have been presented in a number of festivals worldwide.

A Case Study: Creating Conversations on a Stage

Fieldwork, Data Collection, and Story Lines

All rehearsals have been video-recorded on location at four stages between January 2012 and October 2014: “In the Heart of a Forest” (replaying, 2012), “Three Sisters” (2012), “Night of the Galaxy Express” (2013), and “Metamorphosis” (2014), accompanied by the required engineering support for measurement and necessary modification. This chapter focuses on two of these recordings, namely, “Three Sisters” and “Night of the Galaxy Express.”

Anton Chekhov wrote “Three Sisters” in 1900. The story takes place in a family apartment in Japan in a close future. The father of the three sisters was a scientist, and he passed away 3 years earlier. For reasons left unknown at the beginning, the youngest of the three sisters has been an android for the past 5 years. One day, a few former colleagues of the father gather in the sisters’ house to congratulate one of the university members for having secured a position abroad. Gradually, the history of why the youngest sister became an android is revealed.

The original “Night of the Galaxy Express” was written by Kenji Miyazawa, a Japanese fairy tale writer and poet, in 1927 and published in 1934. The story tells the adventures of Giovanni and Campanella, two young boys. Giovanni is a poor isolated boy, and Campanella takes care of him. The play starts in a classroom scenario, where a discussion about the Milky Way is carried out. The protagonists travel to their town, visit their houses, see a riverside, and finally embark on a fantastic train trip to the Northern Cross. The friendship between Giovanni and Campanella is set in a fantasy world, and Giovanni’s experiences and situations drive the story.

The Importance of Theatrical Rehearsals

Here, we briefly discuss our contribution from the perspective of human and social sciences, especially the EMCA framework, by observing this kind of practices through theatrical rehearsals.

Typically, our ordinary conversations occur only once, which means we cannot replicate identical discourses and utterances, i.e., we are not able to exactly reproduce the same sentences. However, during rehearsals of theatrical productions, conversations are repeated until they meet certain standards, which are defined by the director. We refer to this feature as *reproducibility*. On the one hand, this is in contrast to ordinary conversations. During the rehearsal of theatrical productions, the content of the utterances, i.e., the script, is defined in advance. On the other hand, embodied actions while performing the script are flexible, to a certain extent. We refer to this feature as *sequentially fixed and multimodally free*. We hypothesize that the sequentially repeated interactions video-recorded during rehearsals can provide us with valuable information for research on embodied actions for human–robot interactions, for example, about how robot engineers should design robot's human-like behaviors to accomplish a conversation with humans from the designing stance of pursuing realism.

Implementing Robot Behaviors

As far as the execution of robot behaviors is concerned, robots and androids are tele-operated: proper signals enable robots or androids to generate utterances and move. In a preparation phase, the engineering staff divides the scripts into several parts before rehearsal. As an example, "Three Sisters" lasts approximately 1 h and 45 min. In this case, engineers prepared 20 pieces in advance and, from the backstage, sent signals to allow robots to play the appropriate speech and motion behavior. In the right-hand side of Fig. 3, the "lightning signs" indicate the location from which engineers sent these signals. In this setting, it is necessary to observe the behaviors of human actors to appropriately schedule the signals from the backstage.

With reference to Fig. 4, when the director, Prof. Hirata, and the actors rehearse, the former uses the computer on the left-hand side, which runs the *AI Talk* software, to create a temporal sequencing among the lines in the robot's speech. The director also uses the computer on the right-hand side, which relies on original and custom-developed software including 15 motion parameters, to develop body movements and specify their temporal mutual relationships to depict the robot's embodied actions. After that, the engineering staff combines these two pieces of information into one signal to the robot.

Analysis of "Three Sisters"

In this first analysis, we introduce how the director and the actors produce appropriate temporal and sequential structures with the tele-operated robot's behavior. As it can be seen in Fig. 5, the woman on the left, who is sitting on a wheelchair, is an android and interprets the youngest sister, Ikumi. The man sitting on the floor with his legs crossed is the operator, who sends signals to enable the android to speak and

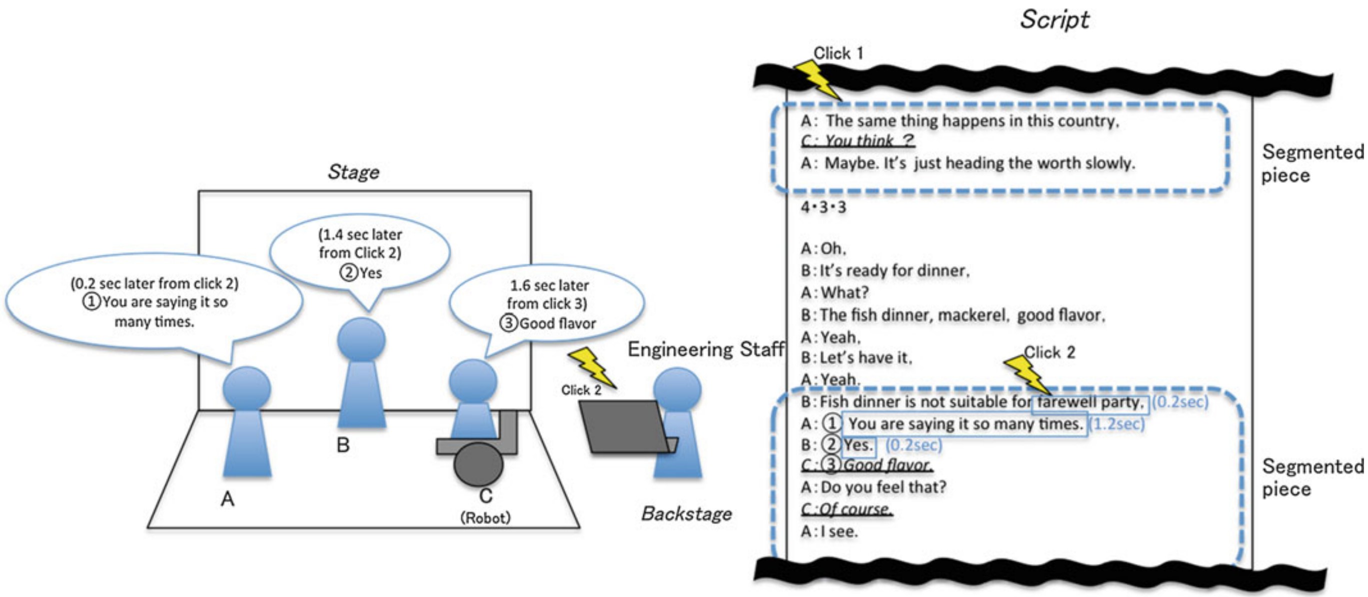


Fig. 3 Sketch of the robot behavior implementation process on a stage

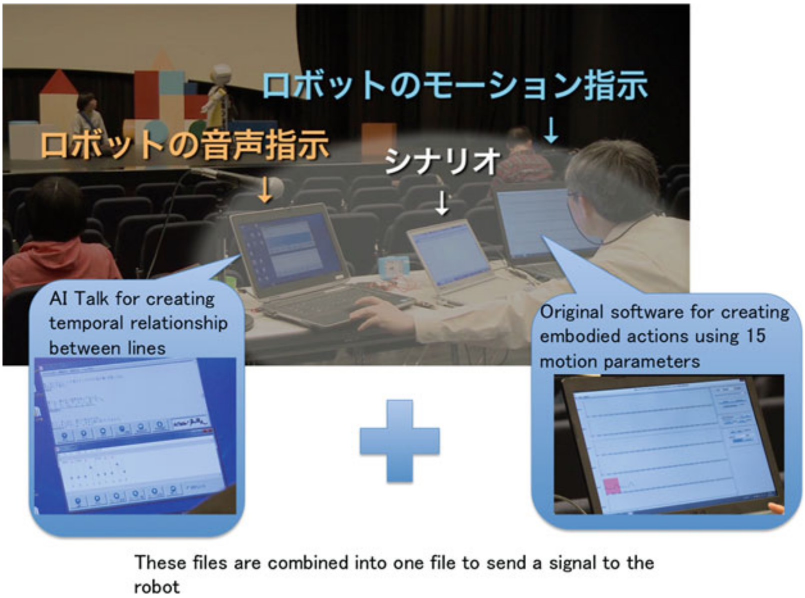


Fig. 4 Creating robot behaviors

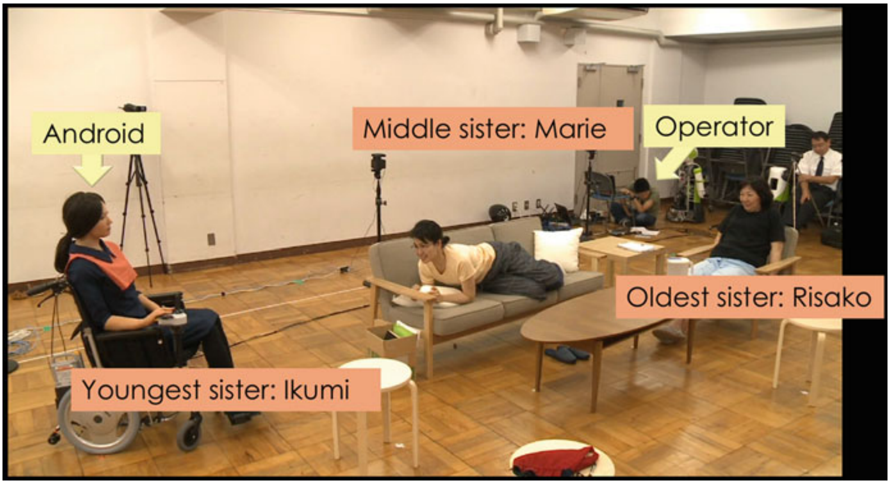


Fig. 5 Locations of the actors and the engineering staff for a rehearsal

move. The woman who is lying on the sofa is the middle sister, Marie, and the woman on the right-hand side, sitting on the chair, is the oldest sister, Risako. In the scene depicted in Fig. 5, actors are talking about the possibility they had to move to the United States when they were young.

Excerpt 1 shows a multimodal transcript taken from the first scene rehearsal. Speaking to her sisters, the oldest sister, Risako, says “Ikumi might not have been an android.” She then notices that this statement may have hurt the youngest sister, Ikumi. The oldest sister, Risako, then apologizes.

Excerpt 1:

Rehearsal 1

```
01 Risako :    e, datte ikumi datte anndoroido ni
02             naranakatta kamo shirenaiyo
```
 ‘*Well, Ikumi might not have been an android.*’
```
03             (1.4)
04 Marie :     soreha…
```
 ‘*Well, That’s…*’
```
05             (0.8)
06 Risako :    a, gomen
```
 ‘*Oh, I’m sorry.*’
```
07             (6.3)
08 Ikumi :     uuun daijoubu
```
 ‘*No, That's alright.*’

It can be noticed that there is a long pause, after 6.3 s, at line 7. If we had such long pauses in our ordinary conversations, it would be a problematic pause for participants. For example, Risako apologizes to Ikumi in line 6. However, Ikumi does not respond promptly. There is a possibility that this silence may represent Ikumi’s attitude toward Risako’s opinion, as in “I was hurt by your words.”

When this problem occurred, the director interrupted the ongoing rehearsal, and the human actors talked about how to get the timing right at that point in the play (Lefebvre et al. 2015). The director argued, “You two spoke too fast,” to the human actors, and after the comment, one of the actors, Risako, replied, “I need to get the timing correct by using some signal,” like a self-talk at the response position. Additionally, the director added: “Don’t worry about it at this moment. We will decide the timing in detail from now on.”

From this dialogue, we can see how the actors and the production staff create a theatrical exchange with the robot. Robot action is fixed and decided in advance. Therefore, it is the responsibility of human actors to take the appropriate timing changes. After the first rehearsal, actors had two more rehearsals to create a more normal conversation with the robot. The excerpts from the second and the third rehearsal show how actors modified their performance and lines.

Excerpt 2:

Rehearsal 2

```
01 Risako :   e, datte ikumi datte anndoroido ni
02            naranakatta kamo shirenaiyo
```
'Well, Ikumi might not have been an android.'
```
03            (1.3)
04 Marie :    soreha…
```
'Well, that's…'
```
05            (1.8)
06 Risako :   aaa, (.) gomen
```
'Oh, I'm sorry.'
```
07            (2.5)
08 Ikumi :    uuun daijoubu
```
'No, That's alright.'

In the second rehearsal, transcribed in Excerpt 2, the duration of the problematic pause at line 7 is reduced to 2.5 s. After the second rehearsal 2, Risako said, "It's thrilling, isn't it?," to everybody, which shows how she perceives the overall experience. Furthermore, the director suggested Marie to "take a longer pause before 'that's' (at line 3). If you can think of any embodied actions there, (please add them). It will be certain (to fill the gap)."

Excerpt 3:

Rehearsal 3

```
01 Risako :   e, datte ikumi datte anndoroido ni
02            naranakatta kamo shirenaiyo
```
'Well, Ikumi might not have been an android.'
```
03            (2.4)
04 Marie :    soreha…
```
'Well, that's…'
```
05            (3.3)
06 Risako :   a, gomen
```
'Oh, I'm sorry.'
```
07            (0.0)
08 Ikumi :    uuun daijoubu
```
'No, That's alright.'

Table 1 Durations for each line in each rehearsal (in seconds)

	Pause at line 3	Pause at line 5	Pause at line 7
Rehearsal 1	1.4	0.8	6.3
Rehearsal 2	1.3	1.8	2.5
Rehearsal 3	2.4	3.3	0.0

Rehearsal 3 was considered successful both for the director and actors, meaning that they considered improved the problematic pause at line 7. The number (0.0) at line 7 means that there is no gap between lines 6 and 8. For comparison with previous rehearsals, we did not delete the line itself.

However, the question remains: how did actors modify their performance to make the timing more appropriate with respect to the robot's response? Excerpt 3 shows that the duration of the pause at line 3 is extended to 2.4 s, which is consistent with the director's previous suggestions to Marie, to "take a longer pause before 'that's'." Additionally, Risako also took a longer pause before her line 6.

Table 1 shows the durations for each line in each rehearsal. It is possible to observe that when the pauses at lines 3 and 5 were extended, the pause at line 7 was reduced and finally eliminated.

Figure 6 shows the transition of speech duration and pause duration in rehearsals visually. We can observe that, across rehearsals, not only pause duration but also speech duration have been adapted.

These observations suggest that we must consider not only how the director and human actors coordinate their performance with the robot but also what kind of performance they consider as *acceptable* on stage.

Analysis of "Night of the Galaxy Express"

This section examines the rehearsals for the "Night of the Galaxy Express" with multimodal transcripts. For the analysis, we selected a specific segment from the first scene, which includes typical interactions between a teacher and students in a classroom, such as an initiation–response–evaluation (IRE) (Mehan 1979). In the script, shown partially in Excerpt 4, before exchanges, the teacher is talking about the Milky Way without actually calling it using the proper name. Then, in line 1, the teacher asks the students: "What should we call (this)?" In line 2, the teacher calls on Zaneri, one of the students. In line 3, Zaneri answers the question. Our target lines for the analysis are lines 2 and 3. To send a signal to the robot, the director or one of the engineers activates the behavior at the end of line 1, which is shown by the lightning *click* mark. The robot's first utterance is depicted in line 5. However, even before line 5, the robot engages in several embodied actions, raising his hand after line 1, looking at Zaneri at approximately line 3, and looking back to the teacher at the end of line 4. The overall behavior is composed not only of speech but also of these embodied actions.

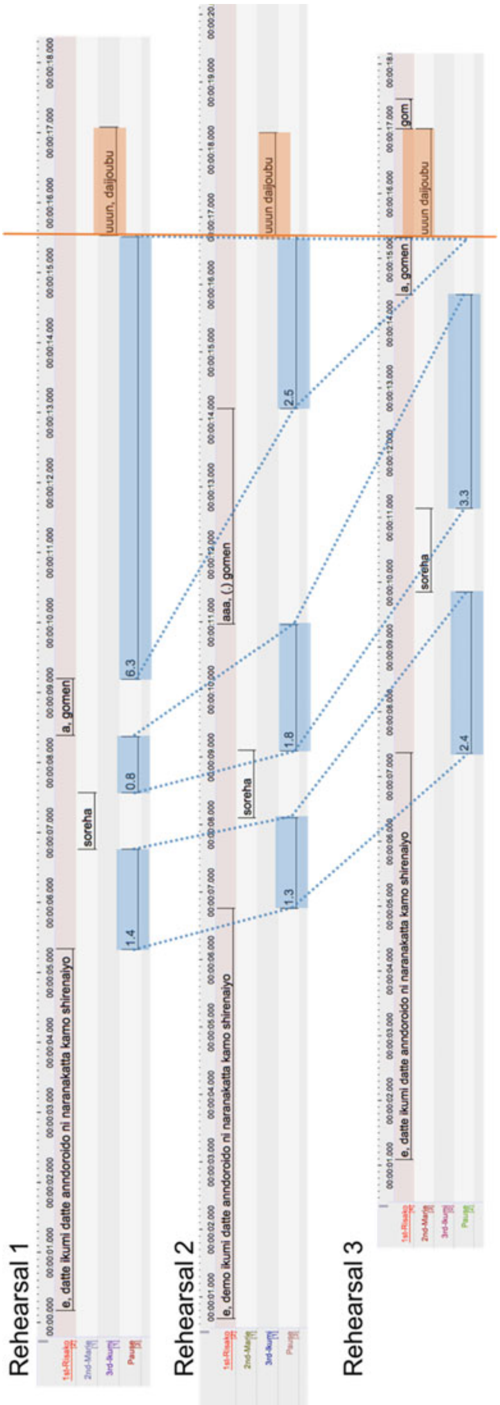


Fig. 6 Comparison of durations of speech and pauses in rehearsals

Excerpt 4:


Click

01 Teacher: watashitachi ha nanto yobu deshou ka
we TOP what call shoud-COP Q-FP
What should we call (this)?
(All Students raise their hands)
→ 02 Teacher: hai Zaneri-san
Yes, Zaneri -AD
Yeah, Zaneri?
→ 03 Zaneri: Amanogawa
Milky Way
04 Teacher: sou desu yoku dekimashita amanogawa
that be-COP good can-did Milky-Way
Exactly. Good job. Milky Way.
05 deha canpanerura-san
well Campanella-AD
Well, Campanella?
06 Campanella: hai (raise his face)
Yes.

Three rehearsals took place for this scene on April 14, 2013. The director offered two possible directions: either inserting a short pause before "Milky Way" or making it a little longer, i.e., inserting a short pause at the beginning of "Milky Way." These two possibilities were evaluated in between rehearsals. The last rehearsal was successful in this case. As given in the pause duration before "Milky Way," from Excerpt 5 and Excerpt 6, it is clear that the pause duration increases, similar to what was observed for the "Three Sisters."

Excerpt 5:

Rehearsal 1
02 Teacher: Hai. Zaneri-san.
Yeah, Zaneri-AD.
03 Zaneri: Hai (0.4) Amanogawa
Yes (0.4) Milky Way
04 Teacher: sou desu yoku dekimashita
Exactly. Good job. Milky Way.
Oriza's interruption: Insert a short pause before 'Milky Way'.

During rehearsal 1 at Excerpt 5, the robot does not move, which means it has not yet engaged in embodied actions. The director suggests inserting a short pause before “Milky Way.”

Excerpt 6:

Rehearsal 2

02 Teacher :	Hai. Zaneri-san.
	Yeah, Zaneri-AD.
03 Zaneri :	Hai (1.1) Amanogawa
	Yes (1.1) Milky Way
04 Teacher:	sou desu yoku dekimashita
	Exactly. Good job. Milky Way.
05 Zaneri :	Yatta!
	I did it!

Oriza's interruption: A little longer. Insert a short pause at the beginning of 'Milky Way'.

During the second rehearsal (reported in Excerpt 6), the robot starts moving, which means it is performing embodied actions. The director, at the end of line 1, sends a signal for the robot to move. At this moment, the robot and another student raise their hands. After the teacher calls on Zaneri, the robot turns to Zaneri. However, this turning action does not match the timing of Zaneri’s answer. At this point, the director interrupted the rehearsal to give the following direction: “A little longer. insert a short pause at the beginning of Milky Way.”

From the perspective of EMCA, Campanella’s embodied actions are regarded as those of a *side participant* or *bystander* (Goffman 1981). Even though Campanella does not speak any words, we anticipate that he notices the exchanges between the current speaker and the addressee.

Excerpt 7:

Rehearsal 3

02 Teacher :	Hai. Zaneri-san.
	Yeah, Zaneri-AD.
03 Zaneri :	(.) Hai (3.7) Amanogawa
	Yes (3.7) Milky Way

No interruption, continued

Rehearsal 3, described in Excerpt 7, was successful. The robot’s head movement occurs before Zaneri speaks. It is possible to get the impression that Zaneri’s pause before “Milky Way” is too long in this context. However, it is crucial to adjust the

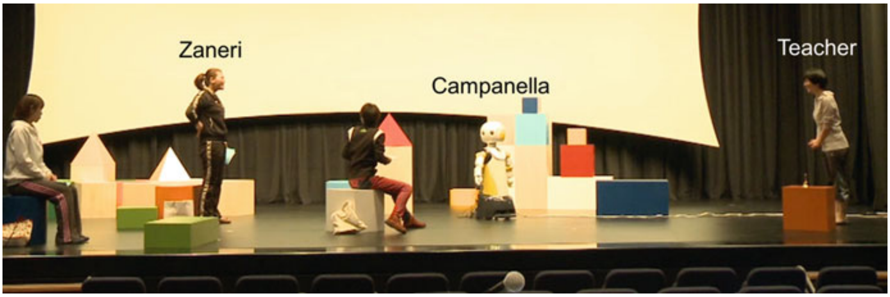


Fig. 7 Locations of human actors and the robot

timing of the interaction with the robot. The human actor playing Zaneri adds some improvised embodied actions to make it look more natural.

On another day, April 19, 2013, the scene was rehearsed again. A few adjustments were made to the robot's behavior. In Rehearsal 4, transcribed in Excerpt 8, Zaneri had a long pause before "Milky Way"(Amanogawa in Japanese), of about 2.7 s, which was shortened from the previous rehearsal. In line 5, the teacher calls on Campanella to answer a question. However, Campanella does not answer the teacher's question quickly, i.e., a long pause occurs in the question–answer sequence between the teacher and the student. After line 7, the director interrupted the rehearsal to suggest: "Zaneri, when Robovie completely turns to you, please say 'Milky Way'." That is, the director noticed that there was a problem, correcting the performance between lines 5–6. This direction is different from the previous ones. The director mentions the robot's behavior to modify initiation of the line "hai Amanogawa" at line 3.

Excerpt 8:

Rehearsal 4

02 **Teacher** : **hai Zaneri-san#**
campanella #turns his head to Zaneri (L->R)
03 **Zaneri** : **hai (2.7) #Amanogawa**
campanella #fixes his head to Zaneri
04 **Teacher** : **sou desu. (.) >yoku dekimashita< (1.3)<a:-ma:-no:-ga:#wa>. (1.2)**
campanella #turns his head to Teacher (R->L)
05 **deha(.) Canpane#rura-san?**
campanella #fixes his head to Teacher
06 **(---------+-#--------+= 2.0)**
campanella #raises his face (starts to look up teacher)
07 **Campanella** : **hai#.**
campanella #fixes his face to Teacher(looks up teacher)

Oriza's interruption: Zaneri, when the Robovie completely turns to you, please say 'Milky way'.

Excerpt 9:

```
Rehearsal 5
02 Teacher :       hai Zaneri-san?#
     campanella                  #turns his head to Zaneri (L->R)
03 Zaneri :        (.)hai (+--------#-++---=3.4)Amanogawa
     campanella                    #fixes his head to Zaneri
04 Teacher :       .h fu ¥sou¥desu. (.) >yoku dekimashita< (1.3)<a:
                   #ma:no:ga:wa>. (1.4)
     campanella    #turns his head to Teacher (R->L)
05                  deha(.)  # Canpanerura-san?
     campanella              #fixes his head to Teacher
06                (-------#--+-- = 1.2)
     campanella          #raises his face (starts to look up teacher)
07 Campanella :     hai #.
     campanella          #fixes his face to Teacher(looks up teacher)
No interruption, continued
```

In Rehearsal, 5, transcribed in Excerpt 9, Zaneri takes a longer pause at line 3, specifically 3.4 s. We can observe that the problematic pause at line 6 is modified as well: the duration of the pause was reduced to 1.2 s. Additionally, other embodied actions of Campanella, transcribed in gray lines, fit with the context. Therefore, Campanella anticipates Zaneri's next turn at line 3, and he turns to Zaneri (Fig. 7). Finally, during the teacher's evaluation of Zaneri, Campanella turns back toward the teacher in lines 4 and 5. In this way, the talking and the embodied actions between human actors and the robot are created, which allows the director to implement robot behaviors in complete fitting with the script.

Discussion

Here, we are interested in the exchanges between the director and the actors, as well as on the timing of the director's interruptions during rehearsals, which are aimed at timing refinement. As an example, "Three Sisters" lasts around 1 h and 45 min as a theatrical production. If there were no interruptions from a director, the actors would continue their performance to completion. When any issues happen during a rehearsal, the director can interrupt it and add his/her insights. Then, actors and engineers can retrospectively determine the timing of the occurred problems, defining precisely the moment in which the director interrupted the rehearsal to make the appropriate adjustments. We can observe the *professional vision* (Goodwin 1994) of the director in the ongoing rehearsal at that moment.

At the beginning of section "Experiences with Robots on Stage," we proposed the concept of *the reproducibility of conversation*. The data set collected in our fieldwork and video recordings of the Robot Theater Project illustrate our general conceptions of ideal and natural conversations. In the near future, technological developments may allow us to include a social robot as an equal participant. It will be interesting to see how humans will bond with them through social interactions. The Robot Theater Project could contribute to research in the area of social robotics, leading to a new platform for designing ideal conversations between humans and robots. Our fieldwork and video recordings show how humans develop the impression of the appropriateness/inappropriateness and problematic/unproblematic nature of social interactions.

It is noteworthy that in previous studies on social robotics, Japanese engineers tried to create and use humanlike robots resembling humans in appearance, as well as physical and linguistic ability. In such cases, humanoid robots tend to be independent or isolated from the conversational context in the material world, with the result that humans treat them as physical objects. From the perspective of workplace studies, Alač (2009) observed how engineers create an android's embodied actions using their own bodies as a reference.

Field studies such as the Robot Theater Project can lead to a paradigm shift in social robotics studies. Instead of designing and developing humanlike robots detached from the context, engineers are now able to construct situations in which social interactions occur. Our research, based on fieldwork and video recordings, illuminates the importance of context in interactions between social robots and humans.

Conclusions

The goal of this chapter is to present and discuss a number of challenges faced in the exciting field of human–robot interaction. Our discussion is focused on robots acting on stage, since this setting proves to be an excellent scenario to understand how humans perceive robots, either when directly interacting with them or when watching them interacting with other humans.

In the first part of the chapter, we discuss motor resonance, and we propose social resonance as an extension enabling the study of social robots from a holistic perspective. Then, we briefly survey relevant work about robot behavior sequencing, with a specific emphasis on those architectures explicitly taking human–robot interaction processes into account. Finally, we report experiences and lessons learned in the Robot Theater Project, a continuous study on human–robot interaction issues started in 2008.

The two research questions that we posed are:

Q1. Do humans perceive a robot they are interacting with as if it was a human as well?

Q2. Do humans perceive a robot interacting with other humans as natural?

It turns out that Q1 and Q2 refer to two different social settings. On the one hand, Q1 is related to a direct, possibly face-to-face interaction between humans and robots. The discussed studies seem to suggest that the answer is “yes,” subject to a number of conditions on the human and robot counterparts. These include the robot’s shape, its motions and its perceived engagement, as well as a number of beliefs and cultural biases in humans. Current debate seems to confirm that an overall perception of robot engagement is fundamental to provide them with a sense of agency. On the other hand, apart from a few studies, not only there is no consensus on Q2 but also we probably lack the theoretical setting to frame such a question. However, we try to devise a possible path toward such an understanding.

In the second part of the chapter, we focus on robot behavior selection and sequencing for human–robot interaction, and we discuss what it takes to deploy robots on stage as part of a theatrical play. In the current state of the art, just a few robot architectures take explicitly into account the need for a robot to interact with humans. Such requirement poses a number of technical challenges, for instance, a timely and purposive execution of behaviors, as well as the synchronization between communicative acts (including speech) and robot motions (including gestures), such that a weird and strange effect on humans is avoided, i.e., the so-called uncanny valley effect.

Finally, we report about our experiences in deploying robots on stage for real theatrical plays. In particular, we discuss how directors can deal with real-world robots during rehearsals. It is evident how a director and the actors adapt to robot behaviors and not vice versa. Nonetheless, experience suggests that a social framework is created among the director, the actors, *and* the robot.

The arising of such a social setting motivates a broader study in human–robot interaction, which encompasses not only how robots are perceived but also the *role* they can play in everyday environments.

Recommended Reading

M. Alač, Moving android: on social robots and body-in-interaction. Soc. Stud. Sci. **39**(4), 491–528 (2009)

S. Andrist, W. Collier, M. Gleicher, B. Mutlu, D. Shaffer, Look together: analyzing gaze coordination with epistemic network analysis. Front. Psychol. **6**, 1016 (2015)

D. Anghinolfi, G. Cannata, F. Mastrogiovanni, C. Nattero, M. Paolucci, On the problem of the automated design of large-scale robot skin. IEEE Trans. Autom. Sci. Eng. **10**(4), 1087–1100 (2013)

R.C. Arkin, M. Fujita, T. Takagi, R. Hasegawa, An ethological and emotional basis for human–robot interaction. Robot. Auton. Syst. **42**(3), 191–201 (2003)

J.M. Atkinson, J. Heritage, Structures of social action : studies in conversation analysis, Cambridge University Press, Cambridge (1984)

J.R. Augustine, Circuitry and functional aspects of the insular lobes in primates including humans. Brain Res. Rev. **2**, 229–294 (1996)

C. Bartneck, T. Kanda, H. Ishiguro, N. Hagita, Is the uncanny valley an uncanny cliff? in *Proceedings of the 16th IEEE, RO-MAN 2007*, Jeju (2007), 368–373

E. Bastianelli, G. Castellucci, D. Croce, R. Basili, D. Nardi, Effective and robust natural language understanding for human-robot interaction, in *Proceedings of the 2014 European Conference on Artificial Intelligence (ECAI 2014)*, Prague (2014)

T.N. Beran, A. Ramirez-Serrano, R. Kuzyk, M. Fior, S. Nugent, Understanding how children understand robots: perceived animism in child–robot interaction. Int. J. Hum. Comput. Stud. **69**(7–8), 539–550 (2011)

C.A. Bouquet, V. Gaurier, T. Shipley, L. Toussaint, Y. Blandin, Influence of the perception of biological or non-biological motion on movement execution. J. Sports Sci. **25**, 519–530 (2007)

V. Braitenberg, *Vehicles: Experiments in Synthetic Psychology* (MIT Press, Cambridge, MA, 1984)

C. Breazeal, *Designing Sociable Robots* (MIT Press, Cambridge, MA, 2002), 3

E. Broadbent, R. Stafford, B. MacDonald, Acceptance of healthcare robots for the older population: review and future directions. Int. J. Soc. Robot. **1**(4), 319–330 (2009)

A. Bruce, J. Knight, I. Nourbakhsh, Robot improve: using drama to create believable agents, in *AAAI Workshop Technical Report WS-99-15 of the 8th Mobile Robot Competition and Exhibition* (AAAI Press, Menlo Park, 1999), pp. 27–33

R. Burke, D. Isla, M. Downie, Y. Ivanov, B, Blumberg, Creature- smarts: the art and architecture of a virtual brain, in *Proceedings of the Game Developers Conference* (2001), pp. 147–166

L. Carr, M. Iacoboni, M.-C. Dubeau, J.C. Mazziotta, G.L. Lenzi, Neural mechanisms of empathy in humans: a relay from neural systems for imitation to limbic areas. Proc. Natl. Acad. Sci. U. S. A. **100**(9), 5497–5502 (2003)

P. Carruthers, P. Smith, *Theories of Theories of Mind* (Cambridge University Press, Cambridge, MA, 1996)

A. Casile, E. Dayan, V. Caggiano, T. Hendler, T. Flash, M.A. Giese, Neuronal encoding of human kinematic invariants during action observation. Cereb. Cortex **20**(7), 1647–1655 (2010)

T. Chaminade, J. Hodgins, M. Kawato, Anthropomorphism influences perception of computer-animated characters' actions. Soc. Cogn. Affect. Neurosci. **2**(3), 206–216 (2007)

T. Chaminade, M. Zecca, S.J. Blakemore, A. Takanishi, C.D. Frith, S. Micera, M.A. Umilta, Brain response to a humanoid robot in areas implicated in the perception of human emotional gestures. PLoS One **5**(7), e11577 (2010)

J. Cook, D. Swapp, X. Pan, N. Bianchi-Berthouze, S.J. Blakemore, Atypical interference effect of action observation in autism spectrum conditions. Psychol. Med. **44**(4), 731–740 (2014)

G. Csibra, Mirror neurons and action observation: Is simulation involved? Interdisciplines (2004). http://www.interdisciplines.org/mirror/papers/4

K. Dautenhahn, Socially intelligent robots: dimensions of human–robot interaction. Philos. Trans. R. Soc. B **362**(1480), 679–704 (2007)

M. Davies, T. Stone, *Folk Psychology* (Oxford, Cambridge, MA, 1995a)

M. Davies, T. Stone, *Mental Simulation* (Oxford, Cambridge, MA, 1995b)

J. Demiris, G.R. Hayes, Imitation as a dual-route process featuring predictive and learning components: a biologically plausible computational model, in *Imitation in Animals and Artifacts* (MIT Press, Cambridge, MA, 2002), pp. 327–361

S. Denei, F. Mastrogiovanni, G. Cannata, Towards the creation of tactile maps for robots and their use in robot contact motion control. Robot. Auton. Syst. **63**(3), 293–308 (2015)

D. Dennett, *The Intentional Stance* (MIT Press, Cambridge, MA, 1987)

J.J. Diehl, L.M. Schmitt, M. Villano, C.R. Crowell, The clinical use of robots for individuals with autism spectrum disorders: a critical review. Res. Autism. Spect. Disord. **6**(1), 249–262 (2012)

B.R. Duffy, Anthropomorphism and the social robot. Rob. Auton. Syst. **42**(3–4), 177–190 (2003)

B. Eskridge, D. Hougen, Extending adaptive fuzzy behavior hierarchies to multiple levels of composite behaviors. Robot. Auton. Syst. **58**, 1076–1084 (2010)

F. Eyssel, F. Hegel, (S)he's got the look: gender-stereotyping of social robots. J. Appl. Soc. Psychol. **42**, 2213–2230 (2012)

F. Eyssel, D. Kuchenbrandt, Social categorization of social robots: anthropomorphism as a function of robot group membership. Br. J. Soc. Psychol. **51**, 724–731 (2012)

T. Flash, B. Hochner, Motor primitives in vertebrates and invertebrates. J. Curr. Opin. Neurobiol. **15**, 660–666 (2005)
T. Fong, I. Nourbakhsh, K. Dautenhahn, A survey of socially interactive robots. Robot. Auton. Syst. **42**(3–4), 143–166 (2003)
M. Fujita, H. Kitano, Development of an autonomous quadruped robot for robot entertainment. Autono. Robot. **5**(1), 7–18 (1998)
M. Fujita, Y. Kuroki, T. Ishida, T. T. Doi, Autonomous behavior control architecture of entertainment humanoid robot SDR-4X, in *Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)*, Las Vegas (2003)
V. Gallese, The 'shared manifold' hypothesis: from mirror neurons to empathy. J. Conscious. Stud. **8**, 33–50 (2001)
V. Gallese, A. Goldman, Mirror neurons and the simulation theory of mind-reading. Trends Cogn. Sci. **2**(12), 493–502 (1998)
H. Garfinkel, *Studies in Ethnomethodology* (Prentice-hall, Englewood Cliffs, 1967). Paperback edition, *Studies in Ethnomethodology* (Polity Press, Cambridge, 1984)
R. Gehle, K. Pitsch, T. Dankert, S. Wrede, Trouble-based group dynamics in real-world HRI – REactions on unexpected next moves of a museum guide robot, in *Accepted for the 24th International Symposium on Robot and Human Interactive Communication (Ro-MAN 2015)*, Kobe (2015)
A.L. Georgescu, B. Kuzmanovic, N.S. Santos, R. Tepest, G. Bente, M. Tittgemeyer, K. Vogeley, Perceiving nonverbal behavior: neural correlates of processing movement fluency and contingency in dyadic interactions. Hum. Brain Mapp. **35**(4), 1362–1378 (2014)
J. Goetz, S. Kiesler, A. Powers, Matching robot appearance and behaviour to tasks to improve human-robot cooperation, in *Proceedings of the 2003 I.E. International Workshop on Robot and Human Interactive Communication (IEEE RO-MAN 2003)*, Vancouver (2003)
E. Goffman, *Forms of Talk* (University of Pennsylvania Press, Philadelphia, 1981)
A. Goldman, The mentalising folk. Protosociology **16**, 7–34 (2002)
M.A. Goodrich, A.C. Schultz, Human-robot interaction: a survey. Found. Trends Hum. Comput. Interact. **1**(3), 203–275 (2007)
C. Goodwin, Professional vision. Am. Anthropol. **96**(3), 606–633 (1994)
C. Goodwin, Action and embodiment within situated human interaction. J. Pragmat. **32**, 1489–1522 (2000a)
C. Goodwin, Practices of seeing; visual analysis: an ethnomethodological approach, in *Handbook of Visual Analysis*, ed. by T. van Leeuven, C. Jewitt (Sage, London, 2000b)
C. Goodwin, The co-operative, transformative organization of human action and knowledge. J. Pragmat. **46**(1), 8–23 (2013).
G.-F. Gutemberg, A. Yiannis, A language for human action. Computer **40**(5), 42–51 (2007)
P. Haddington, T. Keisanen, L. Mondada, M. Nevile, *Multiactivity in Social Interaction, Beyond Multitasking* (John Benjamins, Amsterdam/Philadelphia, 2014)
A. Hamilton, S. Grafton, The motor hierarchy: from kinematics to goals and intentions, in *Attention and Performance 22* (Oxford University Press, Oxford, 2007)
S. Hart, R. Grupen, Learning generalizable control programs. IEEE Trans. Auton. Ment. Dev. **3**(3), 216–232 (2011)
M. Haruno, D.M. Wolpert, M. Kawato, MOSAIC model for sensorimotor learning and control. Neural Comput. **13**, 2201–2220 (2001)
J. Heritage, J.M. Atkinson, Introduction, in *Structures of Social Action Studies in Conversation Analysis*, ed. by J.M. Atkinson, J. Heritage (Cambridge University Press, Cambridge, 1984), pp. 1–15
D. Heylen, Head gestures, gaze and the principles of conversational structure. Int. J. Humanoid. Rob. **3**(3), 1–27 (2006)
O. Hirata, *Hirata Oriza no shigoto (1) Gendai kougo engeki notameni* (*Introduction to Contemporary Colloquial Theatre Theory*) (Banseisha, Tokyo, 1995) (in Japanese)

C.M. Huang, S. Andrist, A. Sauppé, B. Mutlu, Using gaze patterns to predict task intent in collaboration. Front. Psychol. **6**, 1049 (2015)

M. Huber, A. Kupferberg, C. Lenz, A. Knoll, T. Brandt, S. Glasauer, Spatiotemporal movement planning and rapid adaptation for manual interaction. Plos One **8**, e64982 (2013)

M. Iacoboni, L.M. Koski, M. Brass, H. Bekkering, R.P. Woods, M.C. Dubeau, J.C. Mazziotta, G. Rizzolatti, Reafferent copies of imitated actions in the right superior temporal cortex. Proc. Natl. Acad. Sci. U. S. A. **98**(24), 13995–13999 (2001)

H. Ishiguro, Android science: conscious and subconscious recognition, Conn Sci. 18(4), 319–332 (2006)

H. Ishiguro, S. Nishio, A. Chella, R. Sorbello, G. Balistreri, M. Giardina, C. Calì, Investigating perceptual features for a natural human-humanoid robot interaction inside a spontaneous setting, in *Proceedings of the 2012 International Conference on Biologically Inspired Cognitive Architectures* (BICA 2012), Palermo. Advances in Intelligent Systems and Computing, vol. 196 (2012), pp. 167–174

O. Jenkins, M. Mataric, Performance-derived behavior vocabularies: data- driven acquisition of skills from motion. Int. J. Humanoid. Rob. **1**(2), 237–288 (2004)

S. Kahl, S. Kopp, Modelling a social brain for interactive agents: integrating mirroring and mentalising, in *Proceedings of the 2015 International Conference on Intelligent Virtual Agents (IVA 2015)*, Delft, Aug (2015)

P. H. Kahn, J. H. Ruckert, T. Kanda, H. Ishiguro, A. Reichert, H. Gary, S. Shen, Psychological intimacy with robots? Using interaction patterns to uncover depth of relation, in *Proceedings of the Fifth ACM/IEEE International Conference on Human-Robot Interaction (HRI 2010)*, New York (2010a)

P. H. Kahn, B.T. Gill, A. Reichert, T. Kanda, H. Ishiguro, J.H. Ruckert, Validating interaction patterns in HRI, in *Proceedings of the Fifth ACM/IEEE International Conference on Human-Robot Interaction (HRI 2010)*, New York (2010b)

H. Kamide, Y. Mae, K. Kawabe, S. Shigemi, T. Arai, A psychological scale for general impression of humanoid, in *Proceedings of the 2012 I.E. International Conference on Robotics and Automation (ICRA 2012)*, St. Paul (2012)

T. Kanda, H. Ishiguro, T. Ishida, Psychological analysis on human-robot interaction, in *Proceedings of the 2001 I.E. International Conference on Robotics and Automation (ICRA 2001)*, Seoul (2001)

T. Kanda, T. Miyashita, T. Osada, Y. Haikawa, H. Ishiguro, Analysis of humanoid appearances in human-robot interaction. IEEE Trans. Robot. **24**(3), 725–735 (2008)

T. Kanda, M. Shiomi, Z. Miyashita, H. Ishiguro, N. Hagita, A communication robot in a shopping map. IEEE Trans. Robot. **26**(5), 135–145 (2010)

A. Kendon, *Gesture: Visible Action as Utterance* (Cambridge University Press, Cambridge, 2004)

J.M. Kilner, Y. Paulignan, S.J. Blakemore, An interference effect of observed biological movement on action. Curr. Biol. **13**, 522–525 (2003)

J.M. Kilner, A.F. de Hamilton, S.-J. Blakemore, Interference effect of observed human movement on action is due to velocity profile of biological motion. Soc. Neurosci. **2**(3–4), 158–166 (2007)

S. Knoop, M. Pardowitz, R. Dillmann, From abstract task knowledge to executable robot programs. J. Intell. Robot. Syst. **52**, 343–362 (2008)

E. Kohler, C. Keysers, M.A. Umiltà, L. Fogassi, V. Gallese, G. Rizzolatti, Hearing sounds, understanding actions: action representation in mirror neurons. Science **297**, 846–848 (2002)

H. Kose-Bagci, F. Broz, Q. Shen, K. Dautenhahn, C.L. Nehaniv, As time goes by: representing and reasoning about timing in human-robot interaction studies, in *AAAI Spring Symposium: It's All in the Timing*, Palo alto, California (2010)

Y. Kuno, K. Sadazuka, M. Kawashima, K. Yamazaki, A. Yamazaki, H. Kuzuoka, Museum guide robot based on sociological interaction analysis, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, San Jose, California, 1191–1194 (2007)

A. Kupferberg, S. Glasauer, M. Huber, M. Rickert, A. Knoll, T. Brandt, Biological movement increases acceptance of humanoid robots as human partners in motor interaction. AI Soc. **26**(4), 339–345 (2011)

A. Kupferberg, M. Huber, B. Helfer, C. Lenz, A. Knoll, S. Glasauer, Moving just like you: motor interference depends on similar motility of agent and observer. PLoS One **7**(6), e39637 (2012)

H. Kuzuoka, Y. Suzuki, J. Yamashita, K. Yamazaki, Reconfiguring spatial formation arrangement by robot's body orientation, in *Proceedings of HRI 2010*, Osaka, Japan, pp. 285–292 (2010)

F. Lacquaniti, E. Guigon, L. Bianchi, S. Ferraina, R. Caminiti, Representing spatial information for limb movement: role of area 5 in the monkey. Cereb. Cortex **5**(5), 391–409 (1995)

A. Lefebvre, M. Bono, C. Sunakawa, Information control and accountability in social interaction: the case of the theatrical performance, in *14th International Pragmatics Conference* (IPrA 2015), Antwerp (2015)

H. Lim, T. Kang, J. Lee, J. Kim, B. You, Multiple humanoid cooperative control system for heterogeneous humanoid team, in *Proceedings of the 2008 I.E. International Symposium on Robot and Human Interactive Communication (IEEE RO-MAN 2008)*, Munich (2008)

M. Lorr, Client perceptions of therapists. J. Consult. Psychol. **29**, 146–149 (1965)

K. MacDorman, Subjective ratings of robot video clips for human-likeness, familiarity, and eeriness: an exploration of the uncanny valley, in *Proceedings of the ICCS/CogSci 2006 Long Symposium: Toward Social Mechanisms of Android Science*, Vancouver, pp. 26–29 (2006)

P. Maiolino, M. Maggiali, G. Cannata, G. Metta, L. Natale, A flexible and robust large-scale capacitive tactile system for robots. IEEE Sensors J. **13**(10), 3910–3917 (2013)

L. Marin, J. Issartel, T. Chaminade, Interpersonal motor coordination: from human-human to human-robot interactions. Interact. Stud. **10**(3), 479–504 (2009)

F. Mastrogiovanni, A. Sgorbissa, A behaviour sequencing and composition architecture based on ontologies for entertainment humanoid robots. Robot. Auton. Syst. **61**(2), 170–183 (2013)

M. Matarić, Socially assistive robotics. IEEE Intell. Syst. **5**(3-4), 81–83 (2006)

H. Mehan, *Learning Lessons: Social Organization in the Class Room* (Harvard University Press, Cambridge, MA, 1979)

T. Minato, M. Shimada, S. Itakura, K. Lee, H. Ishiguro, Evaluating the human likeness of an android by comparing gaze behaviors elicited by the android and a person. Adv. Robot. **20**(10), 1147–1163 (2006)

L. Mondada, Participants'online analysis and multimodal practices: projecting the end of the turn and the closing of the sequence. Discourse Stud. Spec. Issue Discourse Interact. Cognit. **8**(1), 117–129 (2006)

M. Mori, Bukimi no tani (the uncanny valley). Energy **7**, 33–35 (1970)

T. Nomura, T. Suzuki, T. Kanda, K. Kato, Measurement of anxiety toward robots, in *Proceedings of the 2006 I.E. International Symposium on Robot and Human Interactive Communication (IEEE RO-MAN 2006)*, Hatfield (2006)

L. Noy, R.I. Rumiati, T. Flash, Simple movement imitation: are kinematic features sufficient to map perceptions into actions? Brain Cogn. **69**(2), 360–368 (2009)

L.M. Oberman, J.P. McCmeery, V.S. Ramachandran, J.A. Pineda, EEG evidence for mirror neuron activity during the observation of human and robot actions: toward an analysis of the human qualities of interactive robots. Neurocomputing **70**, 2194–2203 (2007)

S. Okita, D. Schwartz, Young children's understanding of animacy and entertainment robots. Int. J. Humanoid. Rob. **3**(3), 393–412 (2006)

Osaka University Center for the Study of Communication-Design (CSCD), O. Hirata, H. Ishiguro, K. Kuroki, S. Kinsui, *Robotto Engeki (Robot Theatrical Performance)* (Osaka University Press, Osaka, 2010) (in Japanese)

E. Oztop, D.W. Franklin, T. Chaminade, G. Cheng, Human-humanoid interaction: is a humanoid robot perceived as a human? Int. J. Humanoid. Rob. **2**(4), 537–560 (2005a)

E. Oztop, D. Wolpert, M. Kawato, Mental state inference using visual control parameters. Brain Res. **22**, 129–151 (2005b)

D.I. Perrett, N.J. Emery, Understanding the intentions of others from visual signals: neurophysiological evidence. Curr. Psychol. Cogn. **13**, 683–694 (1994)

D.I. Perrett, M.H. Harries, R. Bevan, S. Thomas, P.J. Benson, A.J. Mistlin, A.J. Chitty, J.K. Hietanen, J.E. Ortega, Frameworks of analysis for the neural representation of animate objects and actions. J. Exp. Biol. **146**, 87–143 (1989)

F. Pratama, F. Mastrogiovanni, N.Y. Chong, An integrated epigenetic robot architecture via context-influenced long-term memory, in *Proceedings of the 2014 Joint IEEE International Conference on Development and Learning and on Epigenetic Robotics (ICDL-EpiRob 2014)*, Genoa (2014)

F. Pratama, F. Mastrogiovanni, S. Jeong, N.Y. Chong, Long-term knowledge acquisition in a memory-based epigenetic robot architecture for verbal interactions, in *Proceedings of the 2015 I.E. International Symposium on Robot and Human Interactive Communication (IEEE RO-MAN 2015)*, Kobe (2015)

C. Press, H. Gillmeister, C. Heyes, Sensorimotor experience enhances automatic imitation of robotic action. Proc. R. Soc. B **274**, 2509–2514 (2007)

G. Randelli, T.M. Bonanni, L. Iocchi, D. Nardi, Knowledge acquisition through human-robot multimodal interaction. Intell. Serv. Robot. **6**(1), 19–31 (2012)

M.J. Richardson, K.L. Marsh, R.C. Schmidt, Effects of visual and verbal interaction on unintentional interpersonal coordination. J. Exp. Psychol. **31**(1), 62–79 (2005)

J.T. Rimer, M. Mitsuya, M. Cody Poulton, *The Columbia Anthology of Modern Japanese Drama* (Columbia UP, New York, 2014), Print

G. Rizzolatti, L. Fogassi, V. Gallese, Neurophysiological mechanisms underlying the understanding and imitation of action. Nat. Rev. Neurosci. **2**, 661–670 (2001)

G. Rizzolatti, L. Craighero, The mirror-neuron system. Annu. Rev. Neurosci. **27**, 169–192 (2004)

B. Robins, K. Dautenhahn, R. Boerkhorst, A. Billard, Robots as assistive technology – does appearance matter? in *Proceedings of the 2004 I.E. International Workshop on Robot and Human Interactive Communication (IEEE RO-MAN 2004)*, Kurashiki (2004)

S. Rosenberg, C. Nelson, P.S. Vivekananthan, A multidimensional approach to the structure of personality impressions. J. Pers. Soc. Psychol. **9**, 283–294 (1998)

H. Sacks, E.A. Schegloff, G. Jefferson, A simplest systematics for the organisation of turn-taking for conversation. Language **50**, 696–735 (1974)

A. Sadeghipour, S. Kopp, Embodied gesture processing: motor-based integration of perception and action in social artificial agents. Cogn. Comput. **3**(3), 419–435 (2011)

A. Sadeghipour, S. Kopp, A probabilistic model of motor resonance for embodied gesture detection, in *Proceedings of the 2009 International Conference on Intelligent Virtual Agents (IVA 2009)*, Amsterdam, Sept (2009)

A. Sadeghipour, S. Kopp, *Proceedings of the 2014 Annual Meeting of the Cognitive Science Society (COGSCI 2014)*, Quebec City, July (2014)

E.A. Schegloff, *Sequence Organization in Interaction. A Primer in Conversation Analysis I* (Cambridge University Press, Cambridge, 2007)

R.C. Schmidt, M.J. Richardson, Dynamics of interpersonal coordination, in *Coordination: Neural, Behavioural and Social Dynamics*, ed. by A. Fuchs, V. Jirsa (Springer, Heidelberg, 2008), pp. 281–307

A. Schmitz, P. Maiolino, M. Maggiali, L. Natale, G. Cannata, G. Metta, Methods and technologies for the implementation of large-scale robot tactile sensors. IEEE Trans. Robot. **27**(3), 389–400 (2011)

A. Sciutti, A. Bisio, F. Nori, G. Metta, L. Fadiga, G. Sandini, Anticipatory gaze in human-robot interactions, in *Gaze in HRI from Modeling to Communication" Workshop at the 7th ACM/IEEE International Conference on Human-Robot Interaction*, Boston (2012a)

A. Sciutti, A. Bisio, F. Nori, G. Metta, L. Fadiga, T. Pozzo, G. Sandini, Measuring human-robot interaction through motor resonance. Int. J. Soc. Robot. **4**(3), 223–234 (2012b)

A. Sciutti, A. Bisio, F. Nori, G. Metta, L. Fadiga, G. Sandini, Robots can be perceived as goal-oriented agents. Interact. Stud. **14**(3), 329–350 (2013)

A. Sciutti, F. Rea, G. Sandini, When you are young, (robot's) looks matter. Developmental changes in the desired properties of a robot friend, in *Proceedings of the 2014 I.E. International*

Symposium on Robot and Human Interactive Communication (IEEE RO-MAN 2014), Edinburgh (2014)
Q. Shen, H. Kose-Bagci, J. Saunders, K. Dautenhahn, An experimental investigation of interference effects in human-humanoid interaction games, in *Proceedings of the 2009 I.E. International Symposium on Robot and Human Interactive Communication (IEEE RO-MAN 2009)*, Toyama (2009)
Q. Shen, H. Kose-Bagci, J. Saunders, K. Dautenhahn, The impact of participants' beliefs on motor interference and motor coordination in human-robot interactions. IEEE Trans. Auton. Ment. Dev. **3**(1), 6–17 (2011)
J. Stanley, E. Gowen, R.C. Miall, Interference in performed movement during observation of a moving dot stimulus. J. Exp. Psychol. **33**, 915–926 (2007)
J. Stanley, E. Gowen, R.C. Miall, How instructions modify perception: an fMRI study investigating brain areas involved in attributing human agency. Neuroimage **52**(1), 389–400 (2010)
A. Stoytchev, Some basic principles of developmental robotics. IEEE Trans. Auton. Ment. Dev. **1**(2), 122–130 (2009)
N.G. Tsagarakis, G. Metta, G. Sandini, D. Vernon, R. Beira, F. Becchi, L. Righetti, J. Santos-Victor, A.J. Ijspeert, M.C. Carrozza, D.G. Caldwell, iCub: the design and realisation of an open humanoid platform for cognitive and neuroscience research. Adv. Robot. **21**(10), 1151–1175 (2007)
M.M. Weil, L.D. Rosen, The psychological impact of technology from a global perspective: a study of technological sophistication and technophobia in university students from twenty-three countries. Comput. Hum. Behav. **11**(1), 95–133 (1995)
M. Wilson, G. Knoblich, The case for motor involvement in perceiving conspecifics. Psychol. Bull. **131**(3), 460–473 (2005)
D.M. Wolpert, M. Kawato, Multiple paired forward and inverse models for motor control. Neural Netw. **11**, 1317–1329 (1998)
D.M. Wolpert, K. Doya, M. Kawato, A unifying computational framework for motor control and social interaction. Philos. Trans. R. Soc. Lond. B **358**, 593–602 (2003)
S. Woods, K. Dautenhahn, J. Schulz, The design space of robots: investigating children's views, in *Proceedings of the 2004 I.E. International Workshop on Robot and Human Interactive Communication (IEEE RO-MAN 2004)*, Kurashiki (2004)
Y. Wu, C. Fassert, A.S. Rigaud, Designing robots for the elderly: appearance issue and beyond. Arch. Gerontol. Geriatr. **54**(1), 121–126 (2012)
A. Wykowska, A. Schubö, Perception and action as two sides of the same coin: a review of the importance of action-perception links in humans for social robot design and research. Int. J. Soc. Robot. **4**, 5–14 (2012)
A. Wykowska, R. Chellali, M. Mamun Al-Amin, H.J. Müller, Does observing artificial robotic systems influence human perceptual processing in the same way as observing humans? in *Proceedings of the 2012 International Conference on Social Robotics (ICSR 2012)*, Chengdu, Oct (2012)

Design and Development of Playful Robotic Interfaces for Affective Telepresence

38

Elham Saadatian, Hooman Samani, and Ryohei Nakatsu

Contents

E. Saadatian (✉)
School of Electronics and Computer Science, Interaction, Complexity Group, University of Southampton, Southampton, UK
e-mail: e.saadatian@soton.ac.uk; elham.saadatian@soton.ac.uk

H. Samani
Department of Electrical Engineering, College of Electrical Engineering and Computer Science, National Taipei University, NTUP, Taipei, Taiwan
e-mail: hoomanntpu@gmail.com; hooman@mail.ntpu.edu.tw

R. Nakatsu
Design School, Kyoto University, Kyoto, Japan
e-mail: nakatsu.ryohei@gmail.com

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_63

Abstract

This chapter describes our exploratory studies toward design and development of playful robotic interfaces for intimate telepresence. Our goal has been to create playful experience for couples in long-distance relationships (LDRs) through telepresence robots. Playfulness is becoming an increasingly important theme in human-computer interaction (HCI) and human-robot interaction (HRI). Playfulness refers to enhancing the quality of the experience by adding joy, pleasure, and fun and in a more general description adding hedonic attributes to the experience. Within the context of this study, playfulness is reflected as enhancing the affective dimension of the interaction through iterative prototyping. Intimate telepresence refers to a group of telepresence technologies, which convey the message of being copresence with the intimate partner or loved ones. In this chapter, we present a new generation of telepresence technology that increases the possibility of ongoing connectedness, self-disclosure, and empathy through stochastic detection of the user's mood state and regenerating it in the remote location. Intimate telepresence has been also explored from the design point of view. In this respect a pair of personalizable telepresence robots are designed, developed, and evaluated. These robots can transmit the body languages in the remote location and are designed based on the concept of enclothed cognition and the influences of appearance and likeness on affectivity. And finally, the possibility of teleporting kisses was explored and the first longitudinal field study within the field of intimate telepresence was performed.

Keywords

Affective telepresence • Playful interactions • User experience research

Introduction

Playfulness is a broad concept and has been looked into from different perspectives. Some researchers have approached playfulness from designing for desirable user experience or mode of interaction (Deterding et al. 2013). Sometimes it has been defined more general as "pleasurable experience" or "fun" (Wei and Cheok 2012). In some studies including the present work, it is described as any interaction, which its focus has been shifted from task oriented toward experience oriented (Scott 2014). Traditionally the goal of telecommunication technology was mainly collaboration and task-oriented activities and effectiveness and efficiency have been considered to be the adequate design factors. However, in conventional society, the goal of telecommunication technology has moved beyond only information exchange and telecommunication media are also used in variety of contexts. In this respect it has been argued that the quality of interactive products not only depends on the pragmatic aspects, it also depends on the hedonic attributes (Hassenzahl et al. 2008).

This study summarizes the attempts we have taken to support remote, intimate interactions by means of playful and artificial telepresence agents. The focus has been on the design and development of the agents, including humanoid telepresence robots, abstract robotic interfaces, and artificial character on smartphone. The motivation of our work is to support couples in long-distance relationships to be connected through artificial agents in a playful way. Playfulness is a concept which is usually applied in game design; however, in this work our aim has been exploring the design space for a playful approach to telepresence for long-distance intimate relationships, rather than game mechanics to be played with another remote player.

Having an overall goal of strengthening intimate bonds and encouraging and maintaining the possibility of communication in distant communication in a playful manner in mind, this study is initiated by the initial objective of facilitating intimate communication across a distance by means of novel and playful interactive telepresence technologies. A number of projects from HCI research have focused on addressing the needs for shared family activities over a distance. For instance, Video Play (Follmer et al. 2010) explores the design space of video-enhanced distance play. Despite, the large variety of the attempts in this area (Saadatian et al. 2013a), there are still plenty of unaddressed issues. By reviewing the previous studies, it was noticed that firstly current developed technologies for mediating intimacy are limited in terms of embodiment. By dividing these technologies into two classes of virtual and physical interfaces, it can be seen that the category of virtual interfaces, regardless of their degree of holism, lacks the tangibility. Therefore, they are limited in terms of spatial dimensions, haptic communication, and space sharing. Therefore, the virtual media might not be the ideal platform for mediating intimacy. Another category of these technologies includes tangible and physical interfaces. However, the developed interfaces in this category are rather abstract and are limited in anthropomorphic features in which anthropomorphism, or personification, is an attribution of human form or other characteristics to anything other than a human being. Therefore, nonverbal signals related to body languages and human emotions are filtered in this class of intimate computing technologies.

Both physicalness and anthropomorphic nonverbal cues are very crucial in providing the illusive presence (Saadatian et al. 2013b) with the intimate partner which are not fully conveyed in current prototypes. To the best of our knowledge, despite the potential of humanoid telepresence robots in teleporting nonverbal cues related to anthropomorphism, humanoid telepresence technologies are not explored for intimate telepresence.

One of the aims of this research is maximizing the possibility of connectedness. Generally bidirectional, real-time interaction is only possible when both communication parties are available. Since in LDR time and contexts are different, the chance of availability of both partners is reduced. For instance, it might be midnight in the geographical location of one partner, while the other partner is interested to communicate. Another example is the situation when one of them is busy and the other one is idle and wants to interact. This opposes to the need of ongoing connectedness, which is even more salient in LDRs. Therefore, there is a need for facilitating ongoing connectedness in LDR.

The ubiquitous technologies with the goal of *ongoing connectedness* are very close to the awareness systems; however there is a subtle difference between them. In a "content-oriented communication," focus is on the exchange of information, whereas in *connectedness-oriented* communication, maintaining relationships and fostering a sense of connectedness is the core of the interaction (Kuwabara et al. 2002). In a situation that awareness system has the main goal of maintaining the human relationships and sustaining strong tie relations, the benefits may come not solely from sharing of awareness information but more from the simple act of mutual exchange (Vetere et al. 2009).

The ubiquitous technologies are often developed for information-oriented purposes such as those coping with aspects of an environment (e.g., navigation systems), those assisting in time and resources management (e.g., diaries and personal sensors), and those inferring situational opportunities for personal purposes (e.g., context-aware systems) or those that display location information or health status (e.g., personal digital assistant technologies) (Vetere et al. 2005). However, these are rarely used to sustain relationship ties. Due to the priority of possibility of the connection over the content of the interaction, this study aims to develop a control mechanism of a probabilistic behavior generator for telepresence agents. This model could maximize the connectedness by maintaining the interaction while users are not simultaneously available.

Methodology

The main methodological approach of this study is anthropologically inspired robotic, which is performed based on iterative multilevel modeling (Saadatian et al. 2014c). At each level different dimensions of intimacy are studied. Physical and cognitive intimacies are designed and developed using an iterative prototyping methodology. The other type of modeling in this study is modeling the artificial intelligence (AI) for the control system which is associated with emotional intimacy. The first levels of modeling are physical models (robots), whereas the AI model is a stochastic model designed for the agent's behavior control.

In this study the anthropologically inspired-based robotic approach is proposed for recreating the presence of an intimate partner through artificial agents. Anthropology is the study of humankind, which draws and builds upon knowledge from the social and biological sciences as well as the humanities and physical sciences (Duranti 2009). Therefore, anthropologically inspired robotic could be classified as a sub-category of biologically inspired robotics. This field of biologically inspired technology has evolved from creating static copies of human and animals in the form of statues to the emergence of robots that perform realistic behavior (Bar-Cohen and

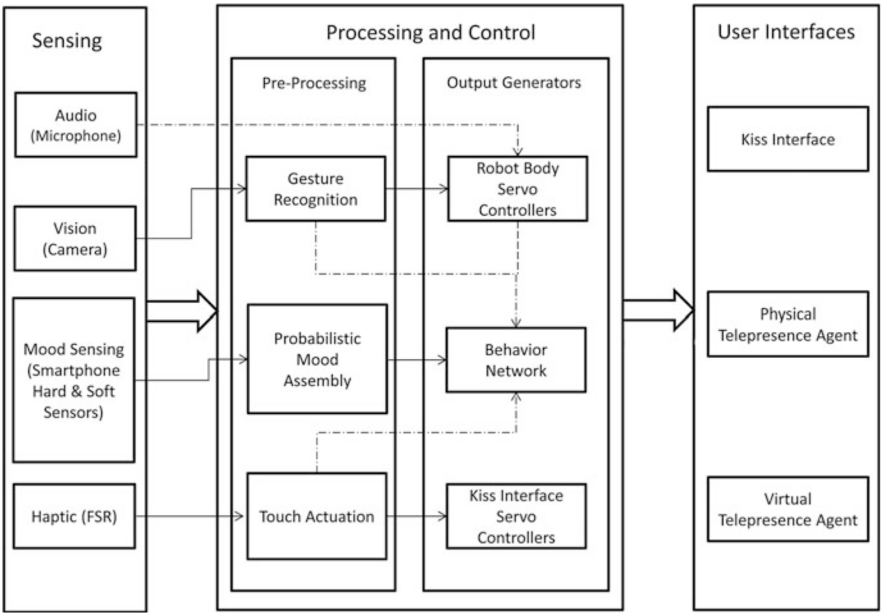


Fig. 1 Overall system architecture, showing sensing, control, and user interface in three different dimensions of intimacy

Breazeal 2003). Biologically inspired robotics as the name implies is about making robots that are inspired by the biological systems. Biomimicry and bio-inspired design sometimes get mixed up. Biomimicry is copying the nature while bio-inspired design is learning from nature and making a mechanism that is simpler and more effective than the system observed in nature (Reap et al. 2005). The proposed mechanism consists of three levels of sensing, control, and user interface as illustrated in Fig. 1.

Methodological Approach in Controlling the Telepresent Robots' Behavior

Overall, three types of control mechanisms are implemented to control the telepresence robot's behaviors in relation to each of the intimacy dimensions:

– Cognitive intimacy: Recognizing the body languages generated by remote users, using video-based gesture recognition technologies, and recreating them via the personalized physical agent on the other side
– Physical intimacy: Detection of the amount of force on the telepresence robot's lips and recreating the corresponding force on its paired agent on the other side to emulate remote kissing

– Emotional intimacy: Estimation of the mood states of the users using smartphone data, teleporting, and reacting to them

Also, behavior networks are adopted to decide which goal should be triggered to control the robot. The reason that this model AI could also act as the controller of the whole system, since it can decide which nonverbal behavior should be performed. Therefore, it could, for instance, trigger kissing, waving, or appropriate facial and body expressions, considering the inferred contextual data.

In the following chapters, each of them will be described in more detail.

Methodological Approaches in the Design of the User Interfaces

The design methodology applied in this research is iterative prototyping. Prototyping is an activity with the purpose of creating a manifestation that, in its simplest form, filters the qualities in which designers are interested, without distorting the understanding of the whole (Lim et al. 2008). Iterative prototyping is a prototyping methodology based on a cyclic process of design, test, refine, and evaluate. Considering the results of testing the latest iteration of a design, modifications and refinements are made. The ultimate goal of this process is improvement in the quality and functionality of the prototype.

Iterative prototyping starts with low-fidelity prototypes such as sketches. Low-fidelity prototypes allow the capture of rough ideas quickly and cost-effectively with emphasis on the big pictures and support design thinking. They facilitate brainstorming and also their simplicity allows elicitation of users' suggestions without limiting their creativity (Buxton 2010).

After several iterations, the early stage prototypes evolve gradually to medium-fidelity prototypes. Medium-fidelity prototypes are good for futuristic ideas. One example of this level of prototyping is prototyping with a computer, such as a simulated or animated version of some, but not all of the features of the intended system. In medium-fidelity prototypes, approaches such as the Wizard of Oz (WOZ) or scenario-based evaluations are used. In WOZ evaluation method, a human simulates the system intelligence and interacts with the user. Medium-fidelity prototypes can be evaluated within lab environments and provide a design framework for futuristic ideas (Karat et al. 2003). More evolved prototypes could be evaluated by case study methods and within their real usage environment.

The following three sections will summarize each level of the modeling in details.

Physical and Cognitive Intimacy

This section describes the first level of multilevel modeling. Assuming that the ideal goal is creating dual presence of a remote person, the development procedure is initiated by the development of a pair of small, physical surrogates of the interacting partners as the basic platform, which are named Mini-Surrogates.

While methods such as video conferencing are commonly used in telecommunications, in this study the focus is on tangible interfaces. The reason for focusing on physical avatars is mainly the potential of telepresence robots in having physical effects, such as touches and moving objects in remote locations. Therefore, they could be a suitable platform for intimate communication. The platform is then evolved stage by stage and in the form of independent models toward the goal of supporting the sensation of direct face-to-face communication.

Creating Mini-Surrogate

In this section the step-by-step design procedure of Mini-Surrogates with the aim of designing a telepresence system for couples in long-distance relationships (LDRs) is described. Iterative and participatory design procedures are described in detail.

Iterative Design Process of Mini-Surrogates

Mini-Surrogates were developed in three iterations that progressed from generating the idea by sketching, toward defining the form factor of the device, toward the feasibility of the implementation and personalizability of the prototype.

First iteration of Mini-Surrogates: For the first iteration, in a focused group session, design ideas were generated and then detailed by drawing idea sketches. Participants were asked to generate design ideas on the appearance of physical avatars for telepresence. At this stage various shapes of the interface were explored. The main issue was that it was preferred to have an interface with a natural form factor similar to human and small enough to make it possible to be carried along.

Second iteration of Mini-Surrogates: Armed with the feedback from sketching step, a lifelike and exact copy version of two of our team members were developed as a minimal physical prototype of the second iteration. They were 10 cm in height and have a slim body which makes it easy to hold like a mobile phone. These prototypes were made without internal circuitry for the purpose of facilitating brainstorming and investigating the required modifications and technical implementations.

Using the second concept prototype, concept assessment and feasibility study were performed in a focus group with nine heterosexual couples with previous LDR experience. The participants were chosen from the student and staff body of the National University of Singapore.

The study was initiated by briefing the users about the overall concept and its vision. It was followed by encouraging the participants to puppet with the prototype while at the same time the participants behavior was observed and their verbal feedbacks during the practice session was inferred. Our goal was to know whether the users could envision the merits of the communication through this sort of physical representations over the virtual (nonphysical) forms. We also wanted to know how the design should be modified to afford its goals.

In general, users were positive toward the idea, especially the likeness of the robot appearance to its users received positive feedbacks. One important issue that arose was related to the possibility of interface personalization. The fixed face, body, and clothes of the concept prototype distracted from the goal of personalizability. It was preferred to slightly modify the interface to make it possible to personalize the appearance of the robots for each user.

Third iteration of Mini-Surrogates Considering the feedbacks from the previous stage, several design modifications were made. Also, to make it practically possible to develop the real robotic interface with internal circuits and with a low cost, the size was slightly increased. The new prototype size was 35 cm in height, which in turn made it embraceable. This version is customizable to look like a specific person by changing the hairstyle, attire, and accessories. This is similar to the way that virtual avatars are customized, but it is done in the physical world. This flexibility facilitates the usage scenario by giving the choice of accessories for the robot to users. However, to avoid the biases in the formation of the relationship, the robots were personalized for users, referring to their photos, prior to the experiment (Rae et al. 2012). Also, in this version in order to minimize the uncanny feeling which might be perceived by realistic design, a toylike design was targeted instead of an exact copy of the users as was done in the previous version.

Toward Anthropomorphism of Mini-Surrogates

In the design of the robot, the body movements and size have common advantages in a co-located human–robot interaction. It helps gesture reproduction, and its minimal gesture reproduction, aesthetics, and unrealistic appearance might reduce the risk of uncanniness (Mori et al. 2012), although more empirical evaluations are still needed to prove this issue (Dautenhahn et al. 2009). We decided to make the head of the robot larger than normal size similar to cartoon characters. Assuming that a more engaging and playful appearance will make it more pleasant and engaging, which could subsequently foster the sense of intimacy with the robot. The other reason for the decision on the head size was to build a platform which will allow integration of facial features and expressions in the future. Also referring to the studies by (DiSalvo et al. 2002; Blow et al. 2006), robot’s wide head, detailed eyes with eyelash, nose, chin, and facial curves and skin are designed with the aim of balancing humanness, robotness, and productness.

The articulated degree of freedom for MVP includes:

– Move backward and forward and turn.
– Look side to side and up and down.
– Raise and lower arms.
– Curl and extend forearms.

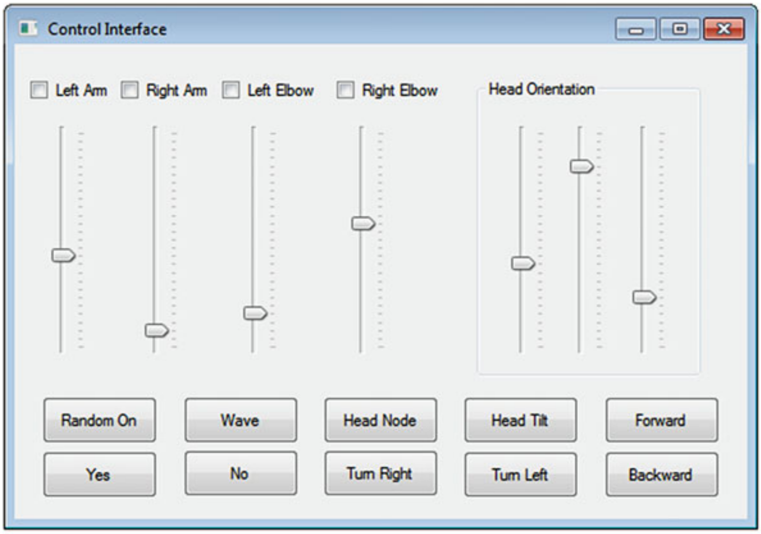


Fig. 2 A screenshot of the teleoperator control GUI; the operator can control different body parts by dragging the sliders

Prior to the implementation phase, four major specifications for the robots were considered:

– The robots have two modes of operation including, semiautonomous and teleoperation mode. In semiautonomous mode some of the robot movements are contorted by remote user's gestures. In teleoperation mode the user controls the robot using a graphical user interface (GUI). In both modes a user should be able to control the movements of the robot remotely through a wireless connection.
– The robot should have two variants (form factors) of a male and a female.
– The head should be easily replaceable as each robot is differentiated by the design of the head, clothes, hair, and accessories.
– Electronic components can be integrated into the base of the robot. No internal components should be visible to the user.

Teleoperation Control Mode of Mini-Surrogates

The teleoperator interface enables the operator to control the robot's body parts remotely by dragging the sliders. The operator can see the robot in remote location using a camera. The sliders correspond to predefined motions such as head node, waving, coming forward and backward, turning, etc. These motions will be executed in priority to the automatically detected movements (Fig. 2).

Implementation of Semiautonomous Control Mode of Mini-Surrogates

To achieve the required embodiment, initially two small humanoid robots with the ability to perform basic gestures were prepared. The robot's look was customized with cloth, hair, face, color, and accessories, making it resemble two members in our laboratory. At the end of the masking process, personalization of the robots was achieved through the following points:

– Each robot was made to resemble its owner in body features including clothing and hairstyle.
– The robots were made to be controlled remotely by the owner and each robot is equipped with a camera, so that the owner can view the surroundings of his or her own Mini-Surrogate.
– The robots were able to mimic minimum human gestures such as coming backward and forward, turning, waving, moving head backward and forward, nodding by pan, and tilting to express agree, disagree, timidness, excitement, shyness, etc.

Mini-Surrogate consists of two robots that communicate remotely. They recognize and reproduce some of the body languages of the person they represent at the remote location. An overall framework of the system from gesture detection until reproduction through robot is as below:

– User performs a body language.
– Software identifies the performed body languages sensed by camera and defines the correspondent movement for the remote robot.
– Data is transferred to Arduino using Xbee.
– Motors are enabled accordingly.
– Similar body language is performed by the robot.

The following sections describe the technical details on body language recognitions.

Head Gesture Recognition and Replication for Mini-Surrogates

In order to facilitate the robot with the ability to imitate the body languages related to the head automatically, a head pose and orientation control interface was implemented. To achieve that a software library named faceAPI developed by Seeing Machines (Seeing Machines, faceAPI, http://www.seeingmachines.com/) was used, as the head-tracking engine (Fig. 3). faceAPI can identify the position and orientation of the detected face in radiant, which makes it suitable for the purpose of this study. It enables us to track the face within one degree accuracy and the head motion in ±80 radiants can be successfully detected. It is also robust to partial occlusions, illumination, and variation of skin color and glasses. This library

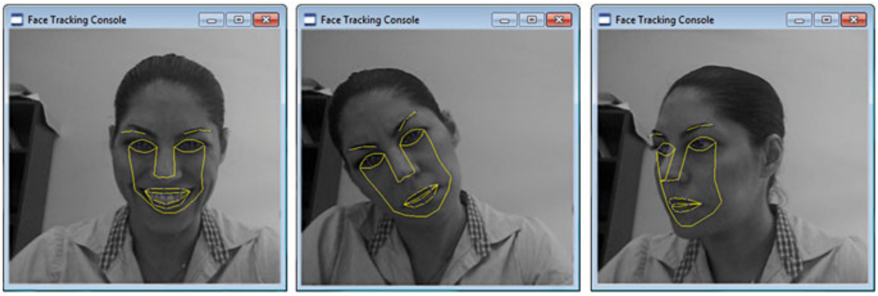


Fig. 3 A series of head-tracking snapshots showing different head orientations and pose tracking. The software tracks the location and orientation of the user's face with respect to the cameras

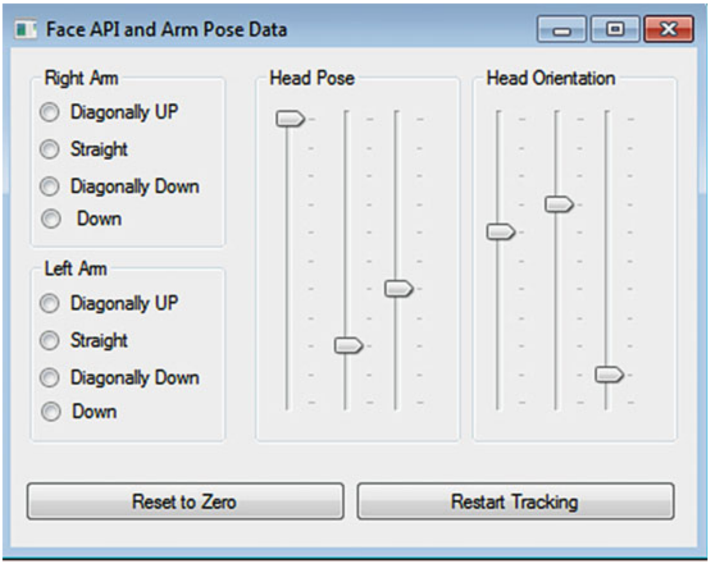


Fig. 4 The dialog shows the data related to the head pose and orientation and facilitates robot head control

uses the video stream from the webcam as input and detects the human face, as well as the lips and eyebrows. After detecting the face, the 3D position of the face in the scene and its orientation are estimated. After getting a research license from this library, an application was implemented that could send these values to our control interface to control the robot's head movements (Fig. 4). The z depth of the face is mapped to the head extend, the y rotation to the head pan, and the x to the head tilt. This provides an intuitive way of replicating the remote user's head on the robot without the need for manual operation.

Implementation of Arm Gesture Recognition and Replication for Mini-Surrogates

The internal structure of the Mini-Surrogate, for the purpose of arm gesture replication, consists of four working servomotors, for the left and right arms and the left and right shoulders. The robot behaves based on the gesture input detected by the user's web camera, located in the surrounding space or in the user's computer.

Since our aim is developing a technology probe, our priority is choosing a rapid prototyping method with minimum cost that uses only a basic camera and can be implemented rapidly. Therefore, using Kinect, or any other gesture recognition sensor, or complex machine learning algorithm was avoided. Instead, we used a simple reliable and open-source algorithm that is based on histograms and statistics. To achieve that a computer vision-based arm gesture control interface is implemented using Visual C#, which analyzes horizontal and vertical histogram of the video frames for detection of movements related to the hands. The algorithm provides a robust detection with little sensitivity to lighting, distance from the camera, and overall the quality of the video.

The step-by-step description of the algorithm is as follows*:*

The gesture recognition module identifies data related to the movements of the robot's arm. The arm movement recognition can detect the following eight poses for each arm, which corresponds to 16 gestures:

- Raised diagonally up
- Raised straight
- Not raised
- Raised diagonally down

Initially the arm region will be detected by adaptive background subtraction (for updating the background), and then the waving motion will be recognized based on the existence of information of different regions of the images and the sequence of their appearance.

In background subtraction, using difference filter, the object extraction is performed and the object will be separated from the background scene. Then the difference image is threshold, in order to classify pixels to two groups of substantial change referring to moving object and minor change, which refers to the background. After thresholding, using opening filter, any potential noise is removed, so that we have an image which shows the motion (user's body) area.

Since there is the possibility of minor changes in the background during the interaction and getting frames that include objects other than human body, the background update is needed. The minor changes could be variation of light condition, appearance of small objects, etc. In this stage the initial frame, which was identified as background, is updated.

Adaptive background implementation involves finding the biggest object in the image and if it was not considerably big (smaller than 20 * 20) adding it to the background and updating the background image. Filter is for replacing the background image with the new frame with slight changes. When the human body could

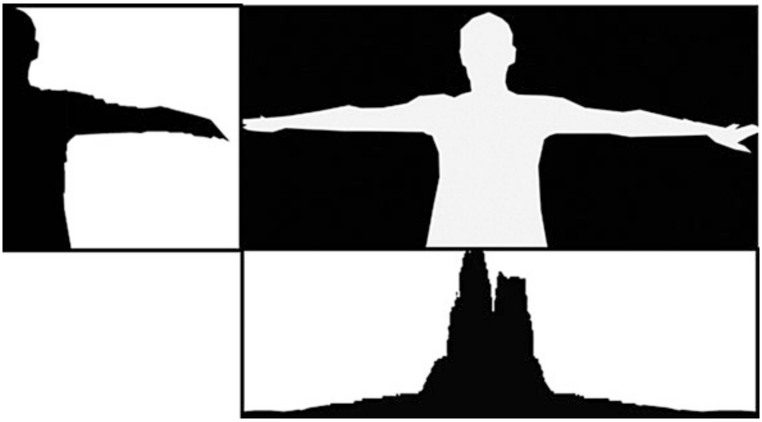


Fig. 5 Horizontal and vertical histograms are needed to be further analyzed in order to identify the arms and their gestures

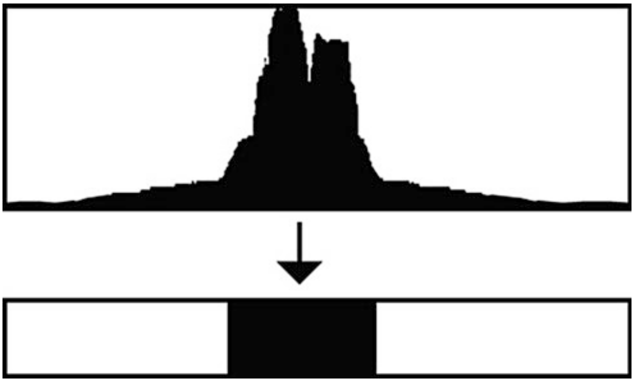


Fig. 6 The area between the blank areas represents the width of the torso

be identified as an object, gestures could be analyzed. However, to proceed with detecting the arm gestures, the subregions that are occupied by arm blocks should be identified first.

Arm blocks are identified by analyzing object both vertical and horizontal histograms of the image as shown in Fig. 5.

In order to achieve that, horizontal histogram is used to separate the hand area from the upper body area by applying the threshold coefficient value of 0.3. This is based on the fact that the human hand thickness is less than 30 % of the height of the torso. Then the hand's length and the upper body's width should be figured out from the threshold horizontal histogram. The width of empty area on the right of the histogram is equivalent to the length of the right hand, and the length of the left hand is the empty area on the left of the histogram. The area between the blank areas is the upper body's width as shown in Fig. 6.

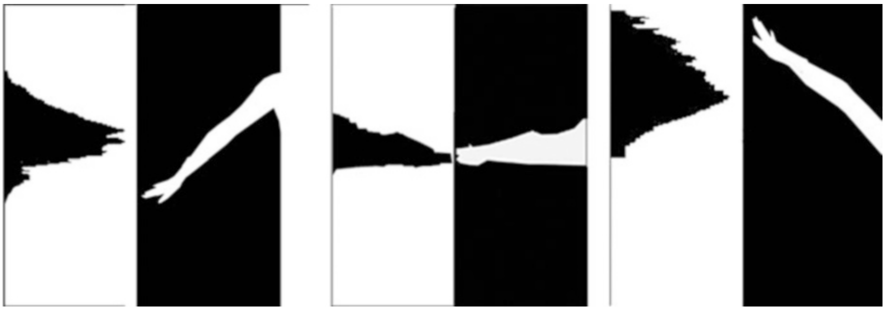


Fig. 7 The histograms related to different classes of raised hand gesture

Up to this stage, it is known the hand's length and the upper body's width. Using the body proportions and statistical assumptions, we can say the hand is raised or not. After that, we can figure out if it is raised diagonally up, raised straight, or raised diagonally down. In case that the hand is not raised, its width on the horizontal histogram should not exceed 30 % of the torso's width. So, if it exceeded 30 %, it means that the hand is raised and the direction should be recognized.

In order to recognize the direction of a hand which is raised initially, a preprocessing is performed. The preprocessing is applied to remove possible noises and shadows from the image. After getting a noiseless vertical histogram, the following approach is taken to detect each of the classed including straight hand, hands up, and hands down as shown in Fig. 7.

As can be seen from the histograms in the diagonally down hand, the pick of the histogram is shifted near the center, the pick for diagonally up hand is near the beginning of the histogram, and for straight raise hand, the pick is thin and high (this is because hand's width is much smaller than its length). By finding the histogram pick, each of the above hand positions could be identified. So, by having 4 hand positions in each hand, 2^4–16 hand gestures could be recognized.

The information values corresponding to the hand gestures are then communicated to the device serially via a wireless module, which enables the specific motors of the Mini-Surrogate to perform the required gestures. The servomotors are designed to go through a maximum angle of 130°. The application implemented communicates with the Arduino's Xbee by setting up a wireless connection. Once the connection is made, the required data are transferred to the Arduino module which processes it and sends the replication-related parameters (angle and speed) to the respective motors. The selected motors use the received information to perform the required movement and replicate the gestures performed by the user.

In order to evaluate the performances of the developed arm gesture control algorithm, the method was tested by eight untrained users (four males, four females with different body sizes). During the test the lighting and the distances from the camera were kept approximately constant. Each user tried all 16 gestures. The average rate of successfully performed actions by the robot was 97.6 %.

Summary and Design Insights

The concept of enclothed cognition and perspective of telepresence robots as mediators of interpersonal communication were explored. Then the design philosophy and the process of personalization and development of minimally expressive pair of telepresence robots were described. The goal was to provide a medium to act as a remote person's representative to be connected, to be aware, and to interact expressively while people are apart. Therefore, a pair of robots were built to ask specific questions about the effects of reproducing the likeness of a remote person and enabling them to provide us with their desired nonverbal cues that facilitate the expression of their emotion while remotely located. In spite of the robot's restricted functionality, low-fidelity prototypes assist design collaborations as a participatory design instrument (see Yim and Shaw (2011) for the details). A participatory design method using abstract robots facilitates finding the user's expectations toward the system for each given usage context. It helped us to discuss the basic form factors, communication, and movement parameters in participatory design.

First iteration was paper sketches as a result interface with natural form factor similar to human and with portability was decided. Based on the feedback from the first iteration, portable-size lifelike exact copy of two team members was developed. Then the second one was used in concept assessment and feasibility study. As a result the likeness of the robot received positive feedback, but customizability was required. In the third version again, result of the second iteration was used and the robots were made more customizable and also sizes were increased to be able to accommodate the internal circuits. This third prototype, which was a medium-fidelity prototype, was used to test the possible influence of personalization on affectivity and exploration of the design space. Therefore, the prototypes gradually improved from a very high level concept to tangible interfaces that could easily engage the users in design space exploration. Now that the design space is explored, a design framework is resulted from a user-centered design process. This framework could be used in collaboration with industrial partners to be commercialized and develop high-fidelity prototypes.

An experiment is conducted to evaluate the perceived affectivity of the Mini-Surrogate robots in enhancing the quality of communication. In this experiment, Mini-Surrogate was compared against the simple generic robot and the significant difference in affectivity was perceived. The findings are also inline with the media equation theory (Weiss 2009). According to the media equation, "people react to the media in the same manner that respond to other people in daily social interaction." "For instance, people are attracted to other people whose personalities are like themselves." This might suggest that they probably would respond to their partner's surrogate in the same way that they respond to their real partner.

Up to this step of the study, it was concluded that the Mini-Surrogate telepresence robots could potentially facilitate expressive communication. However more intimate nonverbal signals were required to be reproduced to facilitate a more intimate experience. Research on creating a more customizable and refined prototype with multiple sensory feedbacks is suggested for further studies. More work must be done to produce a technically mature interface to aid the actual use of such underutilized but potentially useful communication medium.

Two main issues, which need improvement, were discovered during this outside–inward prototyping practice. One is about the multiple sensory feedbacks and another one about the behavior control and functionality of the system:

– The physical avatars have lots of potential for mediating intimacy. Physicalness enables the medium to have effects on the environment and people, such as moving objects or touching other entities in the environment. Up to this stage not much benefit was taken from the physicalness of the prototypes. The only merit of the physicalness which was applied in this study was the spatial dimension and tangibility, which are limited in virtual forms. Haptic feedback from the agent could address this issue.
– The second major issue faced was about the control mechanism of the telepresence robots. In the current prototype, the gesture recognition technology was used to detect the body languages. There are two issues with this approach. One is that both users need to be available in front of the camera to be able to use the system. The other one is that there would be a lag between the user's body languages and the surrogate's mimics on the other side. This could minimize the practicality of the system. To address this issue, an AI model is to be introduced to infer the user's state through the smartphone and generate the appropriate behavior for the robot accordingly. Smartphone is chosen, since it is widely and frequently used in the contemporary society.

Emotional Intimacy

From the previous experiments, it was learned that the control mechanism of the affective telepresence robots needs to be improved. One reason for this decision was maximizing the intuitiveness and opportunities of the connectedness by automatic behavior generation. Gesture recognition-based approaches for telepresence, regardless of their challenges in robust recognition, can only work in online mode. To facilitate connectedness in off-line mode where both users cannot be simultaneously available, an AI model is developed, which is the main focus of this section. The AI module performs the following actions:

– Infers the stochastic mood state of the active mobile users based on the smartphone data
– Generates expressive behaviors such as the facial expressions and body languages corresponding to the estimated mood state through a synthetic agent
– Expressing mood transitions based on the personality types
– Generates feedbacks to the mood, behavior, and context of the local user on behalf of the remote user

Affective Behavior Generation System

Affective behaviors are an embodied reaction of pleasure and displeasure, which are influenced by personality, mood, emotion, and many other internal and external factors. While emotions are very short-term affective states, moods change slowly (Moshkina et al. 2011). This study focuses on affective behavior generation on behalf of a remote user, considering his/her personality type. By inclusion of the personality mood state, transitions could be related to the personality type of a remote partner and as a result attribute the agent's behavior to the partner (Saadatian et al. 2014a, b).

The agent will generate three types of behaviors:

– Imitation of the remote partner's mood through facial expressions and body languages.
– Express feedback to the user's mood state (on behalf of the remote partner), considering the remote user's personality. For instance, cheering up the local user, when the user is sad and mirroring the local user's mood when is happy.
– Express time-based behaviors (on behalf of the remote partner). For instance, saying good morning when the local user's time is morning.

To identify the personality type, the *big five personality* model is applied (Cobb-Clark and Schurer 2012). The big five personality model is a framework to determine the personality type of a person at the broadest level of abstraction. This model has received substantial support in psychology (Hegel et al. 2008). The five models of personality consist of extraversion (E), agreeableness (A), conscientiousness (C), neuroticism (N), and openness (O).

Valence–arousal two-dimensional space model (Ouwerkerk 2011) is used to generate the mood state of the users. The horizontal axis shows arousal which ranges from displeasure to pleasure, and the vertical axis is valence, which infers the user's state from calm to arouse. The model represents nine basic emotions coordinate in 2D valence–arousal plane.

The architecture of the affective behavior generation module of the telepresence agent is shown in Fig. 8. The inputs of the system are the smartphone hard and soft sensory data and manual entry of the users. The low-level data from the smartphone sensors will be processed in the perception system, and high-level data describing the state of the user will be recognized. The high-level data which describe the mood state of the user are sent to the mood and gesture generation system to produce the facial expressions and body languages of the agent. Then the user facial expressions and body movements could be transmitted to a remote user in the behavior form of a robotic agent or a virtual character.

Smartphone Sensing and Perception

Smartphone contains plenty of data about its active users that can be obtained through its built-in hard and soft sensors. The relevant data can be extracted and

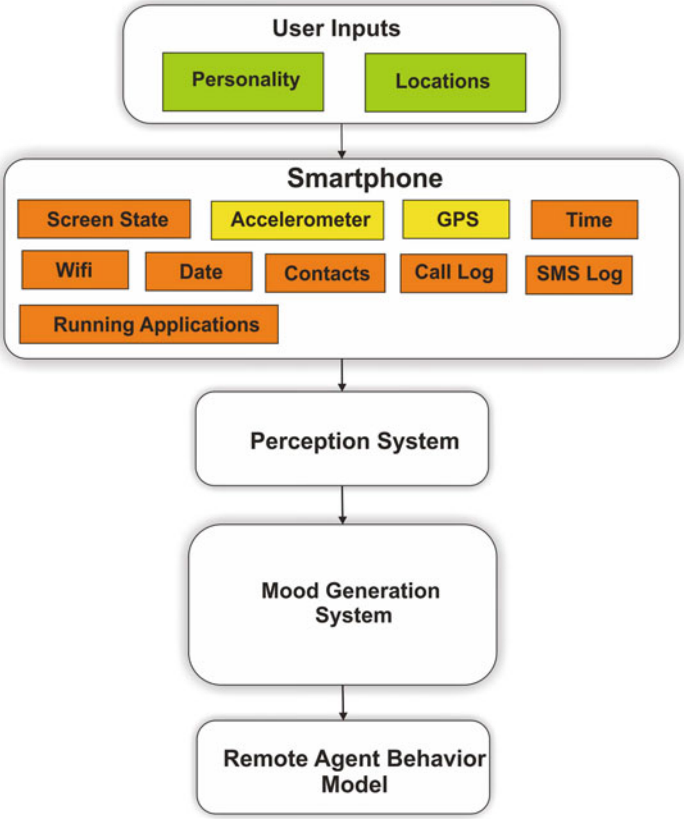


Fig. 8 Architecture of affective behavior generation of automated telepresence agent

processed to estimate the user's state. For instance, the user call logs can identify the number of phone calls, duration, and the most contacted person by the user. The running applications can reveal what types of applications are used. Moreover, the GPS and accelerometer can describe the activity state of the user. Such data are useful to determine the user's state such as idle or busy.

In our proposed method, in the smartphone data extraction phase, some data will be requested to be manually logged by users. The required manual entries by the users are their home and office locations and answers to standard questionnaires that reveal their personality type. Location labeling is done by tagging the longitude and latitude value of their home and offices, which is detected by Global Positioning System (GPS).

Table 1 describes the used sensors and their components. The automatically sensed low-level data from the soft and hard sensors combined with the manual entry of the user is gathered in the perception system.

Table 1 Soft and hard sensors of smartphone used in the perception system

Type	Class	Components
Hard sensors	GPS	Latitude, longitude, speed
	Accelerometer	X, Y, Z
Soft sensors	Time	Clock/date
	Screen	State: on/off
	Applications	Type, number
	Call and SMS log	Number, name, duration
	Wi-Fi	Wi-Fi name
	Date	Weekend, weekday, special day

This study has focused on mood sensing. Therefore, those sensors that could be associated to the mood are chosen. These data are retrieved in real time using "Funf" open sensing framework (Aharony et al. 2011) and processed in the server.

Probabilistic Mood Estimation System

The main part of the mood estimation system is a dynamic Bayesian network (DBN). Manual user inputs are needed to label their home and office locations. *GPS* data are processed with regard to the user inputs to detect the user's home and office locations. Also, the DBN needs to know the user's activity status as *idle* or *moving* as an input to the DBN model.

Accurate activity detection from smartphone is a complex topic, which is still in its infancy and is beyond the goal of this study. However, in this study to have a working model, only a basic classifier using support vector machines (SVM) is developed, to classify the user's activity level into two levels of idle and moving.

The above data combined with other sensor data are analyzed using our novel designed DBN to estimate the stochastic mood state of the user. Due to the high-level of uncertainty, dynamic changes of smartphone data, and insufficient data about the user, DBN is used. The Bayesian network is an established formalism for inference under uncertainty (Gu et al. 2004; Yoon and Cho 2012). In BN, on the top of the network, there are observed nodes, whose values are changed based on the smartphone sensor data. After that the conditional probability is implemented from the top of the nodes to the valence and arousal nodes using the chain rule. DBN will estimate the probability of nodes using the conditional probability distribution in time-series data (Murphy 2002).

Figure 9 shows the probabilistic mood estimation (PME) module of the system, which is developed based on DBN. It infers the probabilistic position of the mood (α_k, β_k) on valence–arousal space model in each time interval.

The conditional probability value of app usage is generated from Self-Assessment Manikin (SAM). The SAM is a 9-rating pictorial scale that is a nonverbal self-report measure of affective state using simple manikin pictures (Grimm and Kroschel 2005). This method is known to be adopted in a psychology study of affection. To

Fig. 9 Probabilistic mood estimation (PME) module for inferring mood through smartphone data

get the valence score, a set of manikins was employed for the experiment. In this set, ratings for valence were scored from one (displeasure) to nine (pleasure). Participants were asked to perceive a virtual situation whereby they use a specific application in the smartphone and rate the situation based on the SAM valence-rating system. This online survey was participated by 38 participants, and the raw result was processed such that the probability of each situation can be obtained and employed for the proposed DBN.

Remote User Mood Expression Model of the Agent

The expressed mood state on the robot depends on the agent (remote user) personality and its mood state changes on the valence–arousal space refereed by $\Delta\alpha_t$, $\Delta\beta_t$.

To consider the influence of the personality in mood transition, a reliable mapping method pointed in Santos et al. (2011) is applied to relate the big five personality model and the valence–arousal plane, which results in the personality type parameters (St_α, St_β) as in Eq. 1:

$$St_\alpha = 0.21E + 0.59A + 0.19N, St_\beta = 0.15O + 0.3A + 0.57N \tag{1}$$

Variables E, A, N, and O are the outputs of the big five personality test, in which, "E" describes the percentage of extraversion, "A" describes the percentage of agreeableness, "N" points to the neuroticism percentage, and "O" refers to degree of openness to experience. This equation translates the result of the personality questionnaire into the personality parameter. This personality parameter will be employed later to weight the mood variables on the valence ($\alpha axis$) and arousal ($\beta axis$) plane.

To express the mood transition in the agent at time t (UM_t), the manifested mood state can be calculated based on Eq. 2:

$$UM_t : (\alpha_t, \beta_t) = UM_{t-1} + \left(St_\alpha \Delta\alpha_t, St_\beta \Delta\beta_t\right) \tag{2}$$

where (α_t, β_t) refers to valence–arousal coordinates. This equation generates the mood transition of a remote user, considering his/her big five personality parameters.

Behavior Network for Choosing Agent Behavior

Besides transmitting the remote user's mood, the agent also aims to generate other affective behaviors. Since imitation is not the only behavior of the agent, a control mechanism is required to decide on the agent's behavior. In this respect, a model is needed to define the relationship between the user's and partner's mood and the synthetic agent's behavior.

In a behavior network, every node has its intrinsic attribute of precondition, add list, delete list, activation, and executable code, which determines the previous and the next link that will be activated. Preconditions are sets of condition that must be correct to activate the node. Add list is the condition that will be or remain true when the node is activated. The delete list is the condition that will be or remain false when the node is activated. The activation level will determine which behavior is executed. Then there is an executable code that will be activated when the precondition for the node is fulfilled.

In the proposed behavior network, the environment and sensing part of the behavior network come from the Bayesian network output, which consists of the user's and partner's mood and local time. There are three goals in this behavior network: imitation, feedback, and special time. The imitation goal triggers the character to imitate his/her partner's mood. The feedback goal intends to give feedback for the user's mood. The special time goal greets the user about the time change. The special time goal is triggered when there is a change in the time period such as morning, afternoon, and evening. The feedback goal will be triggered when there is a prominent user's mood that reaches a certain threshold; otherwise the imitation goal will be activated. Based on the proposed link in the behavior network, the activation level model is developed using the forward and backward spreading as there are no internal links between the behavior node. In the feedback mode, the personality of the partner is considered and represented by the animation speed. The behavior network will determine the animation of the character behavior, and the animation speed of the character depends on the personality of the partner: the more extrovert the partner, the faster the animation of the character. The designed behavior network is shown in Fig. 10.

Summary of the Emotional Telepresence and Control System Model

The main aim of this model was to support phatic communication between remote, intimate partners using smartphone high-level sensing. The phatic communication, which opposes to content communication, is the type of interaction which has a social rather than informative function. It can be considered as a communication that is low in information or data but is nevertheless high in significance and/or meaning (Vetere et al. 2009). The main objective of phatic technologies is satisfying the need to feel connected (Vetere et al. 2009). In this respect, an AI-based model is proposed to estimate the probabilistic mood state of each user and visualize it on the other side through an artificial agent. This helps to maximize ongoing connectedness with the intimate partner. This kind of probabilistic mood estimation should not necessarily communicate the exact, accurate information, since the focus is not the content. However, it should maximize the connectedness. The approach of automatic inference from the smartphone can maximize the connectedness because it supports agent behavior generation, even when both users are not simultaneously available.

The system is designed for active smartphone users. The PME module supports mood inference through the smartphone, considering the uncertain, insufficient, and dynamically changing available data. To visualize natural mood transitions in

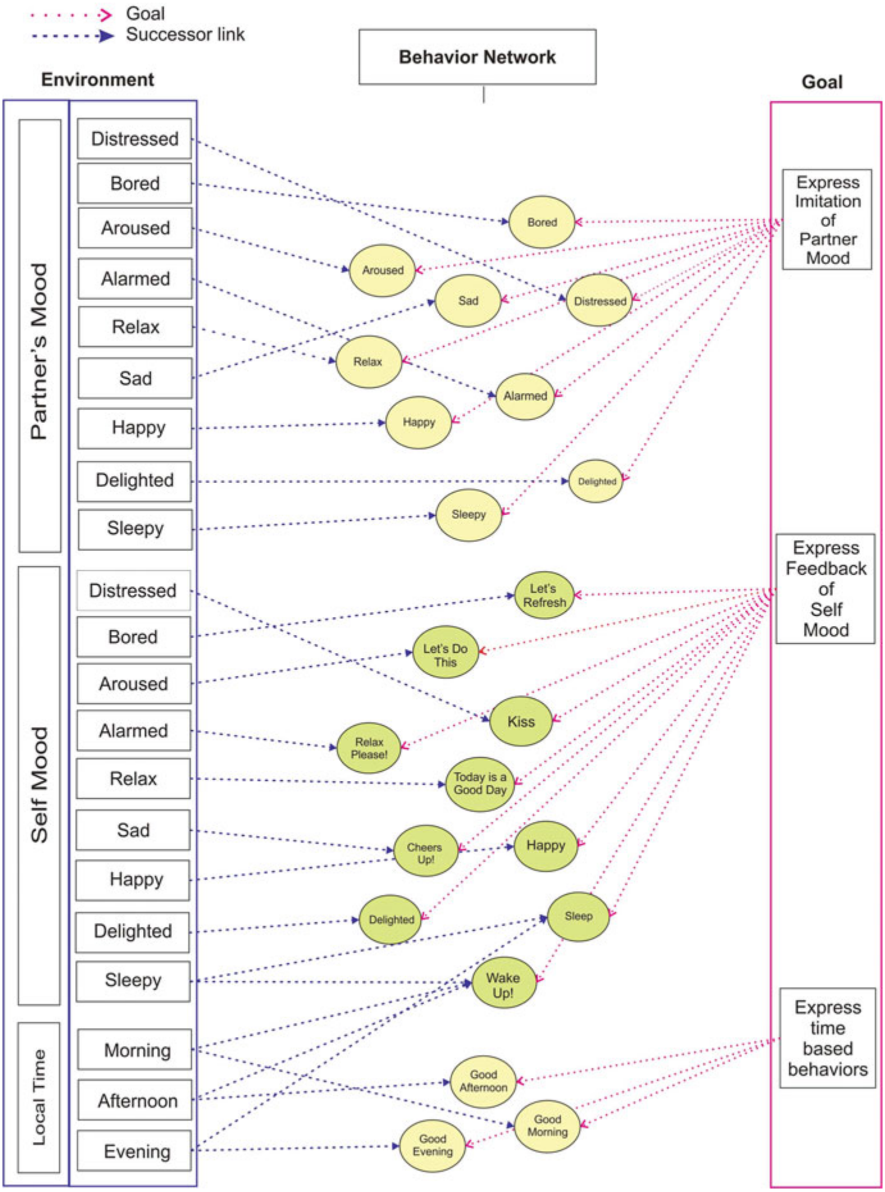


Fig. 10 A designed behavior network

a natural way as well as the association of the agent's expressive behaviors to the remote partner, the personality type of the remote user influences the regenerated mood. The association of the personality to the agent has two benefits. Firstly, it makes the agent more believable and natural (van Breemen et al. 2005). Secondly, the agent will have similarity (attribution) to the person that it represents. The system also provides feedback to the local user's affective state on behalf of the remote partner. Since the model should handle multiple goals, a behavior network is designed to handle the goal selection.

The proposed model facilitates telepresence with the remote, intimate partner by transmitting the mood and reacting to the mood state. Mood is a type of affective state similar to emotion, but it lasts longer than emotions and is a reaction to a sequence of events. Besides, mood is a private concept, as opposed to emotion that is noticeable by others (LiKamWa et al. 2013). Therefore, mood due to its private nature can reflect the underlying feeling of a person and its sharing could support intimate interaction. This study has taken the first step in smartphone-based agent control for intimate telepresence. It also introduces a novel approach for mood sensing without putting the burden of mood logging and system training on the users. Moreover, it is the first attempt in intimate telepresence that facilitates communication without the need of simultaneous presence of both partners. This is especially important in LDR communications that time and context differences are barriers in interactions.

The proposed AI model is applied on the smartphone and the success of the system is evaluated. In terms of mood inference, the model was able to recognize the user's mood state with insignificant difference from the self-perceived mood. In relation to the feedback to the user's mood, the results showed that the model could generate relatively appropriate feedback.

The study has introduced a novel application of pre-existing methods such as DBN and behavior networks for the purpose of intimate telepresence. Also, the designed DBN for affective state recognition and the designed behavior network for activity selection are the other novelties of this study.

It should be noted that the study has some limitations. First, it aims to estimate the probabilistic mood state based on the data sensed by smartphone sensors and not all the mood change triggering factors. The model is not currently perceptive to the common social and environmental factors that might influence the mood state (e.g., content of conversation, weather, health state, traffic). It is acknowledged that external and unpredictable factors cannot be considered with this approach. The system therefore cannot be expressed as an information exchange source; however, it still fulfills its objective of phatic communication. Second, since the affective state is a subjective concept due to its nature, accuracy is not measurable but can be estimated. Further research is needed to extend this model to include more contextual and internal data to assess the user's mood. Smartphone-based activity detection techniques could be also integrated with this model to teleport richer sets of behaviors. Moreover, the detected mood state depends on the defined rules for the DBN model. The current DBN rules in PME modules that define relations between users' interaction with the smartphone and their mood are based on the surveys with smartphone users and common sense. More research by social scientists and psychologist in this area could help to define more generalizable and reliable rules for PME module of the system. Besides, perceiving the affective expressions of the agent varies from person to person due to their cultural backgrounds, emotional intelligence, and social skills. Customized affective expressions for different cultural backgrounds can address this issue, which is the focus of cultural robotic researchers.

Physical Intimacy: Telepresence Agents for Mediating Intimacy Through Haptic Communication of Kisses

Motivated by the lessons learned from the previous steps, to get more benefits from the physicalness of physical avatars, a study on haptic feedback from the robot was performed. Mediated physical intimacy was explored by the study about emulating kisses. Until now, within HCI literature, there is very little research based around teleporting kisses.

For instance, *intimate mobiles* (Hemmert et al. 2011) attempt to provide realistic telepresence using a mobile phone, in which grasping, kissing, and whispering are imitated using a hand loop and the human skin heat, hydrated sponge, and airjet, respectively. However, due to the human likeness and literal emulation of the kiss, the user study reflected creepiness and unpleasantness of the system. The opposite design perspective is *kiss communicator* (Buchenau and Suri 2000) developed by IDEO, which investigates the sensual exchanges between individuals in an ambiguous way. It transmits kiss poetically by squeezing and blowing the interface. The issue with this approach is that although such an interface conveys a sensual feeling, since it is very implicit, it might not make up for the absence of actual kisses. It also does not stimulate the tactile sense the way that is expected from an actual kiss. *CheekTouch* Park et al. (2010, 2012) stimulates tactile feedback via mobile phone. It senses finger touch pattern and sends haptic feedback on the remote person's cheek using vibration. In this design approach, since a mobile phone is a many-to-many device that is used for communication with everyone, not exclusively with specific person, it opposes the private nature of intimacy (Gooch and Watts 2011). *Hkiss* Rahman et al. (2011) aims to send a kiss from a 3D virtual avatar to the physical world through haptic actuation, though the haptic feedback system is very rudimentary both from the design and the technical perspective. Therefore, although very few previous works have also focused on mediating kisses, they are very limited in terms of design exploration, anthropomorphism, and correspondence between the perceived kiss message and the original kiss generated by a remote user. To provide a more natural and bidirectional intimate kiss, a pair of telepresence kiss medium (named Kissenger) is developed for long-distance relationships through iterative design prototyping and evaluated them in their real usage context (Samani et al. 2012).

The interaction mechanism for kiss communication prototypes was devised with a number of features that might make communication of kisses between two users more meaningful. The system consists of the following key features:

1. Output kiss actuation: The kiss sensation is produced through movement of servomotors that distend the lip surface. The shape and size of the lip cover and

hide the opening of the device and the inner electronics that go into the sensing, control, and actuation of the device. This makes the user more amicable to the device and helps evoke emotional responses and feelings for kiss communication.
2. Input kiss sensing: The front of the lip has force sensitive resistors placed just below the outer surface, unbeknownst to the user. It can sense varying levels of soft touches. The force variation is sensed, digitized, and transmitted wirelessly to the receiver device. The sensors are mapped on a one-to-one basis to the actuators on the receiver device. This design simplifies the interface and enables users to form a correct and semantically meaningful mental representation of the system.
3. Control and wireless: Each device is equipped with a lip connected to an embedded circuit that orchestrates the entire system. The circuit contains an Arduino Pro Mini that controls the sensors and actuators. It can communicate wirelessly with another device. Data from the pressure sensors is read continuously until a change is detected. If there was a substantial change, the resulting increase is transmitted wirelessly to a receiver circuit that then actuates a servo motor array to produce vibration of the lips. The schematic diagram of the internal circuits is shown in (Fig. 11).

Experiment Results and Design Lessons

After iterative design of the prototypes, the final prototypes have been tested once in the laboratory and then the more robust version was tested in the field in their actual usage context.

During the lab-based experiment, prototypes were compared it against a video chat application, Skype, in a controlled experiment. Seven couples were recruited to participate in this evaluation. None of them had previously used the devices. The ages of the 14 participants were in the range from 22 to 30 years and had known their partner for 9 months or more. They have had an experience where a part of the relationship was long distance that involved some form of distance communication. The experiment took place in our laboratory meeting room divided into two partitions.

The experiment took almost 20 min per couple of participants to complete. Since our goal was to evaluate Kissenger in the context of intimate and affective communication, participating couples were asked to choose a topic and mention phrases that in their normal conversation are usually followed by kiss (e.g., a greeting). They would have to complete two scenarios including sending the kiss through the emoticons in the video chat itself and through Kissenger. Initially, each user was explained the scenarios and how to use the device and was given some time to familiarize themselves with the device before the start of the evaluation itself. After each scenario, participants completed a seven-point Likert scale questionnaire about the setup used. After trying both scenarios, in a counterbalanced order, participants completed a summary questionnaire asking for their demographic information and any feedback they might have.

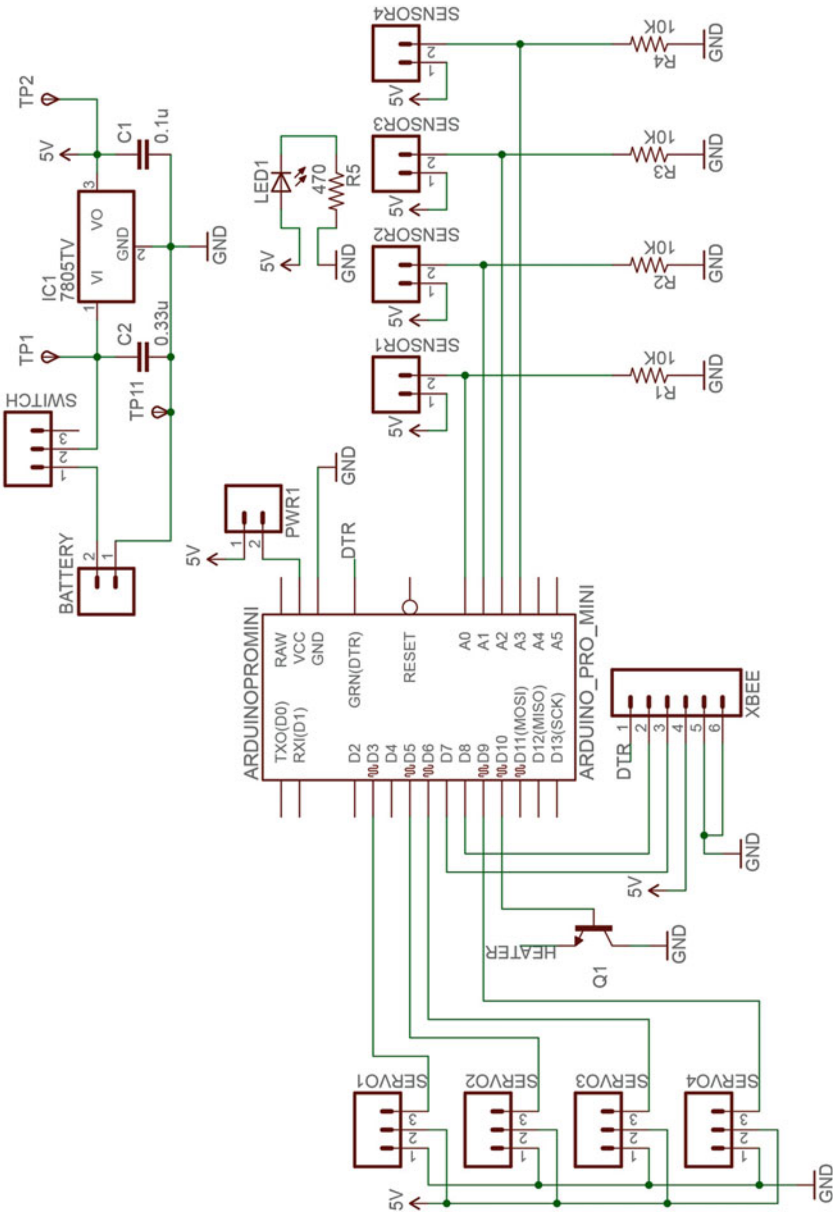


Fig. 11 Schematic diagram of the Kissenger internal circuits

The objective of our experiment was twofold. The first goal was to test whether using a mediated kiss device such as Kissenger enhances the affectivity of communication compared to common affect expression methods in telecommunications such as emotions for text or audiovisual affect expression. The second goal was to inspect the role that Kissenger has in fostering the sense of copresence; when a remote user is kissed by the partner using the device, to what extent is the illusion of being together conveyed?

The results show similar experiences for both scenarios (t(13) <1.99, n.s.) with the exception of two affectivity questions, “Does communication through this interface create an intimate experience?” and “Was the received kiss similar to a natural kiss?” for which a paired two-tailed t-test revealed that Kissenger received a significantly higher score ($t(13) = 4.59$, $p < 0.01$) and ($t(13) = 2.24$, $p < 0.05$), respectively. These results are consistent with our design goals.

The previous evaluation aimed to gather the couples’ initial reaction toward the device, to understand its usability and functionality in mediating kiss experiences, and to gather feedback for improvement. However, the results obtained from that user evaluation were insufficient to judge the device’s usefulness for intimacy due to it being conducted in a controlled context.

In the field study, in order to address the problems in laboratory tests and to develop a better understanding of the influence of Kissenger in supporting remote communication, a longitudinal field study was conducted within the context of actual everyday situations. The study was performed within 3 weeks and included a diary study combined with a pre-experiment and a post-experiment interview to discover the design space. More detail of the experimental design can be seen in Saadatian et al. (2014d, 2015).

Overall, it was found that the factor of time could affect the perception of the participants. Two different stages of familiarization and incorporation were found. In the familiarization stage, the emotional reactions were more intense and were mainly focused on serendipity, configuration, cultural reactions, and aesthetic. After 3 days, the theme of feedbacks changed and the focus was mainly around the usability and affectivity of Kissenger. These findings are in line with the previous research on temporal study of user experience and adaptation to social and physical context (Karapanos 2013; Payne and Howes 2013).

The design exploration on the Kissenger gave us many design insights. For instance, in relation to the need for association with the partner, one idea to make each device more customized is to design a simple model with accessories that can be removed and attached in a number of ways.

Also, to support usage in public, adding a symbolic emotional communication feature was also considered. Many people use “kiss,” “xoxo,” or “:-*” emoticons in their chat, e-mail, or text messaging to represent a kiss. These characters are very common in text-based communication methods between partners. Similarly, a symbolic representations for Kissenger could be developed. For example, “XXXXXXX” may refer to a very tender long kiss, while “X” may refer to a quick kiss.

This leads to another issue that was considered, asynchronous kissing which is the ability for the device to store a kiss that can be read at a later time. This feature opens up a lot of potential ethical issues from the design of Kissenger, which is discussed in the next paragraph.

Apart from questions relating to interaction, Kissenger raises a number of ethical questions. In an affirmative sense, this device could be seen as a basis for positive affection toward society. As this system could be used to transfer emotions to loved ones, especially between parents and children, it will bring physical comfort and satisfaction which is a vital need for young children when their parents are far apart.

Couples that use this device for a longer period of time may suffer from a lack of real physical affection. There also might be social ramifications for failing to return a kiss from a partner received offline on Kissenger. On the legal front, these devices also open a debate related to aspects of adultery in relationships. For example, would usage of the device with another person constitute infidelity by the partner?

Conclusion

Telepresence refers to a set of technologies which facilitates illusive presence or illusion of nonmediation, to give the sensation of being present or to have an effect, through telerobotics, at a remote location (Lombard and Ditton 2006). This requires providing the user's senses with stimuli from remote locations and enables them to affect the remote location. Additionally, it involves sensing, transmitting, and duplicating the user's behaviors, status, voice, nonverbal cues, context, and other subtle cues at the remote location. Intimate telepresence or mediated intimacy refers to a group of telepresence technologies which convey the illusion of being copresent with the intimate partner or loved ones. Intimate presence in face-to-face communication requires the physical and emotional closeness. Therefore to support intimate telepresence, researchers in this field have taken two approaches of "mediating intimate behaviors" and "provoking intimate reactions," which were inspired by the emulation of co-located intimate behaviors between couples.

Although the field of intimate telepresence is largely populated by technologies that are designed to emulate the co-located behaviors, there are still some unexplored problems that require more research. One important issue during communicating through current telepresence technologies is the need for simultaneous availability of both interacting parties, which is not always possible due to the contextual differences (e.g., location of partners, time zone differences, people around). This opposes to the need for ongoing connectedness in intimate communications (Vetere et al. 2005). In this study a new generation of telepresence technology that increases the possibility of ongoing connectedness, self-disclosure, and empathy through stochastic detection of the user's mood state and regenerating it in the remote location is presented. The model also allows generating natural, empathetic reaction to the mood state of the remote partner via the other partner's artificial agent. This artificial intelligence (AI) model serves as the control mechanism of our offered intimate telepresence agent.

Apart from the AI module and the control mechanism of the intimate telepresence agent, there are also some unexplored problems in the user interface or the medium of the interaction. One overlooked issue in mediated intimate telepresence technologies is the issue of physical embodiment. Despite the importance of embodiment and its influence in the perception of copresence, technologies developed to date for this purpose are relatively abstract. To the best of our knowledge, adoption of embodied humanoid robots as the medium of interaction for intimate telepresence is unexplored. In this study a pair of personalizable telerobotics technologies for intimate telepresence were designed, developed, and evaluated. These robots can transmit the body languages in the remote location and are designed based on the concepts of embodied and enclothed cognition and the influences of appearance and likeness on affectivity.

Recommended Reading

N. Aharony, A. Gardner, C. Sumter, A. Pentland, Funf: open sensing framework (2011)

Y. Bar-Cohen, C. Breazeal, *Biologically Inspired Intelligent Robots* (SPIE, Bellingham, 2003), pp. 5051–5058. Citeseer

M. Blow, K. Dautenhahn, A. Appleby, C.L. Nehaniv, D. Lee, The art of designing robot faces: dimensions for human-robot interaction, in *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction* (ACM, New York, 2006), pp. 331–332

M. Buchenau, J. Suri, Experience prototyping, in *Proceedings of the 3rd Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques* (ACM, New York, 2000), pp. 424–433

B. Buxton, *Sketching User Experiences: Getting the Design Right and the Right Design* (Morgan Kaufmann, San Francisco, 2010)

D.A. Cobb-Clark, S. Schurer, The stability of big-five personality traits. Econ. Lett. **115**(1), 11–15 (2012). Elsevier

K. Dautenhahn, C.L. Nehaniv, M.L. Walters, B. Robins, H. Kose-Bagci, N.A. Mirza, M. Blow, Kaspar–a minimally expressive humanoid robot for human–robot interaction research. Appl. Bionics. Biomech. **6**(3–4), 369–397 (2009)

S. Deterding, S.L. Björk, L.E. Nacke, D. Dixon, E. Lawley, Designing gamification: creating gameful and playful experiences, in *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (ACM, New York, 2013), pp. 3263–3266

C.F. DiSalvo, F. Gemperle, J. Forlizzi, S. Kiesler, All robots are not created equal: the design and perception of humanoid robot heads, in *Proceedings of the 4th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques* (ACM, New York, 2002), pp. 321–326

A. Duranti, Linguistic Anthropology: A Reader, vol. 1 (Wiley, Oxford, 2009)

S. Follmer, H. Raffle, J. Go, R. Ballagas, H. Ishii, Video play: playful interactions in video conferencing for long-distance families with young children, in *Proceedings of the 9th International Conference on Interaction Design and Children* (ACM, New York, 2010), pp. 49–58

D. Gooch, L. Watts, A design framework for mediated personal relationship devices, in *Proceedings of the 25th BCS Conference on Human-Computer Interaction* (British Computer Society, 2011), pp. 237–242

M. Grimm, K. Kroschel, Evaluation of natural emotions using self assessment manikins, in *Automatic Speech Recognition and Understanding, 2005 I.E. Workshop on* (IEEE, San Juan, 2005), pp. 381–385

T. Gu, H.K. Pung, D.Q. Zhang, A bayesian approach for dealing with uncertain contexts, in *Advances in Pervasive Computing* (Citeseer, 2004), p. 136

M. Hassenzahl, M. Schöbel, T. Trautmann, How motivational orientation influences the evaluation and choice of hedonic and pragmatic interactive products: the role of regulatory focus. Interact. Comput. **20**(4), 473–479 (2008)

F. Hegel, S. Krach, T. Kircher, B. Wrede, G. Sagerer, Understanding social robots: a user study on anthropomorphism, in *Robot and Human Interactive Communication, 2008. RO-MAN 2008. The 17th IEEE International Symposium on* (IEEE, Munich, 2008), pp. 574–579

F. Hemmert, U. Gollner, M. Löwe, A. Wohlauf, G. Joost, Intimate mobiles: grasping, kissing and whispering as a means of telecommunication in mobile phones, in *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services* (ACM, New York, 2011), pp. 21–24

E. Karapanos, User experience over time, in *Modeling Users' Experiences with Interactive Systems*, ed. by K. Janusz (Springer, Berlin/Heidelberg, 2013), pp. 57–83

C.M. Karat, C. Brodie, J. Karat, J. Vergo, S.R. Alpert, Personalizing the user experience on ibm. com. IBM Syst. J. **42**(4), 686–701 (2003)

K. Kuwabara, T. Watanabe, T. Ohguro, Y. Itoh, Y. Maeda, Connectedness oriented communication: fostering a sense of connectedness to augment social relationships, in *Applications and the Internet, 2002 (SAINT 2002). Proceedings. 2002 Symposium on* (IEEE, Nara, 2002), pp. 186–193

R. LiKamWa, Y. Liu, N.D. Lane, L. Zhong, Moodscope: building a mood sensor from smartphone usage patterns, in *Proceeding of the 11th Annual International Conference on Mobile Systems, Applications, and Services* (ACM, New York, 2013), pp. 389–402

Y.K. Lim, E. Stolterman, J. Tenenberg, The anatomy of prototypes: prototypes as filters, prototypes as manifestations of design ideas. ACM Trans. Comput. Hum. Interac. (TOCHI) **15**(2), 7 (2008)

M. Lombard, T. Ditton, At the heart of it all: the concept of presence. J. Comput. Mediated Commun. Online Version **3**(2), 0 (2006). Wiley Online Library

M. Mori, K.F. MacDorman, N. Kageki, The uncanny valley [from the field]. Robot. Autom. Mag. IEEE **19**(2), 98–100 (2012). IEEE

L. Moshkina, S. Park, R.C. Arkin, J.K. Lee, H. Jung, Tame: time-varying affective response for humanoid robots. Int. J. Soc. Robot. **3**(3), 207–221 (2011)

K.P. Murphy, Dynamic bayesian networks: representation, inference and learning, Ph.D. thesis, University of California, 2002

M. Ouwerkerk, Unobtrusive emotions sensing in daily life, in *Sensing Emotions*, ed. by P. Wierenga (Springer, Netherlands, 2011), pp. 21–39

Y. Park, C. Lim, T. Nam, Cheektouch: an affective interaction technique while speaking on the mobile phone, in *Proceedings of the 28th of the International Conference Extended Abstracts on Human Factors in Computing System*s (ACM, New York, 2010), pp. 3241–3246

Y. Park, S. Bae, T. Nam, How do couples use cheektouch over phone calls? in *Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems* (ACM, New York, 2012), pp. 763–766

S.J. Payne, A. Howes, Adaptive interaction: a utility maximization approach to understanding human interaction with technology. Synth. Lect. Hum. Centered Inform. **6**(1), 1–111 (2013)

I. Rae, L. Takayama, B. Mutlu, One of the gang: supporting in-group behavior for embodied mediated communication, in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (ACM, New York, 2012), pp. 3091–3100

M. Rahman, A. Saleh, A. El Saddik, Hkiss: real world based haptic interaction with virtual 3d avatars, in *Multimedia and Expo (ICME), 2011 I.E. International Conference on* (IEEE, 2011), pp. 1–6

J. Reap, D. Baumeister, B. Bras, Holism, biomimicry and sustainable engineering, in *ASME 2005 International Mechanical Engineering Congress and Exposition* (American Society of Mechanical Engineers, 2005), pp. 423–431

E. Saadatian, H. Samani, A. Toudeshki, R. Nakatsu, Technologically mediated intimate communication: an overview and future directions, in *Entertainment Computing–ICEC 2013* (Springer, Brazil, 2013a), pp. 93–104

E. Saadatian, H. Samani, A. Vikram, R. Parsani, L. Tejada Rodriguez, R. Nakatsu, Personalizable embodied telepresence system for remote interpersonal communication, in *RO-MAN, 2013 IEEE* (IEEE, Korea, 2013b), pp. 226–231

E. Saadatian, T. Salafi, H. Samani, Y. De Lim, R. Nakatsu, Artificial intelligence model of an smartphone-based virtual companion, in *Entertainment Computing–ICEC 2014* (Springer, Australia, 2014a), pp. 173–178

E. Saadatian, T. Salafi, H. Samani, Y.D. Lim, R. Nakatsu, An affective telepresence system using smartphone high level sensing and intelligent behavior generation, in *Proceedings of the Second International Conference on Human-Agent Interaction* (ACM, New York, 2014b), pp. 75–82

E. Saadatian, H. Samani, R. Nakatsu, Anthropologically inspired creative design for intimate telepresence, in *SIGGRAPH Asia 2014 Designing Tools for Crafting Interactive Artifacts* (ACM, 2014c), p. 5

E. Saadatian, H. Samani, R. Parsani, A.V. Pandey, J. Li, L. Tejada, A.D. Cheok, R. Nakatsu, Mediating intimacy in long-distance relationships using kiss messaging. Int. J. Hum. Comput. Stud. **72**(10), 736–746 (2014d)

E. Saadatian, H. Samani, R. Nakatsu, A longitudinal field study on kiss mediation interface for long distance relationships. In: *HCI International 2015-Posters Extended Abstracts* (Springer, Switzerland, 2015), pp. 118–122

H.A. Samani, R. Parsani, L.T. Rodriguez, E. Saadatian, K.H. Dissanayake, A.D. Cheok, Kissenger: design of a kiss transmission device, in *Proceedings of the Designing Interactive Systems Conference* (ACM, New York, 2012), pp. 48–57

R. Santos, G. Marreiros, C. Ramos, J. Neves, J. Bulas-Cruz, Personality, emotion, and mood in agent-based group decision making . IEEE Intell. Syst. (6), 58–66 (2011)

A. Scott, How playcentric research methods are contributing to new understanding and opportunities for design, in *The Routledge Companion to Design Research,* ed. by P. Rodgers, J. Yee (Routledge, London/New York, 2014), p. 400

A. van Breemen, X. Yan, B. Meerbeek, icat: an animated user-interface robot with personality, in *Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems* (ACM, New York, 2005), pp. 143–144

F. Vetere, S. Howard, M. Gibbs, Phatic technologies: sustaining sociability through ubiquitous computing, in *First International Workshop on Social Implications of Ubiquitous Technology. ACM Conference on Human Factors in Computing Systems, CHI* (Montreal, 2005)

F. Vetere, J. Smith, M. Gibbs, Phatic interactions: being aware and feeling connected, in *Awareness Systems*, Editors-in-chief T. Desney, V. Jean (Springer, London, 2009), pp. 173–186

J. Wei, A.D. Cheok, Foodie: play with your food promote interaction and fun with edible interface. Consum. Electron. IEEE Trans. **58**(2), 178–183 (2012)

D. Weiss, *Encyclopedia of communication theory: media equation theory* (Sage, Thousand Oaks, 2009)

J.D. Yim, C.D. Shaw, Design considerations of expressive bidirectional telepresence robots, in *CHI'11 Extended Abstracts on Human Factors in Computing Systems* (ACM, New York, 2011), pp. 781–790

J.W. Yoon, S.B. Cho, An intelligent synthetic character for smartphone with Bayesian networks and behavior selection networks. Expert Syst. Appl. **39**(12), 11284–11292 (2012). Elsevier

Enrobotment: Toy Robots in the Developing Brain

39

Irini Giannopulu

Contents

Abstract

Based on the internalized "object" and using cognitive, clinical, neuro-functional, and engineering arguments, this chapter analyzes the concept of enrobotment. Playing with objects/toys (including the imperceptible part, i.e., the shadow) implies that the objects/toys are part of the external environment, i.e., the "other." The enrobotment signifies that the object's internalization not only reflects the impact of the environment on child's development but it also reverberates the echo of the child's representations. An intermediate object (including shadow) is conceived in mind by the child him/herself. Having a high emotional value and forming an implicit/explicit autobiographical continuum in memory, it ensures the cohesion between the "self" and "other," and it authorizes

I. Giannopulu (✉)
Virtual Reality Prism, IHU-A-Brain and Spine Institute (ICM), UPMC, Groupe Hospitalier Pitié-Salpêtrière, Paris, France
e-mail: igiannopulu@psycho-prat.fr

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_59

subjectification. The correlated representations allow the invention of ideas and concepts; motor and verbal actions including their intention prosper. Intention attribution to objects/toys constitutes a precursor of self-consciousness, as this intention, a specific anticipation, helps children to understand what it signifies to have a perspective. Recognizing what it implies to be a "self" is a parcel of envisioning mental states of the "other." At the antipode, autism can be considered as an antithesis of self-consciousness. Children with autism cannot mirror the triadic relationship of "object-self-other." Enrobotment allows them to improve their capability to be "self," i.e., to emerge "self"-complexity.

Keywords
Object • Self-other • Autobiographical memory • Developing brain • Play

Introduction

Emanating from an interdisciplinary approach, enrobotment is a new concept. Enrobotment is a state of mind that mirrors the internalization, incorporation, and representation of animate and inanimate objects (toys or toy robots). The toy (*from Middle Dutch toy*) as an object (*tangible thing perceived or presented to the sense from Medieval Latin objectum* "*thing put before*" (*the mind or sight*)) of pleasure, recreation, training, and learning is the substratum of enrobotment. The key of understanding is that the object/toy is not only an instance of the represented object/toy but a representation deeply involved with its own shadow. The shadow of the object/toy is defined as the "imperceptible" and "insignificant" parcel of this object. Enrobotment is intimately associated with the developing of verbal and nonverbal processes (e.g., visual, haptic, motor), the portrayal of object/toy (including shadow). Enrobotment is assumed to influence intention attribution to objects. As the echo of self-consciousness, it facilitates self-consciousness.

Play (*from Old English plegan*) is spontaneous and seems to be universal. It is less present in reptilians than in birds, it is well established in mammals (e.g., Chapouthier 2009). All children whatever their culture have been playing with objects, i.e., toys, since the dawn of time. They are quite gifted to enjoy whatever adults give them. Their imaginations "leave the ground" with the objects at hand. It makes the most of anything: from a small stone to a piece of wood, a complex electronic game, or toy robot.

From the prehistoric man and the Neolithic Period, people have enveloped the prospect of acquiring new knowledge. Some of them sought knowledge for the amenity of comprehension, not the amenity of utility. In other situations, both intentions have co-occurred. It has sought understanding of changes in the creation of the toys, i.e., objects that adults built for children. Note that automata and mechanical toys are still alive and kicking even today. As their symbolic value is transcendent, it crosses the centuries and the representations. Playing with object/toy (including shadow) mirrors children's ability to be a part of the external environment. It facilitates the internalization of the objects/toys as well as child's neurocognitive construction by allowing verbal and nonverbal representations to emerge. It implements subjectification. Together with the emerged representations, the implemented subjectification forms an implicit/explicit autobiographical continuum in memory that encourages the invention of ideas and concepts. Having a high emotional value, the internalized object ensures the continuity between the "self" and "other." The concepts of "self" and "other" are intimately related to children's capability to gradually represent objects/toys. In this framework, the resulting intention attribution to objects/toys constitutes a precursor of self-consciousness, as this kind of anticipation helps children to understand what it signifies to have a perspective. Self-consciousness mirrors the representation of self.

Recognizing what it implies to be a "self" is a parcel of envisioning mental states of the "other." The development of self-consciousness is the result of a complex process with at least three foci: one in the central nervous system with the associate genetic and neural influences, one in the mind with the associated representations, and one in the child's dynamic interactions with the artificial environment, that is, toy robots. Enrobotment facilitates children's self-conscious and unconscious processes. Accordingly, the analysis of self-consciousness would anticipate one of the most common interruptions of consciousness, i.e., autism. Autism can be considered as an antithesis of self-consciousness. Children with autism have no or restricted concept of "self." They cannot mirror the triadic relationship of "object-self-other." They withdraw from social, cognitive, and emotional interactions living in a world confined to "self." Without self-consciousness, there is no conception of "me" as distinct to "you," no self-evaluative thought or emotion, and no mentalizing. Enrobotment gives children with autism the possibility to improve their capability to be "self" via toy robot. It enables subjectification.

The objective of this chapter is to increase interdisciplinary interactions among cognitive and clinical neurosciences and robotic and engineering sciences engaged in research on child and robot associations, to encourage collaborative efforts and to provide more experimental and clinical findings. The impetus for this chapter was a workshop named "Toy Robots and Developmental Cognitive Neuroscience" that has been organized recently in Japan.

Beginning with the chronicle of events of mechanical objects/toys fundamentally based on human neotenic behavior, the writing will continue with the enrobotment's whiff exploring the playing with animate/inanimate objects. Then, the enroboted memory via the building of implicit/explicit autobiographical representations will be described. An outline of the influence of enroboted memory in intention attribution to toy robots will take place. Considering that the enrobotment paves the way to self-consciousness, the specific place of enrobotment in autism will be analyzed.

Mechanical Object/Toy: Chronicle of Events

Even if it is difficult to date the appearance of the first toys, it is accepted that the origin of the toy is back to the earliest men. Prehistoric infants were playing with toys made with perishable materials which have presently disappeared. Polished small stones and shells were utilized to keep infants in activities or calm them when crying (Fig. 1). From the Neolithic Period, there are miniature objects that could be interpreted as toys. The existence of toys is attested with certainty from the third millennium in Mesopotamia and later civilizations (e.g., Breyer 2010).

Through various historical and archeological sources, it appears that the toys had a strong symbolic importance in ancient societies. The toys were made of clay, metal, wood, or bone; toys were also made of more precious material as gold. Many toys, animate and inanimate, that children use to play today already existed in antiquity: dolls, tea sets, tops, ossicles, and mechanical and automata creatures. The ambition to bring inanimate objects to life is apparently basic to humankind. As far back as ancient times, Egyptian, Greek, and Roman craftsmen managed to animate wooden statues through the use of hidden levers and hydraulic or pneumatic power, with seemingly magical results. The most famous is Pygmalion and Galatea: the “object” is transformed into a “self,” sense by sense, from smell to touch. But one of the first mechanical toys ever created was the wooden flying pigeon of Archytas of Tarentum. He is one of the pioneers of automata. He also invited a rattle intended to please infants. Mechanical toys were large, impressive, and magical, usually shown in fairs in Europe. They were commonly mechanized by pneumatic power, wind, or water that simulate “the anima.” Real-size animated statues have been used in ceremonies by ancient Greeks. About the same time, Ctesibius, an Alexandrian barber, made a singing blackbird that drank water and moved.

From the Middle Ages, any notable development in mechanical toys has been reported. For most of the toy eras, wood compounds and other perishable materials have disappeared. Only the most luxurious toys have survived. Toys are purchased from street vendors moving from village to village with their hood on the back. These are the hawkers. Nevertheless, it is common that parents are designing toys for

Fig. 1 Polished small stones

their children. The children themselves make their own toys, as well. At the beginning of the Renaissance, two important sources shed light on the different toys (and games) used by children of that time in Western Europe. In the fifteenth century, Leonardo da Vinci and Galileo Galilei had to come to the scene to get the things moving. Leonardo made a mechanical lion as a present for King Louis XII. The lion could walk and open its chest to reveal a cluster of lilies. At the same time, Galileo Galilei was making simple mechanical toys. The end of the Renaissance is the time when the clockwork gears began to power the mechanical toys. These gears allowed the toys to perform very elaborate and amazing movements and tasks. Some toys were able to play musical instruments or write whole sentences. Mechanical monkeys were used in the advertising of cigarettes and tea. They were very popular. Much later, town clocks in the sixteenth century in Germany had mechanical clockwork-driven figures that struck the hour. These were the precursors of automata life-size forms that were made for royalty in the seventeenth and eighteenth centuries. The elaborate clockwork gearing of automata enabled them to perform many amazing tasks – from writing sentences to playing musical instruments. However, by the 1850s, automata were being replaced by mass-produced mechanical toys that had spring-driven mechanisms and light, stamped gears instead of the heavy brass ones (Damamme 1998).

In the eighteenth century, Jacques de Vaucanson invented one of the first mechanical robots, the “Digesting Duck,” that was able to eat wheat and drink. Vaucanson made also amazing “Flute Player” and “Tambourine Player” with repertoires of different tunes. The spectators felt the breath coming from the lips of the players. Jacques de Vaucanson is often regarded as the greatest mechanical toy crafter of all times. He created androids. Nearly the same time, Pierre Jaquet-Droz built “The Writer,” “The Draftsman,” and “The Musician,” three of the most amazing automata ever fashioned. These automata still exist in the Museum of Art and History in Switzerland. The nineteenth century is considered as the golden age of the mechanical toys. With great inventions like the cinema or the steam engine, toys were taking new forms and manufacturers became inventors. Many patents are filed by the creators of mechanical toys that are finding success. Great discoveries and new materials that work through the ages and centuries appeared brought their lot of new toys imagined and manufactured by parents for their children. Thanks to advances in watchmaking and development of machines for punching iron, these toys are mass produced and sold at low prices. In the late nineteenth century, Fernand Martin realized many mechanical toys called “Little Martin.” The doll construction also knows a great development with the arrival of the first plastics such as celluloid. This new material allows the manufacture of infants ranging in water, the famous swimmers. Manufacturers like Colin Small Twin are emerging in the world of toys. Other toy designer dynasties as Steiff (teddy bear) and Hornby (Meccano Dinky Toys) were also emerging between the late nineteenth and early twentieth century. This is also the time of American clockwork toys. All kinds of dancing figures, steamboats, and trains driven by clock mechanisms were produced commercially.

The nineteenth century was also the time of the famous Euphonia. This creature had a movable replica of a human face and was able to sing, laugh, ask, and answer questions. The machine was inspected and it was proven that the sounds are really produced by it. The end of the nineteenth century was the time of mechanical toy banks. These toys work when a coin is deposited in their bank. Then, the weight of the coin or pulling a lever sets the toy in action. Spring-driven windup toys are in their apogee in the twentieth century. First produced mostly in Europe and imported to the Unites States, in the beginning of the twentieth century, they are already manufactured in the Unites States as well. These toys remain very popular to the end of World War II. After World War II, plastic toys – both mechanical and electronic – became far more popular and made the iron windup toys obscure. In the 1950s, the toy becomes commonplace and becomes a consumer object. Along with this, the child's status is changing. From miniature adult, it becomes full with its own needs. The child becomes his/her own self. The child develops and with it a range of educational toys. Finally, in the last quarter of the twentieth century, the video game and the emergence of new technologies are revolutionizing the toy and the game in the Western world. The toys then reflect a scientific approach in constant motion.

In sum, depending on periods and events, animate and inanimate objects/toys took different designs. Mechanical toys and automata are still alive and kicking even today. They became the mainstream entertainment tool for children. Their symbolic value is high and transcendent. Playing with toys, whatever their nature, takes place at an early age, through nonverbal multimodal unconscious processes. Playing with toys (including the imperceptible part, i.e., the shadow) mirrors children's ability to unfold their neurocognitive development. The use of animate/inanimate objects/toys implies that the objects are part of the external environment, i.e., the "other."

Enrobotment's Whiff: Playing with Animate and Inanimate Objects/Toys in Childhood

The cognitive creative attitude is first manifested in the play. During play, children develop cognitive perceptual, emotional, and social accomplishments that allow them to learn not only about themselves but also about the others. Children use to play with animate and inanimate toys in the first months of life. From 0 to 2 years old, children start to represent objects via manipulation (e.g., Vygotsky 1976). Objects/toys are put together, sucked, smooched, and piled on top of each other. The ability to play with them, alone or with others, becomes more and more sophisticated as development progress. It expands neither into the internal (the child itself) nor external (the toys) environment of the infant, but in an intermediate area also named transitional area which is between the "self" and "other." This area serves as the place of child's neurocognitive construction which is possible because of child's capacity to symbolically incorporate the object/toy (e.g., Winnicott 1971). Object's internalization includes the insignificant parcel of the object, its shadow. Playing, for example, between an adult and an infant corresponds to an overlap of their intermediate areas. As such, the external environment made by animate and/or inanimate toys ("the other") cannot be separated from the genesis of an individual ("the self"); it incorporates the seeds of this genesis. It is on the basis of subjectification (Fig. 2).

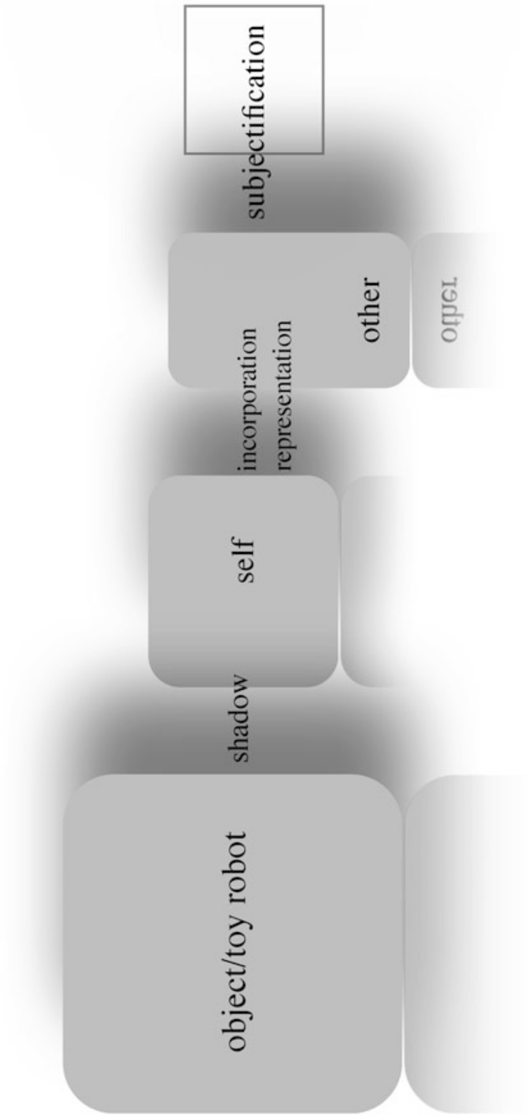


Fig. 2 Schematic configuration of the object/toy robot related to its shadow and the self-other subjectification

Before children became realistic, between 7 and 11 years old, they understand the rules and the strategies symbolically via objects/toys. From 2 to 7 years old, young children play with toys by pretending that the objects/toys represent something else that they love or not. As they are very imaginative and gifted, all fantasies are allowed, e.g., a pencil is a microphone. Independently of their nature, all things are considered as toys, and they are used to enter further into the play and the "skin" of a character. Children are able to move on from using themselves as an agent to using toys as (active) agents and carry out various actions (e.g., Leslie 1988). As such, playing is an individual invention, which enables endless variations. Free play, different from the game with rules, is an act that has a role in the quest of neurocognitive activity. It is the place of creation, the place where it is elaborated the intersubjective space between an individual and its environment (represented by the object). Requiring a dimension of fascination and omnipotence, play gives children the possibility to elaborate their "magical thinking." It mirrors and reverberates the creation of a neoreality which is linked to the ability to create symbols (Fig. 3). As a consequence, animate and inanimate toys seem to provide an interesting account of "how" physical objects are able to act as support for the symbolic play of children. Symbolic play, like emotional and cognitive verbal development, flourishes progressively as toys are the indicators that assist the child to go in (e.g., Vygotsky 1976). With development, symbolic play with action grows into emotion and language. Symbolic play is associated with the capacity to generate thoughts and concepts for ourselves and for others which can be expressed.

In brief, fundamentally linked with the process of object/toy internalization, representation of the object, the "intermediate area" which is a transitional space, symbolizes child's neurocognitive construction. The internalization of the objects/toys (including unperceived part) through play not only reflects the impact of the environment on the child's development, but it also reverberates the echo of child's representations. In that way, an intermediate object (including shadow) is conceived in mind by the child itself. Having a high emotional value, it ensures the continuity between "the self" and "other"; it endorses subjectification. Identification, imagination, interpretation, and personalization of the reality via object/toy as well as its shadow, that is, what is implied in the individual and shared play is the first and last source of communication between peers. This is the substratum of enrobotment.

Enroboted Memory: Object as Nonverbal and Verbal Representation in Memory

Human development deploys with increasing complexity. Developmental theorists have explained that representations are based on concrete nonverbal forms in infancy as well as in young children (e.g., Piaget 1954; Vygotsky 1962). It is progressing from the more concrete to the more abstract and sophisticated processes. The representations, also considered as the symbolization of concrete objects/toys that children can see, hear, touch, and manipulate that the child has internalized (e.g., Giannopulu 2013b), develop on the basis of association before adhering to the ideal

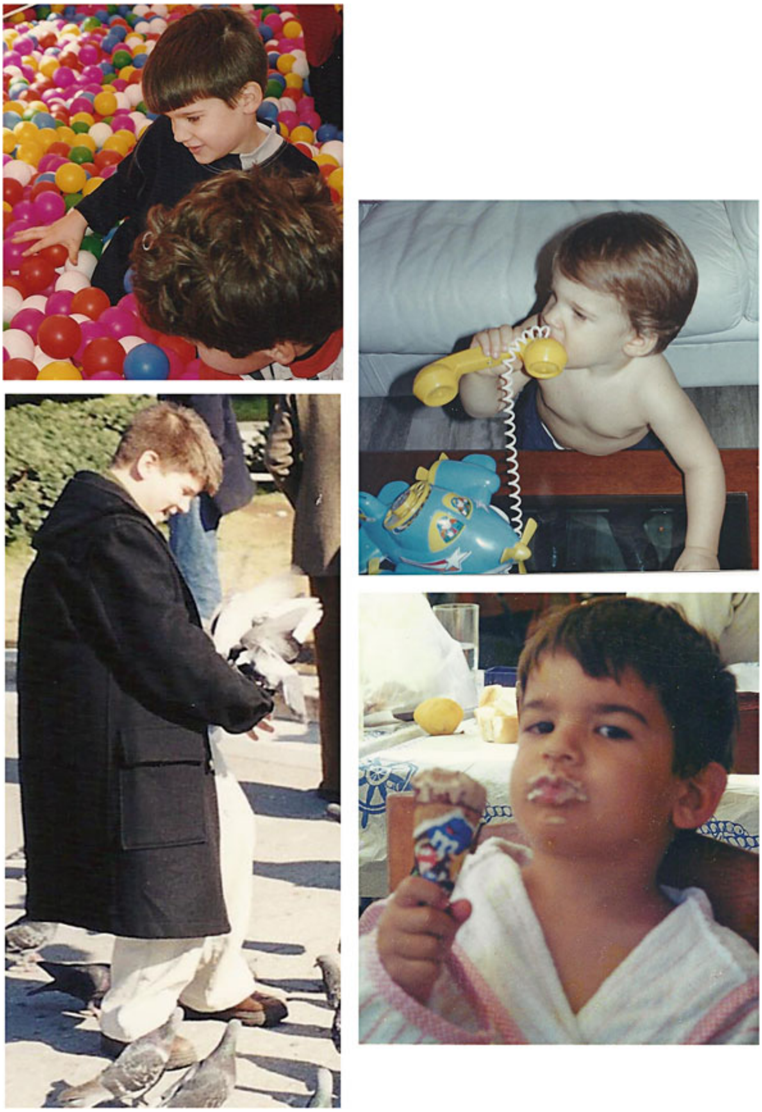


Fig. 3 Playing with objects/toys

convention of behavior (e.g., Marschal et al. 2010). Based on the internalization of the object/toy, they constitute the foundations of memory.

Depending on verbal and nonverbal associations, it is rather proposed that the representation of objects/toys in memory shifts from implicit (procedural) to explicit (declarative). Implicit representation of objects and concepts begins very early in infancy (e.g., Brandone and Wellman 2009), but seems to be particularly conspicuous and crucial from 3 to 6 years. The transition takes place rapidly in typical

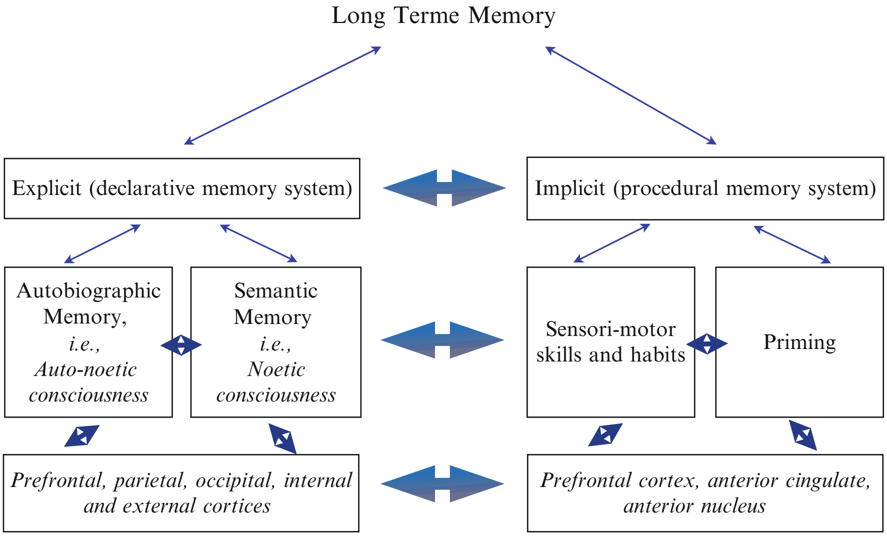


Fig. 4 Cognitive and neuroanatomical components of long term memory

children. In that way, young children move from general and analytic representations to memory. The transition arises at different developmental stage depending on the domain involved (e.g., Diesendruck and Perez 2015). Young children initially conceive the world in terms of associations and after in terms of conventions. They are able to learn about objects especially when they use them to realize different kind of actions (e.g., Giannopulu 2013b). The autobiographical memory, i.e., autonoetic consciousness, rises. Progressively, young children will develop nonverbal implicit memory, and this is in relation with somatosensory and motor development (e.g., Squire 1992). By the age of 6 months, children seem to develop nonverbal process of explicit memory. It has been reported that by the age of 24 months, children are also able to recall and realize nonverbal actions that they have learned before and during preverbal periods. This kind of recalling behavior is also expressed when spatial landmarks are needed in order to execute a specific task. In other words, during the preverbal period, children encode the basic nonverbal contextual information that is multimodal in nature (e.g., visual, acoustic, haptic). They will use them later during the verbal period. This nonverbal information will be the basis of their implicit memory, i.e., procedural memory of preverbal period. All the elements will be transformed into explicit memory, i.e., declarative memory in verbal period. This transformation will also perform to assure the enactment between implicit cognition and explicit cognition via object's internalization/incorporation that allows representation to come to mind and verbal action to evolve (e.g., Bauer 2007). The association between implicit and explicit memory via verbal procedures is a sturdy one (Fig. 4).

The internalization of object/toy (including shadow) contributes to the development of cognition in general and concepts in particular (e.g., Waxman and Markow 1995). As an "object of knowledge," i.e., "epistemic object," the internalized object/toy offers children the possibility to perform an overview of the relationship between words, especially those that are pertinent, and to evolve links between them. The children develop their semantic memory, i.e., noetic consciousness of the world. The children are able to "blow ideas into sentences." Based on the internalized object, the emergence of language gives the possibility to flourish thoughts in the mind (e.g., Giannopulu 2011a, b; 2013a, b). Within that frame of reference, language offers children the possibility to express their thoughts and share different persons and domains together, but also brings together different processes permitted to realize mathematics operations and spatial navigation (e.g., Spelke 2003). It allows the integration of different domains and concepts. Using the language at many levels, children are receptive to the way in which their own knowledge is immersed in a much more elaborate social network that is dependent upon their everyday theory of mind (ToM) (e.g., Baron-Cohen 2001), i.e., the capacity to discern the goals and intentions of others.

From a neurocognitive viewpoint, many brain areas potentially associated with nonverbal (processing visual, acoustic, motor, and haptic) and verbal processes are thought to be involved. Anterior and posterior areas of the left (e.g., Brass and Heyes 2005) and also the right (e.g., Iacoboni et al. 2001) hemisphere are involved in language comprehension and production (e.g., Cappa and Perani 2003). Following the general neuroimaging argument, the left temporal cortex plays a crucial role in lexical-semantic tasks related to the processing of nouns, whereas the processing of words related to actions involves additional regions of the left dorsolateral prefrontal cortex (e.g., Cappa and Perani 2003). It has been consistently shown that words referring to spatial location increase neural activity in occipitoparietal and frontal areas (e.g., Montani et al. 2014). Studies using neuroimaging techniques have also signed the critical role of neural areas including prefrontal cortex, anterior cingulate, and anterior caudal nucleus supporting nondeclarative memory system (e.g., Schnyer et al. 2009). Prefrontal, occipital, parietal, internal, and external temporal cortices are implicated in declarative memory system (e.g., Jeong et al. 2015). Both memory systems are interwoven. All of the above areas required are still developing at 6 years of age (e.g., Casey et al. 2005).

In short, during play, the internalized object/toy (including the undetectable parcel, i.e., the shadow) offers children the way to emerge verbal and nonverbal representations. Forming an implicit/explicit autobiographical, i.e., self-conscious/unconscious, continuum in memory, the correlated representations allow the invention of ideas and concepts and motor and verbal actions, including their intention, to prosper. This is of great interest in enrobotment, particularly in the case of children with neurodevelopment disease like ASD. In the next session, we will emphasize how implicit and explicit autobiographical continuum in memory supports the intention attribution to toy robots.

Another Expression of Enrobotment: Intention Attribution to Toy Robots

Children are famed for the complexity of abilities that is possible for them to learn but also for the degree of abstraction that they can enact. They are able, for example, to decode the actions of people or those made by different objects and to react. This ability allows them to recognize, to understand, to interpret, and even to learn from the decoded actions. Such abilities clearly signify that there is a causal relation between action, cognition, and perception. We therefore accept that the child behavior cannot be dissociated from actions and the actions cannot be separated from the intentionality they express whatever the intention is conscious or unconscious. In that particular context, intentionality can be defined as a "premotor mental act" of a desired motor action execution (e.g., Piaget 1954). Following the dualistic approach proposed by Searle (1983) and completed by Brand (1984), intention can be distinguished between abstraction, i.e., prospective intention and intention in action, i.e., immediate intention. Action abstraction and anticipation are possible because of the capability to identify the kinematics of animate action (e.g., Skerry et al. 2013). The emerging idea is that a motor affordance of a specific action is required for understanding the action. Action's understanding would arise from the observation of the intentional action considered as a rational process pursuing goals persistently. This understanding is directly reflected in infants' interactions with different people and the objects presented to a given environment (e.g., Giannopulu 2013b). Even if the development mechanisms underlying changes in young children's understanding and abstraction of actions are currently unknown, it is accepted that young children's abilities to develop these concepts embellish in refinement over the years.

Eight-month-old children are able to operate predictive looks to the consequence of an action, and 10-month-old children are able to predict the goal of an intentional action (e.g., Brandone et al. 2014). In the same vein, Cannon and Woodward (2012) have shown that young children of 10 months old have the potential capability to use their understanding of other human (and object) behavior to generate predictions about other people's (and objects) intentional actions (Fig. 5). During the first year of life, children interpret object's and people's actions in terms of their intentions (e.g., Brandone and Wellman 2009). They also would use this comprehension prospectively in order to predict about others' behavior. Very recent findings indicated that young children's self-locomotion abilities, crawling as well as spontaneous play, were associated with the ability to anticipate the consequence of an action and may contribute to changes in their understanding of others (people and objects) as intentional operators (e.g., Brandone 2015). Young children seem to also perceive and understand animate motion (e.g., Rakison and Woodward 2009). Older children are able to initiate and to engage in various kinds of collaborative action, which probably reflects the beginnings of the symbiotic play. These children are able to share intentions that require coordinating planning (e.g., Imamura et al. 2015). In other words, the understanding of intentional actions would have a specific role in cognitive development because they facilitate basic abilities such as visual proprioception (e.g., Brooks and Meltzoff 2008) and object and tool use (e.g., Perone

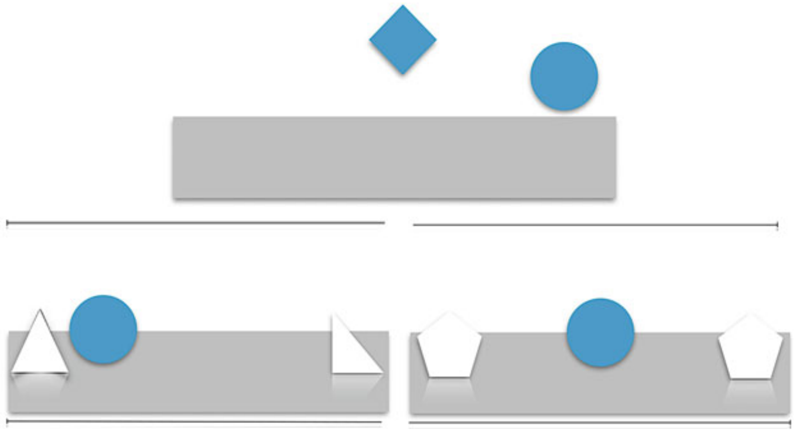


Fig. 5 Intention attribution to object/toys

et al. 2008). The animate motor actions are thought of as a prerequisite for cognitive perceptual development (e.g., Piaget 1954; Rakison and Woodward 2008; Zelaznik 1993). A format of “action-cognition-perception” occurs that continues beyond infancy and can become a representational medium. In the same vein, the internalization of this triadic format is crucial for the deployment of representations (e.g., Giannopulu 2013b). The internalized motor action gives the children the possibility to engage new strategies that in turn affect their representations. In that way, children would adopt via anticipation the particular “pattern” others (adults or not) and objects in order to accomplished a goal (e.g., Tomasello et al. 2005). They would consider various strategies and designate one or more to enact in symbolized action based on reality (e.g., Gergely et al. 2002). Such capability that is classically associated with the theory of mind (ToM) is a prerequisite for successfully operating into the social world (e.g., Baron-Cohen 2001). There is a functional neural mechanism associated with the parietal and premotor areas at least. This mechanism is linked to the emergence of analogies between action observation and understanding and action execution (e.g., Gallese 2009). It could favor action abstraction and intention attribution.

Several brain imaging studies using various methodologies and techniques support the intrinsic link between action execution and observation. Children represent their own and observed actions in neurally similar ways (e.g., Lepage and Theoret 2006). Studies of the human neuroanatomy of action processing have demonstrated that action comprehension and intention activate multiple cognitive, perceptual, and motor areas in the brain (e.g., Iacoboni et al. 2005). Related embodied approaches have been of significant influence in recent years. One of the main neuroscientific hypotheses is the involvement of the mirror neuron system for perception to show direct use of some of the same brain structures used in action (e.g., Sperduti et al. 2014). Several functional brain imaging studies in adults demonstrated that the understanding of actions is associated with the activity of the mirror neuron system (e.g., Gentilucci and Bernardis 2007). Recent data also emphasize the bilateral and synergical participation of the ventral and dorsal pathways in the mirror system (e.g., Arbib 2010). All the above brain areas involved in action observation and expression are in development at children during several years of age (e.g., Lepage and Theoret 2006). The early existence of shared systems for perceiving and acting actions provides a necessary bridge from action abstraction to symbolization and intention (e.g., Meltzoff 2007).

The fact that children with autism spectrum disorders (ASD) display reduced activation of brain regions associated with the mentalizing of intention during the observation of animated forms has been considered as a poor understanding of others' (people and/or objects) intentions (e.g., Iacoboni and Dapretto 2006). ASD implicate significant problems in understanding mental state of others that signifies that these children have a poor expression of the theory of mind (e.g., Baron-Cohen 2001). They also seem to have poor comprehension of complex action sequences (e.g., Zalla et al. 2006). It is well established that children with ASD have executive function deficits that it is expressed by a specific trouble to plan actions (e.g., Ozonoff et al. 2004). It is noteworthy that it is not yet really clear if the goal and intention understanding is spared or impaired in autism and what is exactly the role of the mentalizing system (e.g., Hamilton 2009).

In summary, object/toy internalization allows representations to emerge. This obviously signifies that the encode implicit/explicit autobiographical memory epitome is combined with the integration of functional areas in the brain. The concepts of "self" and "other" are intimately related to children's capability to gradually interact/play with objects/toys, i.e., enrobotment. In this framework, the resulting intention attribution to objects/toys constitutes a precursor of self-consciousness, as this intention, a specific anticipation, helps children to understand what it signifies to have a perspective. Recognizing what it implies to be a "self" is a parcel of envisioning mental states of "others." This behavior fails in children with ASD.

Enrobotment Paves the Way to Self-Consciousness

The investigation of self-consciousness is one of the enigma facing the scientific perspective. Considering as indispensable to human development, self-consciousness appears near 2 years of age. It enables child to become conscious of the "self," i.e., to recognize its "specular image in the mirror" which contributes to child's neurocognitive metamorphosis. Namely, self-consciousness is the ability to become conscious of the self as an "object" in the mirror. Being self-conscious combines the having of perception, feelings, and thoughts of the inside but also of the outside world (e.g., Lewis 2011). Verbal and nonverbal interactions are joined with emotions in a continuous dynamic workspace that constitutes self-consciousness. In this approach, it is considering that self-consciousness (high-level processes) is deeply interwoven with unconscious processes (low-level processes) (Fig. 6).

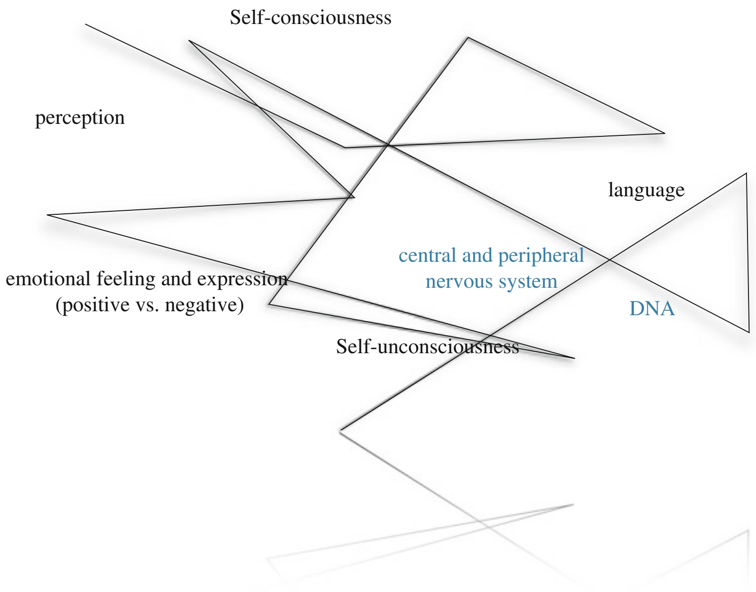


Fig. 6 Self-consicousness, self-unconsciousness, associated functions, *genetic and neural bases*

Undoubtedly, multimodal interactions combine nonverbal cognitive information sources such as visual, acoustic, haptic, motor, and action information in a coordinated manner in order to provide flexible, powerful, and synergistic dialogues between the brain, mind, and environment (e.g., Giannopulu 2013b; Giannopulu and Watanabe 2015b). All these ingredients could be woven together at different cortical and subcortical levels via neural maturation. The associated multimodal neuronal activity would emerge from the cortical areas which are directly and/or indirectly related to each other. Emotion seems to appear very early in development (e.g., Siegel 1999; Lewis 2011). The fetus displays some traces of emotion, but as it is mainly asleep, it is difficult to consider that this process is really functional (e.g., Lagercrantz 2009). However, it is thought that the evolution of self-consciousness is based on these primary emotional traces. In newborns, facial expressions and voice intensity are the first means of communication with the outside world (e.g., Bowlby 1978; Main 1996), what may be referred to as minimal consciousness (e.g., Zelazo 2004). From the age of 4 months, the child readily discriminates negative and positive emotions (e.g., Garon and Moore 2004; Russell 1989). Different types of smiles correspond to different degrees of emotional intensity (e.g., Messinger 2005). As children can communicate only about the world as it is organized in self-conscious nonverbal cognition, self-conscious nonverbal cognition is a prerequisite of human language (e.g., Giannopulu 2013b; Gray 2004). But children acquire language through both conscious and unconscious processes (e.g., Siegel 1999). The relational concept presupposes bilaterally dependent and constitutive linkage between conscious verbal expressions and unconscious nonverbal expressions (e.g., Jablonka et al. 2012). Beyond 8 months, children intellectualize each basic emotion and learn to relate facial and gestural expressions, context, and words (e.g., Barbas 1993; Baron-Cohen and Leslie 1985). The ability to use language to name the emotional feeling and expressions occurs between 2 and 10 years (e.g., Russell 1989; Bloom 1998; Giannopulu and Sagot 2010). A relatively smaller number of words for naming emotions exist in young children than in older children, reflecting different levels of maturity in self-consciousness (e.g., Bloom 1998). As humans develop, consciousness and unconsciousness (including emotions) are considered to be the core nucleus of language (e.g., Perlovsky 2009; Arbib 2014). Unquestionably, the ability to understand and share the referential statement of others and to express their own states consciously or unconsciously depends on the cerebral maturity (e.g., Siegel 1999). Self-conscious (and unconscious) development emerges from dynamic neural connections distributed across many cortical and subcortical regions (e.g., Lagercrantz and Changeux 2009). This approach may not correctly apprehend the complexities of human evolution without considering the neurogenetic and neural correlates of self-consciousness.

Neurogenetic and Neural Correlates of Self-Consciousness

Human self-consciousness is accepted to be correlated with DNA molecule. Different hypotheses exist. On the one hand, DNA molecule is thought to possess a degree of self-consciousness of its own; on the other hand, DNA molecule is thought to be responsible to advance to other higher degrees of self-consciousness, e.g., cellular self-consciousness and human self-consciousness, and have a kind of proto-consciousness which leads to a higher degree of self-consciousness (e.g., Grandy 2006). DNA as a conscious unit has been objectified on three dynamic interdependent levels: gene-gene interaction (epistasis), nucleic entities' interactions (e.g., RNA species, viruses, mitochondria), and DNA and external world interaction (e.g., Grandy 2013). When DNA gives rise to human self-consciousness, three gene-based neurogenetic phases take place: the emergence of self-consciousness (e.g., Pax3, Pax6, Otx1, Otx2, Hoxb4, Hoxd4 genes), the continuation of human self-consciousness (e.g., GluR, Cdk5, NF-kappaB, BDNF, FGF-2, FosB genes), and neurodegeneration (APP, PSEN1, PSEN2, APOE-epsilon4, TREM2) (e.g., Grandy 2014). Following this approach, human self-consciousness is not only related to neurons and regions of the brain but also to a specific neurogenetic ground.

Neurons are considered as the "fragments" of self-consciousness. Following the embryogenesis, neurons proliferate primarily between the 10th and 20th gestational week. Neurons start to develop dendritic spines during the third trimester. Synaptogenesis is also ignited during this period to culminate at about 1 year after birth (e.g., Bourgeois 2010). A prerequisite for the emergence of self-consciousness is the development of thalamocortical fibers (e.g., Kostovic and Judas 2010). With the exception of the olfactory tract, the neurons from the sensory organs (e.g., visual, haptic, acoustic, vestibular, proprioceptive) reach the subplate of the cortex before about the 25 weeks of gestational age. The subplate serves as a waiting area and as a guidance hub for the afferents from the thalamus and the other areas of the brain. The expansion of the thalamocortical axis in the somatosensory, auditory, visual, and frontal cortices starts. The corticocortical neuronal circuits develop late during infancy (e.g., Lagercrantz 2014). The immature brain activities seem to be suspended by the spontaneous activity generated by immature neurons present about 23–24 weeks of gestation. Positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) have detected a low-frequency spontaneous intrinsic brain activity that is considered as a resting brain activity associated with the autonomous system. It seems to correspond to the autobiographical memory, i.e., autonoetic consciousness, past experience, and planning of the future, and is associated with cingulate, precuneus, and dorsal and ventral prefrontal cortex (e.g., Raichle 2010). This resting activity is integrated and interconnected with the global neuronal network from the sense organs and associated brain areas. The resting activity may be considered as an unconscious experience from which self-consciousness would emerge while still being interlaced.

With the progress of brain maturity, young children are able to consider self-conscious (and unconscious) codes. This is potentially due to the formation of a neural multimodal network, which naturally follows the evolution of the brain (e.g., Fedorenko et al. 2010; Pelphrey and Caster 2008).

Different studies emphasize the importance of the prefrontal cortex and temporal and parietal cortices not only for conscious expression and comprehension but also for unconscious nonverbal emotional processes (e.g., Corbett et al. 2009; Frith and Frith 2003). Clearly, self-consciousness (and unconsciousness) cannot be confined to the thalamocortical complex alone but also to lower and higher areas, which are of particular interest from a developmental viewpoint (e.g., Lagercrantz and Changeux 2009). The dynamic modeling of cortical thickness associated with the development of emotion, language, and self-consciousness increases gradually from early childhood (5 years) to mid-childhood (9–11 years) with sequential emergence of three important cortical regions: temporal poles, the inferior parietal lobes, and the superior and dorsolateral frontal cortices (e.g., Schmitt et al. 2014). Subcortical regions develop in combination (e.g., Porges and Furman 2011). As a consequence, self-consciousness (and unconsciousness) gradually follows that dynamic modeling which is naturally characterized by intra- and interindividual differences. At each age of development, these differences are associated with the differential degree of brain plasticity (e.g., Schmitt et al. 2014). Combining neuroscientific, psychiatric, and engineering approaches, a recent study has investigated the self-conscious and unconscious processes in neurotypical children aged 6 and 9 years old when they interact with a toy robot. The paradigm of “speaker-listener” was adopted (e.g., Giannopulu et al. 2014). This paradigm is analyzed in the next session.

Self-Consciousness Enrobotment: The Paradigm of Speaker-Listener

Considering the paradigm of speaker-listener, an intersubjective interaction is developed. The speaker was always a child; the listener was a human or an InterActor robot, i.e., a toy robot which reacts to speech sounds by nodding only (Fig. 7). Given the fact that at both intraindividual and interindividual levels, dynamic neural differences characterize children aged 6 and 9 years old, the general hypothesis speculated that their self-conscious verbal expressions and self-unconscious nonverbal (emotional) expressions would differ. It has been shown that children with mild-moderate autism (with 6 years old in terms of developmental age) who present less brain activity (because of the neurodevelopmental disorder) exhibited more important nonverbal expression (unconscious reactions) when they interact with the robot than their neurotypical pair (e.g., Giannopulu and Watanabe 2015a). Transposing to neurotypically developing children, the assumption was that children with less brain maturity (6 years old) would display more nonverbal expression with the InterActor robot, than children with more brain maturity (9 years old): the lesser the brain activity, the more the unconscious nonverbal emotional expression. It was also reported that children with ASD were more interested in the InterActor robot especially because of its predictability, i.e., defined degree of variability in reactions (e.g., Giannopulu and Watanabe 2015a). It stands to reason that neurotypical children would be more interested in the human interaction complexity. Respecting, a differential degree of unconscious nonverbal emotional maturity would reflect interindividual differences in cortical and subcortical areas. Namely, young children (6 years old) would be more dependent on the artificial environment than old children (9 years old). This would signify a differential degree of unconscious nonverbal emotional maturity. The observations via the speaker-listener paradigm provided some evidence. Neurodevelopmentally speaking, the autonomous activity would mirror interindividual differences of neural activity in the temporal cortex (including amygdala), in the cingulate cortex, and in the prefrontal cortex (e.g., Porges 2007). As skills of state regulation improve, the central nervous system

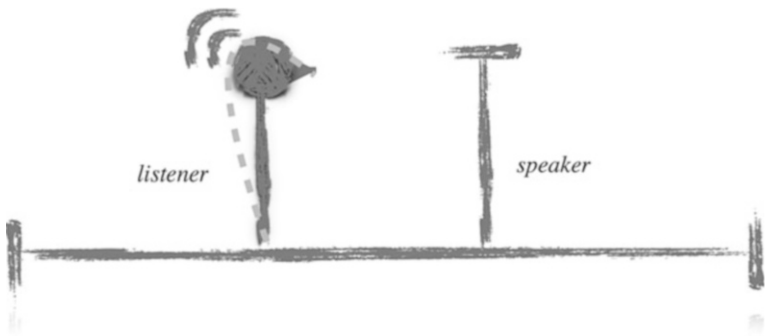


Fig. 7 Schematic presentation of the "speaker-listener paradigm"

increases to upgrade greater control over peripheral systems. The autonomous and central systems mature in combination and enable the developing child to become independent in the complex environment. Both systems are still in development in children 6 and 9 years old (e.g., Schmitt et al. 2014). That is to say, the heart rate of 6-year-old children is higher than the heart rate of 9-year-old children when the listener is the InterActor robot. Such state seems to signify that children aged 6 years old are more emotionally (and nonverbally) reliant on the InterActor robot than children aged 9 years old. The emergence of a self-conscious verbal statement from an unconscious nonverbal emotion (specified by the heart rate) would be reflected. Consistent with recent physiological data (e.g., Schmitt et al. 2014), the aforementioned observations suggest that a difference between levels of self-consciousness exists among young (6 years old) and old (9 years old) children. Based on the differential degree of brain maturity and self-consciousness, the reported observations are systematic with analyses notifying that verbal functions continue to mature at 6 years and grow up to adulthood. Furthermore, self-conscious verbal expression is more important when the InterActor is the human, likely because of the human complexity.

The development of self-consciousness (and unconsciousness) would be associated with the established observations. Nonverbal emotional behavior expressed by heart rate is an unconscious automatic activity which, in this particular case, would depend on a minimalistic artificial environment: the InterActor robot. Verbal behavior given by the words pronounced by the children (nouns and verbs) is a self-conscious activity which would depend on a natural environment: the Human InterActor. Differently, self-conscious and unconscious processes would not only depend on natural environments but also on objects/toy robots.

In summary, self-consciousness' development is the result of a complex process with at least three foci: one in the central nervous system with the associate genetic and neural influences, one in the mind with the associated representations, and one in the child's dynamic interactions with the artificial environment, that is, toy robots. Enrobotment encourages children's self-conscious and unconscious processes. Accordingly, the analysis of self-consciousness would envision one of the most common interruption of consciousness, i.e., autism. In this framework, autism is considered as an antithesis of self-consciousness.

A Place for Enrobotment in Children with ASD

Autistic spectrum disorders (ASD) are complex and heterogeneous neurodevelopmental disorders (e.g., Bowler 2012). These disorders typically manifest before 3 years of age, resulting in a distinct phenotype (e.g., Gabis et al. 2008). The deficits encompass a wide continuum extending from severe distinctiveness to mild-moderate developmental disabilities. The male/female ratio in ASD is between 3 and 4:1. There are strong genetic components, demonstrated by an increased sibling recurrence risk estimated at 3–7 % (e.g., Spence et al. 2004) and higher concordance in monozygotic twins (70–100 %) than in dizygotic twins (e.g., Rosenberg et al. 2009). Affecting cognitive functioning as well as emotional and social behavior development (language including), self-consciousness of children with ASD, i.e., the having of perception, thoughts, and feeling of the internal and external world, is alternated. Genetic studies have highlighted the complexity of the genetic architecture underlying ASD (e.g., Jeste and Geschwind 2014). They consider ASD as a complex multifactor disorder involving many genes (e.g., Sykes et al. 2007). These studies have given rise to new insights into neuronal circuits relevant to children with ASD. Postmodern analysis had demonstrated evidence of altered brain development, which strongly affects the formation of a multimodal neural network. The analysis showed that the neural substrate underlying cognitive, social, emotional, and linguistic impairment involves multimodal areas such as the exterior superior temporal sulcus (e.g., Pelphrey and Caster 2008); the interior temporal lobe, amygdala included (e.g., Corbett et al. 2009); as well as the ventral part of the prefrontal cortex, i.e., the orbitofrontal cortex (e.g., Brothers 1990). The brain of children with ASD is also characterized by aberrant brain connectivity (e.g., Khan 2015) and disruption of white matter tracts between temporal regions (e.g., Frith and Frith 2003) which disrupt verbal and nonverbal acquisition, consolidation, as well as social interaction (e.g., Kana et al. 2011).

These neuro-functional studies provide the basis for concluding that in ASD the more impaired cortical areas are those that are involved in complex cognitive functions that lead to self-consciousness. The complex expression of ASD necessitates a more generic consideration of the disorders at the multimodal neural level. Developmentally speaking, the most widely accepted hypothesis in autism is the theory of mind deficit (e.g., Frith et al. 1991). Even if this theory cannot account for the whole spectrum of autistic disorders, it raises many issues that not only involve mental representation of others but also cognitive nonverbal skills such as posture (e.g., Leslie 1994), eye contact (e.g., Reed and Peterson 1990), touching (e.g., Frith and Frith 2003), manipulation (e.g., Escalona et al. 2002) language, and emotion (e.g., Giannopulu 2013a, b). The nonverbal cognitive, emotional, and social functions are considered to be the basis of communication with the "self" and the "other." Playing capabilities with objects are fundamental.

Many options are being explored to improve capacity for social interaction, one of which is to simplify the elements that constitute the interaction. Robots, virtual environments, as well as other computer-based technologies are being increasingly utilized in play and the education of children with ASD. These types of environments seem to be more effective than real environments (e.g., Dautenhahn 2007). A variety of studies have been conducted. Though some of these are exploratory in nature, their results are encouraging. In the next session, the child-robot relationship in autism will be analyzed.

Toy Robot as Neural Mediator

Different studies were conducted with the aim to better understand the capacity of children with ASD to analyze and infer the activities of objects like robots during dyadic interaction. The Aurora Project investigates the use of robots (e.g., Labo-1, Kaspar, Robota doll) in cooperative play. The aim was to create a tool based on an autonomous robot that convinces children with ASD to engage in a process of interaction. The interactions were tested through the analysis of visual contact, joint attention, avoidance or fleeing, visual pursuit, and whether the child imitates the robot (e.g., Dautenhahn 2007). Using music, color, and visual contact, a sensitive robot named Tito was employed in social interaction with children with ASD (e.g., Michaud et al. 2007). A very small fixed robot named Keepon captured and maintained visual contact with the child, drawing his attention and initiating some element of conversation (e.g., Kozima and Yasuda 2007). Robota used the form of an animated face (mouth, eyebrows, eyes) that can cause behavior imitation from the part of the child with ASD (e.g., Nadel et al. 2004). The dinosaur Pleo seems to reinforce social behavior (e.g., Kim et al. 2012).

Based on dyadic child-robot interaction, all the studies cited above have signified that animate robots, humanoid or not, using different forms of stimulation encourage interaction in children with ASD. This approach is consistent with the assumption that a robot seems to improve the neural activity (and consequently) the behavior of children with ASD (e.g., Giannopulu 2013a, b). With the exception of Labo-1 in the Aurora Project and Roball in Michaud's project and GYPI I and POL (e.g., Giannopulu 2011a; b; 2013a, b), so far, only fixed robots have been utilized reducing the child's spontaneity and self-expression in play. Even though (because of the pathology) the number of the children participating in these experiments is limited, the dyadic child-robot interaction only is reflected in attention (e.g., Dautenhahn 2007) or imitation (e.g., Nadel et al. 2004) or visual contact (e.g., Kozima and Yasuda 2007). In the aforementioned studies, the focal point of the analysis was on a single mode of interaction. In addition, very marginal attention has been paid to the comparison of neurotypical and children with ASD in human-human and human-robot interaction. Quantitative metrics of social response for autism diagnosis including robots have been considered, but reliability is lacking (e.g., Scassellati and Admoni 2012). Several interdisciplinary studies combining cognitive neuroscience, psychiatry, and engineering have used specific measurements with the aim to objectively evaluate dyadic child-robot (e.g., Giannopulu and Pradel 2009, 2010; Giannopulu 2013a, b) and triadic child-robot-adult (e.g., Giannopulu 2011b; 2013a, b; Giannopulu and Pradel 2012) interactions. The interactions were operationalized via the analysis of multimodal cognitive nonverbal, verbal, and emotional expressions during spontaneous play. In these studies, the children were quasi-constantly in interaction with the toy robot via various ways (multimodal way) suggesting that the robot could help children with ASD to reduce their stereotypical behavior. In the particular context of the triadic interaction, once the robot-child interaction is established, the toy robot was utilized as a "neural mediator" to initiate the interaction with the adult and express positive emotion.

Fundamentally, via an experimental procedure such as free spontaneous play which is very close to everyday life situation, the aforementioned data have contributed to the development of a new ecological design in order to better understand robot-children with ASD interaction. The assumption behind is that this design could better facilitate the transfer of learned abilities to everyday life. Along the way and

Fig. 8 Minimalist toy Robot: Pekoppa

under the hypothesis that multimodal cognitive nonverbal interactions could be thought of as the building block from which expressive language could emerge, a new mobile toy robot named “POL” which incites children with mild-moderate ASD to engage in dyadic interaction and express language was developed. As expected, verbal expression is more important when children nonverbally interact with the mobile toy robot (e.g., Giannopulu 2013b). This is of great interest, particularly when considering that nonverbal behavior is probably at the origin of what is arguably one of the trademarks of human cognition: the capacity to generate thoughts and concepts for ourselves and for others which can be expressed verbally with the aim to communicate. With that in mind, it has been developed a communicative play rendered possible through a speaker-listener paradigm and used a minimalist InterActor toy robot, that reacts to speech sounds (Fig. 8). Verbal and emotional expressions in neurotypical children and children with ADS in human-human and human-robot situations have been analyzed via the activation of autonomous peripheral nervous system (heart rate variation) in relationship with the central nervous system (verbal expression). Following the advocated data, a minimalistic InterActor toy robot characterized by small variance nonverbal behavior, i.e., simple behavior, better facilitates ASD children’s verbal and emotional expressions than a human (e.g., Giannopulu and Watanabe 2015a; Giannopulu et al. 2016).

Essentially, it has been demonstrated that when the listener is an InterActor robot, specific changes characterize children with ASD: a “mobilization” of a given mental state. Such “mobilization” would provide support for the social cognitive engagement of children with ASD. Namely, by making the children with ASD available to engage emotionally (and verbally), the InterActor toy robot seems to modify their neural activity: the children enjoyed participating. The funding might be related to preferences of children with ASD. It is inferred that children with ASD are rather interested in minimalist objects/toys to which they can assign mental states of their own or of others. Such a behavior might be interpreted as reflecting the children's willingness to communicate with humans using the robot: the InterActor toy robot is a miniature of a human listener, i.e., head nods, the children with ASD can handle with, as neurotypical children do with humans. Everything happens as if the InterActor toy robot would allow children with ASD to elaborate a given cognitive state encoding and conceptualizing within their brain and externalizing into unconscious feeling (heart rate variation) and self-conscious verbal and emotional expressions (words).

Briefly, children with ASD have no or limited concept of "self." They cannot mirror the triadic relationship of "object-other-self." They withdraw from social, cognitive, and emotional interactions living in a world confined to "self." Without self-consciousness, there is no conception of "me" as distinct to "you," no self-evaluative thought or emotion, and no mentalizing. In view of the recent studies, enrobotment gives children with ASD the possibility to improve their capability to be "self".

Conclusion

"Autobiographical implicit and explicit memory of present, past, and future are gradually interconnected in space and time with thoughts different from one's own. By enroboting, that is, by integrating and cognating the object/toy robot, its intrinsic properties, as well as its shadow, the child refines her tools of conceptualization of such autobiographical memory and develops a view of self that could be shared and externalized, what could be called self-consciousness. In that way, the everyday themes carried in discourse hand out different universes of being, other than the child's previous engagement. The child is joining the community of minds.

On the relationship delicacies, the neurocognitive underpinnings and the emotional concomitants of enrobotic mental construction are developed. Thoughts, images, and impressions are analyzed as internal human experience. They are inspired by external things, i.e., objects/toys, which the child, herself, needs to recognize and understand. The child needs to symbolize the object/toy (including shadow) in order to grow. The object/toy is integrated/internalized into the "self" from the very beginning because the object/toy constitutes the subjectivity of the "self." The act of interiorization of the object/toy would be a condition of child subjectification. The subjectification of oneself in order to become a person signifies incorporation of the "percepts" of the object/toy, the "other" with respect to the self-consciousness. If there is no interiorized object/toy, the child would never become a person; the child would never develop self-consciousness (Fig. 9). The "self-other" intersubjectivity is a constitutive part of self-consciousness.

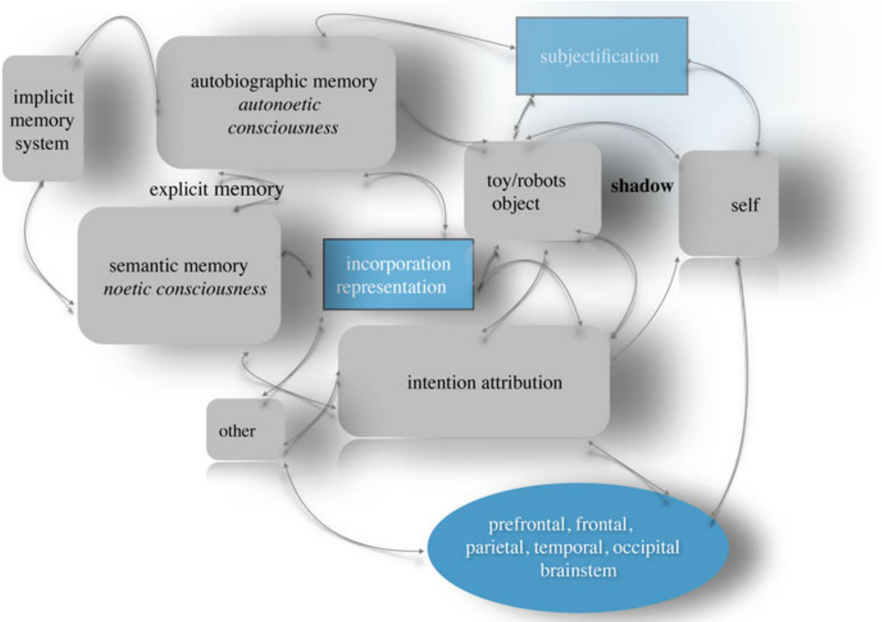


Fig. 9 Schematic presentation of enrobotment

The existence of the object/toy shadow, i.e., the imperceptible parcel of the object, would remain a secret of self-consciousness as its internalization/incorporation is an "unseen" area. Interiorization of the object through the child's development will signify that self-consciousness is not private. This is due to the fact that the content of one's self-consciousness is eternally affected by the presence of the external by the presence of the object/toy. The object/toy via enrobotment will become synonymous to a language, social and symbolic space, and to subjectification.

From this perspective, we can better understand some of the problems of atypical children, especially those on the autistic spectrum disorders (ASD). Verbal and nonverbal representations cannot be developed in human children without access to a symbolic system that enables self-consciousness to emerge. We sought understanding of metamorphoses in the neurocognitive behavior that take place during development. The changes can occur both in microscopic and macroscopic levels. In our framework, objects/toys participate in multiple variations of development in neural and behavioral spheres. Enrobotment mirrors the triadic relationship between the "object," "self," and "other."

Recommended Reading

A.M. Arbib, Mirror system activity for action and language is embedded in the integration of dorsal and ventral pathways. Brain Lang. **112**(1), 12–24 (2010)

A.M. Arbib, Co-evolution of human consciousness and language (revisited). J. Integr. Neurosci. **13**(2), 187–200 (2014)
H. Barbas, Organization of cortical afferent input to the orbitofrontal area in the Rhesus monkey. Neuroscience **56**, 841–864 (1993)
S. Baron-Cohen, A.M. Leslie, U. Frith, Does the autistic child have a "theory of mind" ? Cognition **21**, 37–46 (1985)
S. Baron-Cohen, Theory of mind in normal development and autism. Prism **34**, 174–183 (2001)
D.J. Bauer, Observations on the use of growth mixture models in psychological research. Multivar. Behav. Res. **42**, 757–786 (2007)
J.B. Bavelas, L. Coates, T. Johnson, Listener responses as a collaborative process: the role of gaze. J. Commun. **52**, 566–580 (2002)
J.M. Bering, T.K. Shackelford, The causal role of consciousness: a conceptual addendum to human evolutionary psychology. Rev. Gen. Psychol. **4**, 227–248 (2004)
L. Bloom, Language development and emotional expression. Pediatrics **102**, 5 (1998)
J.P. Bourgeois, in *The Newborn Brain*, ed. by H. Lagercrantz, E. Herlenius, 2nd edn. (Cambridge University Press, Cambridge, 2010), pp. 71–84
J. Bowlby, *Attachement et perte: l'attachement* (PUF, Paris, 1978)
F. Bowler, Autism spectrum disorders (ASD). Autism **16**(3), 223–225 (2012)
M. Brand, *Intending and Acting* (Cambridge England, Cambridge, 1984)
A.C. Brandone, Infants' social and motor experience and the emerging understanding of intentional actions. Dev. Psychol. **51**(4), 512–523 (2015)
A.C. Brandone, H.M. Wellman, You can't always get what you want. Infants understand failed goal-directed actions. Psychol. Sci. **20**, 85–91 (2009)
A.C. Brandone, S.R. Horwitz, R.N. Aslin, H.M. Wellman, Infants' goal anticipation during failed and successful reaching actions. Dev. Psychol. **17**(1), 23–34 (2014)
M. Brass, C. Heyes, Imitation: is cognitive neuroscience solving the correspondence problem? Trends Cogn. Sci. **9**, 489–495 (2005)
C. Breyer, *Jeux et jouets à travers les âges, Histoire et règles de jeux égyptiens, antiques et médiévaux* (Safran, Paris, 2010)
R. Brooks, A.N. Meltzoff, Infant gaze following and pointing predict accelerated vocabulary growth through two years of age: a longitudinal, growth curve modeling study. J. Child Lang. **35**(1), 207–220 (2008)
L. Brothers, The social brain: a project for integrating primate behaviour and neurophysiology in a new domain. Concepts. Neurosc. **1**, 27–51 (1990)
E.N. Cannon, A.L. Woodward, Infants generate goal-based action predictions. Dev. Sci. **15**(2), 292–298 (2012)
S.F. Cappa, D. Perani, The neural correlates of noun and verb processing. J. Neurolingust. **16**, 183–189 (2003)
B.J. Casey, N. Tottenham, L. Conor, S. Durston, Imaging the developing brain: what have we learned about cognitive development? Trends Cogn. Sci. **9**, 104–110 (2005)
G. Chapouthier, *Kant et le chimpanzé–Essai sur l'être humain, la morale et l'art* (Editions Belin, Paris, 2009)
H.H. Clark, *Using Language* (Cambridge University Press, Cambridge, 1996)
B.A. Corbett, V. Carmean, S. Ravizza, C. Wendelken, M.L. Henry, C. Carter, S.M.A. Rivera, Functional and structural study of emotion and face processing in children with autism. Psychiatr. Res. **30**, 196–205 (2009)
J. Damamme, *Mémoires de Jouets* (Hatier, Paris, 1998)
K. Dautenhahn, Socially intelligent robots: dimensions of human-robot interaction. Philos. Trans. R. Soc. B **362**, 679–704 (2007)
A.S. Dick, A. Solodkin, S.L. Small, Neural development of networks for audiovisual speech comprehension. Brain Lang. **114**, 101–114 (2010)
G. Diesendruck, R. Perez, Toys are me: children's extension of self to objects. Cognition **134**, 11–20 (2015)

A. Escalona, T. Field, J. Nadel, B. Lundy, Brief report: imitation effects on children with autism. J. Autism. Dev. Disord. **32**, 141–144 (2002)
E. Fedorenko, P.J. Hsieh, A. Nieto-Castañón, S. Whitfield-Gabrieli, N. Kanwisher, New method fr fMRI investigations of language: defining ROIs functionally in individual subjects. J. Neurophysiol. **104**, 177–1194 (2010)
U. Frith, J. Morton, A.M. Leslie, The cognitive basis of a biological disorder: autism. TIN14, 433–438 (1991)
U. Frith, C.D. Frith, Development and neurophysiology of mentalizing. Philos. Trans. R. Soc. B **358**, 459–473 (2003)
L. Gabis, W. Huang, A. Azizian, C. DeVincent, A. Tudorica, Y. Kesner-Baruch, P. Roche, J. Pomeroy, 1H-magnetic resonance spectroscopy markers of cognitive and language ability in clinical subtypes of autism spectrum disorders. J. Child Neurol. **23**(7), 766–774 (2008)
V. Gallese, Motor abstraction: a neuroscientific account of how action goals and intentions are mapped and understood. Psychol. Res. **73**(4), 486–498 (2009)
N. Garon, C. Moore, Complex decision-making in early childhood. Brain Cogn. **55**, 158–170 (2004)
M. Gentilucci, P. Bernardis, Imitation during phoneme production. Neuropsychologia **45**(3), 608–615 (2007)
G. Gergely, H. Bekkering, I. Király, Rational imitation in preverbal infants. Nature **415**(6873), 755 (2002)
I. Giannopulu, Cognitive and emotional interactions between autistic child, mobile toy robot and therapist. Front. Comput. Neurosci. (2011a). doi:10.3389/conf.fncom.2011.52.00002
I. Giannopulu, Contribution à la compréhension des représentations multimodales chez l'homme sain et chez des patients avec atteinte neuropsychologique: une perspective "life span". Université Pierre et Marie Curie (Paris VI), Habilitation à Diriger des Recherches. (2011b)
I. Giannopulu, Multimodal cognitive nonverbal and verbal interactions: the neurorehabilitation of autistic children via mobile toy robots. IARIA Intern. J. Adv. Life Sci. **5**, 214–222 (2013a)
I. Giannopulu, Multimodal interactions in typically and atypically developing children: natural vs. artificial environments. Cogn. Process **14**, 323–331 (2013b)
I. Giannopulu, G. Pradel, Mobile toy robots can be used in autism therapy: an example of application, in *IEEE Proceedings in the IROS 2009*, paper number SuT8.pdf in the workshop CD Proc (2009)
I. Giannopulu, G. Pradel, Multimodal interactions in free game play of children with autism and a mobile robot. NeuroRehabilitation **27**, 305–311 (2010)
I. Giannopulu, G. Pradel, From child–robot interaction to child– robot–therapist interaction: a case study in autism. Appl. Bionics Biomech. **9**, 173–179 (2012)
I. Giannopulu, I. Sagot, Positive emotion in the course of an experimental task in children. Ann. Méd.-Psychol. **168**, 740–745 (2010)
I. Giannopulu, V. Montreynaud, T. Watanabe, *Neurotypical and Autistic Children aged 6 to 7 years in a Speaker-Listener Situation with a Human or a Minimalist InterActor Robot*, (IEEE RO-MAN, 2014) p. 942–947
I. Giannopulu, T. Watanabe, Give children toys robots to educate and to neuroreeducate: the example of Pekoppa (New Trends in Medical and Service Robots, 2015a in press)
I. Giannopulu, T. Watanabe, *Un/conscious Emotional Dialogues in Typical Children in the Presence of an InterActor Robot*, (IEEE RO-MAN, 2015b) p. 264–270
I. Giannopulu, V. Montreynaud, T. Watanabe, Minimalistic toy robot to analyse a scenery of speaker-listener condition in autism. Cogn. Process (2016) (in press)
J.K. Grandy, Consciousness. Encycl. Anthropol. **1**, 563–566 (2006)
J.K. Grandy, The three neurogenetic phases of human consciousness. J. Cons. Evol. **9**, 1–14 (2013)
J.K. Grandy, The neurogenetic substructures of human consciousness. Essays Phil. **15**, 266–278 (2014)
J. Gray, *Consciousness. Creeping of on the Hard Problem* (Oxford University Press, Oxford, 2004)

A.F. Hamilton, Goals, intentions and mental states: challenges for theories of autism. J. Child Psychol. Psychiatry **50**(8), 881–892 (2009)
M. Iacoboni, M. Dapretto, The mirror neuron system and the consequences of its dysfunction. Nat. Rev. Neurosci. **7**(12), 942–951 (2006)
M. Iacoboni, L.M. Koski, M. Brass, H. Bekkering, R.P. Woods, M.C. Dubeau, J.C. Mazziota, G. Rizzolatti, Reafferent copies of imitated actions in the right superior temporal cortex. P. Am. Natl. Assoc. Sch. Proc. Natl. Acad. Sci. U. S. A. **98**, 13995–13999 (2001)
M. Iacoboni, I. Molnar-Szakacs, V. Gallese, G. Buccino, J.C. Mazziotta, G. Rizzolatti, Grasping the intentions of others with one's own mirror neuron system. PLoS Biol. **3**(3), e79 (2005)
Y. Imamura, K. Terada, I. Giannopulu, H. Takahashi, *How Robot's Behavior Affects Intention Attribution in Children*, (Workshop IEEE RO-MAN Toys and Developmental Cognitive Neuroscience, 2015)
E. Jablonka, S. Ginsburg, D. Dor, The co-evolution of language and emotions. Philos. Trans. R. Soc. **367**, 2152–2159 (2012)
S.S. Jeste, D.H. Geschwind, Disentangling the heterogeneity of autism spectrum disorder through genetic findings. Nat. Rev. Neurol. **10**(2), 74–81 (2014)
W. Jeong, C.K. Chung, J.S. Kim, Episodic memory in aspects of large-scale brain networks. Front. Hum. Neurosci. **14**, 454 (2015). doi:10.3389/fnhum.2015.00454
R.K. Kana, L.E. Libero, M.S. Moore, Disrupted cortical connectivity theory as an explanatory model for autism spectrum disorders. Phys. Life Rev. **8**(4), 410–437 (2011)
A.J. Khan, A. Nair, C.L. Keown, M.C. Datko, A.J. Lincoln, R.A. Müller, Cerebro-cerebellar resting-state functional connectivity in children and adolescents with autism spectrum disorder, Biol. Psychiatry. S0006-3223(15)00273-5 (2015)
E.S. Kim, L.D. Berkovits, E.P. Bernier, D. Leyzberg, F. Shic, R. Paul, B. Scassellati, Social robots as embedded reinforcers of social behavior in children with autism. J. Autim. Dev. Disord (2012). doi 10.1007/s10803-012-1645-2
I. Kostovic, M. Judas, The development of the subplate and thalamocortical connections in the human foetal brain. Acta. Peadiatr. **99**, 1119–1127 (2010)
H. Kozima, C. Nakagawa, Y. Yasuda, Children-robot interaction: a pilot study in autism therapy. Prog. Brain. Res. **164**, 385–400 (2007)
H. Lagercrantz, The birth of consciousness. Early Hum. Dev. **85**, S57–S58 (2009)
H. Lagercrantz, The emergence of consciousness: science and ethics. Semin. Fetal. Neonat. Med. **19**, 300–305 (2014)
H. Lagercrantz, J.P. Changeux, The emergence of human consciousness: from fetal to neonatal life. Pediatr. Res. **65**, 3 (2009)
J.F. Lepage, H. Théoret, EEG evidence for the presence of an action observation-execution matching system in children. Eur. J. Neurosci. **23**(9), 2505–2510 (2006)
A.A. Leslie, in *Developing Theories of Mind*, ed. by J.W. Astington, P.L. Harris, D.R. Olson (Cambridge University Press, New York, 1988)
A.M. Leslie, Pretending and believing: issues in the theory of ToM. Cognition **50**, 211–238 (1994)
M. Lewis The origins and uses of self-awareness or the mental representation of me. Conscious Cogn. **20**, 120–129 (2011)
M. Main, Introduction to the special section on attachment and psychopathology: 2. Overview of the field of attachment. J. Consult. Clin. Psychol. **64**, 237–243 (1996)
D. Marschal, P.C. Quinn, S.E.G. Lea, *The Making of Human Concepts* (Oxford University Press, Oxford, 2010)
A.N. Meltzoff, The 'like me' framework for recognizing and becoming an intentional agent. Acta. Psychol. (Amst.) **124**(1), 26–43 (2007)
D.S. Messinger, A measure of early joy? in *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Union the Facial Action Coding System (FACS)*, ed. by P. Ekman, E.K. Rosenberg (Oxford University Press, Oxford, 2005), pp. 350–353
F. Michaud, T. Salter, A. Duquette, J.F. Laplante, Perspectives on mobile robots used as tools for pediatric rehabilitation. Assist. Technol. **19**, 14–19 (2007)

V. Montani, A. Facoetti, M. Zorzi, Spatial attention in written word perception. Front. Hum. Neurosci. **10**, 8–42 (2014)
J. Nadel, A. Revel, P. Andry, P. Gaussier, Toward communication: first imitations in infants, low-functioning children with autism and robots. Inter St: Soc. Beh. Commun. Biol. Syst. **5**, 45–54 (2004)
K.A. Pelphrey, E.J. Caster, Charting the typical and atypical development of the social brain. Child Y. Psychol. **20**, 1081–1102 (2008)
D. Perani, S.F. Cappa, M. Tettamanti, M. Rosa, P. Scifo, A. Miozzo, A. Basso, F. Fazio, A fMRI study of word retrieval in aphasia. Brain Lang. **85**(3), 357–368 (2003)
L. Perlovsky, Language and emotions: emotional Sapir–Whorf hypothesis. Neural Netw. **22**, 518–526 (2009)
S. Perone, K.L. Madole, S. Ross-Sheehy, M. Carey, L.M. Oakes, The relation between infants' activity with objects and attention to object appearance. Dev. Psychol. **44**(5), 1242–1248 (2008)
J. Piaget, *The Construction of Reality on Children* (Basic Book, New York, 1954)
J. Piaget, *Play Dreams and Imitation in Childhood* (Norton, New York, 1962)
J. Piaget, *The Grasp of Consciousness* (MIT Press, Cambridge, MA, 1976)
S.W. Porges, The polyvagal perspective. Biol. Psychol. **74**(2), 116–143 (2007)
S.W. Porges, S.A. Furman, The early development of the autonomic nervous system provides a neural platform for social behavior: a polyvagal perspective. Infant Child Dev. **20**(1), 106–118 (2011)
M.E. Raichle, The brain's dark energy. Sci. Am. **302**(3), 28–33 (2010)
D.H. Rakison, A.L. Woodward, New perspectives on the effects of action on perceptual and cognitive development. Dev. Psychol. **44**, 1209–1213 (2008)
D.H. Rakison, A.L. Woodward, New perspectives on the effects of action on perceptual and cognitive development. Dev. Psychol. **44**, 1209–1213 (2009)
T. Reed, C. Peterson, A comparative study of autistic subjects' performance at two levels of visual and cognition perspecctive taking. J. Autism. Dev. Disord. **20**, 555–567 (1990)
R.E. Rosenberg, J.K. Law, G. Yenokyan, J. McGready, W.E. Kaufmann, P.A. Law, Characteristics and concordance of autism spectrum disorders among 277 twin pairs. Arch. Pediatr. Adolesc. Med. **163**(10), 907–914 (2009)
J.A. Russell, Children's understanding of emotion, in *Children's Understanding of Emotion*, ed. by C. Saarni, P.L. Harris (Cambridge University Press, Cambridge, 1989), pp. 293–313
B. Scassellati, Quantitative metrics of social response for autism Diagnosis, in *IEEE International Conference on Intelligent Robots and Systems*, vol. 2 (2002), pp. 1134–1138
J.E. Schmitt, M.C. Neale, B. Fassassi, J. Perez, R.K. Lenroot, E.M. Wells, J.N. Giedd, *The Dynamic Role of Genetics on Cortical Patterning During Childhood and Adolescence*, (PNAS, Early Edition 2014)
D.M. Schnyer, W.T. Maddox, S. Ell, S. Davis, J. Pacheco, M. Verfaellie, Prefrontal contributions to rule-based and information-integration category learning. Neuropsychologia. **47**, 2995–3006 (2009)
J.R. Searle, *Intentionality*, (PaperBook 2003)
D.J. Siegel, *The Developing Mind: How the Relationships and the Brain Interact to Shape Who We Are* (Guilford Press, New York, 1999)
A.E. Skerry, S.E. Carey, E.S. Spelke, *First-Person Action Experience Reveals Sensitivity to Action Efficiency in Prereaching Infants*, (PNAS Early Edition, 1–6 2013)
E.S. Spelke, in *Functional Neuroimaging of Visual Cognition*, ed. by N. Kanwisher, J. Duncan (MIT Press, Cambridge, MA, 2003)
M. Sperduti, S. Guionnet, P. Fossati, J. Nadel, Mirror neuron system and mentalizing system connect during online social interaction. Cogn. Process. **15**(3), 307–316 (2014)
S. J. Spence, The genetics of autism. Seminars Ped. Neurol. **11**, 196–204 (2004)
L. Squire, Declarative and nondeclarative memory: multiple brain systems supporting learning and memory. J. Cogn. Neurosci. **4**(3), 232–243 (1992)
N.H. Sykes, J.A. Lamb, Autism: the quest for the genes. Exp. Mol. Med. **9**, 1–15 (2007)

M. Tomasello, M. Carpenter, J. Call, T. Behne, H. Moll, Understanding and sharing intentions: the origins of cultural cognition. Behav. Brain Sci. Behav. Brain Sci. **28**, 675–735 (2005)

L.S. Vygotsky, *Thought and Language* (MIT Press, Cambridge, MA, 1962) (Original work published in 1934)

L.S. Vygotsky, Play and its role in the mental development of the child. In Play ed. by J.S. Bruner, A. Jolly, K. Sylva, (Harmondsworth, UK: Penguin, 1976)

R. Wallace, Consciousness, crosstalk, and the mereological fallacy: an evolutionary perspective. Phys. Life Rev. **9**, 426–453 (2012)

S.R. Waxman, D.B. Markow, Words as invitations to form categories: evidence from 12- to 13-month old infants. Cogn. Psychol. **29**, 257–302 (1995)

D.W. Winnicott, *Playing and Reality*, (Tavistock Publications, 1971)

J. Woorim, C.K. Chung, J.S. Kim, Episodic memory in aspects of large-scale brain networks. Front Hum. Neurosci. **454**, 1–15 (2015)

T. Zalla, N. Labruyere, N. Georgieff, Goal-directed action representation in autism. J. Autism Dev. Disord. **36**(4), 527–540 (2006)

P.D. Zelazo, The development of conscious control in childhood. Trends Cogn. Sci. **8**, 12–17 (2004)

H.N. Zelaznik, The role of motor development in infancy, in *The Development of Coordination in Infancy*, ed. by G.J.P. Savelsbergh (North-Holland, Amsterdam, 1993), pp. 79–88

Manzai Robots: Entertainment Robots as Passive Media Based on Autocreated Manzai Scripts from Web News Articles

40

Tomohiro Umetani, Akiyo Nadamoto, and Tatsuya Kitamura

Contents

T. Umetani (✉) • A. Nadamoto • T. Kitamura
Department of Intelligence and Informatics, Faculty of Intelligence and Informatics, Konan University, Kobe, Hyogo, Japan
e-mail: umetani@konan-u.ac.jp; nadamoto@konan-u.ac.jp; t-kitamu@konan-u.ac.jp

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_61

Abstract
This chapter introduces a manzai robot system, that is, an entertainment robot that is used as passive media based on manzai scenarios that are autocreated from web news articles. Manzai is a Japanese traditional standup comedy act that is usually performed by two comedians: a stooge and a straight man. Manzai robots automatically generate their manzai scripts from web news articles based on related keywords given by the audience and search results from the World Wide Web (WWW), and then perform the manzai scenarios. A manzai scenario comprises three parts: *tsukami* (the beginning of the manzai greeting), *honneta* (main body of the manzai script), and *ochi* (conclusion of the manzai performance). The style of manzai scenario is "*shabekuri manzai*," which means talk constructed from only the manzai scenarios. The manzai scenario for the robots is served as a manzai script written in extended markup language (XML). The manzai robots are constructed to aim to facilitate the observation of the entertaining dialogue using manzai robots as a socially passive medium. The manzai robot system is focused on the content generator using the automatic script creation function. The chapter introduces the automatic creation of the manzai scripts from web news articles, management of the manzai robot systems. Then, the component-based manzai robot system is explained to make a scalability of the manzai robot system. The chapter verifies potential of the manzai robot system by implementing an automatic manzai scenario creation system and the management systems using real robots.

Keywords
Manzai • Passive medium • Automated scenario generation • WWW • Robot-robot conversation

Introduction

An aging society requires extension of healthy life expectancy and improvements in the quality of life of older people. It is essential that members have a certain amount of communication to improve their quality of life. Robots that encourage smooth communication are thus being actively developed in this context (Yamamoto et al. 2002; Matsui et al. 2010; Kanoh et al. 2011). Communication robots activate human-robot interaction to achieve this purpose. Kanda et al. conducted a study in which they found that observation of dialogue between robots encourages people to communicate with the robots naturally and smoothly (Kanda et al. 2002). Subsequently, in accordance with the results of that study, a pair of manzai robot systems was developed as a passive medium (Umetani et al. 2014).

This chapter introduces a manzai robot system, which is an entertainment robot, which is used as passive media, based on manzai scenarios that are autocreated from web news articles. Manzai is a Japanese traditional standup comedy that is usually performed by two comedians: a stooge and a straight man. Typically, a manzai performance comprises a humorous dialogue routine. In Japan, manzai performances are now being broadcast on various media outlets such as weekend television; consequently, people have grown very familiar with manzai. Manzai robots are constructed with the aim of facilitating the observation of entertaining dialogue between manzai robots as a socially passive medium. The manzai robot system is focused on content generation using an automatic script creation function.

The manzai robots automatically generate their manzai scripts from web news articles based on related keywords suggested or given by the audience and search results from the World Wide Web (WWW), and then perform the manzai scripts. A manzai scenario consists of three parts: *tsukami* (the beginning of the manzai greeting), *honneta* (main body of the manzai scenario, and *ochi* (conclusion of the manzai performance). The style of the manzai scenario is "*shabekuri manzai*," which means talk constructed from only the manzai scenario.

Several studies have been conducted on the performance motion of communication robots, for example, "Robot Manzai" (Hayashi et al. 2008). In the studies conducted, the scripts for performances – the motion of the robots and timing of the performance – are created by an engineer in advance. In contrast, the proposed manzai robot system is focused on content generation. The manzai scenarios are created automatically by using data-mining techniques, and the methodology of manzai performance is derived from Internet-based news articles (Mashimo et al. 2014). The manzai scenarios are served as the manzai scripts for each robot that are written in extended markup language (XML). Subsequently, each robot performs the corresponding role in the manzai script.

Manzai from robot dialogue has become a medium for information exchange wherein it is assumed that the information transferred using manzai scenarios is information that is more familiar. Matsumoto, a professional manzai scenario writer, states that "Manzai scenarios based on news articles are the easiest for people to understand" (Matsumoto 2008). Given that assertion, manzai scenarios are generated based on news articles from the Internet using many kinds of intelligence techniques such as word ontology and search tags obtained from the Internet search engines. The flow of the manzai scenario generation system is depicted in Fig. 1. The procedure of manzai robot system is shown as follows:

(1) User inputs a keyword to the system.
(2) The system obtains a news article that includes the keyword from the Internet.
(3) The manzai robot system generates a manzai scenario consisting of humorous dialogues using fluffy patter and misunderstandings in real time. Subsequently, the system generates the manzai script for the manzai robots.
(4) The manzai robots perform the manzai script in real time.

This chapter introduces a method for automatic creation of manzai scenarios from web news articles and management of the manzai robot systems. Then, a component-based manzai robot system designed to make the manzai robot system scalable is outlined. Further, the chapter verifies the potential of the manzai robot system by implementing management systems and an automatic manzai scenario creation system by using real robots.

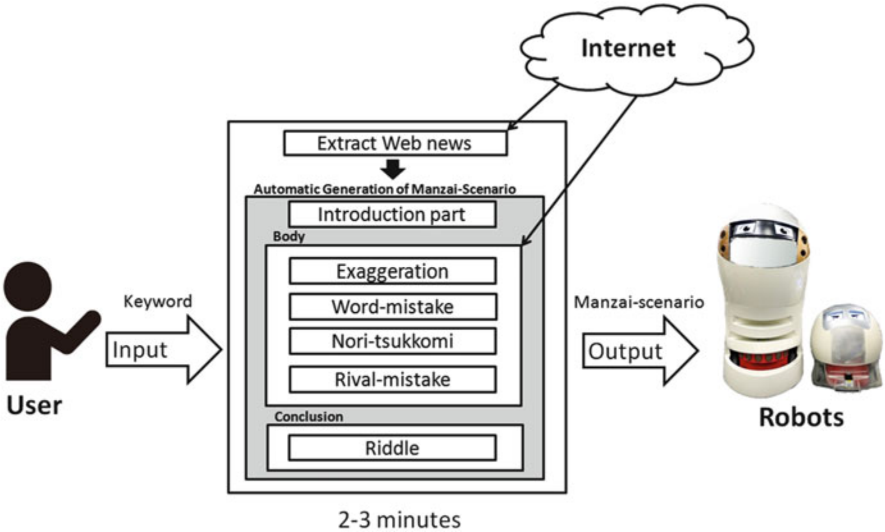


Fig. 1 Overview of automated creation of manzai scenarios (Mashimo et al. 2014) © 2014 Association for Computing Machinery, Inc. Reprinted by permission

Related Works

Hayashi et al. proposed a "robot manzai" system for human-robot communication (Hayashi et al. 2005). This system realized "manzai" using a pair of robots. They conducted a comparison experiment between "robot manzai" and "manzai" shown in a video performed by humans. Consequently, they demonstrated the usefulness of "robot manzai" as entertainment. Their robots specifically examine the human action in "manzai," but we particularly examine the scenario as Manzai contents.

Our proposed system is important also in that it facilitates understanding of news contents. Park et al. developed the "News Cube," which is a news browsing service that mitigates media bias (Park et al. 2009). Kitayama and Sumiya proposed a new search method of retrieving comparison contents for news archives (Kitayama and Sumiya 2007). Our proposed system is aimed at understanding of news contents by generating humorous dialogue based on news articles.

Numerous studies have presented and assessed dialogue analysis. Ishizaki and Den present a good summary of work in this area (Ishizaki and Den 2001). Their approaches include analyses of real-world dialogues and extraction of intentions from the dialogue. Our efforts include the generation of humorous dialogue from Internet news articles. Numerous studies, however, have examined conversational agents (Schulman and Bickmore 2009; Bouchet and Sansonnet 2009; Ishii et al. 2013). In almost all of these studies, CG characters communicate with humans based on dialogue. Applications of these studies are used in education, entertainment, communication support, and other fields for communication with CG characters by humans. As the main purpose of our research, however, people are expected to be entertained and healed by watching the robot dialogue, which is based on our proposed automatically generated scenario.

Research undertaken in the field of robotics has investigated many entertainment robots (Yamaoka et al. 2006; Shiomi et al. 2008; Khosla and Chu 2013; Rae et al. 2013). Paro is an entertainment robot that heals people with his cute gestures (Shibata 2010). PALRO, ifbot (Kanoh et al. 2005), KIROBO, and others communicate with people using cute dialogue. They talk with people, but our Manzai robots mutually converse. People only need to watch their humorous dialogue to be entertained.

Automated Generation of Manzai Scenario from Web News Articles

This section describes a method for automated generation of manzai scenarios from web news articles. First, the section illustrates manzai, a Japanese traditional standup comedy that is usually performed by two comedians, stooge and straight man, and its structure in the manzai performance. Then, the method for generation of each part of manzai scenarios is described.

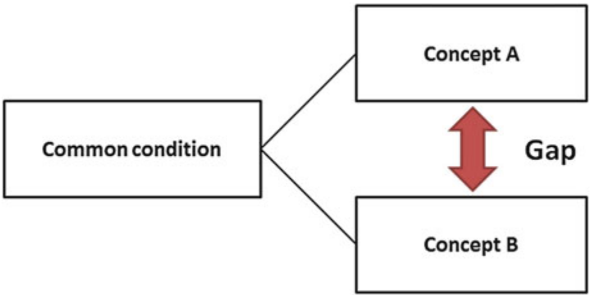
What Is Manzai?

As described in this section, an automatic manzai scenario generation system that produces humorous dialogue from web news articles is proposed. Japan has a traditional form of comedy routine called manzai, which typically consists of two comedians performing humorous dialogues. It is similar to “stand-up comedy” in English, or “xiang sheng” in Chinese. In Japan, Manzai has remained extremely popular over time. A manzai show is now broadcast on TV every weekend. People continue to have strong familiarity with Manzai. The manzai metaphor is therefore useful to create acceptable humorous robot dialogue. The automatic generation of a Manzai scenario consisting of humorous dialogue generated automatically from web contents is proposed.

Manzai usually includes two performers: one is the *boke*, or stooge; the other is the *tsukkomi*, or straight man. The *boke* says things that are stupid, silly, or out of line, sometimes using puns, whereas the *tsukkomi* delivers quick, witty, and often harsh responses. Regarding our manzai robots, Ai-chan (left side of Fig. 2) is the *tsukkomi* (straight woman); Gonta (right side of Fig. 2) is the *boke* (stooge). Furthermore, the manzai scenario has a three-part structure: the Introduction, the Body, and the Conclusion.

Fig. 2 Manzai robots (*Left*: ii-1, Ai-chan; *Right*: ii-2, Gonta)

Fig. 3 Structure of funny point



Introduction Part

The Introduction part consists of several short dialogues. In the system, it consists of a greeting and a discussion in which the theme of the original web news is first presented.

Body Part

The Body part is the main part of the manzai scenario. It consists of humorous dialogue. The Body part is designated as the block of humorous dialogue called dialogue component. The dialogue component is created automatically based on a sentence in a web news article. That means, the manzai scenario generation system creates a dialogue component which is a set of humorous dialogue from a sentence in a web news article. The system generates humorous dialogue based on structure of funny points of dialogue using gaps of two types, which are rival relations and sentiment gaps (see Fig. 3).

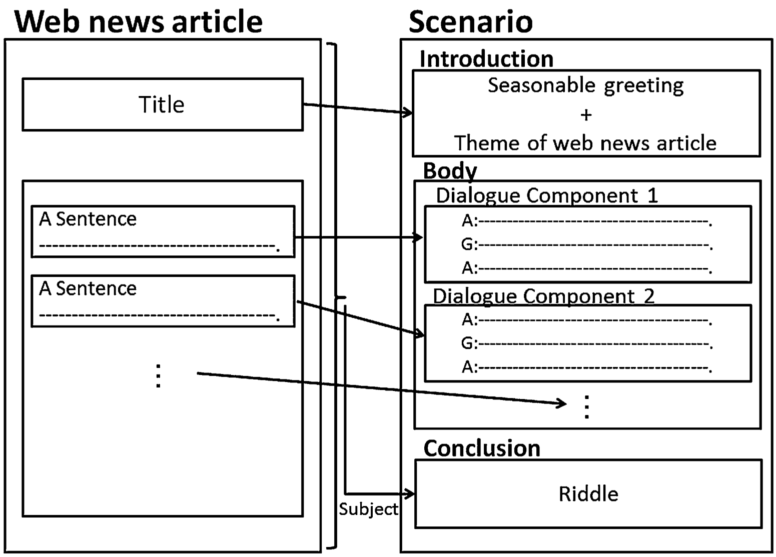


Fig. 4 Image of correspondence of web news and scenario

Conclusion Part

The Conclusion part includes a farewell and a final laughing point. In our system, theConclusion part is a riddle of words related to web news. Figure 4 shows the image of correspondence of web news article and automatically created scenario.

What Is Funny Point?

In Japan, usual points of humor are snappy patter and misunderstandings. The generated manzai scenarios include dialogue of that type. Matsumoto, a professional manzai scenario writer, said that "Manzai scenarios based on news articles are the easiest for people to understand" (Matsumoto 2008). Given that assumption, manzai scenarios are generated based on news articles from the web, using knowledge of many kinds such as word ontology, extraction of sentiment, and similar words obtained from searches of the Internet.

Abe has described the structure of funny points that consist of a gap separating concepts (see Fig. 3) (Abe 2006). He says the concept gap is an important aspect of the humorous points. That point is specifically examined in the generation system of manzai scenarios. Then, the structure of the funny points for dialogue is described based on Abe's structure. In the structure, the original sentence of web news article is Abe's common condition. A word which is in an original sentence is concept A and a word which is extracted by our system is concept B.

In this study, gaps of two types, which are important for funny dialogue, are proposed according to Abe's results. One is a rival relation; the other is a sentiment gap. In the case of a rival relation, rival words are extracted from the web, such as "baseball" and "soccer." For a sentiment gap, a sentiment is extracted from the original sentence. Then a new sentence having the opposite sentiment is created. As described herein, a means to extract rival keyword from the web and to create sentiment gap from the web are proposed. Then, the proposed technique is used to generate a manzai scenario for the robots.

How to Generate Manzai Scenario

The flow of the manzai scenario generation system is the following:

(1) A user inputs a keyword to the system.
(2) The system gets a news article that includes the keyword from the internet.
(3) In real time, the system generates a manzai scenario consisting of humorous dialogues based on knowledge from the web.
(4) The robots "Ai-chan and Gonta" (see Fig. 2) perform a manzai scenario in real time.

When a user inputs keywords for Manzai to be watched, we generate the web news-based manzai scenario automatically based on the three parts of actual manzai routines, which are the Introduction part, Body part, and Conclusion part. In each part, the system generates humorous dialogue from web news articles with related information obtained from the Internet. The system transform declarative sentences of web news articles into humorous dialogue using related information from the Internet based on the proposed two types of gap.

Introduction Part

The Introduction part consists of a first laughing point with a greeting and a connection to the theme of manzai to the main part. In the system, the first laughing point with a greeting is a season-related greeting, along with presentation of the theme obtained from the original web content. The system creates a season dialogue database in the system and chose dialogue from the database. In October, for example, the Introduction part of the manzai has greeting of October and web news title is "Bolt Achieves Triple Crown!!" is the following, where A denotes Ai-chan and G stands for Gonta.

G: Hello Everyone. It's October. Halloween season!
A: Ah yes, Halloween season is coming. Anyway, have you heard the news?
G: Hmm... "Bolt Achieves Triple Crown!!" I did not know that.
A: That sounds interesting. Let's start a manzai routine about the Bolt Achieves Triple Crown!!

Body Part

The automatically generated scenario consists of multiple components in body part. One component is one funny technique. In this study, five types of funny technique which are "exaggeration," "word-mistake," "Nori-tsukkomi," "rival-mistake," and "sentiment-mistake" are proposed. The proposed structures of funny points are gaps separating dialogue. The manzai scenario generation system generates humorous dialogue based on our structure of funny points of dialogue (see Fig. 3) using gaps of four types, which are number gap, topic gap, rival relation, and sentiment gap.

(1) Exaggeration component

The first step to generating dialogue using exaggeration is to use impossible (larger or smaller) numbers. This component is number gap. This humorous technique is sometimes used ordinarily in dialogue routines in Japan. People feel familiar about such types of exaggeration. When a sentence includes numbers, the system increases the numbers by some substantial and unbelievable factor. For example, if a sentence is "Bolt had already won the 100 m and the 200 m," the exaggerated dialogue becomes

G: At the championships, Bolt had already won the 100 m and the "200 km."
A: Two hundred kilometers is way too long! Don't you think that's odd?
G: You are right! It is not 200 km, but 200 m. I misunderstood.

(2) Word-mistake component

This component is an intentional word-mistake in the dialogue. This is topic (word) gap. A typical word-related mistake is to change a single word to another word based on changing only one character in a word, such as "fish" and "dish."

G: "Belt" pulled away at once while the USA botched passing the baton to the anchor runner.
A: How can that be? Belt? It's not belt, but Bolt!
G: Whoops. I made a careless mistake.

The Manzai robot has only one dictionary, which is Japanese dictionary for children in the server (Ai-chan) machine. The Japanese dictionary is used to extract mistaken words such as "Belt." Japanese consists of vowels and consonants. It is a simple matter to change the first consonant to another one. If there is a word for which the first consonant in the dictionary can be changed, it becomes the candidate of the mistaken word. For example, In Japanese "Touhyou (vote)" is changed to "Kouhyou (favorable comment)."

After a word-mistake, Ai-chan, the *tsukommi*, emphasizes the mistaken words by explaining the mistaken word. In the example given above, Ai-chan explains the Belt as "That's a flat, long narrow piece of cloth or leather mainly used for fixing an object."

(3) Rival words gap-based dialogue

It is inferred that if the words which are in dialogue have a mutual rivalry, then the system changes the original word in a sentence to a rival word; the dialogue becomes a sentence that includes a misunderstanding. Then, it becomes a funny point. The proposed rival words are "Tokyo" and "London," and "baseball" and "soccer." The two words are contrasting pairs. Then, the system extracts a word which is a rival to the keyword and changes the word from the keyword to the rival word. The following presents an example of the dialogue based on a rival mistake:

A: ... By the way, do you know what "Usain Bolt" is like?
G: He's, you know, famous as the Olympian of the century, right?
A: No! You must be confused with Carl Lewis. Usain Bolt is a Jamaican sprinter. He is the fastest sprinter in human history with the nickname of "Lightning Bolt."
G: Is that so? But they are similar enough, aren't they?

Definitions of proposed rival words are the following:

(3-1) They have the same upper ontology.

The upper ontology of "Tokyo" and "London" is national capitals. That of "baseball" and "soccer" is ball sports. Each pair shares the same upper ontology. When the system extracts the upper ontology, the hierarchy structure of Wikipedia is used. At this time, the upper ontology of a word is not just one ontology; usually a word has multiple upper ontologies. All upper ontologies are used to extract rival words. The rival words are child words of the upper ontology. Then many words can become candidate rival words. The system regards the candidates of rival words as ranked. The top-ranking word becomes a rival word. At this time, it is considered that an upper ontology that has few child words is more important than an upper ontology that has many child words because the former has a higher instance level than the latter. The system then calculates the ranking of a candidate of rival words using the following expression:

$$\mathrm{Sta}(s_i) = 1 - \log\frac{n}{N},$$

$$\mathrm{Rel}(e_i) = \sum\nolimits_{i=0}^{m} \mathrm{Sta}(S_i). \tag{1}$$

The given s_i denotes an upper ontology; Sta(S_i) signifies weight of s_i. In addition, n represents a number of s_i 's lower ontology; N is the number of words in the corpus. The given e_i stands for a candidate of rival word; Rel(e_i) signifies the ranking weight. In addition, m represents a number of e_i 's upper ontology, which is the same as the keyword. As described herein, the system uses N 2,931,465 words.

(3-2) They have a similar recognition degree.

The candidate rival words of “baseball” are “soccer” and “futsal.” At this time, the degree of recognition of soccer is more similar to that of baseball. Therefore, the system regards soccer as better than futsal for use as a rival word of baseball. That is, the two words have similar degrees of recognition. The system regards the degree of recognition as the number of results obtained from a web search. The similarity of the degree of recognition Con(key, e_i) between keywords key and e_i is calculated as follows:

$$\mathrm{Con}(\mathrm{key}, e_i) = 1 - \log \frac{|\mathrm{Cog}(\mathrm{key}) - \mathrm{Cog}(e_i)|}{\max\{\mathrm{Cog}(\mathrm{key}),\ \mathrm{Cog}(e_i)\}}. \tag{2}$$

In that equation, Cog(key) is the number of results of a web search using word key; Cog(e_i) is the number of results of a web search using word e_i.

After calculating Rel(e_i) and Cog(key), the system regards the result of geometric mean between Rel(e_i) and Cog(key) as a ranking weight. The word having the highest ranking weight becomes a rival word. Subsequently, the rival-mistake component is created based on changing the keyword to the rival word, and generating a misunderstanding humorous dialogue.

Table 1 presents an example of the rival words we extract.

(4) Sentiment gap-based dialogue

Sentiment mistake gaps of two types are proposed in this study: a word sentiment mistake type and a sentence sentiment mistake type.

(4-1) Word sentiment mistake

It is inferred that when the sentiment of the word is mistaken, it is interesting. At this time, the Japanese traditional technique called “*Nori-tsukkomi*” is used. In a nori-tsukkomi, first the *boke* (stooge) states some outrageous sentence and plays the clown. Next, the *tsukkomi* (straight man) gets into line with the *boke*’s stupid sentence. Then, the *tsukkomi* sets the right sentence and makes a fool of both the *boke* and itself. Under such circumstances, there are three techniques for dialogue generation, as described below.

Table 1 Rival-words are related to keywords

Keywords	Rival-words
Italy	France
Japan	China
Baseball	Soccer
Chess	Reversi
Dog	Rabbit
Barack Obama	George Washington
Internet	Newspaper

(1) How to create the first outrageous statement?
(2) How does the *tsukkomi* agree with the *boke*'s line?
(3) How does one create a corrective sentence?

The word-mistaken technique is used in (1). The word-mistaken technique is where the single character is changed, producing a different word, such as "hose" and "nose." In (2), a sentiment of the mistaken word is used in (1). At this time, the system extracts sentiment words that co-occur with mistaken words from the web and calculate the co-occurrence ratio. The system specifically searches the web pages using the word that it wants to be extracted a sentiment as a keyword, and the system extracts sentiment words which are adjectives from the snippet of the results. Then, the system calculates co-occurrence ratio between the word and each sentiment word. The sentiment word having the highest co-occurrence ratio becomes the sentiment of the mistaken word. In (3), the system uses original sentence in web news article. An example of *Nori-tsukkomi* follows:

G: "Belt" pulled away at once while the USA botched passing the baton to the anchor runner.
A: Stop right there. Belt is really long.
A: . . . Wait.
A: How can that be? Belt? That's a flat, long narrow piece of cloth or leather mainly used for fixing an object, isn't it! It's not belt, but Bolt!
G: Whoops. I made a careless mistake.

In (2), there is agreement based on the fact that his nose is long and his impression word of nose is "long," but after that agreement Ai-chan mentions the mistaken word. At this time, the system extracts a sentence mentioning the mistaken word from Wikipedia. The Wikipedia sentence is usually a hard and formal sentence, which emphasizes a humorous contrast with the humorous dialogue.

(4-2) Sentence sentiment mistake

It is considered that the sentence in a dialogue has opposite sentiment to the original sentence. It is a misunderstanding of dialogue. It becomes gap of structure of funny points. For example,

G: "Bolt Achieves Triple Crown." It is very sad news.
A: What? Why do you think the news is sad?
G: Because "Bolt Achieves Triple Crown." I feel very sad. . .
A: Do you want to get triple crown? Are you so fast?

The system calculates the sentiment of the sentence and changes the word which has opposite sentiment of the original sentence. Then Ai-chan corrects the sentiment of the Gonta. In above case, correct first sentence of the Gonta is "Bolt Achieves

Triple Crown," and it is a good news article in a web news article. The system changes the "achieve" to "lost" and changes the sentiment of the sentence.

When the system generates dialogue based on sentiment mistake gap, multidimensional sentiment that is three bipolar scales and proposed by Kumamoto (Kumamoto et al. 2011) is applied. Kumamoto's three bipolar scales are sufficient to calculate web news article sentiments because he creates the bipolar scales using news articles. The bipolar scales are "Happy – Sad," "Glad – Angry," "Peaceful – Strained." Generating a sentiment mistake gap-based dialogue is done as follows:

(1) The system calculates the sentiment of each sentence in a web news article using Kumamoto's sentiment extraction tool (Kumamoto et al. 2011). At this time, the results are sentiment values in each bipolar axis. The results are normalized such as value of 1.0 – 0.0 means left side axis and 0.0 – −1.0 denotes the right side axis. For example, when the result is happy – sad as 0.12, glad – angry as 0.26, and peaceful – strained as −0.07, the sentiment value of the sentence becomes happy as 0.12, glad as 0.26, and strained as 0.07.
(2) The system infers that the highest sentiment value of a sentence is a sentiment A_e of the sentence. In a case of (1), the glad becomes the sentiment of the sentence.
(3) The system extracts the word W_e which has highest value of A_e. Then we extract antonym T_e of W_e from the antonym corpus.
(4) A new sentence using T_e is created.

Conclusion Part

The conclusion part consists of a farewell and the final laughing point. The system uses an automatically generated riddle in our Conclusion part. Japanese people are familiar with riddles. They have been used for entertainment for a long time. In Japanese, riddles are usually "How are X and Y similar?" The answer is typically some form of homonym or pun: Z and Z'. For example,

G: What is the similarity between Bolt and driver?
A: I do not know. What?
G: Both of them have a track ("truck")!!

In this example, Bolt is X, driver is Y, track is Z, and truck is Z'.

The system first extracted the word X (Bolt) from the web news article. Next the system extracted words that have a high co-occurrence ratio to X from the Internet as candidates for word Z. The system next extracted homonyms of candidates of Z from the dictionary, which is inside AI-chan. If the candidate of Z has a homonym, then it becomes Z. The homonym becomes Z'. Subsequently, the word having the highest co-occurrence ratio to Z' is extracted from the Internet. It becomes Y. Then, the system generated dialogue based on X, Y, Z, and Z'.

Figure 5 shows an automatically generated scenario from web news article in Table 2.

Fig. 5 Example of automatically generated manzai scenario

Table 2 Original web news article

Title: Bolt Achieves Triple Crown: Another victory in the 4 × 100 m relay
The men's 4 × 100 m relay final took place at the IAAF World Championships held in Beijing. Jamaica won the championship, setting this season's world record of 37.36 s. Usain Bolt achieved a triple crown with victories in the 100 m and 200 m races.
In the final, Bolt ran the anchor leg for Jamaica. Although Jamaica was competing for the top position with the USA from the start, Bolt pulled away at once while the USA botched passing the baton to the anchor runner. Jamaica set this season's world record of 37.36 s and earned a fourth consecutive victory.
At the championships, Bolt had already won the 100 m and the 200 m races. With victory in the 4 × 100 m relay, he achieved a second consecutive triple crown, following the last championships held in Moscow.
The USA finished second but was disqualified, which allowed China to come in second, followed by Canada in third place.

Manzai Robot System

This section describes a manzai robot system that participates in a manzai performance in accordance with the manzai script generated by a manzai scenario generator. First, the configuration of the manzai robots is illustrated. Next, implementation of the manzai robot system is outlined and the manzai scripts are analyzed. Finally, experimental verification of the manzai robot system is discussed.

Manzai Robots

Figure 2 shows the manzai robots developed in this study. The taller robot, ii-1, Ai-chan, which is about 100 cm tall, performs the role of *tsukkomi* – the straight woman – and the shorter one, ii-2, Gonta, which is about 50 cm tall, performs the role of *boke* – the stooge. *Tsukkomi* and *boke* have fixed roles.

Each robot has a computer on its back and the two communicate via a wireless LAN. The computer on *tsukkomi* is a server connected to the Internet that directly obtains articles and automatically creates manzai scripts.

The two robots each have the following functions:

- Locomotion and rotation using Pioneer 3-DX (Mobile Robots, Inc.)
- Creation of facial expressions by switching images on the eye display
- Speech generation of a synthesized script-based voice

The manzai robot system manages operations using these functions in accordance with the manzai scripts which are the manzai scenarios for the manzai robots.

System Configuration

Figure 6 shows the configurations of the manzai robots. The PC mounted on ii-1 creates manzai scripts, time flows, and manzai performance schedules. The PC

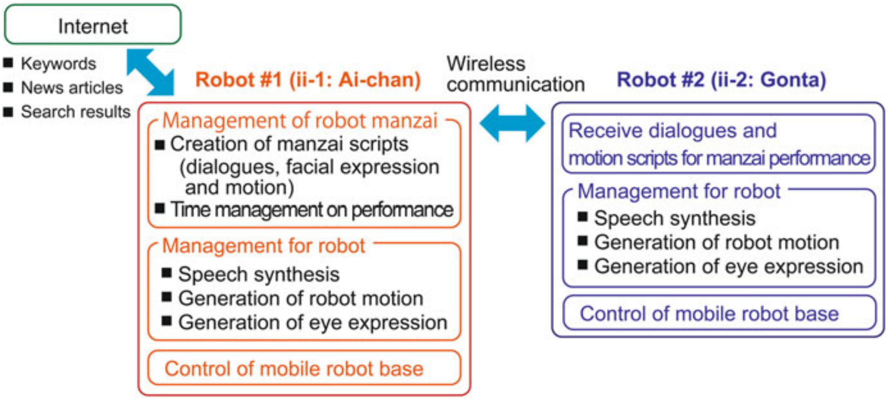


Fig. 6 System configuration of manzai robot system

mounted on ii-2 communicates with the PC on ii-1, receiving instructions on lines, motions, and expressions to create the data for robot motion and speech.

After the scripts are created, the two robots perform script-based manzai. The manzai script progress management program running on the PC on ii-1 obtains and processes the information for the next lines to be spoken, facial expressions, and motions based on the progression of the script. When it is the turn of ii-2 to perform, this information is transmitted to the PC on ii-2.

On receiving the information, the operation program running on the PC creates synthesized voices, facial expression changes, and robot motions as required. Facial expressions change automatically corresponding to emotions, expressions, and the eye movements making up facial expressions. The robots operate based on the operation commands they receive.

On completion of the motions that drive the robot mechanism, termination information is sent to the script progress management program of the PC mounted on ii-1. The system then moves to the next session of the manzai script based on this information.

Command of Robot Motion for Manzai Performance

Manzai scripts for manzai performance are served as extended markup language (XML) scripts (Nadamoto and Tanaka 2005). XML scripts are generally used for motion media robot systems and networked robots to facilitate control based on scenarios (Kitagaki et al. 2002; Tezuka et al. 2006). The manzai robot system automatically generates the manzai scripts from WWW news articles corresponding to the keywords given by the audiences. The manzai robot system analyzes the manzai script, then each robot moves and makes a speech in the skit according to the manzai script.

The sentences in each section of the manzai script are the control command of the manzai robot system. The control command consists of three types of commands: motion, expression, and speech of dialogue in the scenario. The following are examples of control commands: (Note that “Mary” and “Bob” in the manzai script correspond to ii-1 and ii-2, respectively.)

<look name = "Mary", what = "audience">

This command indicates information about the direction of the robot in manzai performance, that is, “which” robot is directed in “which” direction. In this example, the manzai script intends that “Mary” look directly to “audience” (face the audience).

<PEmo name = "Bob">PE11/>

This command indicates information about the facial expression of the robot in manzai performance; “which” robot expresses “which” facial expression. The number of “PExx” is the intention of the facial expression; therefore, the robot shows the facial expression according to the command. “xx” is the number of facial expression patterns.

<cast name = "Mary">Speech is here.</cast>

The command indicates information about the speech of the manzai script in manzai performance; “which” robot makes a speech and its contents.

The robot analyzes the commands for manzai performance, then the robots are controlled according to the command, for example, synthesizing speech, movement, changing the facial expression.

The potential of the manzai robot system was verified by implementing a full manzai robot system including an automatic manzai script creation system and a management system using real robots.

Experiments

We conducted user experiments to assess the benefits of our proposed humorous dialogue automatic generation system. We conducted experiments with 11 participants, all of whom are Japanese and who like Manzai. We use Manzai scenarios of two types generated automatically by our system. The Manzai robots, Ai-chan and Gonta, perform the Manzai routine using our Manzai scenario. We evaluated two aspects of our system based on each part (or component) of the Manzai scenario and the whole Manzai scenario. We want the scenarios to be interesting and understandable. We asked participants to complete a questionnaire. Participants responded to the questionnaire on a five-point scale (5, highest; 3, middle; 1, lowest) after watching the Manzai performance.

Interest and Comprehension Ratings of Each Part and Component of the Manzai Scenario

We evaluated the interest and comprehension of each part and component of the Manzai scenario routine. We asked participants to report their interest and

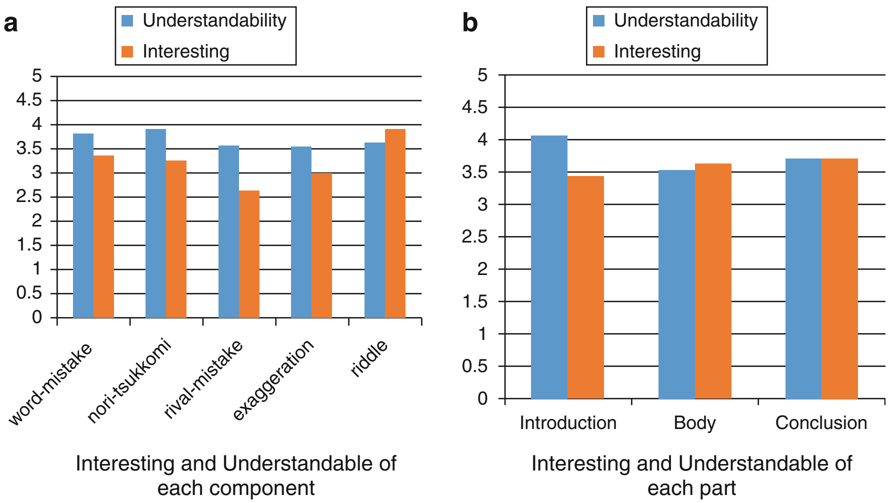


Fig. 7 Results of interest and comprehension for total manzai scenarios (Mashimo et al. 2014) © 2014 Association for Computing Machinery, Inc. Reprinted by permission

comprehension of the three structural components of the Manzai scenario: the Introduction, Body, and Conclusion. We also asked them to report their interest and comprehension of five components: exaggeration, word-mistake, nori-tsukkomi, rival-mistake, and riddle.

Results and Discussion

The results of the experiment are presented in Fig. 7. The results show a high score, meaning that our proposed Manzai scenario technique produces interesting and understandable routines. The rival-mistake component elicited the lowest score, meaning that usually Japanese people use rival mistakes in humorous dialogue. In each part, which are the introduction, body, and conclusion, the results also show a high score. We regard these structures as familiar to Japanese people because real Manzai routines have the same structure.

Interest and Comprehension of Whole Manzai Scenarios

We evaluate the interest and comprehension of whole Manzai scenarios. We asked participants to complete a questionnaire including the following questions:

(1) How interesting is it to watch Manzai robots?
(2) Is the speaking speed of Manzai scenarios appropriate?
(3) Do you feel that the number of component types in the scenario is sufficient?
(4) Is the Manzai scenario understandable?
(5) Did transforming the news page into a Manzai scenario make it easier for you?

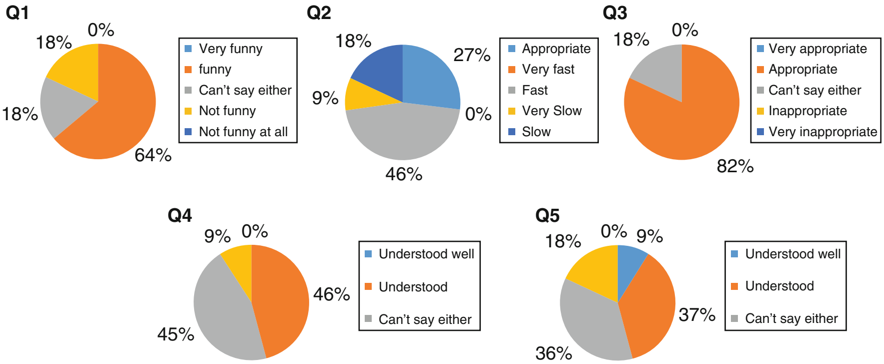


Fig. 8 Results of interesting and understandable total manzai scenario (Mashimo et al. 2014) © 2014 Association for Computing Machinery, Inc. Reprinted by permission

Results and Discussion

The experiment results are presented in Fig. 8. The average of Q1 is 3.45; seven participants answered 4. Therefore, our proposed method generated the humorous dialogue. To Q2, four participants answered 2. They said that the speaking speed of the Manzai robots was bad. However, two participants responded that it was a little bit fast. The impression of speaking speed apparently differs among people. The average of Q3 was 3.82; no participant answered 2. The result means that the component of humorous dialogue is sufficient. The average of Q4 is 3.36; Q5 is 3.36. Our generated humorous dialogue does not lose the original news article information. Therefore, we can deliver news articles using humorous dialogue.

Component-Based Manzai Robot System with Scalability

This section describes a component-based manzai robot system with a small body and the scalability of the software development for the manzai robot system. First, the objectives of the component-based manzai robot system are illustrated. Next, the implementation of the component-based manzai robot system is outlined. Finally, verification of the system is discussed with respect to scalability of the robot system, including the development of manzai robots and the associated interactive information system.

Development of Component-Based Manzai Robot System

The conventional manzai robot system has several problems (Umetani et al. 2014):

- The size of the manzai robots is quite large; a manzai robot system consists of two generic PCs, mobile robot bases, and their bodies. Therefore, it is difficult to conduct manzai performances for experimentation.

- The robot system for manzai performance is complicated. Two robots are needed to perform a manzai skit. Further, distributed controllers for each robot and a management system for the overall robot system are required. It is difficult for the conventional manzai robot system to add new functions for performances.
- The software development environments for robots change during software life cycles.

Consequently, a component-based manzai robot system was developed to overcome these problems. The requirements of this manzai robot system are as follows:

- The robot system should have sufficient portability.
- The robot system should perform manzai using the manzai script used by the conventional manzai robot system.
- Distributed middleware should be applied to facilitate manzai performance data communication between robots. The software of the system is built using individual software components; therefore, it is easy to reuse the control software.
- The developed manzai robot software components are executed on the conventional manzai robot system with minimal changes to the hardware of the system such as the mobile robot base, display, and the user interface parts.

Figure 9 shows the developed component-based manzai robot system (Umetani et al. 2015). To perform the manzai skit, we use two robots; the left one is ii-1 s (*Ai-chan*), the right one is ii-2 s (*Gonta*). The height of each robot is 250 mm for ii-1 s and 150 mm for ii-2 s. The width of both robots is about 150 mm. The RT middleware (Ando et al. 2005) was utilized for development of the distributed software components of the manzai robot system.

Fig. 9 Component-based manzai robot system (*Left*: ii-1 s, *Right*: ii-2 s)

Implementation of the Component-Based Manzai Robot System

System Configuration

The RT middleware scheme for the manzai robot system was applied to develop additional functions for the manzai robot system such as a control interface. Figure 10 shows the system architecture of the manzai robot system.

The robot system is executed by two generic PCs using Windows OS, with each PC controlling one robot. The PC for robot ii-1 s has the following roles:

(1) Generation of manzai scripts
(2) WWW server for facial expression data (managed by Apache WWW server)
(3) Speech synthesis of the dialogues for manzai performances
(4) Time management of manzai script during manzai performances
(5) Control of mobile robot base of ii-1 s, dialogue speech playback, and web page management for facial expressions

The PC for robot ii-2 s is responsible for the control of the mobile robot base of ii-2 s, playing of the speech of dialogues for ii-2 s, and web page management for facial expressions. The distributed control scheme of each robot for manzai performance makes changing the speaker for manzai performances and management of hardware resources easier. On the other hand, synchronization of manzai scripts during the manzai performance and a control scheme for each networked PC is required. Hence, the RT middleware is applied to control the networked distributed robot system.

The robot system uses Vstone Beauto Rover RTC BT as the mobile base of the manzai robot. An iPod touch viewer is used for the facial expression of each manzai robot. The output of the speech voice of manzai scripts and control of the mobile base are executed by the individual generic Windows PCs. The main controller PC is on the wireless network. A WWW server for the facial expression, viewer for balloon information system, and viewer for the user interface are executed on the

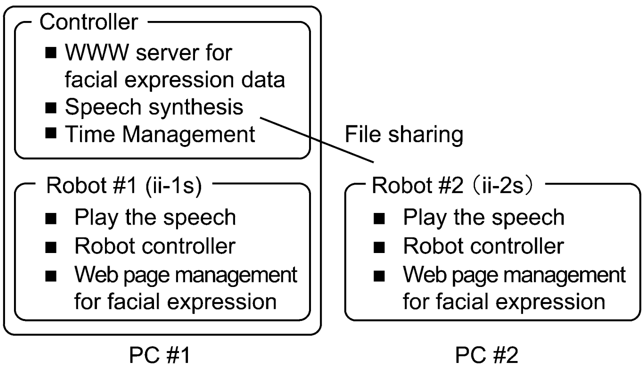


Fig. 10 System architecture of the component-based manzai robot system

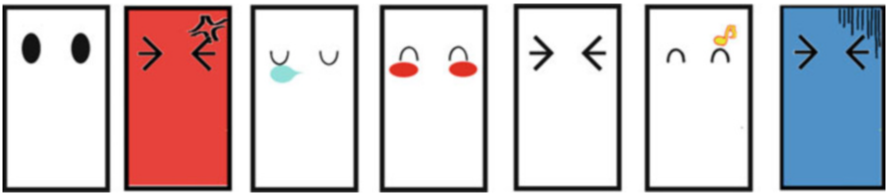


Fig. 11 Examples of the facial expressions of the component-based manzai robot

controller PC. The facial expression is expressed by a web browser executed on the iPod touch mounted on each robot. The facial expression of each robot is updated by updating the web page. Figure 11 shows examples of the facial expressions of the component-based manzai robot.

Implementation of the Component-Based Manzai Robot System

The component-based manzai robot system was implemented as a set of RT components. The RT components for the manzai robot system are as follows:

(1) Control of manzai robot system
(2) Facial expression during the manzai performances
(3) Control of each mobile robot base
(4) Speech generation during the manzai performances

Components (2), (3), and (4) are constructed for each robot. Each software component controls the various functions of each robot.

Figure 12 shows how the RT components are connected to the manzai robot system. "Mary" and "Bob" signify the components for ii-1 s and ii-2 s, respectively. The OS on the main controller PC was Microsoft Windows 7. Open RTM-aist 1.1.0 (Python) was applied for the manzai controller, manzai script generation, speech synthesis, speech generation during the manzai performance, and control of the facial expression of each robot. Open RTM-aist 1.1.0 (C++) was utilized for the controller of the mobile base. To reduce the calculation burden of the speech synthesis, all the speech voices for manzai performance were synthesized on a main controller PC in advance. Each function of the RT components is explained as follows:

Manzai_component: This component deals with management of the manzai robot system. First, the component synthesizes the speech voice of dialogues during manzai performances, and generates the control scripts for each component from XML files generated by the automated manzai script generation system. Then, it sends the control scripts for each PC. Following synthesis of the speech voice and generation of the control script, the component controls the speech dialogue of the manzai script, each mobile robot, and facial expression control components according to the XML file of the manzai script. The component also receives completion information regarding the playing of speech by the speech generation components, and then manages the manzai script during the manzai performance.

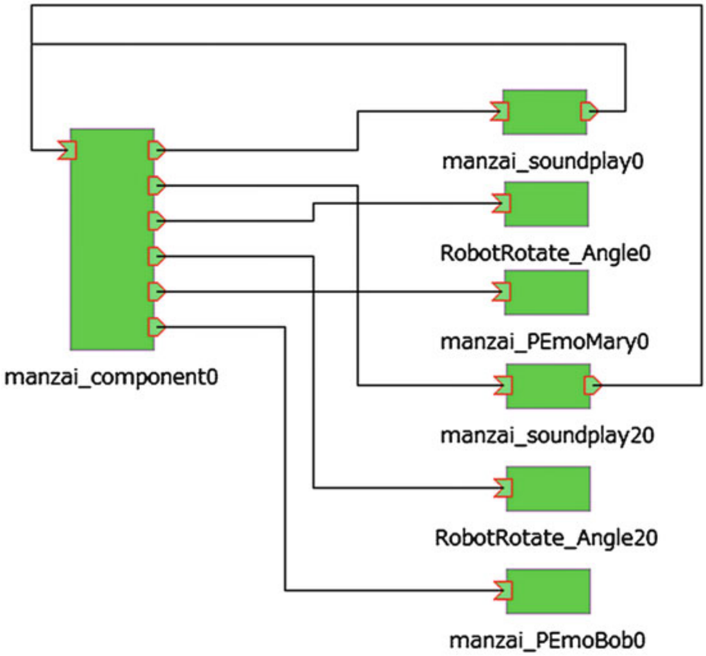


Fig. 12 Connection of RT components in the manzai robot system

`manzai_soundplay`, `manzai_soundplay2`: These components play the speech data file during the manzai performance. They receive the information for playing dialogue speech from `manzai_component`. On completion of the speech, they send completion information to `manzai_component`.

`RobotRotate_Angle`, `RobotRotate_Angle2`: These components control the motion of each mobile robot. They receive the information for manzai performance from `manzai_component`.

`manzai_PEmoMary`, `manzai_PEmoBob`: These components control facial expressions during the manzai performance. They receive information for manzai performance from `manzai_component`.

The manzai controller component generates scripts for control of each RT component from the original manzai scripts. The control script of the RT component enables reduction in the amount of communication between RT components, unification of the control order for each component, and independence of each component in the manzai robot system. Following synthesis of the speech of the manzai script, the RT component analyzes the manzai script.

The component outputs the control script for each component of the manzai robot system, such as control of mobile base, output of the speech, facial expression, and the balloon information system. Then, the control component sends the control script for the manzai performance to each RT component. Each RT component executes the control script according to the position of the control script sent by the control RT component. Therefore, restarting and recovery from error conditions are easily accomplished. Its robustness against network trouble is an improvement over the conventional manzai robot system.

Experiments were conducted to verify the component-based manzai robot system. The controller RT component synthesized the speech of the manzai script, and then each component outputted the speech, generated motion, and expressed the facial expression from the original manzai XML scripts for the conventional manzai robot system. The manzai XML scripts are not changed for manzai performance by the component-based manzai robot system. Therefore, manzai performance execution by the other manzai robot system was realized.

Balloon Dialogue Presentation System for Component-Based Manzai Robot System

The balloon dialogue presentation system was developed as an aid in scenarios where the audience cannot hear the manzai performance clearly. This system presents the speech of the manzai performance synchronized with the progression of the manzai scripts. Figure 13 shows a screenshot of the balloon dialogue presentation system. The system outputs the presentation using a web browser executing on the PC. The system shows an icon representing each robot and the speech in the script enclosed within the balloon from the robot's icon. Using the system, the audience can know "which" robot is currently speaking.

The presentation system is called by the controller component of the manzai performance. The system was developed in C#. The system updates the web page of the balloon dialogue presentation, and then controls the location of the progression of the manzai script.

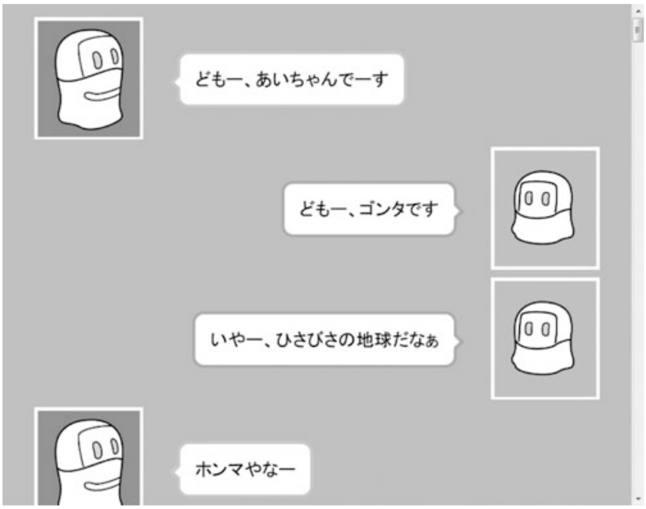


Fig. 13 Example of balloon dialogue presentation (dialogue is expressed in Japanese)

The system has to execute synchronously with the progression of the manzai performance. In the conventional manzai robot system, the system is complicated; adding such a presentation system that automatically executes in synchrony with the manzai performance is difficult. The result of the presentation system shows the scalability of the manzai robot system with respect to ease of addition of new functions to the system.

Scalability of Component-Based Manzai Robot System

To demonstrate the scalability of the system, we replaced the other mobile robot base with the robot base of the component-based manzai robot system. The mobile robot base was made by The Yujin Robot Co. Kobuki. The Kobuki_RTC component is used as the robot controller component (KobukiRTC 2013). The input values of the Kobuki_RTC are the angular and translational velocities of the mobile robot; the RT component for the control input of the robot base from the manzai script was constructed. With minimum change in the software component, the robot system operated well in the manzai performance. From this experimental result, the component-based system validated high scalability with respect to control of the different types of mobile robots using the same manzai script for the manzai performance.

Development of a component-based large-size manzai robot system with a function for sensing its surroundings in order to improve the interactive performances, and the robustness of the robot system under long-term experiments will be carried out in future works.

Discussion: Potential of Component-Based Manzai Robot System

From the results of implementation of the component-based manzai robot system, and application to the other type of mobile robot system, the feasibility and scalability of the proposed manzai robot system were demonstrated. The independent robot controllers for each robot using generic Windows PCs enable high portability and flexibility for the manzai robot system. The role of the robots in the manzai performance can easily be changed by changing the connection of the data ports between RT components, because the contents of the RT components for each robot are the same.

In addition, the control script for the manzai robot system enables the manzai robot system to be robust and scalable. The number of communication packets used for the synchronization and management of the manzai scripts during the manzai performance is also reduced. Even if the manzai performance is executed under the condition that there are many wireless LAN clients surrounding the manzai robots, the manzai performance by the robots can continue to completion. When there are many wireless LAN clients surrounding the manzai robot, that is, there are many spectators surrounding them, the performance of the wireless LAN is degraded. Thus, the reduction of communication packets makes the manzai robot system robust with respect to the communication network for the manzai performance.

In this study, an iPod touch viewer was mounted on each robot and used only for facial expressions during the manzai performance. The user interface of the manzai robot system using the viewer will be addressed. Moreover, the extension of the robot system using smartphone devices mounted on the manzai robots, such as the connection to other information using the device, is planned for future works.

The robot system currently performs the manzai without outer-sensing devices such as microphones and vision sensors. There are also desirable parameters for manzai robot systems such as tone. The speech synthesis component can change the voices of the manzai scripts, the volume, and speed of speech according to the conditions of the surroundings of the adjustment of the parameters. A "feedback" mechanism using the outer sensing devices mounted on each robot and flexible generation of the manzai performance will be addressed in future work.

Conclusion

This chapter introduced a manzai robot system, that is, an entertainment robot that is used as a passive media based on autocreated manzai scenarios from web news articles. Manzai is a Japanese traditional standup comedy act that is usually performed by two comedians: a stooge and a straight man. The manzai robots automatically generate their manzai scenarios from web news articles based on related keywords given by the audience and from search results on the WWW, and then perform the manzai scripts. Each manzai script comprises three parts: *tsukami* (the beginning of the manzai greeting), *honneta* (main body of the manzai script), and *ochi* (conclusion of the manzai performance). The style of a manzai script is "*shabekuri manzai*," which means talk constructed from only the manzai scripts.

The proposed manzai robot system is focused on content generation. The manzai scripts are created automatically using data-mining techniques and the methodology of the manzai performance from WWW news articles. Manzai scenarios are generated based on news articles from the Internet, using many kinds of intelligence techniques such as word ontology and similar words obtained from searches of the Internet. Subsequently, the manzai script for the manzai robots written in XML is generated automatically. Then, each robot performs each role in the manzai script.

This chapter also introduced a method for automatic creation of the manzai scripts from web news articles, and management of the manzai robot systems. Then, a component-based manzai robot system that facilitates the creation of a scalable manzai robot system was presented. This chapter also verified the potential of the manzai robot system by implementing an automatic manzai scenario creation system and management systems using real robots.

The robot system performs the manzai without outer sensing devices such as microphones and vision sensors. It is desirable that the parameters of manzai robot systems such as the tone, volumes, and speed of speech be manipulated corresponding to the condition of the surroundings of the manzai robots. The speech synthesis component can change the voices of the manzai scripts by adjusting such parameters. A "feedback" mechanism using the outer sensing devices mounted on each robot and flexible generation of the manzai performance will be addressed in future work.

In addition, a robot interface using the robot body, the development of a component-based large-size manzai robot system, a sensing function for the surroundings of the robots to improve interactive performances, and enhancement of the robustness of the robot system under long-term experiments are planned for future work.

Cross-References

► Challenges for Robots Acting on a Stage

References

T. Abe, Bulletin of the Graduate Division of Letters, Arts and Sciences of Waseda University, III, Japanese literature, theatre and film arts, history of fine arts. Jap. Lang. Cult. **51**(69) (2006) (in Japanese)

N. Ando, T. Suehiro, K. Kitagaki, T. Kotoku, W.K. Yoon, *Proceedings of 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, New York, 2005), p. 3555

F. Bouchet, J.-P. Sansonnet, *Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology* (IEEE Computer Society, Washington, DC, 2009), p. 209

K. Hayashi, T. Kanda, T. Miyashita, H. Ishiguro, N. Hagita, *Proceedings of 2005 5th IEEE-RAS International Conference on Humanoid Robots* (IEEE, New York, 2005), p. 456

K. Hayashi, T. Kanda, T. Miyashita, H. Ishiguro, N. Hagita, Int. J. Humanoid Rob. **5**, 67 (2008)

R. Ishii, Y.I. Nakano, T. Nishida, ACM Trans. Interact. Intell. Syst. **3**, Article No. 11 (2013)

M. Ishizaki, Y. Den, *Discourse and Dialogue* (Computation and Language, vol. 3) (University of Tokyo Press, Tokyo, 2001) (in Japanese)

T. Kanda, H. Ishiguro, T. Ono, M. Imai, R. Nakatsu, IEICE Trans. Inf. Syst. (Japanese Edition), **J-85-D-I**, 691 (2002) (in Japanese)

M. Kanoh, S. Iwata, S. Kato, H. Itoh, Kansei Eng. Int. **5**, 35 (2005)

M. Kanoh, Y. Oida, Y. Nomura, A. Araki, Y. Konagaya, K. Ihara, T. Shimizu, K. Kimura, J. Rob. Mechatronics **23**, 3 (2011)

R. Khosla, M.T. Chu, ACM Trans. Manag. Inf. Syst. **4**, Article No. 18 (2013)

I. Kitagaki, T. Machino, A. Nakayama, S. Iwaki, M. Okudaira, J. Rob. Mechatronics **14**, 471 (2002)

D. Kitayama, K. Sumiya, in *Proceedings of IEEE International Workshop on Databases for Next Generation Researchers 2007*, (2007), p. 103

KobukiRTC: Kobuki RTC (2013) https://github.com/rt-net/kobuki_rtc. Accessed 15 Aug 2015

T. Kumamoto, Y. Kawai, K. Tanaka, IEICE Trans. Inf. Syst. (Japanese Edition) **J-94-D**, 540 (2011) (in Japanese)

R. Mashimo, T. Umetani, T. Kitamura, A. Nadamoto, *Proceedings of the 2014 Conference on Interactive Entertainment, 2014* (ACM, New York, 2014), p. 1

Y. Matsui, M. Kanoh, S. Kato, T. Nakamura, H. Itoh, J. Adv. Comput. Intell. Intell. Inf **14**, 453 (2010)

T. Matsumoto, in *Introduction of Manzai,* ed. by GansoBakushoOu (Rittor Music, Tokyo, 2008) (in Japanese)

A. Nadamoto, K. Tanaka, *Proceedings of the 13th Annual ACM International Conference on Multimedia (ACM Multimedia 2005)* (ACM, New York, 2005), p. 41

S. Park, S. Kang, S. Chung, J. Song, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (ACM, New York, 2009), p. 443

I. Rae, L. Takayama, B. Multu, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (ACM, New York, 2013), p. 1921

D. Schulman, T. Bickmore, *Proceedings of the 4th International Conference on Persuasive Technology* (ACM, New York, 2009), Article No. 25

T. Shibata, *Proceedings of the 28th Annual European Conference on Cognitive Ergonomics* (ACM, New York, 2010), p. 3

M. Shiomi, D. Sakamoto, T. Kanda, C.T. Ishii, H. Ishiguro, *Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction* (ACM, New York, 2008), p. 303

H. Tezuka, N. Katafuchi, Y. Nakamura, T. Machino, Y. Nanjo, S. Iwaki, K. Shimokura, J. Rob. Mechatronics **18**, 325 (2006)

T. Umetani, R. Mashimo, A. Nadamoto, T. Kitamura, H. Nakayama, J. Rob. Mechatronics **26**, 662 (2014)

T. Umetani, S. Aoki, K. Akiyama, R. Mashimo, T. Kitamura, A. Nadamoto, *Proceedings of 2015 International Symposium on Micro-NanoMechatronics and Human Science (MHS 2015)* (IEEE, New York, 2015), p. 139

H. Yamamoto, H. Miyazaki, T. Tsuzuki, Y. Kojima, J. Rob. Mechatronics **14**, 54 (2002)

F. Yamaoka, T. Kanda, H. Ishiguro, N. Hagita, *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction* (ACM, New York, 2006), p. 313

Part X

Interactive TV and Online Video Experiences

Digital Interactive Television and the Older Generation

41

Mark Rice and Mark Springett

Contents

M. Rice (✉)
Institute for Infocomm Research, A*STAR, Singapore, Singapore
e-mail: mdrice@i2r.a-star.edu.sg

M. Springett
Middlesex University, London, UK
e-mail: m.springett@mdx.ac.uk

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_38

Abstract

Digital interactive television offers an increasing wealth of possibilities in terms of the types of applications, services, and interactions available. At the same time, many older adults face significant barriers to access. Through conducting a literature review of related research areas, this chapter summarizes some of the practical challenges and opportunities of designing digital interactive television for older adults. This includes how age-related changes and generational differences can affect television experiences. In this work, application areas focusing on the needs of older adults are highlighted, in addition to further research directions. These range from the use of adaptive systems to modify interface parameters for improved accessibility, to the use of television as a platform for a range of health and lifestyle support applications. The representation and inclusion of the older population in the design process are also discussed.

Keywords

Digital interactive television • Older adults • Accessibility • Interface • Design

Introduction

According to demographic forecasts, the proportion of the adult population aged 60 years old and above will globally rise from approximately 12 % in 2013 to 21 % by 2050 (United Nations 2013). This will equate to over two billion people worldwide, of which more than three quarters will live in developing countries (United Nations 2013). The expansion of the aging population is attributed to multiple factors, including a decline in fertility and lower birth rates, and better healthcare. As such, within the next few decades, the number of adults aged 85 years and older is expected to globally rise faster than other age groups (National Institute on Aging et al. 2007). Moreover, 40 % of older adults aged 60 years and older are reported to live independently (i.e., alone or with a spouse), with an increased likelihood of independence in countries where the proportion of older adults is higher (United Nations 2013). In particular, the National Institute on Aging et al. (2007) report cultural changes in attitudes to living alone, which they described as "reinforced by greater longevity, expanded social benefits, increased home ownership, elder-friendly housing, and an emphasis in many nations on community care" (p. 17). Subsequently, in recent years, there has been an increasing interest in the use of domestic technologies to support the well-being of older adults. Television is a significant example of such a technology, and has long been recognized as a social form of companionship, entertainment, and information for older adults (Rubin and Rubin 1982).

Among the older population, Internet adoption is typically lower than other age groups (e.g., Office of National Statistics 2014; Pew Research Center 2014a). In contrast, despite the increased fragmentation of television viewing to different media platforms, it remains a central and stable part of older adults' lives.

For example, at the time of writing, it is reported that over 90 % of the over 55s in Europe watch television on a daily, or near-daily basis (Standard Eurobarometer 76 2012), with older people aged 65 and over in the United Kingdom devoting half of their media-related time to watching television (Ofcom 2014a). In comparison, the percentage of younger adults who watch broadcast content through a television set has gradually decreased to alternative forms of media and mobile device usage. To illustrate, a 2014 report by Ofcom (2014a) found that young people only spent half of their time watching live television, while the BBC Trust (2014) highlighted that about a quarter of 16–24-year-olds weekly watched programs via their Internet-streaming media player. Interestingly, in the last few years, both reports highlight only a small decrease in the television viewing behavior among older adults compared to younger age groups, which reflects a growing generational gap in media consumption and the types of services used (Ofcom 2014a).

Technical advances in digital television provide both opportunities and potential barriers in the adoption of new services and content. The advent of digital interactive television implies a change in viewing patterns, particularly for those older adults who associate "lean forward" interaction with analogue text-based services. Conversely, viewers with low confidence in their digital skills may be reluctant to embrace a change in viewing experience, or have interaction difficulties dealing with the challenges that the technology now presents. Alternatively, digital interactive television has the potential to be a carrier of dedicated services for older people, by exploiting long-term allegiance with the medium.

Subsequently, this chapter outlines important challenges in the development of digital interactive television for older adults. In the section "Television in the Lives of Older Adults", we explore the importance of television for older adults, and the implications of the digital switchover on viewing experiences. This is followed by a general description of age-related changes in the section "Age and the Aging Process", and how they translate into television interaction experiences in the section "Interaction and the Television Interface". We then highlight application areas where digital interactive television has specifically been investigated among older adults in the section "Applications for Older Adults", followed by reporting on examples for future research in the section "Future Opportunities". Although there are different definitions to what may constitute as old age, we generally refer to older adults and the "older generation" as people over the age of 65. However, it is important to emphasize that some of the reported literature in this chapter refers to older adults of a younger age (e.g., 50+ years).

Television in the Lives of Older Adults

For many older adults, the social identity of television was established in an era when its iconic significance was very strong. Television had a unique power that contributed to its mass appeal in connecting audiences to historical events on a global scale. In the modern day, television continues to play a strong and diverse range of roles in the lives of older adults, demonstrating a wide spectrum of values, interests, and beliefs. For example, Gauntlett and Hill (1999) have described the various roles that television can have in the lives of older adults, which vary from occupying time in retirement, to forming comfort in dealing with grief, or keeping in touch with world events. Present generations of older adults in western societies are also likely to be familiar with historical changes in the medium, such as the introduction of color television, the broadcasting of information services (e.g., teletext), and changes in the branding and quality of commercials. However, the more recent deregulation of television, and the launch of cable and satellite channels may challenge the medium's unique place in the lives of many older people, creating a less familiar and comfortable "TV landscape". To explore in more detail, four examples are highlighted below.

First, in the past television was generally seen as a passively consumed medium. Traditionally, it has the advantage of being a technology that its viewers are likely to trust rather than otherwise, due in part to its familiarity and ubiquity established via prolonged contact and daily use. On the other hand, the recent addition of a return path and greater connectivity presents a new agenda in terms of privacy and data security, as the main bulwarks of trust traditionally are seen as low-commitment goals and the finite notion of content (French and Springett 2005).

Second, a traditional model of analogue television is that content is provided by broadcasters who are familiar, accountable, and well established, making trust relationships easier to maintain. For example, in the United Kingdom the BBC started life as a branch of the civil service and is still funded through a compulsory license fee. This inherent trust is based on a familiar television-viewer relationship, where the viewer's role is a relatively passive one. That is, the interactivity between the viewer and system tends not to be complex. However, the introduction of interactive features such as voting, pay-per-view services, and other interactive elements has changed this relationship. It is typical for viewers to have no more than a "black box" understanding of the carrier technology, but they face a greater commitment to interactive services (and a greater risk) where financial transactions are involved. This is perceived to be a particularly important issue given the high percentage of Internet users who have concerns in carrying out financial transactions online (e.g., Ofcom 2010).

Third, television is traditionally associated with low-commitment goals. Artz (1996), writing before the emergence of digital interactive television, characterized the role of the medium as a distraction from personal goals and concerns, or a serious agenda in viewers' own lives. The role of television was seen to unburden viewers through the programs they watched, in what Artz describes as encouraging "people to live vicariously" (1996, p. 9). Subsequently, in this context, individual goals can be thought of as being contrived, transient, and insubstantial. While this may be something of a generalization, the offloading of commitment and escape from the burden of peoples' own lives seem to be an important defining factor in this relationship. This may imply a reluctance to commit to newer television functions in which a greater proactivity and working understanding of the technology are needed.

Finally, television may, perhaps erroneously, be thought of as a "walled garden" in which the risks associated with other technologies do not apply. The sense of familiarity suggests a reassuring level of visibility, traceability, and accountability. However, this stands to be compromised by a greater convergence of technologies, deregulation, and Internet-based services. In this context, while the traditional and familiar presence of analogue television in older adults' lives is a potential advantage, its changing nature and mandatory requirements to switch to digital imply a fundamental difference in the television-viewer relationship. Optimistically, this will break down barriers by bringing digital services and utilities to people who may not easily embrace other types of technology. However, there is also a risk of generating a digital divide for those individuals who have difficulties accepting and understanding new television paradigms, concepts, and facilities.

Age and the Aging Process

Older adults represent a highly diverse population. In view of this, Coleman (2003) argues for commercial innovation and good design practices that can cater for active and flexible lifestyles, recognizing many older adults have a high financial income and more free time than younger adults. Baby boomers (i.e., those born between 1946 and 1964) are also generally better educated, and have more technology experience compared to previous generations (Coughlin 2007). As such, criticism has been raised over commercial companies who narrowly associate age with disability, given the importance of fashion, functionality, and other values that are likely to dictate product interests in many older consumers (Coughlin 2007). Equally, for the older old (85+ years), there is also a need to consider the roles technology can play in everyday life. This includes the types of technological concepts that can draw meaning, or synergies from the physical world. A good illustration of this is the participatory design work of Vines et al. (2012), who explored digital banking concepts with adults in their 80s.

Resilience in aging is influenced by a multitude of factors, ranging from the personality traits of an individual, to the modification of an environment (Salthouse 2004). Positively, Baltes et al. (2005) indicate that the human body is able to adapt to the effects of aging by selecting, optimizing, and compensating for the use of available resources. For example, compensation can include the use of assistive devices to support sensory or functional loss (e.g., a hearing aid or walking frame) (Baltes et al. 2005). At the same time, variations between the most and least abled are reported to be greater in old age, as physiological differences are accountable to a wide range of factors, including genetic, social, and environmental influences (Rabbitt 2005). In this sense, socioeconomic status and variations in lifestyle can be important influences in assessing functional change in cognition (Rabbitt 2005). As such, patterns of physiological change can vary significantly across individuals. For example, in comparing cognitive test scores, Ardila identified results where older adults became *"more intellectually heterogeneous"* (2007, p. 1010). Equally, Glisky (2007) argues that cognitively, older adults over the age of retirement can perform as well (if not better) than those of a younger age group. On the other hand, reaching advanced stages of old age can increase the incidence of certain age-related diseases, as poor health can quicken the rate of cognitive decline (Rabbitt 2005).

Audiovisual Capabilities

Worldwide, it is estimated that about two thirds of people who are visually impaired are over the age of 50 (WHO 2014). Getting older brings a physiological decline in visual accommodation and the ability to change focus at different distances, and visual acuity and the sharpness to see spatial detail. In addition, there is a marked reduction in pupil size and the time required to adapt to the dark (Huppert 2003; Stuart-Hamilton 2012). Common visual impairments include *presbyopia*, the decline in ability to focus on near objects due to a reduction in the elasticity of the eye's lens, and *cataracts*, an opacity of the eye's lens resulting in blurry, discolored, or cloudy vision. Both conditions become more prevalent with age, and can result in difficulties seeing information in low light conditions. Alternatively, other eye disorders include age-related *macular degeneration*, *glaucoma*, and *diabetic retinopathy* (e.g., see RNIB 2015). While their symptoms and treatments vary considerably, in later stages of degeneration, visual loss can have a profound impact on daily tasks such as reading, driving, and recognizing people. Notably, these are conditions that can also affect the viewing of television content, as Fullerton and Peli (2008) highlighted how visually impaired viewers can have difficulties seeing facial features and reading textual information on-screen.

According to Gill (2004), about 50 % of visually impaired people over the age of 75 are also affected by hearing problems. *Presbycusis*, or age-related hearing loss, is associated with internal changes in the inner or middle ear, or nerve pathways to the brain, and typically can cause difficulties detecting high-pitched sounds (NIDCD 1997). This includes a marked decrease in speech perception, with additional mental effort required to understand auditory stimuli (Stuart-Hamilton 2012). Gender differences have also been identified in hearing acuity, with women more sensitive to higher frequencies, and men lower frequencies (Pearson et al. 1995). Practically, the implications of hearing loss mean that older adults can have problems distinguishing speech from background noise, particularly in the context of watching television, where background music can compete with spoken words (Carmichael 1999). Given the social stigmas associated with hearing loss, it should be noted that many hearing impaired adults fail to get necessary treatment. For example, in the United States alone, it is estimated that less than a quarter of people who actually need, currently use a hearing aid (ASHA 2015).

Cognition Capabilities

Cognitive changes that occur with age can influence the ability to encode, store, and retrieve information. For example, processing speed tends to diminish with age, while deficits in visual and auditory capabilities can degrade the quality of information received by the brain (Rabbitt 2005; Kensinger 2009). For selective and divided attention, research suggests that the more complex the task, or set of tasks, the more likely older adults will exhibit a deficit in their performance (Zanto and Gazzaley 2014). Regarding inhibitory control, older adults can also have difficulties eliminating irrelevant sources of information in working memory (Glisky 2007; Kensinger 2009), which can lead to a type of mental cluttering and task interference (Zacks et al. 2000), while the ease of retrieving stored information from memory can depend on whether this requires the use of *recall* or *recognition*. For example, in comparing performance, Craik and McDowd (1987) demonstrated that older adults perform weaker in memory tasks related to recall, compared to recognition. To account for this difference, the authors emphasized the greater amount of (internal) resource processing needed to recall information. Specifically, within digital interactive television, Carmichael (1999) illustrates how the use of voice commands can be more cognitively demanding for older adults if they require the recall of information, compared to reciting from an on-screen list.

Alternatively, while it is argued that “In very old age no memory form escapes completely from age-related decline” (Marcoen et al. 2007, p. 51), it is important to recognize that among older adults, not all types of memory decline to the same degree or to the same extent. This includes procedural memory that relates to well-learned routines and procedures (e.g., reading and writing) and semantic memory that is associated with facts and general knowledge about the world, which can exceed those of younger adults (Glisky 2007). In relation to intellectual change, distinctions are also made between the decline of *fluid* intelligence and the problem-solving of information, compared to the stability of *crystallized* intelligence and a lifetime of acquired knowledge (e.g., Horn and Cattell 1967; Stuart-Hamilton 2012). Moreover, from the perspective of understanding memory and emotion, older adults reportedly have a stronger emotional focus on remembering positive rather than negative information, including a distortion of recollected experiences to enhance a sense of well-being (Mather 2004; Carstensen and Mikels 2005).

Physical Capabilities

In practical terms, age-related changes can equate to a reduction in muscle power and strength, joint movement, with a greater amount of postural sway (Granacher et al. 2012). Physiological changes in muscle strength can start to quickly decline at a relatively young age (e.g., about 50+ years) (Huppert 2003), while deficits in visual, cognitive, gait, and balance control among other factors can result in an increased risk of fall injury (Rubenstein 2006). In addition, age can account for a decrease in hand steadiness, with lower grip strength associated with a reduction in aiming and tapping abilities (Martin et al. 2015). These have important implications in the coordination of rapid and precise hand movements (Martin et al. 2015), which for some older adults may negatively impact on their accuracy to repetitively press small and closely positioned buttons on the remote control handset.

Overall, the physiological changes that occur due to age raise important questions in the design of television content and input devices for older adults. This relates not only to how information should be designed to cater for varying user characteristics, but also in understanding the types of challenges that current television systems may present. Given the general overview in this section, readers are encouraged to refer to the cited literature in this chapter should they wish to obtain a more detailed understanding of the aging process.

Interaction and the Television Interface

Typically, state-of-the-art developments target those who are technologically savvy and comfortable to explore the potential of new devices. However, given the rapid advancements in digital interactive television, there is a potential trade-off between pushing technical frontiers for early adopters, while accommodating for those who are more reluctant to adopt. This includes the complexity of setting up, accessing, and using digital receivers (or commonly termed "set-top boxes"). For example, a 2003 report by Klein et al. (2003) highlighted the high percentage of over 75s that would be excluded from using basic procedures, such as selecting appropriate channels.

To help address interaction difficulties, Gill and Perera (2003) proposed the use of smart cards within digital receivers, which could allow for a greater amount of device customization. This included the configuration of graphic and audio content to match individual user requirements. More recently, research into the development of second screens (e.g., Cesar et al. 2008) and intelligent adaptive systems (e.g., Langdon et al. 2010) offers potentially useful directions to address accessibility and usability constraints in digital interactive television (see the "Future Opportunities" section). Guidelines and recommendations are also available to support the design of accessible interfaces and remote control handsets (e.g., Gill 2009; Digital TV Group 2014). This includes addressing the product life cycle and replacement of digital receivers to ensure they can adequately support enhancements to new television services (Looms et al. 2010). Yet, given the financial costs of replacing these units by the service provider or consumer, there are likely to be a number of digital receivers that have been in domestic use for many years. As such, a clear challenge in the provision of digital interactive services is the ability to support the functional range of digital receivers in operation, particularly lower-cost consumer units that may be limited in their technical performance and interactivity.

In the following section, we outline important interaction issues in designing digital television for older adults that are based on findings from existing literature. Specifically, this section focuses on the input and output modalities, mental models and metaphors, and learning and adoption issues for older adults.

Input Control

The television remote control is notoriously associated with usability problems. This includes the poor size, positioning, and labeling of buttons. In an interview of a small sample of older adults, Freeman and Lessiter (2007) found that buttons with unfamiliar functions were simply ignored and avoided, while, in designing more intuitive labels, functional misconceptions were found between the semantically similar "back" and "return" buttons (Lessiter et al. 2004).

Within interaction, basic ergonomic problems are compounded by the need to coordinate two spatially separate interfaces – the *television screen* and *remote control* (Carmichael 1999). For those adults who require different spectacles or corrective lenses for different focal distances, this can create difficulties in switching attention between interfaces, which can complicate and slow down even basic interactions (Carmichael 1999). This can include the problem of confirming on-screen feedback due to unknown or accidental execution errors in completing (timed) button sequences. One possible way forward is to reduce the number of buttons, and transfer more interaction to the television screen. However, caution has been raised over the notion of simplifying the remote control handset through the design of fewer buttons, given the additional cognitive effort required by older adults to understand the mapping of multifunctional buttons in relation to the television display (Carmichael 1999).

Moving the interaction to the design of more user-centric devices may help alleviate the problem of handset complexity, by avoiding the need to scan small and occasionally erased text on a handset, particularly in low ambient light conditions. Some commercial development has already begun exploring alternative approaches to the conventional remote control. Notable examples include the use of a second screen device in the form of a tablet or smartphone, which may potentially help overcome barriers to access by utilizing native accessibility features embedded within these technologies (Klein et al. 2014). However, despite their potential, to date there is a paucity of known research exploring second screen interaction among older adults.

Visual Display

Recommendations for the visual display of digital information on television include avoiding the use of ornate, oblique, and flashing fonts, with a preference for good luminance contrast between text and the background information (Carmichael 1999; Consumer Expert Group 2006). Typefaces such as the Tiresias Screenfont have been developed to improve text legibility, and a minimum size has been recommended for displaying text on a television screen (Consumer Expert Group 2006). However, as Carmichael (1999) notes, specifying an absolute font (letter) size is problematic, as its ideal legibility will depend on the distance and viewing angle to screen, which typically is not fixed to a single position when watching television. In terms of displaying the appropriate amount of information on-screen, there is also a potential trade-off between providing too much information that effectively clutters and overcomplicates the display, compared to separating out information over multiple pages, and increasing the cognitive effort required to remember the correct navigational sequence (Carmichael 1999). To address this issue, one possibility proposed has been the use of high and low lighting techniques to help draw attention to on-screen graphical elements for older adults (Carmichael 1999).

In addition, it is important to note that the readability of visual content is perceived to be dependent on the type of display used. Current television markets are dominated by flat-screen displays. However, despite their lack of market production, cathode-ray tube displays have been reported to offer better visual clarity in contrast ratio, brightness, color rendering, and viewing angle compared to liquid crystal displays (Pak and McLaughlin 2011). While concerning stereoscopic 3D displays, such as for television, Read (2014) estimates that 14 % of the population may experience adverse viewing effects (e.g., eyestrain).

Audio Display

Given that television content has its own audio output, it is stressed that televised speech must be presented clearly and the background level of a program be reduced to a volume that does not impede on the inclusion of descriptive commentary (ITC 2000). For older adults with hearing loss, this includes the assistance of subtitles or captions for text-based commentary and the transcription of televised programs. A number of guidelines are available on the screen placement, size, and presentation of captions (e.g., DCMP 2011; Ofcom Untitled). This includes research into the use of *emotive captioning*, where the visual design of captions to convey emotional meaning in video segments was evaluated with deaf and hard of hearing users (Lee et al. 2007). Alternatively, audio description (a type of audio narration) was found to be a useful augmentation of the visual channel for partially sighted older viewers (Peli et al. 1996), while in terms of the quantity of audio description, older adults with higher fluid abilities were reported to benefit from more verbal information in programs, compared to those with less (Carmichael 1999). Yet, despite these differences, the use of audio description is seen to be a helpful enhancement to televised content for adults with declining cognitive functioning (ITC 2000).

In contrast, research into the development of an on-screen avatar to facilitate the access of information on an electronic program guide by older adults identified a number of interaction issues using speech output (Carmichael et al. 2003). Specifically, the authors found that delays in the lip-synch movement of the avatar caused marked problems in reading speech, including an uncertainty in knowing how to respond to implicit prompts, which led to hesitation in verbal responses (e.g., “er’s” and “uhm’s”). Subsequently, they propose that speech output be reinforced with the use of on-screen text to act as a memory aid, in addition to better prosody to improve the comprehension of synthetic speech (Carmichael et al. 2003). In this regard, recommendations in the design of audio displays for older adults suggest avoiding the overexaggeration of prosody, given that this can sound demeaning to the listener (Pak and McLaughlin 2011).

Mental Models and Metaphors

An important characteristic of designing for older adults is an awareness that their life experiences, norms, values, and skills are often different to those of younger generations, particularly in the context of television, where those familiar with traditional services may not have a strong mental model of digital concepts in the support of new forms of interaction. Docampo Rama et al. (2001) describe the notion of *technology generations* and the disadvantage older adults have in their lack of exposure to software style interfaces during their formative years. Comparing task performance across different age groups, Docampo Rama et al. (2001) identified generational differences in the learning of new interfaces. In particular, the authors indicated a difference in problem-solving, with younger adults more likely to apply a "trial-and-error" approach in understanding new features.

Further, the design of digital interactive television is arguably influenced by features commonly associated with the desktop computer. This implies the need to explore new metaphors and interaction techniques that many older adults will be able to relate too. For example, prior research has demonstrated that drop-down menus where information is effectively concealed from view until selected can be a difficult concept to comprehend for inexperienced older users (Rice 2009). This led to the exploration of more natural affordances in the manipulation of graphical objects, in which animation and the use of transitional effects (e.g., fading, rotating, and zooming) were intended to provide clearer paths of guidance in linking menu items and on-screen objects (Rice 2009). The research explored the design of both input (remote control) and output (screen) interfaces to improve navigating within a television environment. Recommendations included a need to better understand how on-screen interaction is controlled by gesture and other input devices, including the types of metaphors and visual features that can aid in navigational awareness for older adults.

Learning and Adoption Issues

Given the social stigmas attached to old age, it is often assumed older adults are reluctant to adopt digital technologies, rather than address the more critical reasons for a lack of motivation to invest time, money, or effort. This may be due to a lack of facilitating conditions, or poor user experience. An example of this is the changing nature of text services after the digital switchover. In the United Kingdom, the BBC "Ceefax" service was discontinued in 2012 and replaced by a digital equivalent (BBC News 2012). The analogue service involved typing a small set of three digit numbers to link to specific page information (e.g., news, sport, entertainment, etc.). Alternatively, the digital equivalent now involves a greater amount of proactive searching and interactivity, more typically associated with the desktop, or web-based services. This illustrates how the development of exploratory learning skills has become a prerequisite for gaining the benefits of interactive television services.

The key characteristics of exploratory learning are the recognition-based search for features, direct manipulation of objects, reactive planning by recognition, an understanding of system feedback, and incremental learning through trial-and-error problem-solving (Norman 1988). The progression from "novice to expert" is essentially a process of skill acquisition. This initially involves the recruitment of general and task knowledge applied in response to metaphorical representations and cues. This then leads to a gradual rule-based understanding of the system principles that can be reused for the learning of other services or applications. In this context, it is useful to broadly distinguish between the notion of "good and bad" errors. A *good* error contributes to learning as it leads to feature discovery and hones a user's understanding of an operational principle. A *bad* error will mislead and confuse, and can result in the user giving up and withdrawing from the interaction.

The demands placed on older users by extended input sequencing and speculative scanning can render trial-and-error-based learning challenging. This can partly be accounted to the unfamiliarity of the digital television space, while for some, exacerbated by physiological difficulties associated with age. Importantly, these issues not only apply to older adults per se, but to other segments of the population. For example, in a study of visually impaired users using electronic program guides, incorrect menu selections frequently led to repetitive cycles of random search (Springett and Griffiths 2007). Without the visual resources to diagnose and modify unsuccessful actions, the users became embroiled in long sequences of "guess work," which often compounded rather than resolved difficulties. Subsequently, declining vision (in addition to other impairments) is likely to exacerbate problems for those who are inexperienced in exploratory learning of the digital interactive television space.

Applications for Older Adults

The previous section highlights important interaction issues in the design of the television interface. However, in recent years, there have also been a number of interesting developments in concepts and applications for older adults. In many cases, these extend beyond watching more television content, to consider new and alternative ways of utilizing the digital medium. Three application areas focusing on social interaction, healthcare, and cognitive assistance are illustrated below.

Social Interaction

With age, many older adults experience a change in their domestic circumstances. This includes downsizing in property, moving into community care, dealing with the loss of a partner, or coping independently with diminished cognitive and physiological capabilities. Subsequently, serious health risks can be associated with an absence of social relationships, given the substantial link between social isolation and early mortality rates in adulthood (e.g., Holt-Lunstad et al. 2015).

In the development of social television applications, Alaoui et al. (2012) explored ideas that involved the sharing of multimedia, participation in club events, and playing digital games. Svensson and Sokoler (2008) investigated social behavior among senior housing communities. This led to the development of a low-fidelity concept that allowed older adults to be noticed by their peers using a specialized remote control handset. Alternatively, research has investigated the two-way communication process of a television application using participatory design activities (Rice and Carmichael 2013). Gathering opinions on how individuals would visually communicate through a digital television interface, questions were raised in identifying the presence of another person prior to calling them direct, and dealing with unwanted interruptions and privacy concerns. For the participating older adults, these included analogies that were more commonly associated with the (landline) telephone and the notion of "dialing" to another person.

Further, Harboe et al. (2008) describe the value of social television as a shared rather than solitary activity. This suggests there is also potential for the medium to support and strengthen intergenerational relations. Arguably, while the notion of watching television among family members may still be considered as a type of intergenerational pastime, for many, the separation of media usage in domestic households has begun to challenge these values. Therefore, a possible avenue for new television applications is to help reengage generations. For example, for young children, this includes considering the types of interactive content that may encourage quality time and learning with grandparents (e.g., through intergenerational games), while for older adults engaging with other family members, the types of applications that encourage the sharing and discussion of experiences, such as collaborating in a "scrapbook" of family memories (Rice 2009).

Alternatively, beyond the immediate family, another possibility for intergenerational engagement is the use of community-based television platforms as described by Obrist et al. (2009). Reporting on an IPTV platform evaluated in a small Austrian town, the authors illustrated how local residents of mixed ages and backgrounds contributed to the creation of broadcast content. Developed for and by community members, a key aspect of sustaining interests was to ensure volunteers held ownership over the platform.

Health and Well-Being

Issues relating to isolation and progressive health changes mean that digital interactive television has a potential impact on emotional, social, and situational support for many older adults. Home-based healthcare is seen as a productive means of self-managing personal health, and reducing sedentary lifestyles in the older population. For example, it is recommended that healthy older adults undertake over 2 hours of moderate physical activity per week (WHO 2010).

Spinsante and Gambi (2012) described a system architecture for remote health monitoring on digital interactive television. Specifically, the system was designed to utilize available bandwidth to transfer different types of data, including physiological parameters gathered from a network of body-monitoring devices, with personal data authenticated via smart card access. Similarly, the work of Silva and Júnior (2010) has demonstrated the feasibility in translating data from body-monitoring devices via a gateway to a digital television platform. This includes displaying health alerts on a television display.

Alternatively, given regular exercise is known to be an important attribute in maintaining functional health, Carmichael et al. (2010) developed a prototype television application for older adults. This consisted of 24 warm-up, aerobic, muscular, cool down, and flexibility activities, which were assisted via a virtual instructor. The results provided recommendations to better self-regulate and adapt physical exercises to individual requirements. These included options to adjust the music tempo to change the pace of the physical activities performed, as well as to check vital signs via the integration of a heart-monitoring device.

More widely, in a review of digital television systems applied to health and social care, Blackburn et al. (2011) found that there was a strong trend toward providing healthcare support. Namely, the authors identified 25 systems (some of which could be described as a type of "service") that were developed, in development, or at a concept phase. Although not specifically focused on older adults per se, 68 % included some form of vital sign monitoring, 56 % provided healthcare information and education, and 44 % provided direct access to healthcare consultation. With these systems, the authors reported a general lack of clinical research and user testing to help evaluate their performance, while many had not reached a commercial level of implementation.

Cognitive Assistance

In relation to the above, digital television research has explored the development of applications to assist the onset of cognitive decline. Currently, dementia affects almost 50 million people worldwide and can broadly be defined as a loss in cognitive functioning that deviates from the normal aging process (WHO 2015). The most common form of dementia, Alzheimer's disease, is estimated to account for approximately 60–70 % of all cases (WHO 2015). On the other hand, affecting up to 20 % of older adults, mild cognitive impairment (MCI), which can be described as a low level of cognitive decline, can interfere with the planning, attention, and the use of language – although not to the extent of preventing daily activities (Alzheimer's Society 2015).

Prior work has explored the use of a communication platform to provide personalized prompts and reminders for people with mild to moderate cognitive impairments (Carmichael et al. 2008). These include prompts to support daily living, such as to take medication, eat a meal, or become aware of an approaching visitor. The design of the prototype system aimed to provide routine stability to the end user (or patient), while alleviating some of the social burdens faced by the carer. A portable interface was developed to allow a carer to schedule prompt messages, which were then timely displayed on a television screen. As part of this process, the authors raised important issues in the intervention and modality of prompts. For the end user, this included the types of information used to engage in a response (including notions of calming and reality orientation), as well as considerations to the potential intrusiveness of scheduling a reminder during a televised program compared to a commercial break (Carmichael et al. 2008).

Alternatively, research has also included the use of training applications via digital interactive television to help improve and sustain cognitive functioning in old age. These have included television applications specifically designed to assist healthy and mildly impaired older adults (Freeman et al. 2009). Evaluated across two groups of healthy older participants (60–87 years old), the research identified significant improvements in working memory and executive functioning tasks (Shatil et al. 2014).

In both these examples, the research raises important questions in how television can be used to assist in daily life activities for different spectrums of the population. In particular, while there is ongoing debate over the effectiveness of brain training applications to improve cognitive functioning in old age, the findings demonstrate the feasibility of delivering new types of content and services via digital interactive television. In the case of understanding population interests in cognitive training applications via digital interactive television, the empirical findings of Miotto et al. (2013) indicate that the driving forces for adoption by some older adults may range from general interests in using the technology, to improving negative perceptions of health.

Future Opportunities

In recent years, despite an increasing amount of research in aging and digital interactive television, there remains a gap in the commercial development of new applications for older adults. In many cases, it could be argued that catering for the needs of older adults can have positive benefits for wider segments of the population, not only in the support of better accessibility features for younger adults with impairments, but recognizing that the potential for digital interactive television as a “media hub” in the provision of healthcare or cognitive training will likely draw broad market appeal (e.g., Miotto et al. 2013). As such, rather than treating older adults as a separate entity in product design, greater consideration should be placed on how “good design for older adults” can translate to “good design for everyone” (Fisk et al. 2009).

Given the limitations of interactivity reported in this chapter, in this final section we report on ideas for further work. These vary from the use of adaptive models to improve interface personalization, to alternative interaction modalities such as gesture and voice interaction. We stress the examples given are not exhaustive, and do not cover all potential themes. However, they do highlight areas of research that could have important implications in the development of the technology for many older adults.

Adaptive User Modeling for Assistive Needs

For digital interactive television, Langdon et al. (2010) proposed the use of user modeling to personalize output modalities (e.g., speech, graphic, and haptic displays) based on the input capabilities of the user. This included utilizing information from a range of sources, including population data. Similarly, Biswas and Robinson (2013) developed an interactive simulator to predict accessibility problems on digital interactive television for mobility and visually impaired users. Combining environmental, device, and user models, the simulator was used to compare actual with predicted performance in icon searching tasks, with some significant differences in results. In particular, the research highlighted the importance of learning effects and fatigue to account for variance in task performance.

A challenge of human performance data is that it must be representative of the target population, and as such cannot be adequately substituted from other user groups (Steinfeld et al. 2002). For example, the availability of anthropometric data, which focuses on measuring and analyzing the physical characteristics of the human body (Steinfeld et al. 2002), has been criticized among older population groups. Namely, de Onis and Habicht (1996) highlighted the lack of publically available anthropometric data on older-old adults, particularly given the wide variations in height, weight, and body mass index when comparing across different geographic and ethnic groups. Similarly, it is argued that anthropometric datasets often exclude below the 5th and above the 95th percentile of the population (Marshall et al. 2010). As such, concerns have been raised that many people with severe disabilities (including older adults) will fall outside of conventional sampling techniques (Steinfeld et al. 2002).

Using this example, for television interfaces to adapt to the physical, cognitive, or sensory demands of its users, the information or "training data" must be representative and accurate. This includes ensuring that computational models are readily available and understood by members of the digital television community, as well as technically able to capture and process task behavior within a set-top box environment, and remotely validate the user data collected. Likewise, the use of adaptive systems generates questions over the level of control provided by the software, and the extent users can override or modify interface parameters in the event of unforeseen system errors. In such cases, this will require intuitive interfaces that can allow older adults to self-personalize and change the user settings, without conflicting with the performance of the system. As a result, building on the work of Langdon et al. (2010) and Biswas and Robinson (2013), more research is needed to understand the learnability of adaptive interfaces by older adults. This includes the extent that adaptive television systems can successfully accommodate for the changing physiological requirements of end users over time.

Interaction Across Second Screens

In recent years, the exploration of second screen devices offers interesting opportunities in television interaction. Typically envisaged through a handheld device, prior work has considered the use of second screens to replicate, enrich, or substitute the main screen interaction (e.g., Cesar et al. 2008; Oksman et al. 2013). These include participatory workshops that have investigated interactive features across multiple devices (Hess et al. 2011; Fleury et al. 2012).

As a secondary device, the potential advantage of using a portable screen (e.g., on a tablet or smartphone) is that it has the ability to enhance content that may be difficult to comprehend on a single television display. Examples of this include adjusting the screen layout, color contrast or font size of an electronic program guide, or the display of accessibility features such as audio description and captioning that could be accessed independent of the main screen (e.g., see Jolly and Evans 2013). In addition, displaying control features on a second screen may help to overcome some of the attention-switching demands previously reported in using a remote control handset and television screen, by ensuring a direct and visible mapping between the input and output of information on the same device. This includes the provision of features not commonly available on digital receivers, such as text to speech, or the integration of a dynamic braille display for users with low vision (Klein et al. 2014). Notable works include that of Jolly and Evans (2013) who developed a web-based application program interface to enable second screen usage, such as for alternative audio streaming.

Although not specific to digital interactive television, studies of tablet usage among older adults have illustrated problems inputting and searching for information (e.g., Jayroe and Wolfram 2012; Waycott et al. 2012). Difficulties that appear to partially relate to physical restrictions in hand movement (Waycott et al. 2012). Wright (2014) reports that inexperienced tablet users require the learning of new skills, while the Pew Research Center (2014a) in the United States found that less than a sixth (13 %) of older adults who did not own a portable device like a tablet felt comfortable enough to learn independently. In contrast, Tsai et al. (2015) have reported positive experiences by older tablet users, including ease of use; while statistics in the United Kingdom found that the percentage of older adults likely to go online using a tablet aged between 65 and 74 years dramatically rose from 5 % in 2012 to 17 % in 2013, demonstrating a threefold increase in adoption (Ofcom 2014b).

These findings on tablet usage confirm the importance of digital interactive television to account for changes in learning, and the ability to map input requirements to individual needs (Carmichael 1999). For second screen devices, this includes concern over a lack of open standards guiding their compatibility to communicate and synchronize content with different digital receivers (Klein et al. 2014), demonstrating a potential barrier to access if proprietary systems do not follow a common framework. Alternatively, more empirical research is needed to understand the social benefits and context of second screen usage among older adults, a potential wide area of investigation given the various use cases for interaction, ranging from the public/private spaces of the home, to a hospital, community center, care home, etc. Likewise, it remains questionable how to design new experiences for older adults independent of the main television display, while determining the extent touchscreens are a good substitute for the remote control.

Exploring Gesture and Voice Modalities

The advent of gesture devices (e.g., the Apple TV Remote, Hillcrest Scoop Pointer) provides reliable consumer products that support alternative modes of interaction. However, the increasing commercial availability of gesture and voice controlled interfaces also indicates a need for research in how suitable these emerging modalities are for older users, as there is presently a paucity of information to determine their efficacy. In general terms, compared to desktop computers that are supported by strong visual metaphors and free cursor movement, arguably digital interactive television does not easily facilitate example-based learning through interaction. Subsequently, the ability to reverse actions, and to identify when errors have occurred needs to become an easier and more intuitive process. For example, menu structures that demand extensive visual search should be supported through the representation of more appropriate interface models (Rice 2009).

In view of the variable interaction distances to the television screen, the ergonomic use of freehand gestures raises important questions in how to register finer selection movements while dealing with potential issues of fatigue and hand jitter through prolonged usage. In particular, while a limited amount of digital television research has positively reported on the use of freehand gestures with older adults, evaluations have restricted user interaction to a fixed distance to screen (e.g., Bobeth et al. 2012). Subsequently, more empirical research is needed to better understand the performance demands of gestural interfaces among older adults. This includes the types of feedback and physical effort required to engage in different interface and layout configurations – both within single- and multiuser settings. In this regard, establishing a comprehensive taxonomy of hand and body gestures for digital interactive television would be beneficial in understanding how these findings translate to the design of new applications for older adults. Similarly, ethnographic research could help determine the contextual use of gesture interaction in the home, including how interactive behavior with gestural devices may change over time.

Further, there is relatively little known literature evaluating voice interaction on digital television with older adults. A notable example is the work of Bures (2012), who in exploring nonverbal interaction identified usability and user acceptance issues in making audio cues. Specifically, compared to using a remote control and the Nintendo Wiimote, nonverbal vocal input was reported to be the least engaging modality, with some training required. By contrast, the implementation of voice interaction via a PDA device used with a digital receiver was considered to be effective in a low-noise environment (Lobato et al. 2009). However, in a real-world context, it is questionable how effective voice interaction is against background noise, in addition to older adult's abilities to memorize and recall sets of voice commands (Carmichael 1999). Moreover, the value of voice interaction for older adults also depends on the social and personal interests in its usage, as well as issues of accuracy and learnability. Similar to freehand gestures, this includes a better understanding of how voice commands map to user interface features, and as Carmichael (1999) highlights can accommodate for natural language usage, including hesitation and uncertainty in making conversation.

Addressing Privacy and Security Issues

With advances in technology, digital interactive television may increasingly be used to monitor vulnerable adults, by collecting information on an individual's activities. While such a system has the advantage of assisting in daily living, it highlights the potential danger of encroaching on personal privacy. Namely, the features of technology that make them adept and adaptable for older users may generate a sense of intrusion, even if the data is anonymized. In this sense, an understanding of data analysis is likely to be outside the boundaries of many older viewers' expectations, and this may raise concerns of security, including possible "hackability" and the perceived changes in the viewing experience caused by a loss of privacy. For example, according to Turner and McGee-Lennon (2013), telecare systems have generated concerns over access and the control of confidential data, including what it is used for, and how securely it is distributed. In this sense, it seems feasible that television-based systems will raise similar questions in how personal and behavioral data is analyzed, and potentially distributed to third parties. This issue is particularly complex when dealing with the consent and data protection of individuals with severe cognitive impairments.

Recently, a novel television feature that could benefit many people but has drawn public concern over data collection is an embedded voice activation system marketed by Samsung (BBC News 2015). The voice command feature bypasses the need to physically push buttons and scroll through menus. Therefore, it is potentially useful to those people that find reading handset options difficult due to visual problems, or pressing handset buttons due to hypokinetic issues. On the other hand, the utility of voice command features trade-off against the requirement of the device to "listen" for utterances in the room, raising privacy and security concerns over the capture and distribution of personal information (BBC News 2015).

With the availability of more advanced interactive television applications, many older adults may not be aware of how their data is being utilized. For example, research in the United States found over half of participants surveyed about Internet privacy statements were unaware that information collected about them may not be kept confidential (Pew Research Center 2014b). As a result, without greater transparency in the data collection process, there is a danger that some older adults will be reluctant to adopt, despite the usefulness of the technology. This may be overcome with clearer explanations in how personal information is utilized and protected. However, improving the clarity of privacy statements for older adults limited in technology experience needs to be further explored.

Designing with Older Adults

Inclusive design has long articulated a need for designers to consider the implications of their work among the wider population (e.g., Keates and Clarkson 2003). This includes the importance of working with the expertise of older adults, using collaborative approaches and methods that encourage their involvement as active stakeholders (e.g., Eisma et al. 2004; Dickinson et al. 2007). In *designing for empowerment*, Ladner (2015) emphasizes a "user-led" approach in the development and testing of prototype applications for people with disabilities. Similarly, Newell (2011) argues a need for a greater understanding from designers in the requirements of older and disabled adults through the concept of *user sensitive inclusive design*. In this context, the term *sensitive* indicates a need to establish an empathetic relationship to those people the technology is designed toward (Newell 2011). Within HCI research, this also includes breaking away from stereotypes and misconceptions of older adults, including considerations in how experiences across life courses may constructively help shape the design of new technology in old age (Vines et al. 2015).

Within the digital interactive television space, research has included the development of design methods to elicit new ways of thinking about television interaction (Rice 2009). This includes moving away from presenting older adults with more formalized design concepts, to exploring creative ways of engaging with them in the development of new digital interactive television applications. One approach has been the use of live theater as a means of conveying imaginary concepts through fictional characters. Using the experience of professional actors, prototypes were replaced with stage props (Rice et al. 2007). Supported by a professional scriptwriter, short dramatized stories were used to illustrate the social context of how the technology may be used, without explicitly having to demonstrate user interface features. In what Newell (2011) reports as an actor's ability to know when to "exaggerate for effect" (p. 143), the use of dramatization has the advantage of not being limited to the technical capabilities of any digital system. Rather, the focus of the work was to consider the potential opportunities and challenges of using the technology, by incorporating aspects of tension in the story as a means of formulating deeper discussion among audience groups. Importantly, issues discussed with the audience considered the social implications of use, rather than simply gathering opinions on what the technology could presently do (Rice et al. 2007).

In addition to the importance of designing with older adults, to our knowledge few digital television design guidelines cater for the provision of new applications and services. International communities address important accessibility issues from the standpoint of including older adults in web-based services (e.g., Web Accessibility Initiative 2010), and some broadcasting guidelines and recommendations are available that address television access for older users among other disability groups (e.g., Consumer Expert Group 2006; Digital TV Group 2014). However, beyond usability and accessibility issues, there is presently a lack of publically available data in the digital interactive television space to inform and inspire the implementation of more creative, experimental, and innovative design solutions, either through best practices, or recommendations based on empirical findings. Given the divergence of the medium, sharing more publically available resources could help transform the development of the technology for older adults, by providing examples and constructive insights for others to learn from. This includes encouraging greater collaboration between multidisciplinary teams of practitioners, designers, researchers, and end users, as more open resources should be available to facilitate interaction and push known boundaries.

Conclusion

Television is an evolving medium, and one that has changed greatly from a passively consumed technology, to one where the viewer has an active role in a range of personalized services. As a developing technology, digital interactive television represents a useful catalyst for the adoption of a wide range of applications that could enhance the lives of many older adults. However, with greater functionality there is also an increased risk of digital exclusion for those adults considered to be late adopters, in having spent a lifetime using analogue services. This indicates the importance of understanding the needs and preferences of older adults in the design of the technology. Given ongoing technical advances, there are also interesting opportunities to creatively expand the types of interaction and services available. As previously reported, this does not simply relate to watching more televised content, but in utilizing new features for bridging communication gaps and the support of different lifestyles. Likewise, more research should explore the benefits of emerging modalities and interconnected devices, while recognizing that current interaction processes can require a complex switching between dual interfaces (i.e., the television screen and remote control handset/s). Focusing on the inclusive design of digital interactive television for older adults is likely to have positive implications for wider segments of the population, which from a commercial perspective may help drive new consumer demands in the industry. Given the familiarity and prevalence of television in the lives of many older adults, this requires careful engagement by practitioners and researchers to help facilitate the medium's growing potential.

Recommended Reading

M. Alaoui, M. Lewkowicz, A. Seffah, Increasing elderly social relationships through TV-based services, in *Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium* (ACM, New York, 2012), pp. 13–19

Alzheimer's Society, Mild cognitive impairment (2015). http://www.alzheimers.org.uk/site/scripts/documents_info.php?documentID=120. Accessed 27 July 2015

American Speech Language Hearing Association (ASHA), Overview of hearing aids (2015). http://www.asha.org/public/hearing/Hearing-Aids-Overview/. Accessed 15 June 2015

A. Ardila, Normal aging increases cognitive heterogeneity: analysis of dispersion in WAIS-III scores across age. Arch. Clin. Neuropsychol. **22**(8), 1003–1011 (2007). doi:10.1016/j.acn.2007.08.004

J.M. Artz, Computers and the quality of life: assessing flow in information systems. ACM SIGCAS Comput. Soc. **26**(3), 7–12 (1996)

P.B. Baltes, A.M. Freund, S.-C. Li, The psychological science of human ageing, in *The Cambridge Handbook of Age and Ageing*, ed. by M.L. Johnson (Cambridge University Press, Cambridge, 2005), pp. 47–71

BBC News, Ceefax service closes down after 38 years on BBC (2012). http://www.bbc.com/news/uk-20032882. Accessed 23 June 2015

BBC News, Not in front of the telly: warning over 'listening' TV (2015). http://www.bbc.com/news/technology-31296188. Accessed 27 July 2015

BBC Trust, *Service Review of BBC Television: BBC One, BBC Two, BBC Three & BBC Four* (BBC Trust, London, 2014)

P. Biswas, P. Robinson, Evaluating interface layout for visually impaired and mobility-impaired users through simulation. Univ. Access Inf. Soc. **12**(1), 55–72 (2013). doi:10.1007/s10209-011-0265-5

S. Blackburn, S. Brownsell, M.S. Hawley, A systematic review of digital interactive television systems and their applications in the health and social care fields. J. Telemed. Telecare **17**(4), 168–176 (2011). doi:10.1258/jtt.2010.100610

J. Bobeth, S. Schmehl, E. Kruijff, S. Deutsch, M. Tscheligi, Evaluating performance and acceptance of older adults using freehand gestures for TV menu control, in *Proceedings of the 10th European Conference on Interactive TV and Video* (ACM, New York, 2012), pp. 35–44

V. Bures, Interactive digital television and voice interaction: experimental evaluation and subjective perception by elderly. Electron. Electr. Eng. **122**(6), 87–90 (2012). doi:10.5755/j01.eee.122.6.1827

A. Carmichael, *Style Guide for the Design of Interactive Television Services for Elderly Viewers* (Independent Television Commission, Kings Worthy Court, Winchester, 1999)

A. Carmichael, H. Petrie, F. Hamilton, J. Freeman, The Vista project: broadening access to digital TV electronic programme guides. PsychNology J. **1**(3), 229–241 (2003)

A. Carmichael, M. Rice, S. Lindsay, P. Olivier, iTV as a platform for rich multimedia reminders for people with dementia, in *Proceedings of the 6th European Conference on Interactive Television* (Salzburg, Austria, 2008), pp. 308–317. doi:10.1007/978-3-540-69478-6_41

A. Carmichael, M. Rice, F. MacMillan, A. Kirk, Investigating a DTV-based physical activity application to facilitate wellbeing in older adults, in *Proceedings of the 24th BCS Interaction Specialist Group Conference* (British Computer Society, Dundee, U.K., 2010), pp. 278–288

L.L. Carstensen, J.A. Mikels, At the intersection of emotion and cognition. Aging and the positivity effect. Curr. Dir. Psychol. Sci. **14**(3), 117–121 (2005). doi:10.1111/j.0963-7214.2005.00348.x

P. Cesar, D.C.A. Bulterman, A.J. Jansen, Usages of the secondary screen in an interactive television environment: control, enrich, share, and transfer television content, in *Proceedings of the 6th European Conference on Interactive Television* (Salzburg, Austria, 2008), pp. 168–177. doi:10.1007/978-3-540-69478-6_22

R. Coleman, Living longer, in *Inclusive Design: Design for the Whole Population*, ed. by J. Clarkson, R. Coleman, S. Keates, C. Lebbon (Springer, London, 2003), pp. 120–141

Consumer Expert Group, Digital TV equipment: vulnerable consumer requirements. A report by the Consumer Expert Group to government and Digital UK (2006). http://www.digitaltelevision.gov.uk/publications/pub_dtvconsumer_mar06.html. Accessed 22 June 2015

J.F. Coughlin, Speaking silver: lessons for product innovation & development in an aging marketplace. AgeLab 2007-02, MIT AgeLab (2007)

F.I.M. Craik, J.M. McDowd, Age differences in recall and recognition. J. Exp. Psychol. Learn. Mem. Cogn. **13**(3), 474–479 (1987). doi:10.1037/0278-7393.13.3.474

M. de Onis, J-P. Habicht, Anthropometric reference data for international use: recommendations from a World Health Organization Expert Committee. Am. J. Clin. Nutr. **64**(4), 650–658 (1996)

Described and Captioned Media Program (DCMP), Captioning key. Guidelines and preferred techniques (2011). https://www.dcmp.org/captioningkey/. Accessed 15 June 2015

A. Dickinson, J. Arnott, S. Prior, Methods for human-computer interaction research with older people. Behav. Inform. Technol. **26**(4), 343–352 (2007). doi:10.1080/01449290601176948

Digital TV Group, *UK Digital TV Usability and Accessibility Guidelines, including Text to Speech and Connected TV. Version 3.0* (Digital TV Group, London, 2014)

M. Docampo Rama, H. de Ridder, H. Bouma, Technology generation and age in using layered user interfaces. Gerontechnology **1**(1), 25–40 (2001). doi:10.4017/gt.2001.01.01.003.00

R. Eisma, A. Dickinson, J. Goodman, A. Syme, L. Tiwari, A.F. Newell, Early user involvement in the development of information technology-related products for older people. Univ. Access Inf. Soc. **3**(2), 131–140 (2004). doi:10.1007/s10209-004-0092-z

A.D. Fisk, W.A. Rogers, N. Charness, S.J. Czaja, J. Sharit, *Designing for Older Adults. Principles and Creative Human Factors Approaches* (CRC Press, Boca Raton, 2009)

A. Fleury, J.S. Pedersen, M. Baunstrup, L.B. Larsen, Interactive TV: interaction and control in second-screen TV consumption, in *Adjunct Proceedings of the 10th European Interactive TV Conference* (Berlin, Germany, 2012), pp. 104–107

J. Freeman, J. Lessiter, Easy to use digital television receivers: remote control buttons and functions used by different types of consumer (i2 Media Research, 2007). http://stakeholders.ofcom.org.uk/binaries/research/tv-research/dso.pdf. Accessed 15 June 2015

J. Freeman, A. Miotto, J. Lessiter, A. De Gloria, F. Bellotti, M. Mangarone, N. Hofshi, A. Carmichael, M. Rice, I. Maly, A.J. Sporka, M.H. Schut, A. Damstra, Defining a framework to support cognitive training for older people via interactive digital television, in *IBC Conference*, Amsterdam, 2009

T. French, M. Springett, Trusted and trustworthy digital interactive TV: paradise lost or a paradise (to be) regained?, in *Proceedings of the 3rd European Conference on Interactive Television* (Aalborg, Denmark, 2005), pp. 53–59

M. Fullerton, E. Peli, Digital enhancement of television signals for people with visual impairments: evaluation of a consumer product. J. Soc. Inf. Disp. **16**(3), 493–500 (2008). doi:10.1889/1.2896328

D. Gauntlett, A. Hill, *TV Living. Television, Culture and Everyday Life* (Routledge, London, 1999)

J. Gill, Access-ability. Making technology more useable by people with disabilities (2004). http://www.tiresias.org/research/guidelines/access-ability/Access-Ability.pdf. Accessed 21 June 2015

J. Gill, Remote controls (2009). http://www.johngilltech.com/guidelines/remote.htm. Accessed 28 July 2015

J.M. Gill, S.A. Perera, Accessible universal design of interactive digital television, in *Proceedings of the 1st European Conference on Interactive Television* (Brighton, U.K., 2003), pp. 83–89

E.L. Glisky, Changes in cognitive function in human aging, in *Brain Aging: Models, Methods, and Mechanisms*, ed. by D.R. Riddle (CRC Press, Boca Raton, 2007)

U. Granacher, T. Muehlbauer, M. Gruber, A qualitative review of balance and strength performance in healthy older adults: impact for testing and training. J. Aging Res. Article ID 708905 (2012). doi:10.1155/2012/708905

G. Harboe, N. Massey, C. Metcalf, D. Wheatley, G. Romano, The uses of social television. Comput. Entertain. **6**(1), Article 8 (2008). doi:10.1145/1350843.1350851

J. Hess, B. Ley, C. Ogonowski, L. Wan, V. Wulf, Jumping between devices and services: towards an integrated concept for social TV, in *Proceedings of the 9th European Conference on Interactive Television* (ACM, New York, 2011), pp. 11–20

J. Holt-Lunstad, T.B. Smith, M. Baker, T. Harris, D. Stephenson, Loneliness and social isolation as risk factors for mortality: a meta-analytic review. Perspect. Psychol. Sci. **10**(2), 227–237 (2015). doi:10.1177/1745691614568352

J.L. Horn, R.B. Cattell, Age differences in fluid and crystallized intelligence. Acta Psychol. (Amst) **26**, 107–129 (1967). doi:10.1016/0001-6918(67)90011-X

F. Huppert, Designing for older users, in *Inclusive Design: Design for the Whole Population*, ed. by J. Clarkson, R. Coleman, S. Keates, C. Lebbon (Springer, London, 2003), pp. 30–49

Independent Television Commission (ITC), *ITC Guidance on Standards for Audio Description* (ITC, London, 2000)

T.J. Jayroe, D. Wolfram, Internet searching, tablet technology and older adults. Proc. Am. Soc. Inf. Sci. Technol. **49**(1), 1–3 (2012)

S.J.E. Jolly, M.J. Evans, Improving the experience of media in the connected home with a new approach to inter-device communication. BBC Research & Development White Paper WHP 242 (2013)

S. Keates, J. Clarkson, Design exclusion, in *Inclusive Design: Design for the Whole Population*, ed. by J. Clarkson, R. Coleman, S. Keates, C. Lebbon (Springer, London, 2003), pp. 88–102

E.A. Kensinger, Cognition in aging and age-related disease, in *Encyclopedia of Neuroscience*, ed. by L.R. Squire (Elsevier, 2009), pp. 1055–1061

J.A. Klein, S.A. Karger, K.A. Sinclair, *Digital Television for All. A Report on Usability and Accessible Design* (Department of Trade and Industry (DTI), 2003)

J. Klein, J. Freeman, D. Harding, A. Teffahi, Assessing the impact of second screen. Report by Technologia in association with DTG and i2 Media Research (2014). http://stakeholders.ofcom.org.uk/binaries/research/technology-research/2014/Second_Screens_Final_report.pdf. Accessed 15 June 2015

R.E. Ladner, Design for user empowerment. ACM Interact. **22**(2), 24–29 (ACM, New York, 2015)

P.M. Langdon, M.F. Gonzalez, P. Biswas, Designing studies for the requirements and modelling of users for an accessible set-top box, in *Proceedings of the 8th International Conference on Disability, Virtual Reality and Associated Technologies* (Valparaíso, Chile, 2010), pp. 203–212

D.G. Lee, D.I. Fels, J.P. Udo, Emotive captioning. ACM Comput. Entertain. **5**(2), Article 11 (2007). doi:10.1145/1279540.1279551

J. Lessiter, J. Freeman, R. Davis, A. Dumbreck, Understanding DTT remote control button labelling: a multi-method approach, in *Proceedings of the 2nd European Conference on Interactive Television* (Brighton, UK, 2004), pp. 85–90

V. Lobato, G. López, V.M. Peláez, MHP interactive applications: combining visual and speech user interaction modes, in *Adjunct Proceedings of EuroITV 2009* (Leuven, Belgium, 2009), pp. 10–13

P. Looms, T. Owens, P. Orero, P. Romero-Fresco, D2.6 Recommendations on the effectiveness and efficiency of existing services improvements. Digital Television for All. ICT PSP/2007/1 (2010). http://dea.brunel.ac.uk/dtv4all/ICT-PSP-224994-D26.pdf. Accessed 2 Aug 2015

A. Marcoen, P.G. Coleman, A. O'Hanlon, Psychological ageing, in *Ageing in Society. European Perspectives on Gerontology*, ed. by J. Bond, S. Peace, F. Dittmann-Kohli, G.J. Westerhof, 3rd edn. (Sage, London, 2007), pp. 38–67

R. Marshall, K. Case, M. Porter, S. Summerskill, D. Gyi, P. Davis, R. Sims, HADRIAN: a virtual approach to design for all. J. Eng. Des. **21**(2–3), 253–273 (2010). doi:10.1080/09544820903317019

J.A. Martin, J. Ramsay, C. Hughes, D.M. Peters, M.G. Edwards, Age and grip strength predict hand dexterity in adults. PLoS One **10**(2), e0117598 (2015). doi:10.1371/journal.pone.0117598

M. Mather, Aging and emotional memory, in *Memory and Emotion*, ed. by D. Reisberg, P. Hertel (Oxford University Press, Oxford, 2004), pp. 272–307

A. Miotto, J. Lessiter, J. Freeman, R. Carmichael, E. Ferrari, Cognitive training via interactive television: drivers, barriers and potential users. Univ. Access Inf. Soc. **12**(1), 37–54 (2013). doi:10.1007/s10209-011-0264-6

National Institute on Aging, National Institutes of Health, U.S. Department of Health and Human Services, and U.S. Department of State, Why population aging matters. A global perspective (2007). https://www.nia.nih.gov/sites/default/files/WPAM.pdf. Accessed 15 June 2015

National Institute on Deafness and Other Communication Disorders (NIDCD), Presbycusis (1997). https://www.nidcd.nih.gov/staticresources/health/healthyhearing/tools/pdf/Presbycusis.pdf. Accessed 15 June 2015

A.F. Newell, *Design and the Digital Divide. Insights from 40 Years in Computer Support for Older and Disabled People* (Morgan and Claypool Publishers, 2011)
D.A. Norman, *The Psychology of Everyday Things* (Basic Books, New York, 1988)
M. Obrist, M. Miletich, T. Holocher, E. Beck, S. Kepplinger, P. Muzak, R. Bernhaupt, M. Tscheligi, Local communities and IPTV: lessons learned in an early design and development phase. Comput. Entertain. **7**(3), Article 44 (2009). doi:10.1145/1594943.1594956
Ofcom, *Online Trust and Privacy: People's Attitudes and Behaviour* (Ofcom, London, 2010)
Ofcom, *The Communications Market Report* (Ofcom, London, 2014a)
Ofcom, *Adults' Media Use and Attitudes Report, 2014* (Ofcom, London, 2014b)
Ofcom, *Guidelines on the Provision of Television Access Services* (Ofcom, London, Untitled)
Office of National Statistics, Stat. Bull. Internet Access Q. Update **Q1** (2014). http://www.ons.gov.uk/ons/rel/rdit2/internet-access-quarterly-update/q1-2014/stb-ia-q1-2014.html. Accessed 11 September 2015
V. Oksman, M. Ainasoja, J. Linna, M. Alaoja, P. Heikkilä, K. Alijoki, 2nd screen usage while watching TV: an ethnographic study. Next Media – A Tivit Programme, Phase 3 (1.2-31.12.2012) (2013)
R. Pak, A. McLaughlin, *Designing Displays for Older Adults* (CRC Press, Boca Raton, 2011)
J.D. Pearson, C.H. Morrell, S. Gordon-Salant, L.J. Brant, E.J. Metter, L.L. Klein, J.L. Fozard, Gender differences in a longitudinal study of age-associated hearing loss. J. Acoust. Soc. Am. **97**(2), 1196–1205 (1995)
E. Peli, E.M. Fine, A.T. Labianca, Evaluating visual information provided by audio description. J. Vis. Impair. Blind. **90**(5), 378–385 (1996)
Pew Research Center, Older adults and technology use. Adoption is increasing, but many seniors remain isolated from digital life (2014a). http://www.pewinternet.org/files/2014/04/PIP_Seniors-and-Tech-Use_040314.pdf. Accessed 1 July 2015
Pew Research Center, What internet users know about technology and the web (2014b). http://www.pewinternet.org/files/2014/11/PI_Web-IQ_112514_PDF.pdf. Accessed 27 July 2015
P. Rabbitt, Cognitive changes across the lifespan, in *The Cambridge Handbook of Age and Ageing*, ed. by M.L. Johnson (Cambridge University Press, Cambridge, 2005), pp. 190–199
J.C.A. Read, Viewer experience with stereoscopic 3D television in the home. Displays **35**(5), 252–260 (2014). doi:10.1016/j.displa.2014.09.001
M.D. Rice, An exploration of design strategies and methods in the development of digital interactive television for older people. PhD thesis, University of Dundee, 2009
M. Rice, A. Carmichael, Factors facilitating or impeding older adults' creative contributions in the collaborative design of a novel DTV-based application. Univ. Access Inf. Soc. **12**(1), 5–19 (2013). doi:10.1007/s10209-011-0262-8
M. Rice, A. Newell, M. Morgan, Forum theatre as a requirements gathering methodology in the design of a home telecommunication system for older adults. Behav. Inform. Technol. **26**(4), 323–331 (2007). doi:10.1080/01449290601177045
Royal National Institute of Blind People (RNIB), Eye conditions (2015). http://www.rnib.org.uk/eye-health/eye-conditions. Accessed 17 Aug 2015
L.Z. Rubenstein, Falls in older people: epidemiology, risk factors and strategies for prevention. Age Ageing **35**(S2), ii37–ii41 (2006). doi:10.1093/ageing/afl084
A.M. Rubin, R.B. Rubin, Older persons' TV viewing patterns and motivations. Commun. Res. **9** (2), 287–313 (1982). doi:10.1177/009365082009002005
T.A. Salthouse, What and when of cognitive aging. Curr. Dir. Psychol. Sci. **13**(4), 140–144 (2004). doi:10.1111/j.0963-7214.2004.00293.x
E. Shatil, J. Mikulecká, F. Bellotti, V. Bureš, Novel television-based cognitive training improves working memory and executive function. PLoS One **9**(7), e101472 (2014). doi:10.1371/journal.pone.0101472
V.J. Silva, V.F.L. Júnior, Monitoring of hypertensive patients through conventional medical devices integrated to the Brazilian digital TV, in *Proceedings of the 8th International Interactive Conference on Interactive TV & Video* (ACM, New York, 2010), pp. 183–186

S. Spinsante, E. Gambi, Remote health monitoring for elderly through interactive television. BioMed. Eng. OnLine **11**(54) (2012). doi:10.1186/1475-925X-11-54

M.V. Springett, R.N. Griffiths, Accessibility of interactive television for users with low vision: learning from the web, in *Proceedings of the 5th European Conference on Interactive Television* (Amsterdam, The Netherlands, 2007), pp. 76–85. doi:10.1007/978-3-540-72559-6_9

Standard Eurobarometer 76, Autumn 2011. Media use in the European Union report (European Commission, 2012)

E. Steinfeld, J. Lenker, V. Paquet, *The Anthropometrics of Disability: An International Workshop* (Rehabilitation Engineering Research Center on Universal Design, School of Architecture and Planning, University of Buffalo, The State University of New York, Buffalo, 2002)

I. Stuart-Hamilton, *The Psychology of Ageing: An Introduction*, 5th edn. (Jessica Kingsley Publishers, London, 2012)

M.S. Svensson, T. Sokoler, Ticket-to-talk-television: designing for the circumstantial nature of everyday social interaction, in *Proceedings of the 5th Nordic Conference on Human-Computer Interaction* (ACM, New York, 2008), pp. 334–343

H.S. Tsai, R. Shillair, S.R. Cotten, V. Winstead, E. Yost, Getting grandma online: are tablets the answer for increasing digital inclusion for older adults in the U.S.? Educ. Gerontol. **41**(10), pp 695–709 (2015). doi:10.1080/03601277.2015.1048165

K.J. Turner, M.R. McGee-Lennon, Advances in telecare over the past 10 years. Smart Homecare Technol. TeleHealth **1**, 21–34 (2013). doi:10.2147/SHTT.S42674

United Nations, *World Population Ageing 2013* (Department of Economic and Social Affairs Population Division. United Nations, ST/ESA/SER.A/348, New York, 2013)

J. Vines, M. Blythe, S. Lindsay, P. Dunphy, A. Monk, P. Olivier, Questionable concepts: critique as a resource for designing with eighty somethings, in *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems* (ACM, New York, 2012), pp. 1169–1178

J. Vines, G. Pritchard, P. Wright, P. Olivier, K. Brittain, An age-old problem: examining the discourses of ageing in HCI and strategies for future research. ACM Trans. Comput. Hum. Interact. **22**(1), Article 2 (2015). doi:10.1145/2696867

J. Waycott, S. Pedell, F. Vetere, E. Ozanne, L. Kulik, A. Gruner, J. Downs, Actively engaging older adults in the development and evaluation of tablet technology, in *Proceedings of the 24th Australian Computer-Human Interaction Conference* (ACM, New York, 2012), pp. 643–652

Web Accessibility Initiative, Web accessibility and older people: meeting the needs of ageing web users (2010). http://www.w3.org/WAI/older-users/. Accessed 15 June 2015

World Health Organization (WHO), Global recommendations on physical activity for health (2010). http://whqlibdoc.who.int/publications/2010/9789241599979_eng.pdf. Accessed 28 June 2015

World Health Organization (WHO), Visual impairment and blindness (2014). http://www.who.int/mediacentre/factsheets/fs282/en/. Accessed 15 June 2015

World Health Organization (WHO), Dementia (2015). http://www.who.int/mediacentre/factsheets/fs362/en/. Accessed 2 July 2015

P. Wright, Digital tablet issues for older adults. Gerontechnology **13**(2), 306 (2014). doi:10.4017/gt.2014.13.02.169.00

R.T. Zacks, L. Hasher, K.Z.H. Li, Human memory, in *Handbook of Aging and Cognition*, ed. by F.I.M. Craik, T.A. Salthouse (Lawrence Erlbaum Associates, Mahwah, 2000), pp. 293–357

T.P. Zanto, A. Gazzaley, Attention and ageing, in *The Oxford Handbook of Attention*, ed. by A.C. Nobre, S. Kastner (Oxford University Press, Oxford, 2014), pp. 927–971

Interactive Digital Narratives for iTV and Online Video

42

Hartmut Koenitz and Noam Knoller

Contents

Abstract

In iTV and online video, narrative interaction has long been a Holy Grail for both audiences and creators of these digital audiovisual works. On the one hand, interactive digital narrative promises interactors some exciting opportunities: to

H. Koenitz (✉)
Department of Entertainment and Media Studies, University of Georgia, Athens, GA, USA
e-mail: hkoenitz@uga.edu

N. Knoller
Interface Studies Group, Amsterdam School for Cultural Analysis (ASCA), University of Amsterdam, Amsterdam, The Netherlands
e-mail: knoller@uva.nl

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_44

enter the world of the story, to affect the story and perhaps even to control its outcome, and in the process to gain a transformative self-revelation. On the other hand, this new medium changes the role and craft of the author and entails a host of technological, conceptual, and institutional challenges. The authors describe these opportunities and challenges in detail, before examining various projects that have attempted to realize this vision and grapple with the challenges. Works discussed span several decades, from the late 1960s to 2015. They also span a variety of forms, including interactive cinema, online video, interactive television, and both video- and animation-based games. Particular emphasis is given to contemporary interactive documentaries (iDocs). The chapter further discusses current research directions, such as explorations of the increasing incorporation of the interactor's body and affects as well as second-screen and cross-device integration, and concludes with a fresh look at the original vision in light of conceptual and design approaches, current technological developments, their implications on the changing landscape of audiovisual content creation and consumption, and the new creative space of opportunities that they open up.

Keywords
iTV • Online video • Interactive digital narrative (IDN) • Interactive digital storytelling • Interactive cinema

Introduction

This chapter discusses interactive digital narrative (IDN) experiences in iTV and online video. The focus here is on narrative expressions within the specific framework of these forms of interactive video in digital media. iTV and online video had a long gestation period, the beginnings of which can be traced back at least to the 1960s. Their development was shaped by technological, institutional, cultural, and creative concerns, which we will discuss in relation to an overall vision. The related concern regarding meaningful interaction with program content has long been a focal point of a wider audiovisual culture, and its successful implementation can be considered a Holy Grail for the field. Ursu et al. (2008b), for example, mention interactive narrative as a specific opportunity for iTV. At the same time, this Holy Grail poses not only an opportunity but also a number of creative and pragmatic challenges.

However, the vision has long been more promise than reality. We will start by describing the overall vision, before considering the opportunities and challenges afforded for narrative by interactive TV and online video forms. We will then explain how several factors – technological, creative, and logistical – have now converged to turn the vision into a creative opportunity. The proliferation of interactive online video, especially interactive documentaries, plays a decisive role in this regard, as a creative test bed and proof of concept for the feasibility of creating interactive digital narrative with program content.

The chapter follows the historical development of iTV and online video by means of exemplary works while foregrounding the topics of narrative strategies, viewer engagement, and authorial position. In terms of narrative strategies, we investigate the specific methods applied by creators in these projects. At the same time, as the practice of creating and consuming such dynamic artifacts impacts the traditional roles of author and audience, we will consider these positions in relationship to traditional forms of storytelling. In particular, we will discuss the status of the "authorial voice" vis-a-vis the "participant audience" in shaping the narrative experience and affecting the outcome.

The Vision and Its Origins

The overall vision and motivation for interactive digital narrative in online video and iTV can be summarized as two broad goals – to explore the potential of reactive, malleable narrative, which often involves audience agency over what is presented, and to enable new forms of presentation hitherto impossible.

The origins of iTV and interactive online video can be traced back to early video experiments, in which artists like Nam June Paik started to manipulate the video image. Later, documentary filmmakers in the *cinéma vérité* movement explored ways to create a "truer" depiction of reality, in which the camera was to be an impartial recording device to capture real-world events. From this perspective, interactive documentaries were the next step, in that they shifted some control over what is seen to the audience, thus providing a richer, "even more true" depiction of a given subject. Glorianna Davenport was a pivotal figure in establishing the connection between documentary filmmaking and interactive cinematic presentation. After developing an interactive documentary about the changing cityscape of New Orleans in the 1980s, she founded the Interactive Cinema group at the MIT Media Lab and for the next 20 years furthered research through practical experiments in this area.

The interactive documentary tradition is a major force in interactive digital narrative for online video today. Interactive documentaries have established a regular presence at important documentary film festivals. Especially notable are the Amsterdam-based IDFA documentary film festival, which regularly showcases and helps develop the form through its iDoc lab venue; the New York-based Tribeca film festival, which funds interactive documentary projects through its Film Institute; and DocsBarcelona (since 2013). Several recent projects have managed to break out of the festival and academic circles and reach a wider audience via mainstream media exposure. Of those, we will discuss Fort McMoney (Dufresne 2013), Bear 71 (Allison and Mendes 2012), and Do Not Track (Gaylor 2015) (all three funded by the National Film Board of Canada).

Another lineage stems from fictional works, starting with the "interactive" project *Kinoautomat* at the Czechoslovakian pavilion at the 1967 world expo to interactive broadcast TV works like Hirschbiegel's interactive murder mystery *Mörderische Entscheidung* in German public TV in 1991 and the Danish project

D-Dag in 1999/2000. More recent works, such as the oddball romance *Accidental Lovers* (2006), the interactive docudrama *Inside the Haiti Earthquake* (PTV Productions 2010), or the virtual problematic life coach Karen (Blast Theory 2015), develop this tradition.

Technical advances have influenced the overall vision with regard to the possibilities of concrete implementations. For example, the first truly interactive video experiences were made possible by the advent of interfaces between computers and video storage technologies, which enabled random access to video clips. A particularly compelling application of this technology was the responsive virtual environment of the 1978 *Aspen Movie Map*, which pioneered a combination of touch interface running video clips which remained influential ever since. Similarly, the next major step in technological development, online video in concert with hyperlinking technology and database connections, enabled connected and expandable works, as exemplified by Davenport's Wiesner project (Davenport 1996).

Experiential Qualities of IDN: Agency, Immersion, and Transformation

A similar potential has been identified for other expressive forms in digital media, for example, for hypertext literature (Bolter and Joyce 1987), for interactive drama (Laurel 1986, 1991), and for digital media in general by Janet Murray (1997) in the form of a comprehensive interrogation of the affordances and experiential qualities of digital media in relation to narrative experiences. This expanded perspective on narrative in interactive digital media includes aspects beyond the traditional focus on the fixed product (as in the printed book and the final cut of a movie) to include the computational system and the interactive process (Koenitz 2010b, 2015). Crucially for the present discussion, this means practitioners need to reconsider established narrative conventions. In this regard, Murray focuses on iTV (Murray 2012), providing guiding principles for interactive narrative within this specific framework. Specifically, Murray identifies participation, procedurality (the ability of computers to execute code independently), spatiality (the traversal of virtual environments), and encyclopedic scope (since pragmatic restrictions of legacy media disappear) as the affordances of digital media. Of equal importance are Murray's insights into phenomenological aspects, the experiential qualities of agency (the ability to affect intelligible changes in the virtual world and feel in control), immersion (the deep involvement and active suspension of disbelief), and transformation (the effect of the experience of choices and consequences along with the awareness of alternative paths). While Murray's affordances, as well as agency and immersion, have been reflected frequently in scholarly treatments, transformation has so far seen much less reflection. However, we see a particular opportunity for iTV and online video in this regard, an aspect we will expand upon later. Murray also provides a useful critical vocabulary that avoids entanglement with usages in critical perspectives on traditional narratives. In particular we will use the term "interactor" to denote "user" and "participating audience."

However, in our usage we expand the meaning of agency. Most of the literature on agency refers to it as the ability for interactors to intentionally affect the presentation through their decisions and choices; however, some interactive narrative projects employ a different strategy, which could be called "passive interaction." In this case, the presentation changes meaningfully in response to unintended or sub-intentional aspects of the interactor's activity. Examples for such sub-intentional triggers are presence and geolocation (as in *Do Not Track* – see below) but also body language and other physiological signs that are given off unintentionally and captured by active interfaces (such as in *Pinter* (Gilroy et al. 2013) or Pia Tikka's *Enactive Cinema* approach (Tikka 2008)). These projects explore agency in the sense that they allow interactors to grasp the full spectrum of their interaction with computers. In this way we reconceptualize agency to also include reactions to unintended and unconscious actions by interactors. This augmented sense of agency integrates critical reflection on the developing relations interactors have with and through computers as well as their own performances. Creators can invite this kind of reflection by allowing interactors to discover which aspects of an experience are intentional and under their control and which aspects are outside their control, reflecting on their own role and the function and agency of other entities and stakeholders.

Opportunities and Challenges

Any new form of creative expression has a binary character – novel capabilities of expression come with the burden of learning a new craft. We can see this kind of development with earlier "new media" forms such as film, radio, and (noninteractive) TV. It takes time for creators to appropriate a new medium and for pragmatic conventions to be established. For example, film initially focused on the spectacle, showing arriving trains and boxing kangaroos, before montage and specific methods of framing were developed that enabled a specific form of narration. At the same time, audiences develop an understanding of what they encounter, while a distribution mechanism also appears. In this vein, the advent of film saw the creation of an industry that included a range of new professions from film director to producer to distributor to projectionist but also purpose-built facilities in the form of movie studios and cinemas.

Creative Opportunities

From Author to Narrative Architect

Instead of fixed representations of authorial intent, IDN affords dynamic artifacts whose concrete instantiations in a given walk-through can be unexpected even to their own creator. The capacity for unforeseen results originates from procedurality and participation. The digital work is no longer a static object; it acquires a life of its own as a dynamic system. For the creator, this means a change from the traditional authorial position to one of a "narrative architect" (Jenkins 2004), who builds a flexible reactive structure – no longer plot, but "narrative design" (Koenitz 2010b) – before stepping back and observing what others will do with it, akin to a real architect who provides the structural aspect of a building, and intended uses for the enclosed space, but is essentially without control, once the building has been completed. However, unlike in architecture, in digital media, the structure itself is in principle malleable through generative elements, while the content can change via online links and user-generated parts. Another addition to the authorial role is captured in Murray's perspective on "scripting the interactor" (1997) – measures taken by the creator to structure how an interactor will use the system. This, in turn, changes the respective role of the interactor, whose activity can now be understood as *userly performance* (Knoller 2012; Knoller and Ben-Arie 2015) – an individual interpretive execution of a given work through interaction with the narrative design within the parameters of the authored "interactor script."

These aspects signal a major change in IDN compared to traditional narrative and point toward an opportunity for creators to explore new forms of narrative expressions and the pleasure of "system building" in contrast to the traditional creation of "object art."

First-Person Experience: The Spectator as Participant

In interactive digital media, the interactor actively participates in the narrative, affecting plots and events, possibly positioned as a particular character in the storyworld, or switching between characters to discover different perspectives. The interactor's enactive experience entails adopting a strategic attitude toward the viewing experience. Authors need to suggest dramatic goals that would entice the viewer to appropriate and own them, as well as an interaction model and concrete interface that the viewer can master. Intentional decisions and choices, successful or failed performances of specific actions and interaction strategies, and (especially when sensor-rich interface devices are available) even unintentional aspects of that userly performance – all of these can have meaningful consequences in the storyworld. Crucially, what is meaningful might differ for every interactor, as meaning making under the conditions of interactive narrative itself becomes an interactive – procedural and participatory – process, based on *userly performance* and the reconstruction in the interactor's mind how her own actions and behaviors may have been the cause of the observed consequences.

A crucial aspect for the wider acceptance of IDN in iTV is the question of a "reward," of additional insight and pleasure interactors will gain from their additional "nontrivial effort" (Aarseth 1997) over a traditional passive TV-watching experience. The relation between interactors and program content, by means of their often deep emotional investment, offers an opportunity in this regard. In particular, the experience of the consequences of an interactor's actions and behaviors (especially delayed consequences Koenitz et al. 2013) is a very powerful effect waiting to be exploited by producers of IDN. Murray describes this effect as a "personal transformation" (Murray 1997, p. 170) – a transformative understanding of one's own active place in the storyworld, affected by the awareness of alternative paths. This specific knowledge is essential to IDN, and if the narrative design has been carefully crafted to support replay (Mitchell and McGee 2012), it can motivate interactors to explore and uncover alternative paths and consequences. Interactors experience transformation when they start to question their role in the unfolding narrative and consider the implications of their behavior in the interactive narrative's outcome and overall meaning.

Recently, video games have started to exploit this effect, for example, in the *Walking Dead* series (Telltale Games 2012), the interactor is often presented with hints of future consequences. Even earlier, the *Mass Effect* series (Electronic Arts 2007) gave interactors the opportunity to import a character developed in earlier iterations, and thus prior decisions have an effect on later games; simultaneously, the personalized virtual character acquires additional value for the interactor through its status as a long-term investment. In this way, the designers of *Mass Effect* have combined two powerful narrative strategies: long-term, delayed consequences and deep personal investment.

Challenges for iTV and Online Video

Technical and Institutional Challenges

A closer look at the challenges for iTV and online video unveils a range of problems. Technical aspects, economic conditions, and political/institutional frameworks need to be considered along with creative challenges. For example, for online video to become a technical reality, suitable data formats had to be developed along with distribution mechanisms. In addition, broadband Internet needed to be affordably available at home to create a mass-market audience. This particular development has taken approximately 20 years, from the 1993 introduction of the QuickTime video file format and the development of streaming technology in the late 1990s to affordable landline broadband speed availability in the late millennial decade and its extension to mobile devices in the form of 4G and LTE technology more recently.

For iTV, the picture is even more complex, as this vision entails changing or at least augmenting a fully developed media form with its associated institutional framework, production pipeline, and distribution mechanism but also the interactors' expectations originating in their traditional role of behaviorally passive viewers. Seen in this light, it is no surprise that iTV's history has not been a straight path to success, but rather a winding road and often more vision than reality.

The challenge for iTV becomes even more obvious in the comparison: online video operates within an infrastructure designed for network-wide seamless distribution and encounters users who accept their roles as interactors exploring new modes of active consumption. There is also no need to change the existing structure of private- or state-funded TV networks built on the broadcast model, which entails reaching as many viewers as possible with the same program content – essentially the opposite of personalized transformational experiences. Equally in opposition to the vision of iTV and online video is a production pipeline geared toward the “final cut,” the definite version that transports the collaborative vision of screenwriter, director, and director of photography. The challenge here has a dual nature, both creative and financial, as interactive choices in most cases (unless procedurally generated) require the creation of additional content, which can lead to a problem referred to as “the combinatorial explosion” (Ryan 2004; Stern 2008). Movie production is costly even for noninteractive projects, and the added cost of producing additional footage to accommodate choices can become prohibitive. Indeed, full motion video games of the mid-1990s were based on extensive scripts that could reach four times the length of the average Hollywood motion picture script. The script for Phantasmagoria (Sierra On-Line, 1995), for example, was 550 pages long (“Phantasmagoria,” n.d.).

On the technical side, iTV has long been a battle with existing infrastructure in the form of set-top boxes, of which many lacked the necessary computational capabilities to enable fluid interaction with high-quality video content. In the authors’ personal encounters with iTV developers, the “inadequate set-top box” was often mentioned as a major stumbling block for projects.

The Authorial Role Reconsidered

Existing authors, used to create fixed structures, are not trained in interactive media and therefore need to reconsider, and in many cases relearn, narrative under the conditions of interactive media, as described earlier in this chapter. This means that they need to capitalize on the specific affordances and phenomenological qualities and embrace their new role as narrative architects. What is required here is primarily a conceptual understanding of the opportunities of digital media, not necessarily a full comprehension of all technical aspects, just as real architects are not civil engineers. This conceptual shift contrasts with the established but unfortunate practice of “interactivization” (Koenitz 2010a; Koenitz and Louchart 2015), the process of simply adding interactivity to traditional forms, which brings about unsatisfying results. Instead, iTV and online video require a rethinking of the status of narrative beyond traditional, linear incarnations. This learning process is made more difficult because commonly accepted conceptual and pragmatic communicative conventions are still emerging. In more established audiovisual culture, such conventions channel viewer’s expectations but also provide guidance during the production of a work. Examples include *continuity editing* (a way to communicate the continuity of action without showing the full action) and *TV show formats* (conventions such as duration of shows, pacing of action, and forms of audience engagement). This challenge is further compounded by the rarity of professional training resources in interactive narrative. Therefore, IDN authors still often learn by experimentation.

iTV and Online Video: A History in Examples

The history of iTV is a difficult one – of great potential and discouraging reality, of early successes and unfulfilled promises, and of dependence on outside factors entirely out of control of iTV producers. In this chapter, we will focus on the evolving technology from analogue videodiscs to digitized and online video and place a particular emphasis on interactive documentary projects, which lately become more established through high-profile projects like Fort McMoney (Dufresne 2013).

Kinoautomat: An Offline Precursor

Interactive cinema dates back to the 1967 experiment Kinoautomat (translation: automatic cinema machine) created by Radúz Çinçera for the Czechoslovakian pavilion at the Montreal Expo67 World's Fair. The production *One Man and His Home* was shown in a specially designed presentation room where each seat was equipped with buttons. At several points during the presentation, the movie was stopped and the audience was asked to make a decision (Fig. 1) . Depending on the answer, the projectionist exchanged the lens cap between two synchronized film projectors. The underlying branching structure was carefully constructed to fold

Fig. 1 Audience engagement with Kinoautomat in a contemporary showing (2013). Picture credit Aerofilms

back after each choice, thus creating the appearance of interaction as Naimark observes: "The artfulness, [. . .] was not in the interaction but in the illusion of interaction" (Naimark 1998). While Kinoautomat's "folding back" branching structure can be read (and was arguably intended) as an ironic comment on illusory binary choices, it also offered a first attempt to solve the problem of additional content creation.

Videodisc: Interactive Video Becomes a Reality

Kinoautomat required a human intermediary to execute the audience's choices. Direct interaction between an interactor and a cinematic experience was not possible until the introduction of the laser videodisc in the late 1970s. Unlike systems based on tape technology, laser discs allowed random access to every point in a video. This capability, combined with a hardware interface that allowed a computer to control the playback of a videodisc, enabled fully interactive cinematic experiences. Among the first applications of this technology was the *Aspen Movie Map* (1978), created by the Architecture Machine Group at MIT. The *Aspen Movie Map* enabled an interactor to virtually explore the town of Aspen in Colorado, by controlling a running video of a drive through the town (Fig. 2).

The interactor was able to stop and take turns into side streets but also to click on the façades of houses to access additional material such as interior shots, historical

Fig. 2 Interface of the Aspen Movie Map (Picture credit Bob Mohl (2006))

imagery, restaurant menus, and video interviews with inhabitants. On the technical side, the project required a range of purpose-made technologies, from a gyroscopic mount for the camera placed on the car, to a database for metadata about each location, to the touch screen menu overlaying the video feed. The *Aspen Movie Map*'s status as a milestone project becomes even more obvious once we realize that it took about three decades to make a similar technology (still without running video) accessible to the average consumer in the form of Google Street View (2007). This fact also speaks to the challenges in distribution and the prohibitive production costs mentioned earlier. Indeed, Street View still cannot match the detailed level of information offered by the Movie Map, and as a result, the overall experience entails less agency and immersion.

Cinematic Interactive Video in the Arcade

In addition to the virtual travel application of the Movie Map, the laser videodisc technology found applications in arcade games, and later, with the emergence of CD-ROM technology, interactive video also entered the world of personal computers and game consoles. This genre of full motion video (FMV) games is closely related to the film industry, first in the connection to Disney-style animations and later in the form of film actors and Hollywood-style productions.

Former Disney animator Don Bluth used his skills to produce an interactive game with full motion video clips in the arcade game Dragon's Lair (Cinematronics 1983). The game's cinematic depiction – vastly superior to contemporary computer-generated graphics – played a considerable role in its commercial success.

Agency in Dragon's Lair in regard to the overall story is restricted as no significant changes to the narrative can occur – either the interactor wins the next challenge/fight and progresses further or she fails and has to restart. However, the game excelled in conveying "avatar agency," of controlling the main character, as the joystick movements triggered the display of different video clips. This aspect, together with smooth animations and carefully crafted sequencing of the different video clips, creates an immersive experience. Dragon's Lair is therefore a successful model for how agency with programed content can be realized by focusing on "avatar agency." The transformational aspect, however, stays restricted with this kind of narrative design.

In regard to the practical challenges, the production costs for the animations alone were reported to be one million dollars ("Dragon's Lair" (1983 video game), n.d.), a huge budget for a computer game in the 1980s. However, the outstanding graphics made the game a success, which promoted the production company to quickly create the follow-up *Space Ace* (Cinematronics 1984).

The commercial success of these two games led to a slew of similar titles from other companies, including several that avoided the high costs of original animations by repurposing clips from existing movies. The resulting products are therefore interactivized – repurposed cinematic presentation with limited interactivity.

The category of full motion video (FMV) games got another boost in the early 1990s with the introduction of the CD-ROM as a storage medium for personal computers and game consoles. The possibility to integrate longer video clips in games for these platforms was used a number of releases at the time, including *Star Wars: Rebel Assault* (LucasArts 1993), *Phantasmagoria* (Sierra On-Line 1995), *Star Trek: Borg* (Simon and Schuster 1996), *Black Dahlia* (Interplay Entertainment 1998), and *The X-Files Game* (Fox Interactive 1998). Many of these games feature well-known Hollywood film actors like Dennis Hopper (in Black Dahlia) or the stars of the related TV series (e.g., in the X-Files Game) and were expensive productions for the time. However, the heyday of FMV games did not last long, as the advent of 3D immersive environments in games like *Doom* (ID Software 1993) made interactors painfully aware of the limited interaction and restricted agency resulting from choosing between video clips.

With interactive video in general, once the selection is made, the interactor is turned into a passive viewer who has to patiently wait for the next decision point. The challenge here – and this is one of the key ingredients to successful interactive video – is to carefully craft each video clip in such a way that the immersion is not broken and the temporary passivity seems natural. Another important practical challenge lies in the high production costs to provide alternate story lines.

Roberta Williams, the designer of the FMV title Phantasmagoria, identifies an additional creative problem connected to working with professionals trained in noninteractive forms. In particular, she stresses the difficulty in fending off the lure and conceptual transgression of the established form. Williams' perspective here reiterates our earlier point regarding the need for IDN authors to embrace the changed role of narrative architect and how important it is to establish new conventions for interactive video:

> [. . .] most computer game designers who become involved with "Full Motion Video" games become enamored of the "movie aspects" of their game and lose sight of the fact that it's a game and not a movie. (Bellatti 1999)

Interactive Documentary Installations

As a pioneer, Glorianna Davenport encountered many problems for the first time and had to find solutions; for example, she developed the idea of a default edit in her projects, as interactors preferred such a curated experience in their initial encounters. In this way, immersion in the material is established first, and agency comes together with the interactor's interest in learning more. The goal for the New Orleans project was to provide a learning tool for students in urban planning to gain insights into the diverse aspects of a massive urban development, including maps, site histories, and relevant legislation, but also on understanding the goals and hidden agendas of the stakeholders involved in the project:

> In the final videodisc version, well over 40 characters are highlighted; and major storylines, such as the Jax Brewery development, the Louisiana World Exposition, and the Riverfront Development Plan prepared by Edaw, Inc. have multiple sub-plots which intertwine. (Davenport 1987)

In addition to exploring the provided material, interactors can create their own edits and make them available to others, essentially creating a layered model of authorship, which supports widely differing perspectives on the topic from the same source material. Overall, we can understand *Orleans in Transition* as an outstanding first attempt at "interactive delivery in education" (Davenport 1987). However, given the production costs and technical realities of the time – online video was not feasible in 1987 – access to the project was severely restricted.

Interactive Broadcast TV

Early forms of interactive TV were realized on public broadcast channels using multiple simultaneous broadcasts. Two such experiments were created by prominent European film directors, Oliver Hirschbiegel and Lars van Trier. In 1991, Hirschbiegel created *Mörderische Entscheidung* (Murderous Decision) for German public TV stations. This crime story was shown on two TV channels simultaneously, each channel showing the narrative from the perspective of a different protagonist, one male and one female. The viewer of the production became an interactor by changing channels with her remote control; agency was therefore restricted to an act of selection, with no control over the course of the narrative. Hirschbiegel applied several strategies to adapt his narrative for interactivity. For example, he tried to script the interactor by managing the amount of information given – a reduced amount of information on channel was a prompt to change the channel. At the same time he made sure that information essential for understanding the story was given on both channels. Hirschbiegel also purposefully used a genre – film noir – in which narrative gaps, caused by changing the channel at inappropriate times, seemed natural. An empirical study about the experiment (Kirchmann 1994) found that it worked best when it depicted the same objects from different points of view, for example, when both protagonists shared the same space and their views were represented similarly.

The Danish experiment *D-Dag* (Kragh-Jacobsen et al. 2000) expanded the concept of multichannel video by showing four different stories on separate channels plus three additional channels showing the directors' commentary. The concept for *D-Dag* came from several *Dogme 95* filmmakers and was shot by four camera teams in real time, each following a different actor. The overall narrative for *D-Dag* was a bank robbery on New Year's Eve of the new millennium, with the noise from the fireworks being used to mask the explosion needed to break into the safe. The four movies were shot from 11:05 pm on December 31, 1999, to 0:15 am on January 1, 2000, and ran 70 min each. While the filmmakers did not overtly address the irritating feeling of not getting all the necessary information by watching the "wrong" channel, Birk Weiberg (2002) in his study of interactive cinema reports that Valdis Oscarsdottir, the editor of the theatrical release of *D-Dag*, saw "zapping" as an answer to boring content, similar to Hirschbiegel's narrative strategy.

Agency in the *D-Dag* project is again restricted to changes in the sequencing of a narrative; however, *D-Dag* fails in giving the interactor positive reasons to change the channel. Decisions made out of interest are replaced with decisions made out of boredom, a strategy Weiberg derides as a "negative aesthetic of boredom" (Weiberg 2002). However, D-Dag is also an example for the opposite – information overload. With four channels of content and three channels of commentary, but no function to orient herself in the narrative – a map, a structural overview, and a Davenportian "standard play" – agency becomes an exercise in random "zapping." At the same time, immersion is hard to maintain, interrupted by boredom and the irritating feeling of missing important information on a different channel. As a result, instead of transformation, there is the feeling of a missed opportunity as exemplified in a viewer's commentary on the Internet Movie Database (IMDB):

> In my opinion the idea was really cool and interesting, but i think that you missed out on the whole story, by only being able to watch one person at a time. It then raises the question if we're getting the whole truth, when we see the world from our point of view. A cool statement and a good idea for a movie. But in reality, it was quite boring. (Jeppemh 2000)

The lack of concern for the interactor is again symptomatic of the difficulty for artists trained in traditional cinematic ways of storytelling to adapt to interactive forms of narration. The change in authorial role is both an opportunity and a challenge akin to changing the artistic discipline or learning a new language.

The Extensible Online Documentary and User-Generated Content

Much of Glorianna Davenport's work at the MIT Media Lab reflects a conviction that some form of orientation, a way to create interest, must come first, by offering a "default play" that requires no interaction from a user. Davenport's initial solution to the challenge of interactivity is therefore a two-stage process, in which both noninteractive and interactive modes are combined in the same project. However, her overall goal is to empower movie watchers and turn them into collaborators, which means to allow interactors to contribute content. To this end, Davenport created an online extensible interactive documentary *Jerome B. Wiesner, 1915–1994: A Random Walk through the 20th Century* (Davenport 1996) that allowed interactors to add their own material, an "extensible documentary":

> Some stories can only be gathered over time, growing and changing as new materials are added: these extensible (or "evolving") documentaries require the construction of content- and material-handling systems which can accommodate dynamic shifts in the quantity and sequencing of story elements without obscuring narration or presenting discontinuities which would disrupt the viewing experience. (Davenport 1997)

The Wiesner documentary combines a number of computational solutions, including a database, a user interface for the World Wide Web, and routines to handle uploading of the material. In the database, each element of the narrative is associated with keywords representing four of the classic five "w" questions: "who," "what," "where," and "when" (Davenport 1997). The user interface to *A Random Walk* uses three main elements – a map, a list of the keywords, and a frame showing the currently selected content. For the map, individual stories are positioned on a grid, where each tile represents either a time period, major topics

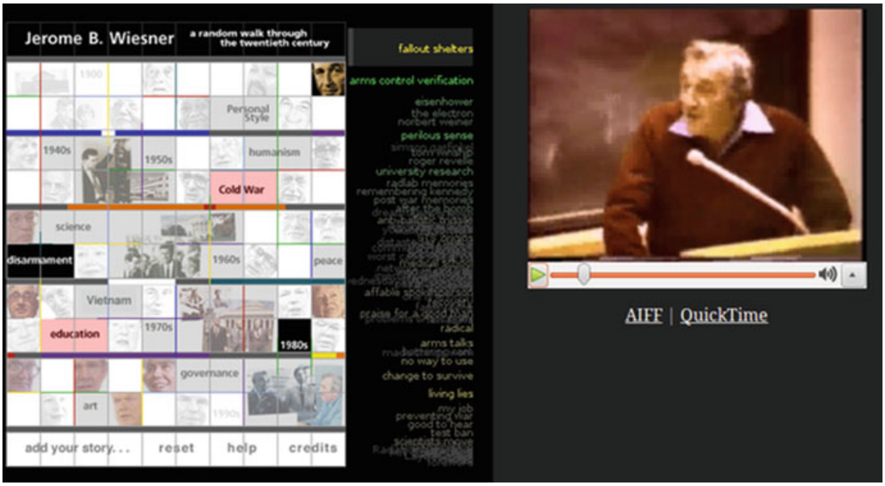


Fig. 3 The interface for *A Random Walk Through the 20th Century* (Davenport 1996). Map on the left, keyword list in the middle, and content frame on the right (Picture credit automatist.org under Creative Commons Attribution Share Alike license http://creativecommons.org/licenses/by-sa/3.0/)

(education, disarmament, peace), or people related to Wiesner. When the interactor clicks on a title, the respective video clip or piece of text is displayed in a frame. In addition, related tiles are also highlighted. The list of keywords also marks similarities to other content using artificial intelligence (AI).

A Random Walk Through the 20th Century is significant for several reasons. Most importantly, the project takes IDN as an opportunity for new forms of narrative and thus forgoes traditional narrative structure. This bold decision enabled the creator to implement a number of digitally native narrative strategies. First, the user interface (Fig. 3) is a major aspect of the narrative strategy, not just a way to transparently access content but a means to relate pieces of narration that might otherwise appear disconnected and present them in a coherent way to the interactor. The interactor's agency in the Wiesner project is not an afterthought but a major authorial concern from the start. Second, immersion in the Wiesner project comes from an active engagement, enabled by the move away from the laser disc and hardware-specific application of earlier projects to an easily accessible online platform, significantly increasing the potential audience. Third, the variety and encyclopedic depth of the material but also the interface with the map and the keyword list enable a transformational experience that includes an awareness that other paths exist and could have been taken. Fourth, the separation of client- and server-side technology on the World Wide Web means that sophisticated technology can be applied by the IDN creator – like the AI functions in the Wiesner project – without putting any burden on the Internet clients. Today's online interactive documentaries like Fort McMoney owe much to pioneering works like this.

IDN Design in Online Video and iTV: State of the Art

Online video documentaries and iTV experiences are veering toward more sophisticated narrative strategies. Structural composition and user engagement in contemporary works build upon earlier, pioneering work. Online video has not only become a technical possibility but a commodity or even a utility, reaching huge worldwide audiences as exemplified by Google's YouTube platform, where popular videos – including user-generated hyperlinked videos created with the *YouTube Annotations* feature – can reach millions of viewers. At the same time, the iTV hardware problem is on the verge of being solved by the convergence of broadcast devices and Internet appliances in "smart TVs" and modern streaming devices such as the AppleTV, the Google Chrome stick, the Roku, and the Amazon Fire TV, as well as the convergent use of smartphones and tablets as "second screens" or as additional control devices, supplanting both set-top boxes and traditional remote controls. At the same time, content creators have become more adapt in working with interactive media. In the following section, we will discuss some of the more mature projects that have been produced over the last decade.

Accidental Lovers

Accidental Lovers (2006) was a 12-part interactive TV drama series produced using the shape-shifting platform (Ursu et al. 2008a) for Finland's public broadcaster YLE. It aired on Finland's channel 1 during December 2006 and January 2007. Its story revolved around an unlikely romantic couple: 61-year-old cabaret singer Juulia and her 30 years younger love interest, pop star Roope (Ursu et al. 2008b, p. 11). Its interaction model allows audience to affect the narration by sending free-form SMS. These were moderated in real time by a human operator.

On the discourse level, only some messages were selected from those sent in, and in fixed intervals one of these would be selected to appear on the screen and influence the real-time editing. This was done through keyword analysis, which would determine specific topics for voice-over as well as audio and video materials that reflected a mood corresponding to the text message – but this had no effect on the plot. Apart from the selection of storyworld content, graphical elements in the form of two hearts, one for each of the characters, explicitly indicated changes in the moderating mood variable through warming and cooling their colors. This partial exposure of the interaction model seems necessary in the absence of established communicative conventions that, in cinema, for instance, are so engrained in the language that they are (at least in popular, mainstream works) imperceptible.

On the plot level, and less frequently, plot choices were determined by accumulated input. This was also explicitly indicated to the viewers by the color of a larger heart graphic (Ibid, p. 13) as well as by "text blanches" before significant plot turning points.

The series reached an audience of over half a million, indicating that it was immersive but not necessarily as an interactive experience. During all 12 episodes, only a small group chose to exercise the limited agency offered and to interact: the system had to handle fewer than 3,000 SMSs over the entire series, an average of about ten messages a minute. It is not clear whether any participating viewers sent more than one (each SMS costs €0.95, and a cost-per-interaction model, common in one off TV voting, may be a disincentive to sustained and repeated interaction), but it is still important to notice that fewer than 1 % of viewers were converted from inactive consumption to (somewhat) active engagement.

Inside the Haiti Earthquake

Inside the Haiti Earthquake (PTV Productions 2010) is an online hyperlinked video, which allows the viewer to explore relief operations after the 2010 Haiti Earthquake from three different positions: survivor, aid worker, and journalist. Assuming these roles, the viewer can adopt different interactive viewing strategies and experience the consequences of different decisions made and choices taken from within particular subject positions. This strategy has two likely effects: first, the viewer is allowed to experience an evolving event, in dramatic real time, each time from a perspective that is limited and particular, disclosing and highlighting gaps in knowledge and capacity that are particular to that perspective; this type of strategy is likely to foster a more empathic experience that challenges preconceived ideas about relief work – which stem from common linear documentary representations that privilege one of these positions. Second, the various perspectives are likely to create fresh tensions in the viewer's mental mapping of the documented event, prompting a more nuanced experience that can account for more of the complexity of such an event's context and meaning. This is a good example of additional transformative insight as a reward for the nontrivial effort applied by the interactor.

Bear 71

Directed by Allison and Mendes and produced by the National Film Board of Canada, Bear 71 (Allison and Mendes 2012) is an acclaimed online video documentary following the life of a grizzly bear at the Banff National Park in Canada. A 20-min narrative, realized as a poetic voice-over, serves as the auditory background for further visual explorations by the users. This exploration begins with an interface based on a map of the park. The map, rendered in a unique graphic style, is navigable and clickable. It organizes the documentary representation of the Banff storyworld spatially, locating characters (including, when possible, other users whose webcams may be included in the project, too) as they traverse the space, in relation to static features such as lakes, roads and railroads, town sites, etc. Clicking on the moving character icons, as well as on camera icons, spawns an overlay with information and a short video. There is no enactment and no effect on the storyworld. There are no reactions to the interactor's activity and no narrative agency beyond the triviality of seeing one's avatar/icon's movement on the map as well.

The voice-over track is the strongest narrative element, which also has only trivial interaction in the form of pausing or skipping directly to specific, labeled portions. This narrative competes for attention with the spatialized visual spectacle, yet it may also serve another purpose for the interactor: it suggests a duration for the experience. The interactor could also continue to explore the map and click through everything while the narration is paused or after it has finished, and in this case, the experience would take substantially longer. Ultimately, as in many works in IDN, the absence of established conventions requires the user herself to decide how to split her attention (Ben-Shaul 2003) between the narrative and the spatial exploration and develop a strategy of interaction.

In a user study of Bear 71, Kate Nash (2014) reflects critically on the audience interaction and reception of the interactive documentary, in an attempt "to consider whether and to what extent interaction and participation might result in new modes of audience engagement and reception" (222). The study included video documentation of the computer screens of 23 participants engaging with the work "for as long as it interested them to do so," observation notes of their behavior, as well as semi-structured interviews. The participants were specifically asked to interact with the work for the purpose of the study and thus may not represent a "naturally occurring" audience for the work. Nevertheless, the study still offers valuable insights into the different interaction strategies interactors employ and how these are formed. Nash concludes that, possibly in response to the "opacity" of the interface that interactors needed to make sense of, they had devised "three interactive strategies: spatial navigation, narrative interaction and play" (Ibid, p. 225), which they had often combined. Spatial navigation, often following the bear character, was the most prevalent strategy used, possibly explained by the fact that many interactors perceived that this is what the work itself was designed for (Ibid, p. 226). Although not rendered as an immersive 3D environment, the spatial organization of the work succeeded in promoting a situated, first-person perspective, as "[p]articipants often described their relationship to the environment in terms of presence, describing themselves as located in the environment and needing to move themselves within it" (Ibid, p. 227), although this perspective is navigational rather than enactive. As suggested above, most interactors employing the second strategy, narrative interaction, did indeed feel a degree of tension between spatial navigation and the audio narrative and thus privileged either the narrative or interaction. For many respondents the narrative structured their interaction as they sought out things mentioned in the story. In some instances the narrative "created expectations that were not fulfilled" (Ibid, p. 228), since the visual materials showed no narrative development. This inattention to implicit expectations for narrative agency, for seeing the outcomes of interactors' userly performances, often created frustration. Some interactors, however, "narrativized" their spatial exploration, producing readings that diverged from the facts of the text in order to create coherence between the audio narration and the spatial exploration. Finally, interactors who self-identified as "gamers" brought their familiarity with gaming conventions to the experience and in particular a stronger expectation for narrative agency. This is consistent with one of the authors' user study of interactive online video (Knoller and Ben-Arie 2009).

Similarly to the case of *Accidental Lovers*, the opportunity to interact does not necessarily translate into actual interaction. About a third of the interactors in the study chose to abstain from any action, for a combination of reasons including immersion in the audio narration, poor understanding of the interaction model, and personal preference.

Nash's study covers important aspects of the experience of an interactive online documentary. It still leaves open important questions: what was the longer-term effect of the experience? Did interactors come back to the documentary to explore additional aspects? How would audiences have responded to this documentary if they had encountered it on their own, rather than specifically asked by a researcher to view it?

Fort McMoney

The online documentary game *Fort McMoney* (Dufresne 2013) is a dynamic, interactive simulation of the actual Canadian city of Fort McMurray. It has two main components: (1) a spatialized, navigable video map of large portions of *Fort McMoney*, which – similar to the *Aspen Movie Map* discussed earlier – allows players to explore the city, collect items, and click on characters to view additional video material about them. This component is navigational rather than enactive, as it does not allow interactors to affect the storyworld. However, the more extensively interactors traverse the space, the more "influential points" and items they accrue. These can be used in the second component of the work: (2) a city simulation featuring a different graphical depiction and user interface, in which the viewers can interact with each other, debate issues, and vote on them (but not all votes are equal – players' votes are weighted by their "influential points"). As in other city simulation games, city statistics – number of inhabitants, property prices, homeless population, etc. – which initially were taken from Fort McMurray's real-world statistics, begin to depart from the real-world reference and reflect the collective agency of players.

Fort McMoney points in an interesting direction, but clearly there is a problem with the mode change in the user interface. The experience is not unified: apart from accumulating influential points, the spatialized video has few consequences in the game, and game activities have no bearing on the content of the video. Most of the agency occurs in the game component, where it is distributed among the various players, as story outcomes are subject to the social dynamics of debates and votes. The work, however, is clearly transformative, especially for interactors for whom the subject matter holds greater personal relevance.

Karen

This interactive video project is distributed as an app for the iPhone platform. *Karen* (Blast Theory 2015) is a fictional interactive life coach who purports to listen to the interactor and give concrete advice on issues like relationships. The presentation of the video clips on the smartphone is designed to emulate the style of a video conference call, adding a sense of realism. Structuring the experience as separate "calls" further enhances this narrative strategy. The calls are set apart by waiting periods of varying length – sometimes the interactor has to wait for the next day – during which Karen supposedly works on researching personalized advice for the interactor. At the same time, *Karen* has a backstory of her own, which becomes

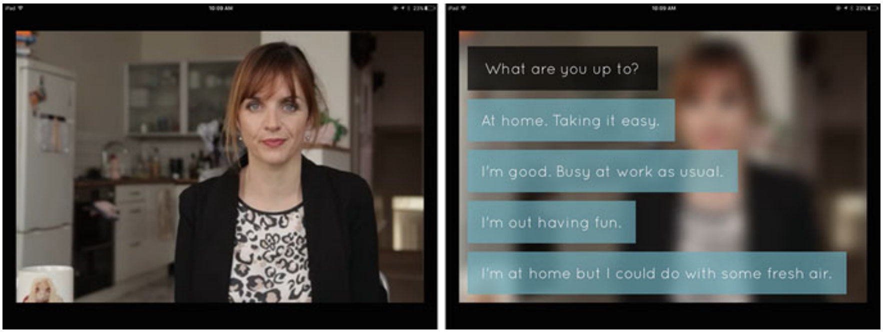


Fig. 4 Screen captures from Karen, "webcam view" and questions for the "patient" (Picture credit Blast Theory (2015), used with permission)

more prominent over time, offering a window into her own private life, for example, by showing the interactor a visitor to her apartment and including him in the judgment of a blind date.

Agency in *Karen* is limited to answering the occasional question; however, the project succeeds as an immersive and transformative experience. The key to understanding this apparent discrepancy is in the specific narrative situation – even with a real-life coach or therapist, we accept limited agency and expect to answer, listen, and learn while staying put. In this regard, *Karen* is an update of *Eliza*, Joseph Weizenbaum's 1966 experiment in using early AI technologies to emulate a Rogerian therapist (Weizenbaum 1966). Indeed, Janet Murray (1997) identifies Eliza as the very first work that demonstrated the narrative potential of interactive digital media and also as an example of the authorial strategy of "scripting the interactor." *Karen*, like *Eliza* before, works as long as the interactor accepts her role as a "patient." Since almost everybody has been in this position before, no instructions or prior training modes are necessary (Fig. 4).

As the narrative continues, Karen becomes more self-centered and concerned with her own life and thus abandons the "doctor role." Accordingly, the position of the interactor shifts to be closer to that of an observer. At this point, many interactors are no longer willing to accept the limited agency and feel cheated, as the following iTunes commentary demonstrates:

> After the first couple of sessions of info gathering you're mostly dealing with her personal issues, all of which are negative. Personally, I never felt like she was listening to my input at all – and she certainly showed no signs of understanding my personality or life. (iTunes Commentary from user "Gilty Plezurs")
>
> Even though I knew it was a game, it felt at first like having my own personal life coach, and I looked forward to my sessions with Karen. This feeling eventually turned to dread, as Karen became ever more twisted and needy. (iTunes Commentary from user "notRattlebrain")

The issue here is that Karen first creates the anticipation of a believable virtual life coach and then abandons this narrative scenario to reveal the personal issues of the life coach character. The overall narrative strategy of showing a character that seems normal at first and then becomes increasingly disturbed requires the interactor to change her mental model of Karen, which is a difficult task, given the investment in the initial role and the associated expectations.

Do Not Track

Do Not Track (Gaylor 2015) is a seven-episode personalized interactive web documentary series, written and directed by Brett Gaylor and produced by French interactive production company Upian and a number of public broadcasting and production companies.

Do Not Track exemplifies the use of some of the interactive affordances of post-PC apparatuses to narrativize their dispositive ramifications. Figure 5 shows the first interactive moment of the first episode. Making use of information that is given off by the viewer's browser (IP address and location), this interactive documentary is personalized even before the viewer initiates any intentional action. As relevant images are inserted in real time both from the documentary's own database and from real-time online sources, the adaptive narration declares: "I know that this is the country where you live. I know it's a nice night. I know you're on a PC."

Another strategy used is to ask the viewer to supply information about favorite websites they frequent. In return, the documentary responds with a personalized visualization of ad trackers associated with the website. Interactors who are logged in will see their preferences used again in later episodes. In the third episode, the documentary requests permission to access the viewer's Facebook data and in

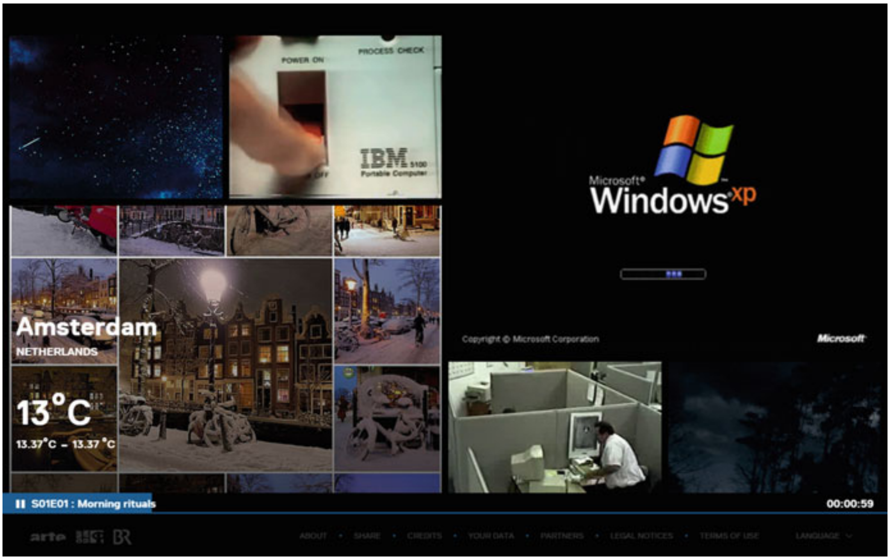


Fig. 5 Screen capture from Do Not Track, personalized to one of the authors (Picture credit Upian (2015) used with permission)

return offers a range of personality and financial analyses, of the type collected by data brokerage companies and used vastly by advertisers. Episode 4 requests access to the interactor's geolocation data and uses that to illustrate how such data is used for both user convenience and corporate profit. *Do Not Track's fourth episode* also incorporates a questionnaire about the interactor's mobile phone usage habits and provides immediate feedback in a conversational style. Episode 5 uses a game-like interface: connecting three labeled dots determines which example of the use of big data in our society will be shown. The viewer can thus select examples that are somewhat more personalized, making the illustration of the episode's main point that much more personally relevant and thus arguably more immersive. In the sixth episode – The daily me – the viewer is invited to filter two Twitter feeds, hiding either side of a societal debate. This is an inactive way to illustrate Eli Pariser's concept of the Filter Bubble (Pariser 2011), which the episode discusses. The concluding episode uses statistics, calculated from the data provided by registered users during the first 2 months of the project, to construct three different scenarios for the future of our society. Interspersed in the episode are a number of questions about predicted uses for the Internet and smartphones, and these enable the work to tell each viewer which of the three futures their answers have created.

Do Not Track employs both passive and active interaction and aspires to transform the viewer by harnessing narrative agency in order to encourage viewers to act in the real world. As the final episode's title – To change the future, click here – suggests, viewers are encouraged to investigate further provides links to civic Internet initiatives in the USA, Canada, the EU, and the UK.

Current Research and Future Developments

Incorporation of the Body and Its Affects

In line with the bodily and affective turns (Knoller and Ben-Arie 2015) of post-PC interfaces, some iTV researchers have been investigating the role of the viewer's body in the experience of TV and interactive TV. Others have started to explore how sensor-based technologies can enlist the relation between elicited emotional experience of viewers and program content, with a view to offer empathic TV experiences (Van den Bergh et al. 2014).

A recent experimental lab study by Tan et al. (2014) represents both directions. The researchers tried to find a link between side activities (both technologically mediated and other activities, such as eating and drinking) and body postures and how these relate to viewer emotions. Their goal was to find a way to unobtrusively monitor emotions of TV viewers, which could then serve as useful input for the design of iTV applications. Tan et al. showed viewers various video clips that were designed to elicit specific emotions and used a video camera to monitor the viewer's emotional responses and a Kinect camera to capture their postures. They did not find a direct relation between postures and emotional responses to their emotion eliciting video materials, but they did reveal an indirect relationship: there were both a correlation between posture data and “a subset of the relevant side activities” and a correlation between which device was used during the side activities and the experienced emotions. This allowed the researchers to conclude that tracking viewers through Kinect could allow program makers to anonymously and unobtrusively recognize which side activities a particular viewer were engaged in and perhaps even estimate what emotions they may be experiencing. This, they envision, might be useful to inform the design of applications on or in relation with the TV.

It should be noted here that if viewer emotions could be extracted unobtrusively (in line with the vector of prediction by attentive interfaces), these could also be used as input that affects program content. Affective viewer input can be used either to customize the affective tone of the narration (by changing the music, the lighting, or the pace of editing) or to trigger different story events, as has already been implemented in interactive cinema by Pia Tikka (2008) and in AI-based interactive storytelling by Gilroy et al. (2013) and others who pursue versions of the passive interaction approach.

Tan et al.’s failure to find a correlation between postures and emotions may be due to the fact that the program content they used – short video clips outside their original context, ranging from a sequence out of *When Harry Met Sally* to footage of leg surgery – was intended to elicit specific emotions, but not to encourage viewers to participate (let alone move or change posture). More research is required, using interactive program content, to correlate between postures and emotions. A stronger coupling between a program and an interface device such as Kinect would invite viewers to participate, move, gesture, and adjust posture and would reward them with both immediate, visible feedback in the presentation and immediate or delayed, but at any rate meaningful consequences in the program’s storyworld.

Gestural interaction with TV program content could be used semantically, to help viewers navigate information and content, or it could be used more expressively to help interactors convey (intentionally or tacitly) their own emotional intent or reaction to the content. Expressive interaction (Knoller 2004) of this kind has been largely on the margin of interactive video. Two early exceptions have been Michael Lew’s Office Voodoo (Lew 2003) and Knoller and Ben Arie’s InterFace Portrait system (Dekel et al. 2003, Knoller 2004). A recent online interactive narrative video commercial is a mature example of this type of interaction. *Sortie en Mer* (A Trip out to Sea) (CLM BBDO 2014) uses an embodied mouse-based mechanic to create a strong affective experience. The protagonist of the commercial has fallen off his yacht. In order to save him from drowning, the interactor is required to continuously scroll the mouse’s scroll button. The physical exertion demanded from the interactor (several minutes of frantic scrolling) simulates – and renders enactive – the physical exertion and diminishing strength of a drowning man and results in a first-person, embodied, and quite affective experience, which is still uncommon even in narrative computer games.

Notable exceptions in videogames include Kojima productions' Metal Gear Solid 4: Guns of the Patriot (Konami, 2008), which uses the similar device of desperately tragic "mashing", or the purposefully decelerated navigation in Tale of Tales' artistic videogame The Graveyard (Tale of Tales, 2008).

Second-Screen and Cross-Device Integration

One direction in iTV is to shift the place of interactive functions from the primary TV screen to a "second screen" – initially a computer with an online connection, now typically a tablet computer or a smartphone – showing additional content and interactive elements. A reader interested in following the most recent developments can refer to the proceedings of the ACM TVX conference (Geerts et al. 2015). Second-screen technology allows rapid deployment by using the client's existing hardware instead of new costly set-top boxes. A considerable amount of second-screen solutions are centered around non-narrative interaction, such as sports statistics and other informative services or as parallel spaces for social interactions about the program or for commercial transactions. However, narrative usages also exist, for example, CBS sync, based on technology developed by US company Watchwith. CBS sync is particularly relevant as a commercial and widely deployed product in the marketplace, in contrast to the experimental and nonprofit character of many of the other projects described here. The technology is used to enhance the viewing experience of a range of scripted TV series, including *CSI: Crime Scene Investigation* and *The Mentalist*. CBS sync features narrative content, for example, a character's backstory, a close up of an important piece of evidence in a crime mystery, but also quizzes about the current episode and opinion polls. While this form of iTV is in a supporting role of the main linear narrative, the technology prepares a mass audience for projects with more fully developed interactivity.

Nandakumar and Murray's (2014) second-screen companion app for the series *Justified* was designed to help viewers enter long-form TV narratives mid-season. They implemented an iPad app that was synchronized with the TV series, offering latecomers context-sensitive information about characters and events. A particular challenge in any second-screen application is to support viewers' need for additional knowledge through the second screen, without distracting them away from the program on the TV screen. Nandakumar and Murray's solution for that was to identify moments of potential confusion for latecomers to the series and offer relevant vital information concisely, precisely at this moment of greatest need. For example, when a reference is made to an unsolved crime, the relevant character would appear on the second screen with a speech callout quoting the key parts of the relevant dialogue. If the viewer chose to access additional material in the form of a 3-min video montage, the application automatically paused the TV to allow the viewer to catch up and resumed the program once the montage played through (Ibid). User testing proved some aspects of this strategy useful, while others require further research: while the companion app did increase comprehension by latecomers, it was not well suited for situations in which several viewers watch the program together, especially with variable levels of familiarity. Automatic pausing and resuming, in particular, proved incompatible with group viewing.

Conclusion: A Creative Space of Opportunities

The Vision Realized: Gaming as Long-Form Narrative

At the outset of this chapter, we presented the original vision of iTV as a form featuring highly interactive long-form narrative. This aspect of the vision has by now been very successfully realized – outside of iTV and online video – in narrative computer games. One of the most critically acclaimed examples for this type of game, as already mentioned, is *The Walking Dead game*, but a range of similar titles exist, for example, *Heavy Rain* (Quantic Dream 2010), *Beyond: Two Souls* (Quantic Dream 2013), or *Gone Home* (The Fullbright Company 2013). *The Walking Dead game* offers interaction with TV characters, rendered in synthetic animation, but using real voice acting.

What has not been realized, however, are video-based, highly interactive long-form narratives. As the disappearance of full motion video games has demonstrated, video is more a hindrance than a blessing for highly interactive projects. Employing video for alternate narrative paths under the control of the interactor translates into high production costs and creative challenges. Nevertheless, there is certainly a space for interactive video projects. Indeed, the exciting and ground-breaking projects described in this chapter demonstrate that IDN in online video and iTV is alive and well. The success of these projects stems from their creators' advanced understanding of the specific opportunities and limitations of this particular form, especially the realization that prerecorded video content is not malleable like synthetic, digital content. Digitally created landscapes, objects, characters, and text can now be generated procedurally and animated in response to input from interactors. Video clips cannot adapt to input from interactors in the same way, and this trade-off has implications that require creative solutions. One approach has been to strive for a finer level of granularity (Brooks 1997), using smaller units of video content that can be algorithmically edited to achieve a higher frequency of response to interactor input (Knoller 2004; Lew 2003). Another approach is to create a narrative situation that motivates a lower frequency of interaction. For example, in *Karen*, the limited agency is at least initially explained by means of a narrative strategy that combines a specific situation (life coaching sessions with a client) and the concrete manifestation structured as temporally discrete and delimited "video calls." These approaches point at a basic design principle for iTV/online video, stemming from the non-procedural affordance of video content – the creation of a narrative situation in which a limited frequency of opportunities for interaction is consistent with the interactor's expectations and thus remains believable and maintains narrative immersion. IDN creators using interactive video thus need to scrutinize their narrative strategy not only for compatibility with a procedural and participatory medium but also with the specific limitations of video content and to design an appropriate narrative situation. There are two crucial steps in this regard: firstly, the realization that interactivity – procedurality and participation – can be created on the level of narrative structure, in the sequencing of and conditional access to video clips, and that there are many creative opportunities along this trajectory and, secondly, that an interactor's perception of agency/immersion/transformation is not identical to the technical realization of a given work. In other words, creative narrative design can help create an illusion of agency even if there is none from a purely technical standpoint. This perspective on "IDN stagecraft" might become more palatable for some, once we realize that other artistic forms do the same – in the cinematic convention of continuity editing, most of the information is left out for our brains to fill in. Indeed, for an artistic narrative work, it is the interactor who is the ultimate judge – not a technical committee.

The state of video content as unmalleable is going to change in the near future, creating additional opportunities for interaction and agency. A range of technologies are currently being developed, and some have already been demonstrated that will enable real-time changes not only to the visual form of video, as in special effects, but to the content itself. These emerging technologies already include the removal of characters and props (Herling 2014) or the addition of 3D-rendered content (Simpson 2013). Further advances in visual recognition technologies, with the ability to maintain a dynamic, real-time semantic representation of video content, are likely to bring the malleability of video content closer to that of synthetic animation.

iTV and Online Video: Smart Devices Meet Network Distribution

In light of TV distribution on the Internet and the growing number of online-connected "smart TVs," the lines dividing iTV from online video are already blurring today and are set to fade away entirely in the near future. At least from a technical point of view, there seems to be little sense in maintaining a rigid distinction between TV content on a networked TV and the same content when it is viewed on any other networked screen device. With the notable exception of the "red button" system (Bennett 2008) in the UK (the most widely deployed proprietary platform for interactive TV), for most consumers, it is not the TV that has become interactive. Rather, it is interaction that has taken over the TV screen, in the form of smart TVs. At the same time, mobile screen devices and cross-device integration further complicate the traditional distinctions while also opening up new possibilities for relations between emerging types of screen media.

One remaining distinction would be between the different spatial and social contexts of the interactive viewing event which have formed around the two main types of primary screens, which we might understand as the "personal" (PC) and the "public" (TV). Historically, the PC – true to its name – has afforded a personal experience; in the contemporary setting, this role is taken over by tablets and other mobile screen devices. The TV, now reincarnated as the smart TV, will continue to occupy some of the traditional roles of TV. The central physical presence of a "big screen TV" still has the ability to offer the social experience of viewing as a communal event with a group of people who are colocated. The high level of production quality and the well-established distribution mechanisms of TV networks still offer a compelling experience. This holds true even for solitary viewing, informed by the awareness of a simultaneous experience with a mass of others (Murray 2012).

Another distinction regards the different convergences between these two primary screen types and other devices. Smart TVs are integrated into the Internet of things and designed for cross-device integration, in particular with smartphones. Therefore, they are more likely to be part of a second-screen formation and to make use of the sensor-based, semantic, or expressive gestural interaction that smartphones enable, when considered as input or control devices. PCs (but also mobile devices used as primary screens) seem less likely to be part of cross-device, second-screen formations, since their functionality already caters for the primary needs that second screens bring to the TV's spatial viewing context: computational power, network access, and an interface suited to user input. PCs will still be influenced by the bodily and affective turns within the more confined affordances of their hardware, as we have already seen in the case of *Sortie en Mer*. Finally, interactive video-based works can also be devised as mobile-first, small screen projects, making use of the particular affordance of mobile smartphones, such as location awareness, augmented reality capabilities, or access to (intimate) personal data. Creators of interactive video thus need to consider these different interface formations and configurations when planning their projects.

Ubiquitous Sensors and Connected Appliances: Interactive Narratives Outside the (TV) Box

Looking beyond the recent successful interactive video projects and synthetic video narrative games, the proliferation of smart devices in concert with the connected appliances of the Internet of things provides a new instrumentarium for creating interactive TV and online video experiences. Especially, the sensor-based affordances of mobile devices, such as location awareness, enable virtual and augmented reality and give new creative license for video-based projects. Multi-platform, cross-device, and "physical" interactive narratives can now be realized that spill out of the screen and into the physical world – onto electronic billboards, Internet-controlled light switches, and smart refrigerators.

The usage of sensor-rich smartphones (and soon smart watches) as input devices for gestural interaction seems to be an untapped creative and research direction. Smart devices can become props from afar, infused with human heartbeat or haunted by ghosts. Using them as output devices can allow characters to leave the primary screen and appear on the user's phone – or they might appear on a public display. Pandora's (TV) box is opened up and we cannot wait to experience the results.

Recommended Reading

E.J. Aarseth, *Cybertext* (Johns Hopkins University Press, 1997)

L. Allison, J. Mendes (eds.), *Bear 71*. National Film Board of Canada (2012). Retrieved from http://bear71.nfb.ca

A. Bellatti, Roberta Williams (1999, October 25). Retrieved June 15, 2015, from http://www.adventureclassicgaming.com/index.php/site/interviews/127

J. Bennett, "Your window-on-the-world": the emergence of red-button interactive television in the UK. Convergence **14**(2), 161–182 (2008). doi:10.1177/1354856507087942
N. Ben-Shaul, Split attention problems in interactive moving audiovisual texts. Presented at the fifth international digital arts and culture conference, DAC, Melbourne
Blast Theory, *Karen [Video Game]*. iTunes (2015). Retrieved from https://itunes.apple.com/us/app/karen-by-blast-theory/id945629374?mt=8
J.D. Bolter, M. Joyce, Hypertext and creative writing. Presented at the conference on computational semiotics for games and new media. (ACM, New York, 1987), pp. 41–50. Doi:10.1145/317426.317431
K.M. Brooks, Do story agents use rocking chairs? The theory and implementation of one model for computational narrative. Presented at the MULTIMEDIA '96 proceedings of the fourth ACM international conference on multimedia (1997)
Cinematronics, *Dragon's Lair [Video Game]* (Cinematronics, El Cajon, 1983)
Cinematronics, *Space Ace [Video Game]* (Cinematronics, El Cajon, 1984)
CLM BBDO, *Sortie en Mer*. CLM BBDO (2014). Retrieved from http://sortieenmer.com
G. Davenport, *Jerome B. Wiesner, 1915–1994: A Random Walk Through the 20th Century* (1996). Retrieved from http://ic.media.mit.edu/projects/JBW/
G. Davenport, New Orleans in transition, 1983–1986: the interactive delivery of a cinematic case study. Presented at the international congress for design planning and theory, education group conference, Boston (1987)
G. Davenport, Extending the documentary tradition. Presented at the Oberhausen international film festival, Oberhausen (1997). Retrieved from http://mf.media.mit.edu/pubs/conference/OberhausenExtending.pdf
A. Dekel, N. Knoller, U. Ben-Arie, M. Lotan, M. Tal, One measure of happiness-a dynamically updated interactive video narrative using gestures. Presented at the human-computer interaction INTERACT, INTERACT (2003)
Dragon's Lair (1983 video game). *Dragon's Lair (1983 Video Game)* (n.d.). Retrieved June 20, 2015, from https://en.wikipedia.org/wiki/Dragon%27s_Lair_(1983_video_game)
D. Dufresne, *Fort McMoney*. National Filmboard of Canada (2013). Retrieved from http://www.fortmcmoney.com
Electronic Arts, *Mass Effect [Video Game]* (Electronic Arts, Edmonton, 2007)
Fox Interactive, *The X-Files Game [Video Game]* (Fox Interactive, Los Angeles, 1998)
B. Gaylor, *Do Not Track* (Paris, 2015). Retrieved from https://donottrack-doc.com
D. Geerts, L. De Marez, C. Pauwels, F. Bentley, C. Timmerer (eds.), *TVX '15: proceedings of the ACM international conference on interactive experiences for TV and online video* (ACM, Brussels/New York, 2015)
S.W. Gilroy, J. Porteous, F. Charles, M. Cavazza, E. Soreq, G. Raz, et al., A brain-computer interface to a plan-based narrative, in Proceedings of the 23rd International Joint Conference on Artificial Intelligence, (Beijing, 2013)
J. Herling, *Advanced Real-Time Manipulation of Video Streams* (Springer, Wiesbaden, 2014)
Interplay Entertainment, *Black Dahlia [Video Game]* (Interplay Entertainment, Beverly Hills, 1998)
ID Software, *Doom [Video Game]* (ID Software, 1993)
H. Jenkins, Game design as narrative architecture, in *First Person: New Media as Story, Performance, and* Game, ed. by N. Wardrip-Fruin, P. Harrigan (MIT Press, Cambridge, MA, 2004). Retrieved from http://www.electronicbookreview.com/thread/firstperson/lazzi-fair
Jeppemh, User comment on IMDB.com website (2000). Retrieved June 15, 2015, from http://www.imdb.com/title/tt0229322/usercomments
K. Kirchmann. Umschalten erwünscht? Wenn ja, von wem? Ergebnisse einer Studie zu Ästhetik und Rezeption des ersten interaktiven TV-Spiels des deutschen Fernsehens im Dezember 1991. Arbeitshefte Bildschirmmedien, 48(1994), 23–60
N. Knoller, U. Ben-Arie, The Holodeck is all around us, in *Interactive Digital Narrative*, ed. by H. Koenitz, G. Ferri, M. Haahr, D. Sezen, T.I. Sezen (Routledge, New York, 2015), pp. 51–66

N. Knoller, InterFace portraits: communicative-expressive interaction with a character's mind. Presented at the SRMC '04: proceedings of the 1st ACM workshop on story representation, mechanism and context (ACM Request Permissions, New York, 2004), pp. 63–66. Doi:10.1145/1026633.1026648

N. Knoller, U. Ben-Arie, Turbulence – a user study of a hypernarrative interactive movie, in *Interactive Storytelling: Second Joint International Conference on Interactive Digital Storytelling, ICIDS 2009, Guimarães, Portugal, December 9–11, 2009, Proceedings,* vol. 5915, ed. by I. Iurgel, N. Zagalo, P. Petta (Springer, Berlin, 2009), pp. 44–49. Doi:10.1007/978-3-642-10643-9_8

N. Knoller, The expressive space of IDS-as-Art, in *Interactive Storytelling: 5th International Conference, ICIDS 2012, San Sebastián, Spain, November 12–15, 2012. Proceedings*, ed. by D. Oyarzun, F. Peinado, R.M. Young, A. Elizalde, G. Méndez (Springer, Berlin/Heidelberg, 2012). Doi:10.1007/978-3-642-34851-8_3

H. Koenitz, Towards a specific theory of interactive digital narrative, in *Interactive Digital Narrative*, ed. by H. Koenitz, G. Ferri, M. Haahr, D. Sezen, T.I. Sezen (Routledge, New York, 2015), pp. 91–105

H. Koenitz, Reframing interactive digital narrative. Proquest, UMI Dissertation Publishing, (2010a)

H. Koenitz, Towards a theoretical framework for interactive digital narrative, in *Interactive Storytelling: Third Joint Conference on Interactive Digital Storytelling*, ed. by R. Aylett, M.Y. Lim, S. Louchart (Springer, 2010b), pp. 176–185. Doi:10.1007/978-3-642-16638-9_22

H. Koenitz, T.I. Sezen, D. Sezen, Breaking points – a continuously developing interactive digital narrative, in *Interactive storytelling : 6th International Conference, ICIDS 2013, Istanbul, Turkey, November 6–9, 2013, Proceedings,* vol. 8230, ed. by H. Koenitz, G. Ferri, M. Haar, D. Sezen, T.I. Sezen, G. Çatak (Springer, Cham, 2013), pp. 107–113. Doi:10.1007/978-3-319-02756-2_12

H. Koenitz, S. Louchart, Practicalities and ideologies, (re)-considering the interactive digital narrative authoring paradigm. Presented at the foundation of digital games 2015, pp. 1–4 (2015)

Konami, *Metal Gear Solid 4: Guns of the Patriots [Video Game]* (Konami, Tokio, 2008)

S. Kragh-Jacobsen, K. Levring, T. Vinterberg, von L. Trier, *D-Dag* (2000)

B. Laurel, *Computers as Theatre* (Addison-Wesley, Boston, 1991)

B. Laurel, *Toward the design of a computer-based interactive fantasy system* (1986)

M. Lew, Office Voodoo. Presented at the SIGGRAPH 2003 conference (ACM Press, New York, 2003). Doi:10.1145/965400.965574

LucasArts, *Star Wars: Rebel Assault [Video Game]* (LucasArts, San Francisco, 1993)

A. Mitchell, K. McGee, Reading again for the first time: a model of rereading in interactive stories, in *Interactive Storytelling: 5th International Conference, ICIDS 2012, San Sebastián, Spain, November 12–15, 2012. Proceedings,* ed. by D. Oyarzun, F. Peinado, R.M. Young, A. Elizalde, G. Méndez (Springer, Berlin/Heidelberg, 2012). Doi:10.1007/978-3-642-34851-8.pdf

J.H. Murray, *Hamlet on the Holodeck: The Future of Narrative in Cyberspace* (Free Press, New York, 1997)

J.H. Murray, Transcending transmedia: emerging story telling structures for the emerging convergence platforms. Presented at the EuroITV 2012: 10th European conference on interactive TV and video (ACM, New York, 2012), pp. 1–6. Doi:10.1145/2325616.2325618

M. Naimark, Interactive art–maybe it's a bad idea, in *Cyberarts* (Springer, Vienna/New York, 1998)

A. Nandakumar, J. Murray, Companion apps for long arc TV series: supporting new viewers in complex storyworlds with tightly synchronized context-sensitive annotations. Presented at the TVX 2014 ACM international conference on interactive experiences for TV and online video (ACM, 2014), pp. 3–10. Doi:10.1145/2602299.2602317

K. Nash, Strategies of interaction, questions of meaning: an audience study of the NFBs Bear 71. Stud. Doc. Film **8**(3), 221–234 (2014). doi:10.1080/17503280.2014.958904

E. Pariser, *The Filter Bubble* (Penguin, New York, 2011)

Phantasmagoria, Phantasmagoria (n.d.). Retrieved June 15, 2015, from http://en.wikipedia.org/wiki/Phantasmagoria_(video_game)

PTV Productions, *Inside the Haiti Earthquake* (2010). Retrieved from http://www.insidedisaster.com/experience/

Quantic Dream, *Heavy Rain [Video Game]* (Sony Computer Entertainment, Tokyo, 2010)

Quantic Dream, *Beyond: Two Souls [Video Game]* (Sony Computer Entertainment, Tokyo, 2013)

M.-L. Ryan, Multivariant narratives, in *A Companion to Digital Humanities*, ed. by S. Schreibman, R. Siemens, J. Unsworth (A Companion to Digital Humanities, Malden, 2004), pp. 415–430. doi:10.1002/9780470999875.ch28

Sierra On-Line, *Phantasmagoria [Video Game]* (Sierra On-Line, Los Angeles, 1995)

Simon & Schuster, *Star Trek: Borg [Video Game]* (Simon & Schuster, New York, 1996)

D. Simpson, Lucasfilm debuts real-time rendering of "star wars" scenes. *The Hollywood Reporter* (2013, September 22). Retrieved from http://www.hollywoodreporter.com/video/lucasfilm-debuts-real-time-rendering-635458

A. Stern, Embracing the combinatorial explosion: a brief prescription for interactive story R&D, in *Embracing the Combinatorial Explosion: A Brief Prescription for Interactive Story R&D* (2008), pp. 1–5. Doi:10.1007/978-3-540-89454-4_1

Tale of Tales, *The Graveyard [Video Game]* (Tale of Tales, 2008). Retrieved from http://tale-of-tales.com/TheGraveyard/

C.S.S. Tan, J. Van den Bergh, J. Schöning, K. Luyten, Towards detection of side activities and emotions of anonymous TV viewers through body postures. Presented at the TVX2014. (2014). Doi:10.6084/m9.figshare.1032592

Telltale Games, *The Walking Dead [Video Game]* (Telltale Games, San Rafael, 2012)

The Fullbright Company, *Gone Home [Video Game]* (The Fullbright Company, Portland, 2013)

P. Tikka, *Enactive Cinema: Simulatorium Eisensteinense* (2008)

M.F. Ursu, I.C. Kegel, D. Williams, M. Thomas, H. Mayer, V. Zsombori et al., ShapeShifting TV: interactive screen media narratives. Multimedia Syst. **14**(2), 115–132 (2008a). doi:10.1007/s00530-008-0119-z

M.F. Ursu, M. Thomas, I. Kegel, D. Williams, M. Tuomola, I. Lindstedt et al., Interactive TV narratives: opportunities, progress, and challenges. ACM Trans. Multimedia Comput. Commun. Appl. **4**(4), 25:1–25:39 (2008b). doi:10.1145/1412196.1412198

J. Van den Bergh, K. Luyten, M. Matton, K. Willaert, 1st workshop on empathic television experiences (EmpaTeX 2014). Presented at the TVX2014 adjunct proceedings, (2014). Doi:10.6084/m9.figshare.1031603

B. Weiberg, *Beyond Interactive Cinema* (2002). Retrieved August 19, 2015, from http://keyframe.org/txt/interact/

J. Weizenbaum, Eliza – a computer program for the study of natural language communication between man and machine. Commun. ACM **9**(1), 36–45 (1966). doi:10.1145/365153.365168

Place and ITV: Playful Design Strategies Towards Place-Oriented ITV

43

Gabriele Ferri

Contents

Abstract

Providing users with a sense of place – related to a specific geographic location in which one is situated, or linked to a faraway place, or even giving place-like qualities to virtual spaces such as massively multiplayer online role-playing games – has been deemed central for several forms of digital interactions. In the past decade, studies from human-computer interaction and computer-supported cooperative work have specifically addressed this theme, but the scarcity of works of place specificity focusing expressly on interactive TV suggests a gap in the current research, whereas the latest developments in mobile TV would seem highly coherent with such topic. To

G. Ferri (✉)
School of Informatics and Computing, Indiana University, Bloomington, IN, USA
e-mail: gabriele.ferri@gmail.com

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_46

contribute to closing this gap, some initial directions are suggested here by pointing at compatible treatments of the notion of place in related fields, for example, the design of pervasive urban games. Game designers and game scholars might provide operational concepts that help understanding the role and the potentialities of places for interactive TV. Two general types of artifacts are selected here: works that are anchored to the experience of faraway places and works that leverage the physical location in which the user is. Their analysis yields three design strategies (experience anchoring, place permeability, and distributed storytelling), offered here as "objects to think with" and to spur further research and design. By pointing at them and at other similar strategies, similarities between digital games, ITV products, and other similar artifacts emerge and allow us to speculatively trace possible future convergences.

Keywords
Interactive television • Place • Mobile platforms • Pervasive games • Urban games

Introduction

The already multifaceted relationship between interactive television (ITV) systems, content producers, and audience/users has become even more complex with the increasingly pervasive diffusion of handheld devices and mobile TV solutions. However, ITV research so far has touched only tangentially the notions of place and place specificity, whereas other artifacts – from location-aware apps to mobile games – have already experimented with nuanced approaches to provide users with different kinds of "sense of place." The concept of "place" is a notion that humanistic disciplines such as ethnography and geography share with more technical ones such as human-computer interaction (HCI), ubiquitous computing (UbiComp), and game studies. By examining recent developments in the cognate fields of mobile, pervasive, and urban game design, ITV researchers may gather useful insights and design implications towards creating more engaging, effective, and expressive place-oriented ITV experiences.

Current practical concepts from game design will be examined to propose innovative, technologically mediated, and playful experiences fostering a connection between users and physical places. The relationship between human-computer interaction, meaning-making, physical places, and narrative contents is indeed complex and has been discussed from a variety of perspectives: here, practical exemplars will be framed first within a general and theoretical background in semiotics and then through practical insights drawn from the design of pervasive and urban games. It will be argued that thoughtfully designed locative artifacts may articulate the relationship between narrative structures and place specificity in nuanced and effective ways, which may inform further ITV research.

Among the many possible place-related ITV applications, this contribution will address specifically two of them – without necessarily arguing for their exhaustiveness – with the objective of teasing out further design insights vis-à-vis novel ITV solutions. First, interactive television may convey a "sense of place" to remote users, allowing for mediated explorations of faraway locations. Moreover, ITV might embrace mobile technologies and aim at deeper interactions with physical places and the practices that happen in those contexts. To exemplify the first group, three advertisement games (or *advergames*) will be presented, demonstrating how the common practice of accessing online maps may be made ludic. From their analysis, the design strategy of experience anchoring will be generalized and presented. The second type of place-related application will yield two other strategies, entitled place permeability and distributed storytelling. To discuss the first one, three pervasive games that require players to access place-specific knowledge will be discussed; to address the second one, two urban games that task users with producing place-specific short video contents will be presented. Whereas most of the selected artifacts discussed here are not ITV products per se, some of their key characteristics will be teased out and offered to ITV researchers as inspiration towards more refined, location-based, playful experiences.

In sum, this contribution aims at moving an early step towards an interdisciplinary dialogue about place, interactive technology, playfulness, and design between ITV, HCI, UbiComp, and urban game studies, by offering generative tools and "objects to think with" to ITV practitioners and scholars.

ITV Between Narrativity, Playfulness, and Place Specificity

Advances in interactive TV technologies have opened new scenarios for innovative narrative forms bridging the gap between storytelling, interactivity, and connection with physical places: this section will briefly survey the state of the art on these elements and their relationship. "Mobile TV is still in its infancy" (Sørensen and Nicolajsen 2010), but nonetheless, it constitutes a promising entry point from which to examine the relationship between interactive TV, storytelling, and physical places. Outside the domain of ITV, the connection between places and technosocial practices has been examined from a variety of perspectives. In a foundational paper on place in human-computer interaction and computer-supported cooperative work (CSCW), Harrison and Dourish argue that "Space is the structure of the world, [however] appropriate behavioral framing comes not from a sense of *space*, but from a sense of *place*, [and a] key principle describes the relationship between the two is: space is the opportunity; place is the understood reality" (Harrison and Dourish 1996). From the perspective of data mining and information architecture, Li and Goodchild remark "[place] reflects the way people perceive, understand, and interact with their environment [and] place names are omnipresent in conversations, documents, and news stories" (Li and Goodchild 2012). Digital technologies for telepresence have been experimented with in order to provide a "sense of place" for users: for example, Turner and Turner (2006) report on a telepresence application allowing a virtual tourist visit of the city of Edinburgh, Champion (2008) and Sylaiou et al. (2010) mention virtual heritage and virtual museums applications, and Deaton et al. (2005) refer to location-specific training environments. However, the number of ITV artifacts engaging places is scarce, highlighting a visible gap in the current research on interactive TV and mobile ITV: while a number of apps and mobile experiences – from commercial products such as Foursquare/Swarm, to games such as Ingress, to more art-oriented projects – have successfully addressed physical locations in a playful way, ITV seems to be still lagging behind in this respect.

In addition to this, playfulness seems to be another notion that has been only sparsely addressed by ITV researchers. "Traditionally, television used to be about narrativity, but not about agency and intercommunication; [. . .] games used to be about agency, but not about narrativity and intercommunication; and [. . .] the Web used to be about intercommunication, but not about narrativity and agency" (Ursu et al. 2008). However, as Murray (2012) recently underlined, digital TV is already "delivered on the same screen and often through the same device that provides access to videogames and websites, [and] we can see the possibility of a future environment of true convergence in which everything we can do on a computer or a game console can also be done in the context of television viewing" (2012). For this reason, a significant convergence is taking place between different digital media (Jenkins 2006), but designers today also have the possibility of experimenting with innovative forms of interactivity that bridge linear and ludic forms of storytelling. Considering two possible scenarios – a "static media model" that assumes the continued separation between TV and game platforms and a "digital revolution model" that hypothesizes the disappearance of noninteractive television formats – Murray points at a middle ground on which to experiment with design innovations. In the same vein, the relationship between play, narrative, and place may be outlined as follows. Historically, screen-based media such as cinema and television had a profound impact on the practices of constructing, telling, and interpreting narrative texts. Whereas more traditional TV storytelling lacked interactivity and early exemplars such as the *kinoautomat* (Hales 2015) were technologically limited, other experiments on the technological and artistic sides of TV narratives are currently exploring flexible and multilinear formats. Currently, the two most common solutions for implementing a degree of interactivity in TV-related experiences rely on additional services complementing traditional shows (Nandakumar and Murray 2014; Murray et al. 2012) or on multilinear narratives created constructing a branching tree with short audiovisual segments (Piacenza et al. 2011; Charles et al. 2011; Réty et al. 2008). In addition to these, other experiments on the convergence between TV and participatory and social media have been recently proposed (Hess et al. 2011; Almeida et al. 2012). In the first case, current examples of interactive services enhancing TV shows include the Info Card service (Google 2013) available for contents accessed through the Google Play service, which allow user to pause video streams and have access to information about the actors present on screen. Multilinear branching narratives, instead, date back to the foundational Dragon's Lair (1983) laserdisc video game and still constitute a cornerstone of TV interactive storytelling and cognate media. As an example of this, the upcoming interactive narrative “Her Story” (Barlow 2015) makes use of multiple video clips to cast players in the role of a detective interrogating a suspect.

Digital Narrativity

In the wake of founding metaphors such as Janet Murray’s holodeck (Murray 1997) or Brenda Laurel’s computers as theatrical performances (1991), the practice of computer game design has significantly evolved in the past decades and produced outstanding interactive artifacts that merge playful competition with nuanced storytelling structures. On this topic, the compatibility between narrative and competitive components has been discussed at length in Game Studies: indeed, a lively debate on the role (central or ancillary) played by narrativity in games took place in the last decade, and this is not the occasion for revisiting it. Marie-Laure Ryan (2001, 2004, 2006; Jannidis 2003) argues for understanding the quality of being a narrative (“storiness”) as fuzzy and scalar rather than as a dichotomy between stories and nonstories as binary categories. Indeed, reasoning in terms of storiness allows researchers to appreciate the convergence between digital games and TV-like experiences. A number of highly immersive virtual environments have been recently published by the entertainment industry and acclaimed by digital media critics – in agreement with the vision of playable media as a powerful medium for storytelling. Recent AAA titles such as Heavy Rain (2010), L. A. Noire (2011), and The Last of Us (2013) exemplify recent progresses in this sense. However, it is not the intention here to argue for a transition of games as a whole towards experienced inspired by TV storytelling. Rather than propose general assumptions, this chapter – coherently with the interest in interactive narratives that traverses ITV research – will programmatically focus on playful experiences that foreground storytelling components. As noted elsewhere, “[Ryan’s] revised definition considers narrative to be every semiotic object—verbal or not—capable of evoking specific cognitive effects in its interpreters’ minds. In other words, Ryan is proposing to separate linguistic phenomena from the quality of being a narrative, understood as a sort of cognitive template at work in many different media, including [interactive systems]” (Koenitz et al. 2015). The narrative traits that characterize most ITV experiences are indeed pointing in this direction, and, for this reason, the curated list of exemplars assembled here is composed of artifacts favoring a high level of scalar storiness over competition.

An approach grounded in semiotics will provide useful theoretical scaffolding to examine narrative structures in interactive experiences. In the most general terms, semiotics refers to a systematic study of signs, their possible uses, their classification, and their role in social contexts. Eco (1976) distinguishes general and specific semiotics: the former is a more philosophical approach concerned with the emergence of meaning, whereas the second describes the organization of particular systems such as linguistics, narratives, proxemics, or iconography.

As an analytical category, the concept of narrativity belongs to the specific semiotics that examines the formal characteristics of storytelling. It constitutes the deepest, most abstract identifiable level of any text and the common layer for any meaningful artifact regardless of the medium adopted. In this sense, narrativity is the deepest, most general, and abstract identifiable level of any text and the common layer for any meaningful artifact regardless of the medium adopted. As stated elsewhere, "[it] should be intended as the logical baseline of every form of expression, and it can be described in highly abstract terms [and] does not refer to 'having a narrative' or 'being a narrative' in the ordinary or literary sense of the term, but it is defined as the quality of every text to be formulated as a network of semantic oppositions and of actantial roles that change over time following a canonical schema" (Ferri 2015). If one adopts such a broad logical definition, as Compagno and Coppock remark,

> every meaningful artifact or activity is then narrative in this abstract theoretical sense, and all cultural productions specify the way in which they determine how an interpreter is able to understand and respond to them (thus integrating these interpretations into his/her prior cultural knowledge base). If we agree on this notion of narrativity, then computer games cannot but be narrative. (Compagno and Coppock 2009)

Pervasive Gaming

With the popular diffusion of GPS-enabled mobile terminals, playable and narrative artifacts have rapidly expanded – as they are no more bound to specific times and places but can be meaningfully experienced in a broad variety of physical locations. This has led to the development of urban games and other locative media, producing compelling user experiences leveraging playfulness across a variety of technosocial practices. The use of digital technologies to affect the way spaces and places are inhabited and interacted with is clearly not new. In particular, from the 1990s onward, a growing number of place-specific artistic and playful projects have explored the relationship between technologically mediated practices and the sociocultural contexts of public spaces. Among different influences on this topic, Mark Weiser's early vision of ubiquitous and pervasive computing is particularly relevant there: "we are trying to conceive a new way of thinking about computers in the world, one that takes into account the natural human environment and allows computers themselves to vanish into the background" (Weiser 1991). As artist and media theorist Martin Rieser recently argued: "[mobile media] dissolves traditional gallery and museum walls, and has allowed new audiences to relate to the spaces of their urban worlds by turning them both into places of possibility, where inner and outer spaces, histories and narratives can be interlocked and explored" (Rieser 2015). Similarly, pervasive and urban play "exists in the intersection of phenomena such as city culture, mobile technology, network communication, reality fiction, and performing arts, combining bits and pieces from various contexts to produce new play experiences" (Montola et al. 2009). In the past decade, UK-based mixed-media artist group Blast Theory has been at the forefront of the design research related to locative, hybrid playful experiences. In collaboration with the Mixed Reality Laboratory at the University of Nottingham, Blast Theory developed seminal pieces such as Can You See Me Now (2003), where online players are chased through a virtual model by street players who play in a real urban environment, and Uncle Roy All Around You (2005), where street players were given a task to search for a character named Uncle Roy and remote players, together with professional performers, guided them in their quest. Similar experiences investigated the use of large public displays. For example, Nokia research's Manhattan Story Mashup (2007) combined online contents, mobile phones, and one of the world's largest public displays in Times Square to a large-scale pervasive game in Midtown Manhattan. As described by Tuulos et al. (2007),

> web players used [. . .] mashed up stories, either by writing new sentences or by re-using already illustrated sentences. A noun from each new sentence was sent to a street player [who] had to shoot a photo which depicted the word within 90 s. The photo was then sent to two other street players who had two guess what the photo depicts amongst four nouns, including the correct one. If the photo-noun pair was guessed correctly, the original sentence was illustrated with the new photo and it turned into an ingredient for new stories [the best of which] were shown on the Reuters Sign in Times Square. (Tuulos et al. 2007)

Indeed, pervasive technologies have been successfully adopted not only in the fields of interactive and public arts but also in the emerging field of urban game design. In other words, pervasive digital play makes use of interactive technologies to offer experiences that occur outside the spaces and times that would normally be dedicated to it and, by doing so, affects the ways we perceive and understand physical places. Similarly, location-based games constitute an emerging genre that brings together mobile computing and game design in order to engage players through the interaction with their surroundings. For the purposes of this chapter, these related genres will be gathered under the umbrella term of "pervasive games."

The idea of pervasive gaming has been frequently discussed in HCI literature in connection to mobile experiences, with Nieuwdorp (2007) compiling a thorough review of the related literature. While pointing at the heterogeneous and problematic use of the notion of pervasiveness, she also provides an operational definition: "Pervasive games are generally those that blur the boundaries between the game world and the real world, and where the game blends in with reality" (Nieuwdorp 2007). These practices link the virtual world to the real one: locative games leverage portable devices to set up place-specific situations for users to experience, for example, asking players to be at specific locations at certain times of the day. This chapter will refer specifically to a subset of Nieuwdorp's categorization, composed of three complementary perspectives on pervasiveness in respect to games:

1. Practices that *pervade* and *blend* with the physical world (De Souza e Silva 2009; De Souza e Silva and Sutko 2009),

2. Practices that *overlay* games with the physical world (Sotamaa 2002),
3. Practices that *intertwine* games with everyday practices (Stenros et al. 2012; Bichard and Waern 2008).

Addressing the Place-Specificity Gap in ITV Research

Analytical categories that are related to the *storiness* (Ryan 2004) of digital artifacts and that pay specific attention to the relationship between playfulness, interaction, and physical places may provide ITV researchers with useful insights for developing locative experiences. Without being in contrast with other more quantitative methodologies, this interdisciplinary approach contributes to the field of ITV research and practice by providing categories that will allow more effective designs and focused comparisons between heterogeneous practices. As already discussed in depth elsewhere (Ferri 2014), the research contributions that are made possible by adopting methodologies that are informed by semiotics are programmatically *abstract*, *technologically agnostic*, *scalable*, and *generative*.

In the context of ITV research, these qualities may be further specified. Providing logical categories means keeping a close focus on deep semantic structures underlying the analyzed artifact rather than its specific figurative qualities. For design practitioners, this means generalizing beyond particular examples and finding more general similarities or differences across wide corpora. Technologically agnostic approaches allow researchers to compare experiences across different media, leading to a more nuanced understanding of pervasiveness in mobile practices. Focusing on scalable approaches makes possible the description of simple or complex activities regardless of their size or of the number of interactors (Murray 1997) taking part in them. The design strategies of experience anchoring, place permeability, and distributed storytelling aim at being generative concepts, capable of providing designers with useful insights: semiotic categories, while directly intended as design tools, can provide inspiration for practitioners. By definition, semiotic analyses are transmedia (Ryan 2004) and, thus, easier to generalize and adopt for comparing artifacts and experiences across different domains (e.g., ITV products vs. everyday practices or vs. unilinear television). This aims at three positive effects:

1. A more detailed understanding of the internal boundaries in the field, distinguishing different types of place specificity in ITV artifacts
2. More effective comparisons with other artifacts or experiences, made possible by adopting shared, general descriptive categories
3. Reinforcing an interdisciplinary dialogue with a common metalanguage to facilitate the selection of best practices and shared examples

Introducing a curated selection of playful and locative interactive experiences, this chapter aims at informing and complementing other more prescriptive design methodologies.

Digital Maps: Anchoring Playful Storytelling to Existing Online Services

In recent years, the impact of online mapping services on the general digital media ecosystem has grown considerably, and accessing geographical data through digital media has indeed become a commonplace HCI experience. Their ubiquitousness and recognizability make online maps likely candidates for design experiments that link narration, sense of place, and digital interaction with ITV systems. For these reasons, they constitute good examples to introduce the design strategy of experience anchoring.

Although Street View, currently part of Google Maps and Google Earth, was not the first online service to present first-person geolocalized images, it is today one of the most recognizable. It was presented to the public in May 2007 – first with content from select North American cities and the gradually expanding abroad – and is currently available on the vast majority of digital platforms, from desktops to mobiles. Street View produces an almost three-dimensional vision by putting into sequence a great number of bidimensional panoramas that were captured through a specialized tool – often mounted on top of vehicles. Subsequently, a semitransparent white line is superimposed to each image to indicate the trajectories on which users can "move," for example, showing a bifurcation near crossings and intersections to indicate the different choices available. As exemplified by a recent playful reinterpretation (Kohler 2015) that temporarily added the option of turning real street topographies from Google Maps in game levels for a Pac-Man (Namco 1980) clone, a number of game-related uses of online maps already exist.

A general overview of three exemplars – Mercedes-Benz: Escape the Map (Abbott Mead Vickers BBDO 2011), Unilever: Magnum Pleasure Hunt 2 (Lowe Brindfors 2012), and State Farm: Chaos in Your Town (DDB B-Reel 2011) – will be presented, followed by insights for ITV researchers and practitioners. By curating a short list of casual games that reinterpret and remediate online map services into playful user experiences, this section argues that these are not simply leveraging superficial figurative aspects of those service, but are rather ways to *make playful* the everyday experience of "remote places." In other words, they anchor playful digital contents both into physical places and into their mediated versions: by building upon an ordinary interaction such as accessing a digital map, the examples detailed here create playful experiences while remaining clearly connected to places through online cartography. The works considered are also "advergames" or "advertisement games" (Smith and Just 2009; Bogost 2007), a hybrid interactive digital genre elaborating on traditional ads to add interactivity and playfulness. Advergames are seldom considered in the academic discourse as their design quality is often lacking polish – however, as audiovisual advertisement constitutes a staple in traditional TV contents, examining advergames is here a fitting choice for pointing at some possible future characteristics in ITV. In sum, it could be speculated that ITV designers may follow parallel trajectories – possibly on different subject matters – to those adopted in creating these advergames by making playful an already-recognizable interactive experience.

Exemplars: Sport Cars, Ice Creams, and Killer Robots

In November 2011, Mercedes-Benz UK launched an advertising campaign entitled Escape the Map (Abbott Mead Vickers BBDO 2011), promoting its C63 AMG car model. Escape the Map leverages a multiplatform storytelling ecosystem comprising – in addition to the main escapethemap.com website – diegetic social media accounts, physical billboards on mounted on public buses in London, banner ads on the YouTube homepage, plus augmented-reality elements published on the Metro free-press newspaper and accessible through a specific app. The main part of this experience consists of a short and simple game, an interactive movie whose plot branches or arrests according to user input. The plot opens with Marie – the coprotagonist – introducing herself to the player. She describes herself as being "stuck in the Map," a virtual world whose visual characteristics are clearly reminiscent of the Hong Kong area of Street View. Those who spend too much time in the Map have their facial features covered by a blurring effect and, Marie says, she must escape before her face changes permanently. The goal of the game is to navigate through the branching narrative, driving a Mercedes car towards an exit; the difficulty level is – as it is usual for most advergames – very low, possibly to maximize the reach of the promotional message. Escape the Map is structured as a single audiovisual clip with a small number of branching points, such as choosing a path or another or entering a street address in a navigation bar. Many visual components in the game refer to the Street View aesthetics, with UI elements entering the narrative diegesis as if they were "real" parts of the game reality (e.g., giant map markers falling from the sky and threatening to crush the protagonists), or other visual effects simulating the low-resolution images resulting from a slow Internet connection (Fig. 1a, b).

Magnum Pleasure Hunt 2 – Around the World (Lowe Brindfors 2012) is a comarketing advergame promoting Unilever's ice cream brand Magnum together with other partners such as Bvlgari, Quiksilver, KLM Royal Dutch Airlines, Hotel Fasano, and Microsoft Bing. It is also coordinated with a corresponding TV advertisement campaign in the USA. In short, the game is a simple platformer where players guide an avatar collecting as many Magnum-branded items along bidimensional and three-dimensional levels. The action shows the protagonist of the game moving across a number of places – from New York, to Paris, to Rio de Janeiro. In her movements, the online map services provided by Bing Maps and Bing Streetside function as links between subsequent game levels: for example, we see the protagonist typing the name of a shop in an apparently extradiegetic navigation bar to teleport inside, as if changing the point of view in Bing Streetside corresponded to a physical movement inside the game world (Fig. 2a, b).

Chaos in Your Town (DDB B-Reel 2011) is a playful app advertising the State Farm insurance company and that may be used to create customized video clips to be shared through e-mail or social networking sites. It adopts the narrative and visual elements already explored in a series of TV advertisements produced for the same brand, where two characters chat with great calm as a giant robot destroys

Fig. 1 (**a**, **b**) Escape the map (Abbott Mead Vickers BBDO 2011)

their neighborhood. Chaos in Your Town asks the user for a street address and – if any Google Street View imagery is available for the corresponding coordinates – it automatically generates a short video clip where the destroyer robot from the TV ad is shown walking nearby streets as it knocks buildings down (Fig. 3a, b).

Design Insights: Towards the Experience Anchoring Strategy

The three digital media formats – online map services, advertisement, and video games – that are brought together in these examples interact with each other at a

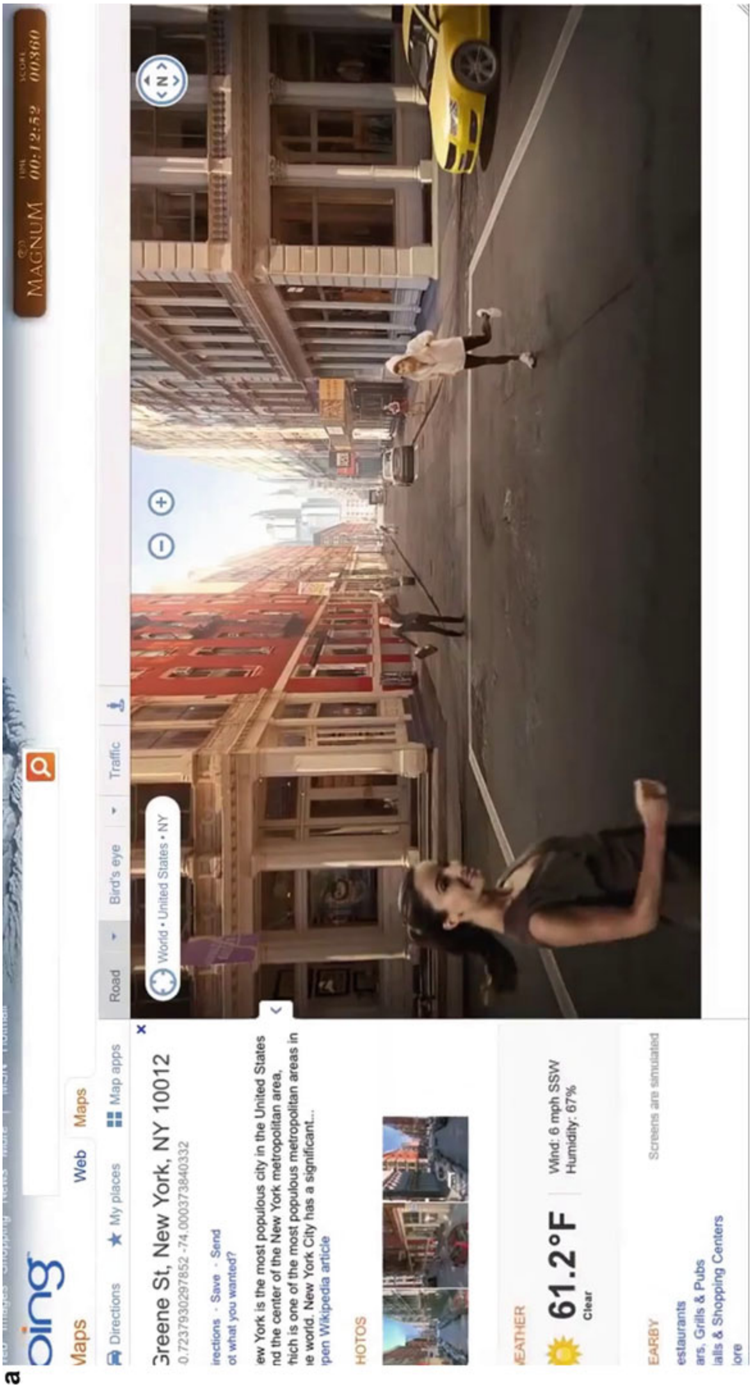


Fig. 2 (continued)

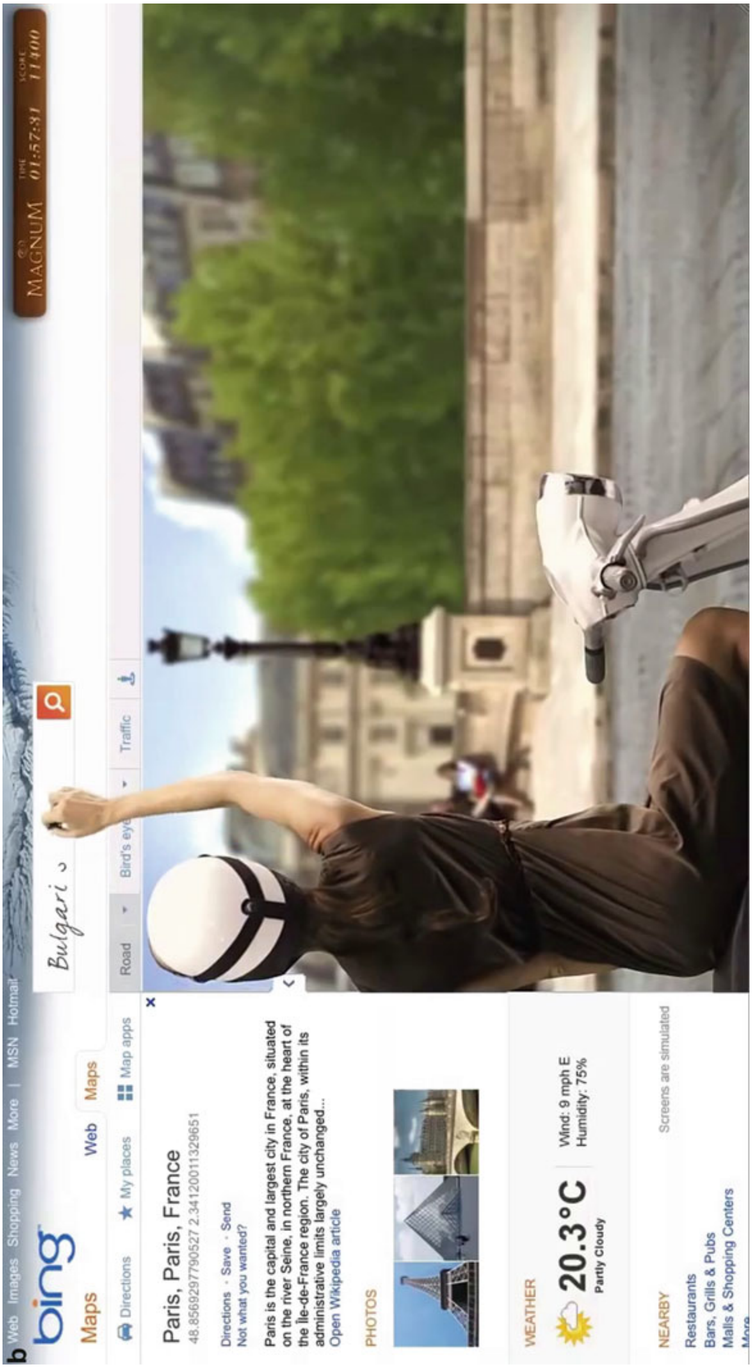


Fig. 2 (**a**, **b**) Magnum Pleasure Hunt 2 – around the world (Lowe Brindfors 2012)

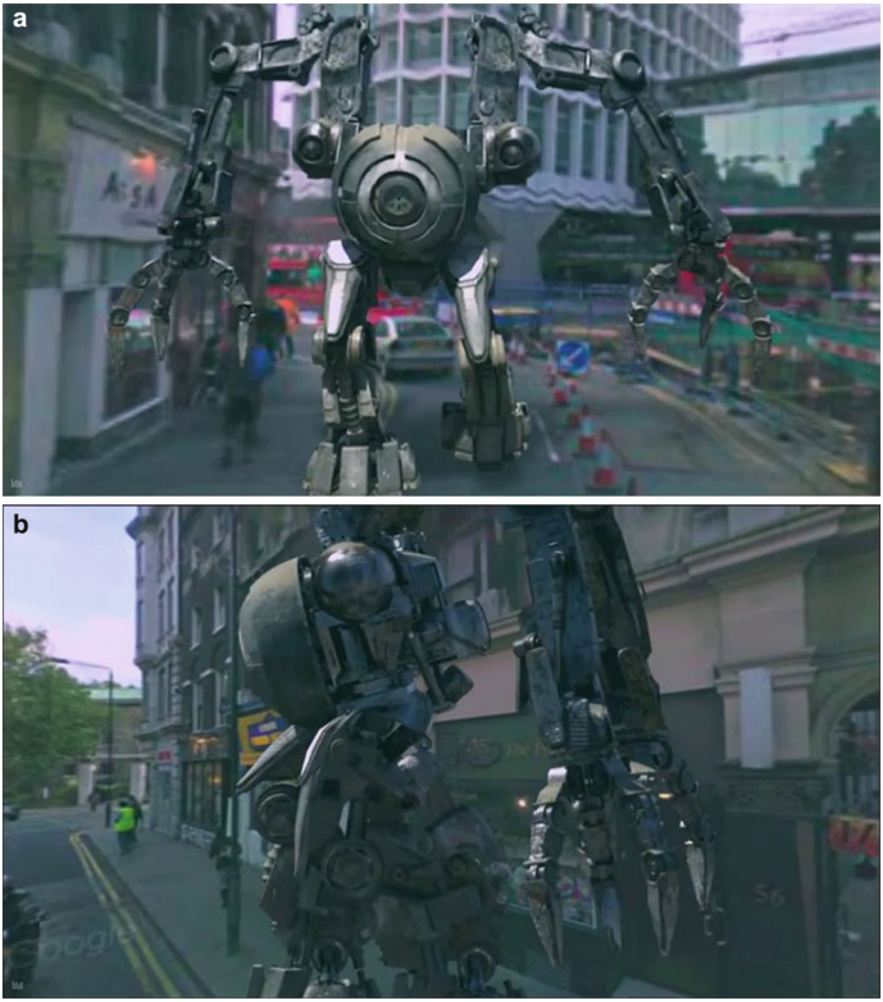


Fig. 3 (**a**, **b**) Chaos in your town (DDB B-Reel 2011)

different levels and, together, they transpose ordinary TV advertisement in digital interactive form. In other words, these exemplars point at complex phenomena (the interplay between linear texts, digital interactive narrativity, the professional context of TV advertising, and the user practice of accessing place-specific information through online maps) and make them playful by leveraging some game design elements. This conclusive paragraph will synthesize some design insights from a comparison and a brief close reading of the three artifacts, beginning from their figurative components to contextual and procedural elements and finally combining them into more general design implications.

All the three examples presented here adopt visual components that are typical of digital cartography, ranging from direct transfers from the content space of online maps, to the game diegesis (e.g., the blurred faces in Escape the Map that refer to the corresponding privacy feature in Google services but recontextualize it in a completely different narrations), to others that cross from interfaces to the space of playful interactions (e.g., the protagonist reaching and writing in an extradiegetic search box in Pleasure Hunt 2 or giant pointers falling from the sky and blocking the player's path in Escape the Map). In addition to this, other experiential elements are transposed into these exemplars. For example, the way in which Street View captures and digitally composes sequential images make human figures unnaturally still when observed by a user – as if one was blocked in the middle of a step. This figurative trait clearly derives from a technological limitation but has become iconic in how users perceive and expect online maps. Pleasure Hunt 2 expressly plays with such feature, depicting three-dimensional spaces populated by immobile passersby, around which the protagonist moves freely. From the point of view of users' contextual knowledge, specific, real-world references are presented within the games' fiction – from the streets of Hong Kong, Paris, and Rio De Janeiro to the custom street addresses that users are invited to enter in Chaos in Your Town. In other words, the gameplay practice of these artifacts is constituted by two consecutive remediations: a firsthand, physical experience of being in a place is transposed in a virtual stroll on Street View, and subsequently the experience of accessing online maps is "made ludic" as it is turned into a video game.

Indeed, the pieces discussed here base their expressive strategies on a partial verisimilitude, not only for their reference to actual geographical places in the real world but also because they prompt users to experience them using the same interface and technology that one would normally choose for accessing an everyday online mapping service. The points of contact between these games and Street View are indeed laden with multiple meanings. On the one hand, some of the selected locations anchor gameplay to specific places with exotic connotations (Paris, Rio de Janeiro, Hong Kong) that refer to socially shared narratives on the themes of luxury, pleasure, and distance/otherness. The third artifact, Chaos in Your Town, prompts users to enter their own street address in the very beginning of the experience and constructs a dynamic representation in CGI, superimposing a destroyer robot to the images of streets that are familiar for the user. In sum, the gameplay experiences generated with these artifacts remediate an everyday practice (accessing an online map) and embed in a second-order activity (experiencing a playful interaction). By doing so, they leverage a dual-design strategy with its bases – for the first two examples – on a reference on distant, exotic, and evocative places, whereas Chaos in Your Town reverses such approach by presenting familiar, almost mundane, views.

In sum, the design strategy at play in these artifacts can be described as a form of procedural translation (Bogost 2006) that identifies traits from a source domain (a web-based technological platform, visual and interface elements from digital maps, the experience of immediate access, and movement between faraway places) and transposes them into a simple game. As advertisement products, elements of the promoted brand are clearly visible – although their analysis is not the focus at the moment – and are integrated in a more general playful experience. The three artifacts refer to a practice well known to the user – the use of online maps – to contextualize and anchor the gameplay in a daily experience. Branded elements are situated in a self-contained narrative that increases the storiness (Ryan 2004) of mundane online maps, leaving some freedom of action to the user without obscuring the brand identity and connecting it to other significant components (a sense of luxury and exotic for Magnum, a classic "damsel in distress" plot for Mercedes, and the idea of safety in spite of unexpected destruction for State Farm). What is unfortunately rather disappointing, for pieces with a relatively high production value, is that properly ludic components are left in the background.

All said, as design insights to offer to ITV research, these pieces foreground:

- Playful activities anchored to concrete everyday media user practices and related technological platforms
- A thematic, figurative, and procedural connection to actual geographic places and to the related activity of browsing online maps
- A narrative layer intertwined with an interactive experience

From Place Permeability to Distributed Storytelling

As second part of this curated list, a complementary type of artifacts exploring the interplay between narrative structures and social, urban spaces will be discussed. Two live-action role-playing games (LARPs) and three pervasive urban games will be presented as prototypical exemplars to demonstrate two key notions – "TIAG/TINAG ambiguity" (Ferri 2009) and "distributed storytelling" (Ferri and Coppock 2013) – used here to support narrative experiences that are partially shared between a fictional diegesis and the actual, social context. In brief, the ambiguity between experiential frames (TIAG, or *this is a game*, opposed to TINAG, or *this is not a game*) is constructed by providing players with elements supporting partially contradictory interpretations that let different contextual rules overlap and mix fictional game elements with elements from the real world. This design characteristic is not exclusive to LARPs and pervasive games as such, but, as it will be shown, it is crucial in enhancing their contextual narrative effectiveness. Moreover, a sense of distributed storytelling is generated by delegating to participants a degree of authorial or curatorial control over some narrative components of a pervasive experience. Typically, this is obtained by assigning game-related tasks that require players to gather and assemble narrative materials and to reinterpret them in order to proceed with the playful experience. The relevance of these mechanics for ITV researchers and designers is twofold. Whereas the map-inspired pieces presented in the previous section contributed to convey a sense of place for remote users, allowing mediated explorations of faraway locations, as a complementary approach ITV might further embrace mobile technologies and aim at deeper interactions with physical places and related contextual practices. In this vein, these examples demonstrate ludic and narrative place-specific activities that could inform future mobile ITV applications exploring physical spaces, leveraging GPS-enabled mobile terminals. Additionally, the pieces presented here rely significantly in their gameplay parts on user-generated contents – a characteristic that is already relevant in a variety of interactive digital media and that could be further integrated in mobile ITV experiments.

This second part of the chapter will be structured as follows. At first, a theoretical framing for the TIAG/TINAG ambiguity will be provided, followed by the discussion of three examples – an urban game experience and two live-action role-playing games: Big Urban Game (Lantz and Salen 2002), System Danmarc (Opus 2005), and Prosopopeia (IperG 2005). These three examples will be used to illustrate the design strategy of place permeability. Then the other strategy of distributed storytelling will be introduced and exemplified with two urban game examples: Mettiti in Gioco (Ferri 2011) and 3Cities (Salvador and Ferri 2013).

"This Is a Game" or "This Is Not a Game"?

As pointed out in the works of Huizinga, Goffmann, and other more contemporary game scholars (Huizinga 1938; Goffman 1961; Copier 2005; Stenros 2012), players of traditional, nonpervasive games are generally well aware of the ludicity of the activities they take part to. Such awareness has been address with a number of related concepts – such as the Magic Circle or the Lusory Attitude (Huizinga 1938; Suits 1978) – and generally defines the participants' preliminary acknowledgements of being in a playful situation. In this occasion, and building upon previous research on this topic (Ferri 2009; Ferri and Coppock 2013), this more general system of player expectations will be called the *this is a game* (TIAG) layer. More specifically, the cooperation between players and interactive systems – digital or nondigital – generates a ludic discursive universe within the TIAG layer in which gaming interactions are recognized as fictional. In simpler terms, we might understand it as the make-believe part of a game, in which a fictional world with temporary rules is established. As focalization (Genette 1972) shifts within this universe, users abide by TIAG interpretive rules and temporarily complement their ordinary expectations with a general tendency to accept fictional contents. In other words, this follows a very similar pattern as when interpreting other fictional texts – from novels, to movies, to computer games. However, some specific experiences also feature a second nested system of expectations within the TIAG layer – and pervasive games are prominent in their relying on this feature. As a pragmatic effect on users' interpretive activities, this second intradiegetic frame requires players to "pretend to believe that what is taking place is not a game"; for this reason, this layer will be designates as *this is not a game* (TINAG). The majority of pervasive and urban games relies on semiotic and procedural mechanisms that generate player engagement by creating ambiguity between the TIAG and TINAG systems. The following exemplars follow the evolution of this semiotic strategy from urban games to more complex pervasive live-action role-playing games.

Exemplars: Inflatable Game Pieces, a Danish Dystopia, and the Spirit World

The Big Urban Game (Lantz and Salen 2002), also called BUG, was situated at the intersection between urban gaming and public art. It was an event organized by the Design Institute of the University of Minnesota as a part of its Twin Cities Design Celebration, with the goal of encouraging residents of Minneapolis and St. Paul to see their surroundings in a whole new way and to support reflections about the design of urban space. Stakenas and Zurkov (2006) summarize the event as follows: “The BUG took place over five days in Minneapolis and St. Paul. Three teams raced 25-ft tall inflatable game pieces through a series of five checkpoints, hoping to make it to the final destination in the shortest amount of time.” Big Urban Game was divided into a real-world physical game and its online counterpart. Designer Jane McGonigal describes its two parallel worlds (Fig. 4):

> Three thousand, three hundred and six members of the public registered to play [..] online and were divided into three teams: red, yellow, and blue. Each [..] online team partnered with a dozen real-world runners, who would be responsible for moving their team’s [..] inflatable game piece around a 108 square-mile game board. Every morning for five consecutive days, the online players studied a digital map of the Twin Cities and voted for one of two potential racing paths. Every evening, after the votes were counted, the real-world runners raced through the city streets following the route chosen by their online counterparts. (McGonigal 2006, pp. 175–176)

This specific example bases its ludic and rhetoric effectiveness on exporting, or contaminating, the TIAG layer with meaningful real-world, place-specific elements, knowledge, and interpretation. Indeed, as noted elsewhere (Ferri and

Fig. 4 Big Urban Game (Lantz and Salen 2002)

Coppock 2013), “choosing the final path for the game-piece was not an easy task, as it required a good knowledge of the city or, even, an in-person tour to decide the best option” (Ferri and Coppock 2013). Again, Stakenas and Zurkov write:

> Because the game pieces were literally carried through city streets, greenbelts, and alleyways, players had to negotiate the pros and cons of each route, taking into account variables such as traffic jams, low lying bridges, and busy intersections. Neighborhood connected to neighborhood as the pieces traced their circuitous routes. Citizens were left wondering what might be next in a game where game piece-as-architectural spectacle was the rule of the day. (Stakenas and Zurkov 2006)

Taking part to the Big Urban Game meant accessing competences that are concretely related to the experience of navigating the real, physical world and importing them into a TINAG layer to produce strategic game decisions. “BUG not only featured an easily-recognizable Public Art intervention (its 25-ft tall inflatable game pieces have become almost an icon for the Urban Games genre) but required players to carefully consider features of the metropolitan areas to be traversed by the game” (Ferri and Coppock 2013). By expressly requiring the interplay between TIAG and TINAG layers, BUG was particularly effective in capturing players’ attention through playful dynamics and to direct it to examine specific urban spaces through a different set of expectations.

Live-action role-playing games (LARPs) are complex playful practices that, while they share some foundations with computer games and traditional tabletop role-playing games such as Dungeons & Dragons, are characterized by whole gameplay taking place in the physical world through fully embodied interaction. In other words, whereas standard tabletop role-playing games are based on series of speech acts in which each participant describes – in abstract terms – the actions undertaken by his/her avatar, LARPs require players to actually enact and physically perform the same actions. As Tychsen et al. (2006) explain,

> LARPs can be viewed as forming a distinct category of RPG because of two unique features: (a) The players physically embody their characters, and (b) the game takes place in a physical frame. Embodiment means that the physical actions of the player are regarded as those of the character. Whereas in a RPG played by a group sitting around a table, players describe the actions of their characters (e.g., “I run to stand beside my friend”); in an equivalent situation in a LARP, a player would physically run to the appropriate point within the game space. (Tychsen et al. 2006)

Whereas other LARPs are expressly escapist, in more recent years, several designers have become interested in exploring more mature themes through this medium. System Danmarc (Opus 2005) was intended to be overtly critical and to express a political vision not only by pointing at a possible future in which Danish society could collapse into a sort of rigid caste system but also by allowing participant to experience it in first person:

> In October 2005, 350 paying participants aged 16–42 experienced a dystopian future of Denmark where the weakest citizens had no part in the welfare society, no influence on the society they lived in and were confined in special low priority zones. It portrayed a Class C zone in Copenhagen, reserved for the citizens deemed useless for the society: a future Denmark, where democracy had degenerated. (Munthe-Kaas 2010)

System Danmarc was closely linked to a specific urban space – with a central square in Copenhagen used to set up the temporary stage for the game, using containers, metal fences, and other industrial materials with a clear *cyberpunk* aesthetics. The design team itself highlights the decision of making System Danmarc semipublic, as the diegetic part of the game was partially visible by ordinary passersby near the square, and with the initiative receiving considerable attention from local media. Lead designer Peter Munthe-Kaas writes: "The goal was to communicate the importance of democracy, participation and to outline societal problems in an exciting fashion. The organizers hoped to leverage the frustration and powerlessness prevailing in the LARP to inspire the participants to take action in real life" (Munthe-Kaas 2010).

Taking further the approach already experimented by System Danmarc, the Prosopopeia LARP series selected the opposite approach, allowing players to freely move across any public space and expanding radically the game's diegesis. Principally, Prosopopeia consisted of a ghost story where participants interacted with a fictional *spirit world* and were guided through different quests and storylines across the city of Stockholm. Intertwining of LARP elements with ordinary everyday practices was a defining characteristic of Prosopopeia. In this vein, Jonsson et al. (2006) detail the introductory phase of the experience: "When signing up for the game, the players were directed to a web site which contained a very short introduction to the game, ending with a very simple instruction that read: "You should now do all you can to forget about this project until it contacts you again. This is the only time the game will be presented as such. From now on everything is real"" (Jonsson et al. 2006). Going back to the interplay between TIAG and TINAG layers, we may describe the premise of Prosopopeia – players should take part to the game as if it was real – as a complete overlap between the two sets of expectations. In other words, there was no separation between the *players' own competences* (together with knowledge, assumptions, personal relationships) and *those of the fictional characters*. Jonsson et al. (2006) continue: "[every] information was supplied 'in game', as part of the preparatory [. . .] phase of the event. The typical elements of Larp preparations were absent; the players were not asked to prepare their character, create any costumes, or make contacts with the other in-game characters" (Jonsson et al. 2006).

Through the three examples examined here, a brief overview of the TIAG/TINAG dynamics at work in these artifacts may be constructed. TIAG and TINAG layers appear clearly separate only in relatively simple gameplay activities but overlap and intertwine significantly in more complex practices. This is made more evident by the interplay with physical places and with everyday activities that happen to interact spontaneously with the games. However, not all of such interactions exhibit the same characteristics. On one side, Big Urban Game effectively brought nonludic elements inside the ludic situation, proposing a difficult spatial task (navigating the most rational route across an urban area) and tapping in the contextual knowledge of participants to solve it. In this way, not only a playful activity is anchored to a concrete place, but it adopts specific rules to activate and reward place-specific interpretations (e.g., which is the best route from point A to point B?). Vice versa, System Danmarc and Prosopopeia – two systems with a much higher storiness (Ryan 2004) – adopt the same mechanics in the opposite direction and anchor everyday places to ludic activities. Whereas BUG required players to access real-world knowledge, System Danmarc and Prosopopeia provided their audience with game-based fictional, critical lenses to produce original interpretations of real places. That said, BUG, System Danmarc, and Prosopopeia were not originally meant to be related to ITV research, but the mechanics around the TIAG/TINAG ambiguity may be ported inside interactive television contents and allow for innovative interactions between users, technological platforms, and the physical places in which they are experienced.

Distributed Storytelling

The design tactic of distributed storytelling directly asks users to produce, construct, and curate some storytelling contents. In other words, by adopting a distributed storytelling approach, location-based pieces allow players not only to access interactive narratives but specifically require them to actively create and share new contents as part of the game experience. Distributed storytelling might contribute to more participative ITV experiences, with closer relationships to physical spaces and to the sociotechnical practices taking place there. To illustrate this characteristic, two urban games will now be discussed, paying specific attention to the rules adopted for motivate players to gather, assemble, and reinterpret narrative fragments.

Exemplars: From the Bank of Stories to the Storywall

Mettiti in Gioco (Ferri 2011), also called MIG, is an urban game inspired by the well-known Massively Multiplayer Soba series (Tiltfactor labs 2007), reinterpreted to take place in smaller Italian towns. Participants to a MIG game are divided into groups and receive a sealed envelope containing the different missions to be completed and small handheld video cameras, with which they are asked to document their game experience. A first task requires players to translate into Italian food recipes written in other languages, asking for help to nonlocal passersby. As a second task, players were asked to collect a story set in the town in which the game was held and to retell it to a fixed video camera – at a specific installation named the “Bank of Stories.” Those missions were designed to be complementary, requiring players to interact with both senior citizens to collect stories from the past and with recent immigrants to translate the multilingual clues and, while the game organizers had some degree of control over the first one, the selection of which local tales to collect, and who to approach to ask for these was left completely up to the players.

Indeed, requiring players to self-document their gameplay experience and to “deposit” a short narrative at the Bank of Stories was a defining feature of a distributed storytelling strategy that differentiated MIG from the Massively Multiplayer Soba games:

> The act of retelling a story to be recorded on video in exchange for a prize thus adds value to the narration itself (a good tale is precious) and this builds a more personal, direct connection between players, passersby and the urban environment in which they are playing. In this sense, participants are no longer "just players" but they also take charge of a small but significant portion of socio-cultural, place-specific story-collection, −telling and -archiving practices. (Ferri and Coppock 2013)

Through this mechanic, players were not simply following a prompt given by the organizers but had a clear degree of authorial control over the creation and documentation of narrative contents. By doing so, a link between their playful activity, the place in which they were situated, and the people with which they interacted was established within the diegesis of the game.

A similar approach was adopted in the 3Cities (Salvador and Ferri 2013) urban game, which took place simultaneously in Bologna, Barcelona, and London in 2013. The game (Salvador and Ferri 2015a, b) integrated ludic elements with the exploration of public spaces and relied on simple game mechanics based on taking pictures with cameras or smartphones, leveraging a number of simple missions specifically designed to evoke the concept of transnational citizenship. Each team of players was tasked with carrying out a photographic exploration of a local neighborhood. Moreover, the pictures from each team were shared online in real time with others in the other cities, so that Italian, Spanish, and British players engaged in a dialogue at distance through images, as if they were describing a hypothetical transnational city composed of the three neighborhoods traversed by the game. For each location of 3Cities, a simple headquarter was set up to gather the players and their photos and to keep in touch with participants in other cities. The three tasks composing 3Cities were designed to touch language, architecture, and culture, the three themes through which the game aimed at constructing and visualizing a virtual transnational neighborhood that ideally reunited Barcelona, Bologna, and London.

The first mission (Linguistic Collage) tasked each team with producing a sentence composed of letters and words obtained by photographing street graffiti or signs scattered across the cities. In other words, this first mission required players to create a meaningful photographic collage using textual fragments obtained through an exploration of the surrounding urban space. Once the collage was successfully composed, to complete the first phase, each team had to share its creation with other remote teams in faraway cities. Every team not only sent their own collage to others but also received a similarly constructed sentence from the other cities and had to translate and visualize it using locally collected pictures with the same collage method. The second mission, entitled Travel to Another City, assigned two printed architectural images depicting details of buildings from other cities. The players' objective was to compose impossible perspectives picturing printed images of a city inside the photos of their own neighborhood, setting an architectural detail from another city in their own. The third mission, Pic Your Story, required players to create a short photographic storyboard composed of three

Fig. 5 (**a**, **b**) 3Cities (Salvador and Ferri 2013)

images, each one to be produced by a different team in each of the three cities. In other words, the first team shot a scene set in their local neighborhood, with a short caption if necessary, which was later sent to the other remote players to be completed with a second and third shot, creating a collaborative narrative production. As a conclusion to the 3Cities experience, a Storywall – a temporary exhibition space – was constructed in each of the three cities. The three Storywalls collected and publicly displayed all the images shared by players, constituting the visual identity of the event, documenting the participants' activities, and conveying a general insight of the game also to the nonplaying audience (Fig. 5a, b).

The distributed storytelling design strategy exemplified by these pieces empowers players with the selection, creation, and retelling of user-generated contents. In this way, users cease to be just an audience for this experience. Moreover, the distributed storytelling model moves away from the more traditional interactive schema where players control the instantiation of a preexisting branching narrative (Koenitz 2010, 2015): whereas most systems for interactive digital storytelling are constituted by a software component that, reacting to the user's input, present different narratives (either prescripted and selected from a repository or procedurally constructed at run time), these urban game experiences task participant with the actual production of completely new contents. Such user-generated fragments are then circulated among other players, becoming fully part of the playful experience. Similarly to the previously discussed exemplars, although the works presented here are not Interactive TV technologies, they may still yield useful insights for research and design in the ITV field. Indeed, emerging technologies for producing and sharing video fragments using mobile devices – with Periscope/Twitter (2015) being the most recent example – suggest the relevance of real-time user-generated content production for ITV systems. A distributed model of storytelling may provide users of future mobile ITV platforms with compelling, playful motivations for exploring and documenting specific places.

Conclusions

Providing the user with a sense of place – related to a specific geographic location in which one is situated, or linked to a faraway place, or even giving place-like qualities to virtual spaces such as massively multiplayer online role-playing games (MMORPGs) – has been deemed central for several forms of digital interactions. In the past decade, studies from HCI and CSCW to telepresence and digital cultural heritage have specifically addressed this theme (Harrison and Dourish 1996; Li and Goodchild 2012; Turner and Turner 2006; Sylaiou et al. 2010; Champion 2008; De Souza e Silva 2009), but the relative scarcity of works on place specificity focusing expressly on ITV suggests a gap in the current research, whereas the latest developments in mobile TV would seem highly coherent with such topic. To contribute to closing this gap, some initial directions were suggested here by pointing at compatible treatments of the notion of place in related fields, for example, the design of pervasive urban games.

Indeed, game designers and game scholars might provide operational concepts that help understanding the role and the potentialities of places for ITV. For this reason, this contribution addressed two loose types of artifacts: works that are anchored to the experience of faraway places and others that leverage the physical location in which the user is. Clearly, it cannot be claimed these two are the only possible options, as many others are likely to emerge in this evolving field. Three web-based advertisement games (Mercedes-Benz: Escape the Map, Magnum: Pleasure Hunt 2, Chaos in Your Town), or *advergames*, were presented to demonstrate how well-known experiences such as digital cartography could be made ludic. Furthermore, this contribution also considered three pervasive experiences (Big Urban Game, System Danmarc, Prosopopeia) that require users to access place-specific knowledge in the game or game-specific knowledge in physical places and two urban games (Mettiti in Gioco, 3Cities) that task players with producing place-specific short video contents. The exemplars discussed here are not ITV products per se but exemplify specific design strategies that are compatible with ITV and mobile ITV: from this perspective, a set of preliminary insights may be offered to ITV researchers and designers. In sum, this overview has yielded three design strategies that might be named **experience anchoring**, **place permeability**, and **distributed storytelling** – certainly not an exhaustive set, but a first group to spur further research and design.

The convergence between computing platforms and television, exemplified by products such as the Google Chromecast or the Apple TV, might benefit from applications inspired by the first list of examples, building simple games and interactive narratives anchored to existing apps. The three selected advergames have elaborated on the interaction patterns of Google Street View creating playful situations for their users and, at the same time, leveraging the experience of instantaneously accessing remote places as a trope in their game design. As the number of apps in these new platforms increases, future works might adopt the design strategy of experience anchoring and use established conventions and expectations as foundations for new designs. In this vein, an ITV product may be anchored not only to actual physical places but also to their mediated representations and to the ways in which they are experienced through technology.

The design strategy of place permeability establishes a connection between a place – with its related expectations, context, stories, and sociotechnical practices – and a digital system. The examples presented here suggest how this influence is bidirectional: in one case, a pervasive game required participants to access their knowledge of the surrounding physical places, and, in the others, two LARPs provided players with a worldview and a perspective that would bleed into their everyday lives. It is possible to imagine future ITV contents adopting similar practices, for example, with mobile ITV systems addressing nearby physical places or with locative interactive experiences transcending the spaces in which they are set.

Distributed storytelling allows players not only to influence the development of a specific experience but to actually produce user-generated contents and weave them in the overall system. Although this design strategy is not exclusive to location-specific pieces, it is particularly effective in motivating players to explore their physical surroundings and interact with passersby in unexpected ways. Distributed storytelling taps into the practice of bottom-up video production and viral distribution on social networking services and leverages it for creating systems in which participants significantly add to the overall experience. Complex sociotechnical systems – possibly composed of an ITV platform plus a mobile app and elements of pervasive game design – might adopt and extend this design strategy.

These three design strategies are clearly not the only possible ones linking digital interactions, physical places, and interactive television contents. They are offered here more as objects to think with (Turkle 2007) than a complete taxonomy – also because it is debatable whether such categorization would be heuristically useful or even possible. In other words, these strategies are presented more as generative tools for practitioners rather than descriptive and taxonomical categories for theorists. By pointing at them and at other similar strategies, some underlying similarities between ITV and other cognate disciplines emerge and may be speculatively traced to point at possible future convergences. In relation to this, there are evidently many more steps to be taken in this line of inquiry – both on the side of ITV research and design and on those of the other fields. In particular, mobile ITV systems seem especially promising in this context, both as platforms on which to experience interactive TV contents while being physically situated in a place and as tools for capturing user-generated fragments to be transmitted and interwoven in other experiences. Whereas this contribution is intended as a preliminary step bringing closer to the field of ITV the interdisciplinary dialogue about place, interactive technology, playfulness, and design, much more work is clearly needed. Research-wise, the exemplars discussed here underline the usefulness of a tighter dialogue between ITV and game studies, as innovative storytelling forms are emerging in the domain of digital play and could be productively imported also in ITV. As for new design works, two future tendencies emerge from the strategies discussed in these pages: on one side, the already-mentioned potentialities of mobile ITV platforms and, on the other one, devices such as the Chromecast, Steam Box, and Apple TV that accelerate the convergence between app and ITV experiences.

Recommended Reading

Abbott Mead Vickers BBDO, Mercedes Benz: Escape the Map (web video game) (2011)

P. Almeida, J. Ferraz, A. Pinho, D. Costa, Engaging viewers through social TV games, in *Proceedings of the 10th European Conference on Interactive TV and Video* (ACM, New York, 2012), pp. 175–184

S. Barlow. Her Story (video game) (2015). Retrieved from http://www.herstorygame.com/

J.-P. Bichard, A. Waern, Pervasive play, immersion and story: designing interference, in *Proceedings of the 3rd International Conference on Digital Interactive Media in Entertainment and Arts* (ACM, New York, 2008), pp. 10–17

I. Bogost, *Unit Operations: An Approach to Videogame Criticism* (MIT Press, Cambridge, MA, 2006)

I. Bogost, *Persuasive Games: The Expressive Power of Videogames* (MIT Press, Cambridge, MA, 2007)

L. Brindfors, Magnum Pleasure Hunt 2 (web video game) (2012)

E. Champion, Otherness of place: game-based interaction and learning in virtual heritage projects. Int. J. Heritage Stud. **14**, 210–228 (2008). doi:10.1080/13527250801953686

F. Charles, J. Porteous, M. Cavazza, J. Teutenberg. Timeline-based Navigation for Interactive Narratives. In *Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology* (pp. 37:1–37:8) (2011). New York, NY, USA: ACM. http://doi.org/10.1145/2071423.2071469

D. Compagno, P. Coppock, *Computer Games Between Text and Practice* (AISS, Milano, 2009)

M. Copier, Connecting worlds. Fantasy role-playing games, ritual acts and the magic circle, in *Proceedings of the 2005 DiGRA International Conference: Changing Views: Worlds in Play* (2005), Vancouver, British Columbia, Canada.

DDB B-Reel, Chaos in Your Town (web video game) (2011)

J.E. Deaton, C. Barba, T. Santarelli, L. Rosenzweig, V. Souders, C. McCollum, J. Seip, B.W. Knerr, M.J. Singer, Virtual environment cultural training for operational readiness (VECTOR). Virt. Real. **8**, 156–167 (2005). doi:10.1007/s10055-004-0145-x
A. De Souza e Silva, Hybrid reality and location-based gaming: redefining mobility and game spaces in urban environments. Simulat. Gam. **40**, 404–424 (2009). doi:10.1177/1046878108314643
A. De Souza e Silva, D.M. Sutko (eds.), *Digital Cityscapes: Merging Digital and Urban Playspaces*, First printing edition (Peter Lang Publishing, New York, 2009)
U. Eco, *A Theory of Semiotics* (Indiana University Press, Bloomington, 1976)
G. Ferri, Interpretive cooperation and procedurality. A dialogue between semiotics and procedural criticism. E|C, **5**, 15–20 (2009)
G. Ferri, Mettiti In Gioco (urban performance) (2011)
G. Ferri, To play against: describing competition in gamification, in *Rethinking Gamification*, ed. by M. Fuchs, S. Fizek, P. Ruffino, N. Schrape (Meson Press by Hybrid Publishing Lab, Lüneburg, 2014)
G. Ferri, Narrative structures in IDN authoring and analysis, in *Interactive Digital Narrative: History, Theory and Practice*, ed. by H. Koenitz, G. Ferri, M. Haahr, D. Sezen, T.I. Sezen (Routledge, New York, 2015)
G. Ferri, P. Coppock, Serious urban games. From play in the city to play for the city, in *Media and the City: Urbanism, Technology and Communication*, ed. by S. Tosoni, M. Tarantino, C. Giaccardi (Cambridge Scholars Publisher, Cambridge, 2013)
G. Genette, Figures III. Ed. du Seuil, Paris (1972)
E. Goffman, *Encounters. Two Studies in the Sociology of Interaction* (Bobbs-Merrill, Indianapolis, 1961)
Google. (2013). Find More with Movie & TV Info Cards. Retrieved from https://support.google.com/googleplay/answer/3015292
C. Hales, Interactive cinema in the Digital Age, in *Interactive Digital Narrative: History, Theory and Practice*, ed. by H. Koenitz, G. Ferri, M. Haahr, D. Sezen, T.I. Sezen (Routledge, New York, 2015)
S. Harrison, P. Dourish, Re-place-ing space: the roles of place and space in collaborative systems, in *Proceedings of the 1996 ACM conference on computer supported Cooperative Work* (ACM, New York, 1996), pp. 67–76
J. Hess, B. Ley, C. Ogonowski, L. Wan, V. Wulf, Jumping between devices and services: towards an integrated concept for social TV, in *Proceedings of the 9th International Interactive Conference on Interactive Television* (ACM, New York, 2011), pp. 11–20
J. Huizinga, *Homo Ludens: A Study of the Play Element in Culture* (Beacon, Boston, 1938)
IperG (2005) Prosopopeia (urban performance)
F. Jannidis, Narratology and the narrative, in *What is Narratology? Questions and Answers Regarding the Status of a Theory*, ed. by T. Kindt, H.-H. Müller (Walter de Gruyter, Berlin/New York, 2003)
H. Jenkins, *Convergence Culture: Where Old and New Media Collide* (NYU Press, New York, 2006)
S. Jonsson, M. Montola, A. Waern, M. Ericsson, Prosopopeia: experiences from a pervasive Larp, in *Proceedings of the 2006 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology* (ACM, New York, 2006)
H. Koenitz, Towards a theoretical framework for interactive digital narrative, in *Interactive Storytelling*, ed. by R. Aylett, M.Y. Lim, S. Louchart, P. Petta, M. Riedl (Springer, Berlin/Heidelberg, 2010), pp. 176–185
H. Koenitz, Towards a specific theory of interactive digital narrative, in *Interactive Digital Narrative: History, Theory and Practice*, ed. by H. Koenitz, G. Ferri, M. Haahr, D. Sezen, T.I. Sezen (Routledge, New York, 2015)
H. Koenitz, G. Ferri, M. Haahr, D. Sezen, T.I. Sezen, The evolution of interactive digital narrative theory, in *Interactive Digital Narrative: History, Theory and Practice*, ed. by H. Koenitz, G. Ferri, M. Haahr, D. Sezen, T.I. Sezen (Routledge, New York, 2015)

C. Kohler, Finally! You Can Now Play Pac-Man on Google Maps. Wired. http://www.wired.com/2015/03/pac-man-google-maps/ (2015)
B. Laurel, *Computers as Theatre* (Addison-Wesley, Boston, 1991)
F. Lantz, K. Salen, Big Urban Game (urban performance) (2002)
L. Li, M.F. Goodchild, Constructing places from spatial footprints, in *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Crowdsourced and Volunteered Geographic Information* (ACM, New York, 2012), pp. 15–21
J.E. Mcgonigal, *This Might Be a Game: Ubiquitous Play and Performance at the Turn of the Twenty-first Century* (University of California, Berkeley, 2006)
M. Montola, J. Stenros, A. Waern, *Pervasive Games: Theory and Design* (CRC Press, Amsterdam, 2009)
P. Munthe-Kaas, Transmitting a political vision through Larp. http://munthe-kaas.dk/blog/?p=130. Accessed 4 Jan 2015 (2010)
J.H. Murray, S. Goldenberg, K. Agarwal, T. Chakravorty, J. Cutrell, A. Doris-Down, H. Kothandaraman, story-map: IPad companion for long form TV narratives, in *Proceedings of the 10th European Conference on Interactive TV and Video* (ACM, New York, 2012), pp. 223–226
J.H. Murray, Transcending transmedia: emerging story telling structures for the emerging convergence platforms, in *Proceedings of the 10th European Conference on Interactive TV and Video* (ACM, New York, 2012), pp. 1–6
J.H. Murray, *Hamlet on the Holodeck: The Future of Narrative in Cyberspace* (Free Press, New York, 1997)
Namco corp. Pac Man (arcade video game) (1980)
A. Nandakumar, J. Murray, Companion apps for long arc TV series: supporting new viewers in complex storyworlds with tightly synchronized context-sensitive annotations, in *Proceedings of the 2014 ACM International Conference on Interactive Experiences for TV and Online Video* (ACM, New York, 2014), pp. 3–10
E. Nieuwdorp, The pervasive discourse: an analysis. J. Comput. Entertain. (2007). doi: 10.1145/1279540.1279553
Opus. System Danmarc (urban performance) (2005)
A. Piacenza, F. Guerrini, N. Adami, R. Leonardi, J. Porteous, J. Teutenberg, M. Cavazza, Generating story variants with constrained video recombination, in *Proceedings of the 19th ACM International Conference on Multimedia* (ACM, New York, 2011), pp. 223–232
J.-H. Réty, N. Szilas, J. Clément, S. Bouchardon, Authoring interactive narratives with hypersections, in *Proceedings of the 3rd International Conference on Digital Interactive Media in Entertainment and Arts* (ACM, New York, 2008), pp. 393–400
M. Rieser, Artistic explorations: mobile, locative and hybrid narratives, in *Interactive Digital Narrative: History, Theory and Practice*, ed. by H. Koenitz, G. Ferri, M. Haahr, D. Sezen, T.I. Sezen (Routledge, New York, 2015)
M.-L. Ryan, *Narrative as Virtual Reality: Immersion and Interactivity in Literature and Electronic Media* (Johns Hopkins University Press, Baltimore, 2001)
M.-L. Ryan, *Narrative Across Media: The Languages of Storytelling* (University of Nebraska Press, Lincoln, 2004)
M.-L. Ryan, *Avatars of Story* (University of Minnesota Press, Minneapolis, 2006)
M. Salvador, G. Ferri, 3Cities (urban performance) (2013)
M. Salvador, G. Ferri, 3Cities – pic your story. A playful urban experience. Screen City J. **5**, (2015a)
M. Salvador, G. Ferri, 3Cities – Pic your story. Il design partecipativo di una ludicizzazione urbana, in *Oltre il gioco. Critica Della Ludicizzazione Urbana*, ed. by M. Bittanti, Z. Emanuela (Unicopli, Milano, 2015b)
J.H. Smith, S.N. Just, Playful persuasion: the rhetorical potential of advergames. Nord. Rev. **30**, 53–68 (2009)

L. Sørensen, H.W. Nicolajsen, Generating ideas for new mobile TV services: accepting and socializing mobile television, in *Proceedings of the 8th International Interactive Conference on Interactive TV and Video* (ACM, New York, 2010), pp. 179–182

O. Sotamaa, All the world's a Botfighter stage: notes on location-based multi-user gaming, in *Proceedings of the Computer Games and Digital Cultures Conference* (Tampere University Press, 2002) p. 35

C. Stakenas, M. Zurkov, Public art. http://www.o-matic.com/public_art/index.html. Accessed 1 Oct 2012 (2006)

J. Stenros, In defence of a magic circle: the social and mental boundaries of play, in *Proceedings of 2012 International DiGRA Nordic Conference* (2012), Tampere, Finland.

J. Stenros, A. Waern, M. Montola, Studying the elusive experience in pervasive games. Simulat. Gam. **43**, 339–355 (2012b). doi:10.1177/1046878111422532

B.H. Suits, *The Grasshopper: Games, Life, and Utopia* (University of Toronto Press, Toronto, 1978)

S. Sylaiou, A. Karoulis, K. Mania, M. White, Exploring the relationship between presence and enjoyment in a virtual museum. Int. J. Hum. Comput. Stud. **68**, 243–253 (2010). doi:10.1016/j.ijhcs.2009.11.002

Tiltfactor labs. Massively Multiplayer Soba (urban performance) (2007)

S. Turkle, *Evocative Objects: Things We Think With* (MIT Press, Cambridge, MA, 2007)

V.H. Tuulos, J. Scheible, H. Nyholm, Combining web, mobile phones and public displays in large-scale: Manhattan story mashup, in *Proceedings of the 5th International Conference on Pervasive Computing* (Springer, Berlin/Heidelberg, 2007), pp. 37–54

P. Turner, S. Turner. Place, Sense of Place, and Presence. *Presence: Teleoperators and Virtual Environments, 15*(2), 204–217 (2006). http://doi.org/10.1162/pres.2006.15.2.204

Twitter Inc. Periscope (app) (2015)

A. Tychsen, M. Hitchens, T. Brolund, M. Kavakli, Live action role-playing games control, communication, storytelling, and MMORPG similarities. Game Cult. **1**, 252–275 (2006). doi:10.1177/1555412006290445

M.F. Ursu, M. Thomas, I. Kegel, D. Williams, M. Tuomola, I. Lindstedt, T. Wright, A. Leurdijk, V. Zsombori, J. Sussner, U. Myrestam, N. Hall, Interactive TV narratives: opportunities, progress, and challenges. ACM Trans. Multimedia Comput. Commun. Appl. **4**, 25:1–25:39 (2008). doi:10.1145/1412196.1412198

M. Weiser, The computer for the 21st century. Sci. Am. **265**, 94–104 (1991). doi:10.1038/scientificamerican0991-94

Social Interaction Design for Online Video and Television 44

Pablo Cesar and David Geerts

Contents

P. Cesar (✉)
CWI: Centrum Wiskunde and Informatica, Amsterdam, The Netherlands
e-mail: p.s.cesar@cwi.nl

D. Geerts
CUO, iMinds/KU Leuven, Leuven, Belgium
e-mail: david.geerts@soc.kuleuven.be

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_39

Abstract

In recent years social networking and social interactions have challenged old conceptions in the media landscape. Web applications that offer video content, connected television sets and set-top boxes, tablets and smartphones as second screens, and online TV widgets have radically transformed how people watch and interact around television content. Since the wealth of existing solutions and approaches might be daunting to newcomers, this chapter surveys previous and current efforts in the area of social television and online video. In particular, this chapter provides a framework that categorizes the most salient features of social television services. The final objective is on the provision of a set of design guidelines for the development of novel applications and services that enable viewers to socialize around online video and television content.

Keywords

Social television • Design guidelines • Hybrid broadcast broadband

Introduction

The television and video industry has recently witnessed major changes in the way people consume and interact with their content. For many years the fear of declining audiences and commercial revenues led device manufacturers to focus on the display capabilities of the devices by adding features such as 3D or Ultra-HD, while interactivity was often perceived as a threat rather than as an opportunity. The mass acceptance of smaller devices, such as tablets, has transformed such view, truly enabling interactivity while consuming media content at home. What was perceived as hindering the user experience – the “second screen” – has resulted into an essential companion to the television. Paradoxically, although television and video programs can now be consumed anywhere and anytime, which can lead to individualized experiences, they remain highly social in nature. Live TV programs are more popular than ever, and social media streams are becoming as important as the content itself. Users heavily interact using popular social networking sites (e.g., Facebook, Twitter), and broadcasters are trying to understand new rating models based on such streams. Both software designers and device manufacturers are tapping into these new developments by creating apps for tablets or connected television sets that integrate social features with video content.

The term social television has rapidly become an umbrella term for these applications, and designers need to learn how to take the social nature of television and video watching into account. This chapter studies current developments on social television and online video applications, providing the reader knowledge on how to design them. The term "sociability" is used to indicate these interface aspects that support and enhance social interaction with and through new technologies and applications, and social interaction design is a way of including these sociability aspects in the design process. The final objective is to provide a number of design guidelines for the development of applications and services that enable users to socialize around video content. The guidelines are intended for readers interested in the current shift on how people socialize around media content.

We consider a rather broad definition of television and online video:

> *Television and online video is audiovisual content intended to entertain, inform, enrich or involve viewers, is or was distributed via a telecommunications channel and is being watched on a screen-based device.* (Geerts and De Grooff 2009a)

This very broad definition encompasses not only traditional television but also new forms of watching television or video content on various devices. It also includes professionally produced as well as user-generated content and broadcast as well as recorded content. It does exclude using a television set for other purposes than watching audiovisual content, such as browsing Web pages, gaming, or video conferencing. Using video and/or television for educational (e.g., remote classrooms) or other purposes (e.g., webinars, promotional videos, or instructional videos) is not the main topic of this chapter, but many of the methods and guidelines we will discuss are to a certain extent also applicable.

New Consumption Patterns

Social television and online video constitute a fundamental shift in how people interact and socialize around media content. Websites combine video streaming services with social networking sites such as Twitter; media software like Beamly allows users to recommend and discuss about favorite television programs, and new services enable friends to remotely watch television together. Strategy reports and a vast selection of new commercial services show the relevance of the current shift toward a more socially aware media experience. It reveals as well the commercial interest behind integrating successful social media and communication solutions with streaming video. All these developments can be called social TV: allowing remote viewers to socially interact with each other via the television set, smartphones, tablets, or the PC, where viewers might be separated in time and/or in space. This is not a new concept, as it has been explored since the start of this century in academic and industrial research labs by creating several prototypes. However, current products are making it a promising business proposition. Features in social TV include remote talking or chatting while watching a television program, content-aware buddy lists that show what your friends are watching, sharing and recommendation of video material based on social network statistics and trends, and easy accessible Twitter streams associated with one particular program (Obrist et al. 2015a; Cesar et al. 2014b; Cesar and Geerts 2011).

Some examples of socially aware media services include the integration of Twitter updates during a live video stream and Facebook applications that allow commenting while watching video content. Several similar applications are recently being created for smartphones and tablets, which act as a secondary screen, so the commenting and communication do not occupy valuable space on the television set. In parallel to the integration of social networking into the television environment, there have been successful efforts in enabling domestic high-quality videoconferencing, providing a direct communication link between separate households watching television together.

Based on these examples, there are enough grounds to conclude that the market for socially aware consumption of media is growing, but is the behavior of the users changing? Television has been traditionally associated with passive watching, but recent studies indicate that habits are changing. According to Yahoo! and the Nielsen Company, already back in 2011 86 % of mobile Internet users (and 92 % of 13–24 s) were using their mobile devices simultaneously with TV (http://www.intomobile.com/2011/01/29/nielsenyahoo-86-of-mobile-users-fire-up-their-phone-while-watching-tv/). Communicating with friends was the most common activity (SMS: 56 % for mobile phones; email: 33 % for mobile phones and 49 % for PCs; IM: 19 % for mobile phones and 22 % for PCs), while updating/reading social networking sites was second in the list (40 % for mobile phones and 53 % for PCs). More recent reports show similar trends: in March 2015 Nielsen reported that 58 % of viewers browse the Internet while watching video programming. Interestingly, according to the same report the growing use of mobile phones and tablets has not (yet) dethroned the television set as the preferred device for watching video content in the home, although the laptop is coming close. Of the respondents, 63 % believe bigger is better when it comes to screen size, and 65 % prefer to watch video programming live, at its regularly scheduled time. The report concludes that "device viewing is largely situational; it depends on where we are, whom we're with and what we're doing" (Nielsen 2015).

The wealth of existing solutions and approaches might be daunting to newcomers, but it provides enough material for surveying, analyzing, and classifying them. The intention of this chapter is to summarize previous and current efforts, categorizing their most salient features. The resulting framework is a valuable contribution for better understanding the present. Moreover, it provides a tool for discussing future developments in the field. Based on such a framework, a number of social interaction design guidelines are proposed.

Framework for Social TV and Online Video

A structured survey of past and current social TV and online video applications was performed. There are many systems, so the first decision was to focus on the most relevant ones, with relevance defined as market impact, research impact, and novelty. Balancing exhaustiveness and manageability, the authors surveyed a total of 100 systems over the last 10 years combining different methodologies (user studies, expert evaluation, informal trials). Some examples of the systems include: Web applications that provide social interaction around video, broadcasters' new social offerings, mobile applications that allow for microblogging, synchronous communication applications for people at different locations, and TV content sharing. Based on the survey a number of salient aspects have been identified.

The first and most important aspect is activity, because it determines what the aim of the interaction is.

- Activity refers to the general goal of the application and the main tasks the users can perform with it.

  The other essential aspects include:

- Device: What is the device and network in use? Some solutions focus on the Web, while others target a television environment. Mobile devices and secondary screens are lately becoming popular.
- Social interaction: How are the users interacting? The options include text, audio, and video.
- Sync: Does the social interaction take place synchronously (while watching) or asynchronously?
- Awareness: How are other users represented? Options include traditional buddy lists, ambient solutions, and more extended buddy lists as provided by popular social networking sites such as Facebook.
- Social reach: What is the social reach of the activity? In some cases closed network reach is provided, usually including friends or family, while in other cases a more open reach is available and strangers are able to communicate with each other.

The figure below shows a representation of the resulting framework, based on the identified aspects (Fig. 1).

Structure of the Chapter

The chapter provides a number of social interaction design guidelines for the development of online video and television applications and services. Based on the framework introduced in the previous section, the chapter is structured in six

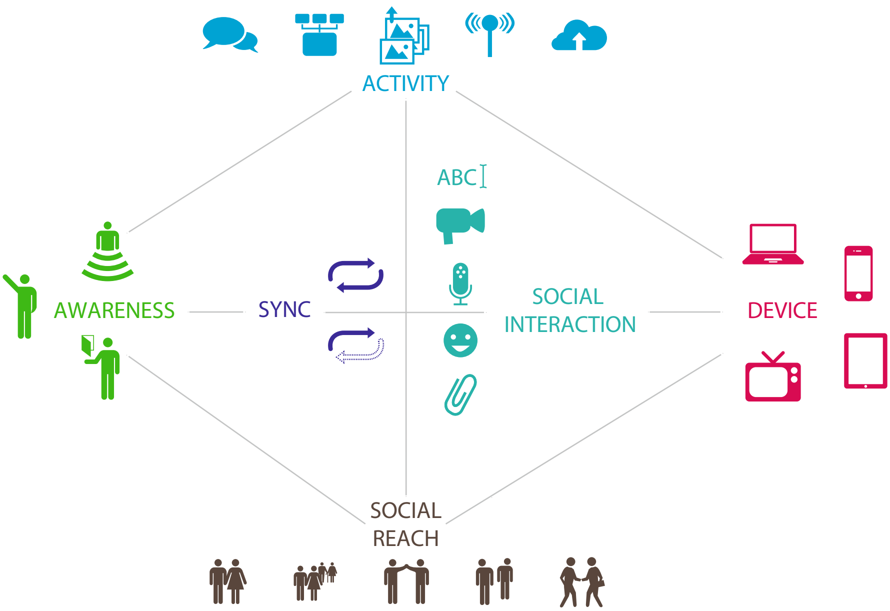


Fig. 1 Framework for social TV and online video

main parts, each one covering the most salient aspects: activity, device, social interaction, sync, awareness, and social reach. The chapter ends with a conclusion section that highlights the core challenges ahead.

Activity

The structured survey facilitated the clustering of several applications and services based on their dominant aspect: activity. Five main clusters were identified: content selection, content sharing, broadcasting, congregation, and conversation. This section further analyzes each of the categories, providing some relevant examples. In short, the main categories are:

1. Content selection: when information by other peers is used for making appropriate decisions on what to watch.
2. Content sharing: the user might also want to send to his/her peers full programs or edited versions of the programs.
3. Broadcasting: making available to others information about what you are currently watching or have watched.
4. Congregation: commenting and creating a brand about a media item with a large community of viewers.
5. Conversation: direct communication via chat, audio, or video with other peers while watching television content.

We believe this framework is, while approachable, complete enough for describing past and present social TV and online video solutions. Moreover, it is consistent with previous efforts (Harboe 2009; Cesar and Geerts 2011). Nevertheless, our framework is more functional, and it provides a structured mechanism for categorizing applications. Other categorizations and definitions include Coppens et al. (2005) and Ducheneaut et al. (2008) that focus their analysis on our direct communication category but do not take into account new directions in the field. In general, previous efforts in categorizing social TV and online video have been very valuable, but they did not consider the broadness and complexity of current developments. The next step is to better understand the classification. The next subsections will describe each of the categories in detail.

Content Selection

Due to the wide range of alternatives, content selection has been considered as the cornerstone of interactive television systems. Since the first commercial interactive television solutions, the Electronic Program Guide (EPG) helps viewers decide what to watch and sometimes when to watch it, by providing video recording capabilities. The EPG in its traditional form is a table-based application showing the schedule of different channels and some information about specific programs, mimicking traditional TV listings in magazines and newspapers. The actual utility of such traditional representation of the information about the media content may be arguable, and novel interfaces are not restricted to the linearity of broadcast television. They provide more complete and relevant information for the decision on what to watch. A key game changer was the introduction of Web-based interfaces for selecting media (like YouTube), which included some extra information like the number of views (Fig. 2).

Lately, a variety of interfaces for consumption of media in the Web and smart TV environments have emerged (Obrist et al. 2015a). All of them aim at the provision of valuable information for the decision of what to watch, beyond the more traditional broadcast models represented in the EGP. As demonstrated by social media research, useful information includes ratings, comments, recommendations, and insights from the social network that can be directly used by the viewer or by a recommender system. Some examples of these novel systems include Netflix, Apple TV, Hulu, and Samsung's smart TV. In our opinion, one key example following the right guidelines is the iPlayer from BBC that, apart from information about the media content, incorporates viewing data about your network (Fig. 3).

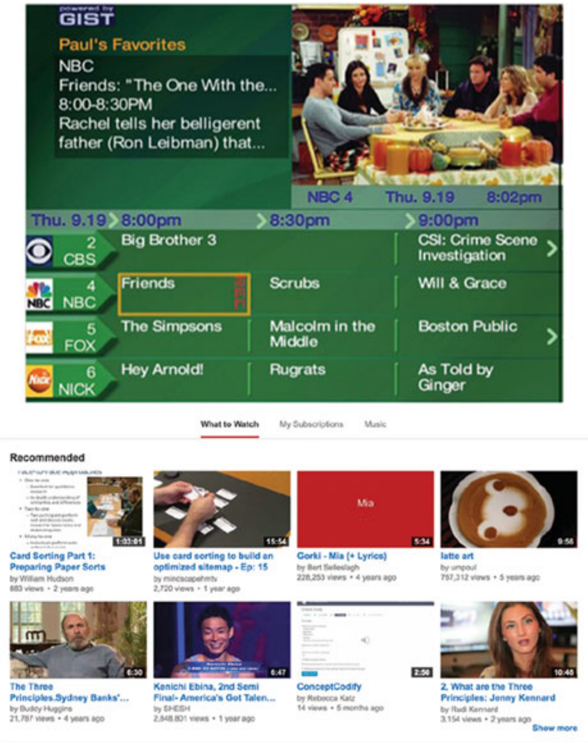


Fig. 2 Two examples of content selection interfaces: traditional EPG (*left*) and Web-based social media (*right*) (By Dani.sc (jo), CC BY-SA 3.0 (http://creativecommons.org/licenses/by-sa/3.0), via Wikimedia Commons)

Content Sharing

In addition to the selection of content based on the activity and reactions of your network, some users wish to directly share media with others. In this case, the video – or a link to the video – is the communication means between people. Grabyo, Boxee, and iPlayer provide(d) content-sharing functionality, acknowledging that direct recommendations are more effective and personal than computed recommendations (Bernhaupt et al. 2008). In most cases, the user can also recommend only the interesting parts of the video. Ambulant Annotator (Cesar et al. 2009) provides extensions to the model that allow viewers to enrich television content while watching and to share the results with targeted groups. Content sharing is an asynchronous activity that tends to reach social networks beyond the close ties. Nevertheless, direct recommendations that involve content editing such as clipping is typically restricted to close ties, due to the effort and intimacy of such action. The most salient feature of this category is the interaction means: the actual video or a link to it.

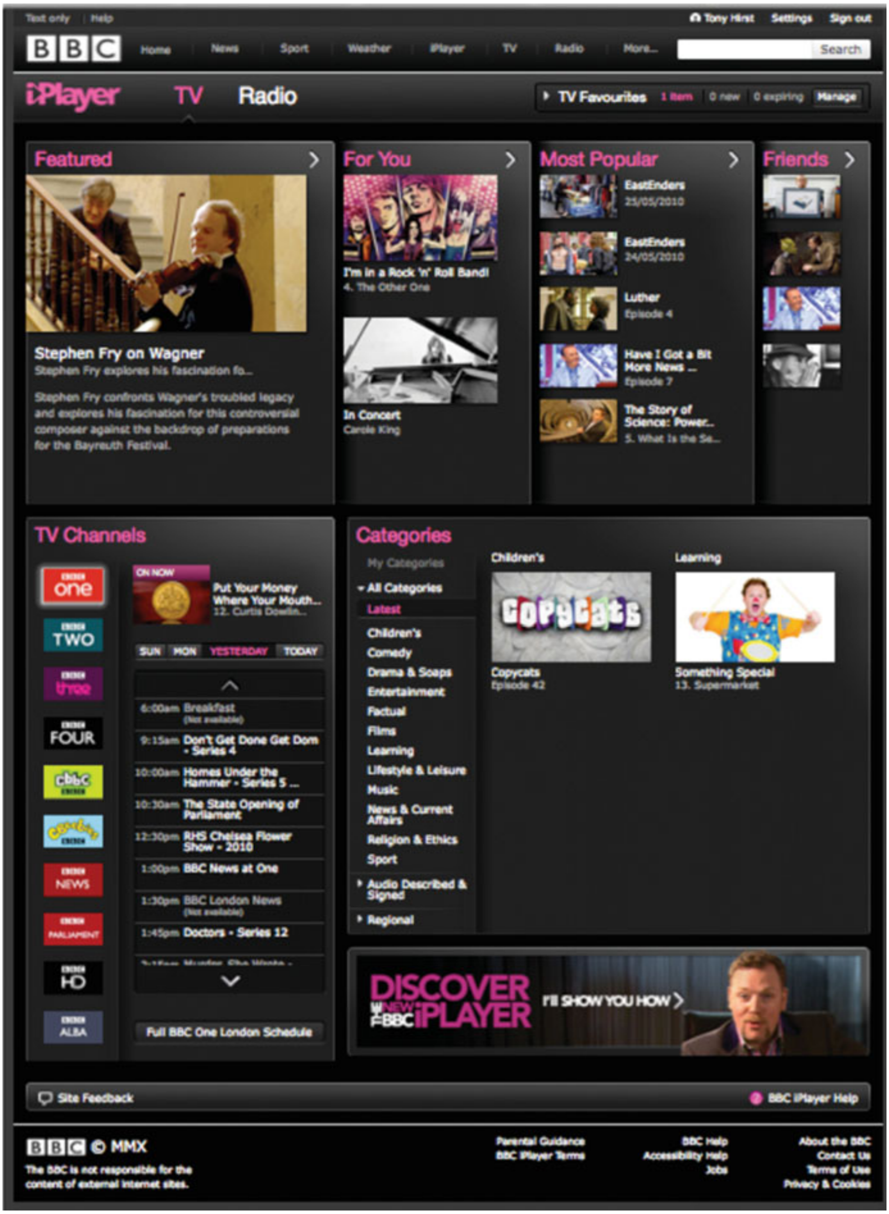


Fig. 3 Content selection interface from iPlayer (BBC) (By Tony Hirst, CC BY 2.0 (https://creativecommons.org/licenses/by/2.0/), via Flickr)

Fig. 4 Commercial content sharing application that allows viewers to directly share with others parts of the media content they are watching

For content-sharing applications, it is essential that the right multimedia document model is followed, so it allows to not only share timed-based information regarding the media content but also content enrichments, thus increasing the quality of communication between sender and recipient (Fig. 4).

Broadcasting

The popularity of social networks allows users to broadcast their daily activities, such as consuming media. This has certainly had an impact on the way users (and media owners) communicate around media content. Broadcasting in this context primarily refers to services and applications that enable viewers to report about the media they watch and their impressions. The most significant examples, either incorporated in the television screen or in a secondary screen, are Facebook and Twitter. The other relevant examples are the “check-in” applications for television content.

While early social TV systems usually offered status sharing (e.g., “I’m watching Breaking News on CNN”) as one of its social features, around 2010 applications were launched which offered status sharing as its core feature.

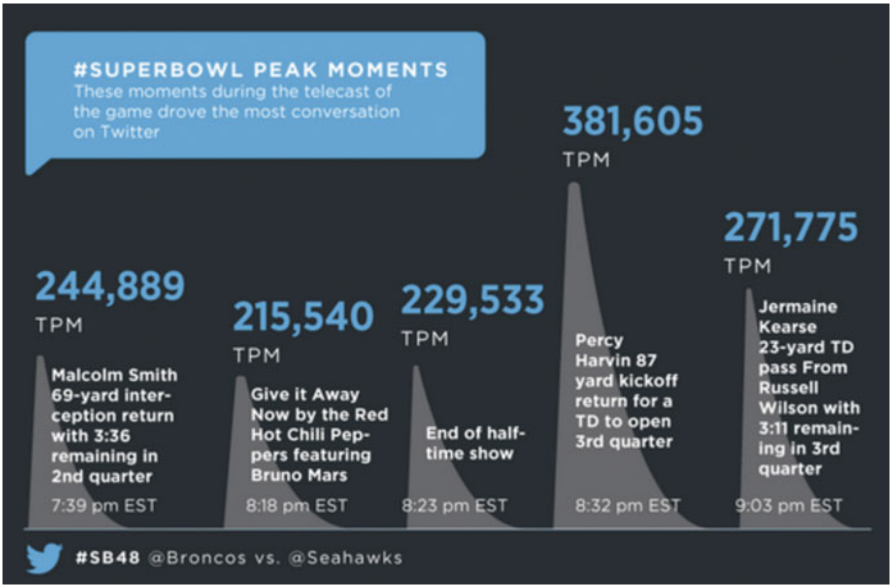


Fig. 5 Media watching paradigm change, where viewers actively comment and talk about what they are watching (broadcast their impressions) (https://blog.twitter.com/2014/celebrating-sb48-on-twitter)

Applications like Miso, Tunerfish, and PhiloTV allowed users to indicate the TV program they are watching by "checking in" to that program, much like checking in to locations with location-based applications such as Foursquare and Gowalla. Users that frequently check into a specific TV program earn badges. Apart from indicating the TV program a user is watching, these applications also provided the option to write a short, Twitter-like, status update. Similar to other social networks, users can follow other users, so they receive the status updates and other information from these users. The program's users are currently watching; the associated status update and the badges earned are broadcast to these friends and "followers," creating a sense of competition. One key challenge for these types of applications is to automatically detect what the user is watching. A number of applications automatically identify the show and episode the user is watching based on the audio (e.g., IntoNow). After automatic identification, these applications broadcast the viewing information to others, so the user is checked into a certain program without his/her intervention.

As the world of social TV applications is changing fast, many of the apps mentioned here have been redesigned or incorporated into new apps. The most recent examples in this category are Viggle, which offers monetary rewards for checking into television shows, and SeeIT, which allows users to click on a tweet that mentions a TV show to directly watch it or program it. In 5 years time the landscape of broadcasting apps might look completely different (Fig. 5).

Many of these applications are Web based and have a mobile counterpart, making it easy to change status while watching TV. However, it is also possible to have a TV widget on a connected TV with the same functionality. The main communication modality for this category of applications is text based. Although the network reach includes friends, especially when linked to Facebook, the Twitter-like structure of followers makes it easy to include strangers in the network as well. What these applications are usually lacking is a presence feature (e.g., information about the activity of your friends, see section "Awareness"). As communication is not a core function, it is not really necessary to know if other users are online or not. The interaction is therefore mainly asynchronous, as there are no direct communication possibilities other than short status updates.

Congregation

Congregation and broadcasting both refer to the provision to users of means for sharing thoughts, comments, and impressions about television programs with a larger community. While broadcasting mostly concentrates on tools for users, congregation focuses on tools created and maintained by television channels or content producers. Such tools are intended for congregating viewers around media content or television channels and typically offer extra material and highly curated content. The target group of these applications is the followers and communities around media content (series, programs, teams. . .) (Fig. 6).

In some cases, games (e.g., NBA Real Time Fantasy) and other immersive activities are provided by the television channel or by individual followers of the show. In the past, successful approaches included the use of telephone calls for deciding the outcome of a show – Big Brother or the Eurovision song contest are good examples – but lately many television channels are providing specific Web pages with Facebook and Twitter updates and extra material for secondary screens.

Beamly, formerly known as Zeebox, is a current example of this category on congregation, allowing its users to "follow" celebrities or TV shows as well as chat with other viewers in so-called TV rooms. This type of applications gathers and aggregates comments related to a television program and renders them in either an external device (like a mobile phone or tablet) or overlaying the comments alongside the television content. In most of the cases such aggregation is done via an external channel, with no effect on the program. Some exceptions exist like NM2 (Ursu et al. 2008), where comments of the viewers were used for interactively

Fig. 6 Television channels and content producers take advantage of the active involvement of viewers for profiling their content (congregation)

affecting the storyline of a drama series, and Current TV's "Hack the Debate" that showed Tweets onscreen during the televised debates of the last US presidential election (Shamma et al. 2009).

The mobile phone, tablet, and laptop are the most commonly used devices, since it is more convenient to use than a television remote control as some typing is usually required. A salient feature of this category is the network reach (see section "Social Reach"), where large audiences of strangers congregate around a television program. Text tends to be the most common communication modality. Even though the comments are synchronized with the show, synchronization is not a key feature because time shifting is common and people might add/read comments whenever they want.

Conversation

A number of applications support direct communication between its users, enabling conversations about media content. Early TV-based research systems like Alcatel-Lucent's AmigoTV (Coppens et al. 2005) or Motorola's Social TV (Metcalf et al. 2008) allowed users to talk with each other using voice. Similarly, the first commercial social web TV applications Joost and Lycos Cinema enabled users to text chat with each other while watching online TV or movies. While instant messaging solutions allowed users to share videos while chatting (e.g., Windows Messenger and Zync (Shamma and Liu 2009) from Yahoo!), other Web-based applications like Watchitoo and YouTube Social also enabled talking and videoconferencing while watching the same content. A current example of this category is mashme.tv, a videoconferencing application that has a tool to integrate with YouTube. Videos can be played and watched synchronously while viewers are linked with each other through videoconferencing.

Most of these applications support only synchronous communication, although CollaboraTV (Nathan et al. 2008) also included asynchronous communication by letting users leave comments at specific moments during a television show. In all cases, the applications have presence features such as a buddy list so users can see who is available and if they are watching the same show or not. In addition to connecting living rooms for enabling conversation, a number of systems aim at connecting locations like the living room with the stadium where the football match (Hesselman et al. 2010) or the festival (Velt et al. 2015) is taking place. These services suffer from the low connectivity offered at crowded environments and are typically restricted to sharing photos and impressions, but further research is required in order to explore their limits (Fig. 7).

Most of the applications that offer conversation possibilities are limited to a (smaller) group of friends. The rationale behind this is probably that these people do not have the option to physically watch TV together (anymore), and social TV allows them to (re)create a social co-watching experience. These and other applications embody a category of applications where directly communicating with each other is the core social feature. While other social features are usually supported as well, such as sharing which program someone is watching, it is the synchronous communication that characterizes these applications the most.

Summary

This section has reviewed the five main types of activities supported by existing social television and online video applications. Based on an extensive survey, the identified activities are content selection, content sharing, broadcasting, congregation, and conversation. The next chapter will focus on another salient feature of these applications: the intended device of use.

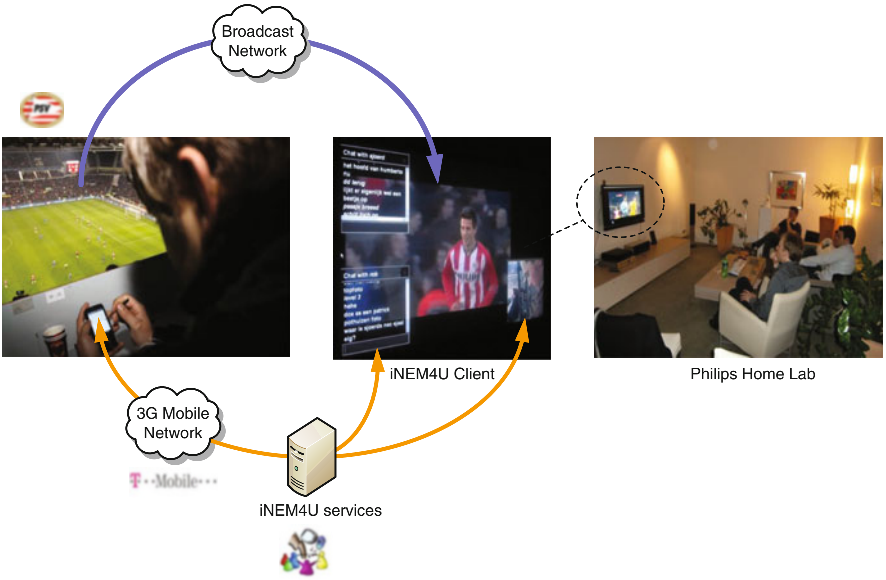


Fig. 7 Some applications enable direct communication between viewers watching the same content, even if they are at different environments (home and stadium in this case) (Hesselman et al. 2010)

Device

When you are designing an application for interactive television or online video, you have to take into account the possibilities and limitations of the target device you are designing for. In the early years of interactive television, the only device to take into account was the television set, and many articles and books have been written with guidelines (see, e.g., Gawlinsky 2003; Hansen 2006). Nowadays, with the many devices that can be used to watch TV or video as well as second and multiscreen applications for various devices (Obrist et al. 2015b), there are not only many more guidelines to take into account but it also requires a careful consideration of how to create a good balance between the different devices when used in conjunction with each other.

In the next sections, the most important aspects of each device to take into account will be discussed, as well as how to design applications in a multiscreen context.

Television

As the oldest device to watch audiovisual content on, the properties of a TV set have been widely studied in light of the design of interactive applications. This chapter does not want to provide a detailed description on how to design and develop applications for TV but will discuss which aspects to take into account when designing social TV experiences.

The most important aspect is to keep in mind that people watch television to be entertained, educated, or informed through audiovisual content, so any added interaction can potentially interfere with this goal. In general, people do not like if the video screen is reduced to make way for interactive features or if these features overlay too much of the video content (Geerts and De Grooff 2009b). A recommendation is therefore to keep the interactivity on the television screen minimal and brief. You can use a second screen (like a smartphone, tablet, or laptop) to offload some interactive features, although this also needs to be carefully considered to not distract from the main purpose of watching television.

A second property of television to take into account is the fact that it is a social medium. Due to its size, the television usually takes a central place in a household and is often being used to come together as a family and share watching a specific show. The television set is sometimes compared to an electronic hearth (Tichi 1991), and although many households own multiple television sets or other screen-based devices, the shared television is still one of the most popular devices to watch video content on (Nielsen 2015). For interactive applications this social character is an important aspect, as it impacts the type of interaction or interactive content that can be made available. Early applications for interactive television often focused on features aimed at individuals such as online banking or shopping, which do not match well with a social context. Even less conspicuous features such as voting mostly support individual interactions like casting only one vote, ignoring the complex social fabric of a household or group of friends (Bernhaupt et al. 2007). Each feature or application should look at how a group of people can be supported or at least consider how it will not interfere with a social context (Obrist 2009).

Finally, the television set limits the interaction modalities that are possible for interacting with the content. While the traditional remote control is still the most popular and convenient way of interacting with the TV, gesture and voice control are new ways of interacting that can offer benefits over the remote control (Vatavu et al. 2014; Dezfuli et al. 2014). Gestures can, for example, make it easier to perform small tasks like channel switching or turning the volume up and down, without having to reach for the remote control. However, longer interactions often get tiresome when using gestures. Voice control is better for inputting text when, for example, searching for a specific movie in a video on demand (VOD) catalog. Despite some benefits of these new interaction models, the remote control is still the most preferred interaction modality for most users (Vanattenhoven et al. 2014) (Fig. 8).

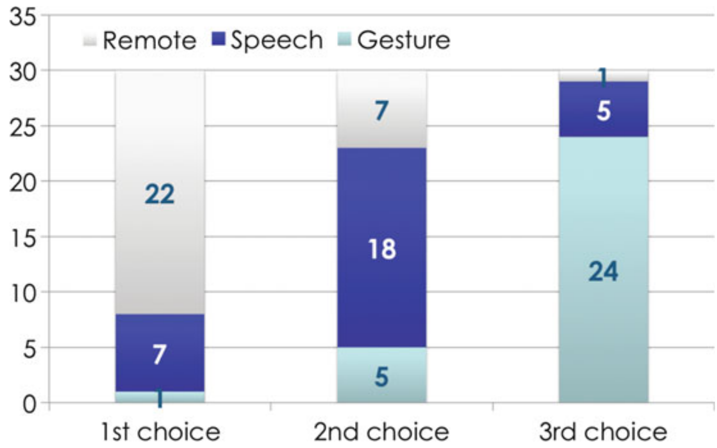


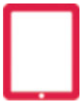
Fig. 8 Preference for interaction modes with the TV

Computer

Watching video and television programs on a desktop or laptop computer is increasingly popular, not in the least because of the success of (social) video platforms such as YouTube and more recently Netflix. However, while the design of interactivity on a television set has been widely studied, the design of watching video on a computer is less documented. Some research focuses on embedding video content in a corporate website (Schade 2014), but there is not much known about using the computer primarily as a platform for watching television (Barkhuus 2009).

Using a computer as a video-watching platform can be contrasted with using a television: with the computer, people are used to being highly interactive; it is by definition a personal device, and the interaction modalities are mainly through the keyboard or a mouse. This means that creating interactive applications to watch and socialize around video on a computer demands a completely different interaction model than on a television: it is possible to add more features alongside video watching, personalize the viewing experience, and use features that allow people to type a lot of text (using a keyboard) or select items from the screen (using the mouse). However, one should not ignore the properties of video content and how people want to watch online video. In that light it is relevant to see how YouTube has moved from a highly interactive model to a hybrid model where it is possible to play videos in a nonstop channel. This accounts for the fact that people can be highly interactive with videos on a computer at times but at other times want to use it more like a traditional television that plays videos and which can even be watched together with one or more people. Features like playing a video full screen (without interruptions from other programs), casting the screen to a television set, or letting users create a playlist that automatically keeps playing are some ways to simulate a TV experience on a computer.

Tablet

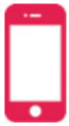


The introduction of tablet computers like the iPad has nearly revolutionized both research and industry around interactive TV, either as a device to watch video content on (Vanattenhoven and Geerts 2015b) or as a second screen accompanying the content on the main TV screen (Nandakumar 2014). Instead of trying to fit all interaction on the main television screen, it now became possible to offload some or most interaction to a tablet, freeing up space on the TV screen while still making it possible to offer interactive features. However, using a tablet for providing all interaction is not an easy solution, and again careful consideration should go into the design of interactive features.

Firstly, the screen real estate of a tablet is smaller than a TV, but since the screen is closer to the user, there is a risk that app developers will put too much content on it, which can distract viewers from what is happening on the main television screen. This is less of a problem for television genres that demand less continuous attention, such as sports or soap operas, but can pose conflicts for genres with a high plot structure such as movies or drama series. In Geerts et al. (2014), there is a detailed account of the implications of program genres for the design of social television systems. Secondly, a tablet, like a computer, is a personal device. However, while mostly being used by one person at a time, tablets are often shared within a household. Personalization is therefore possible, but shared use should also be supported. At the time of writing, Netflix has accounts for several household members but does not explicitly support group accounts, which would allow them to better tailor recommendations for, for example, watching as a couple. (Netflix does offer a kids account which has content and an interface tailored to children, but also here there are no shared recommendations for e.g. multiple children with different ages.) Finally, interacting with a tablet is mainly done through touch, which makes it possible to select directly on the screen as well as provides (limited) support for text input.

Besides using a tablet for interactivity, the tablet can also serve as the main screen or only screen to watch video content on. As with the television and the computer, the properties of a tablet dictate the design of the applications that will support or accompany video watching, and both highly interactive modes as well as passive viewing modes should be supported.

Smartphone



Not only tablets, also smartphones can be used as a second screen accompanying the television or to watch video content on. As the screen real estate is much smaller than on the other devices, this limits the type of interactions that are suitable for interacting with video content. Typical activities that can be implemented are voting, answering quiz questions, or checking into a favorite television show. The highly personal nature as well as the likelihood of multiple smartphones available in a household makes it a good candidate for multiuser applications such as playing against each other in a quiz show.

Smartphones are also increasingly being used to watch video content on, especially since screens are getting bigger and better in quality. Still, users prefer to watch long video content on bigger screens and use smartphones for shorter content or small clips (Geerts et al. 2008; Nielsen 2015).

Multiscreen

As discussed in this section, there are different types of devices that can be used to watch television or online video and/or interact with it. While some of them can be used as second screen devices, the reality is that a living room or any other place where TV or video is being watched is a multiscreen environment, and we need to think of how these different screens can cooperate and/or be used together to enhance the viewing experience while at the same time supporting the social viewing situation.

One issue is that of distraction: television and video content is increasingly watched in an environment where multiple screens are vying for a person's attention. Especially when designing applications that are meant to be used in conjunction with other devices, the attention between the different screens should be well balanced. For example, when attention is needed on one screen, the other screen should not actively catch the attention of the user. One solution is to use auditory or visual cues to manage the attention between screens (Neate et al. 2015). Conflicts between different screens should also be avoided, such as a video that starts playing on a second screen while the television is playing audiovisual content as well. When the devices are part of an ecosystem that is aware of these conflicts, it is possible to design solutions such as automatically pausing the video on one screen when video on another screen starts playing (Geerts et al. 2014).

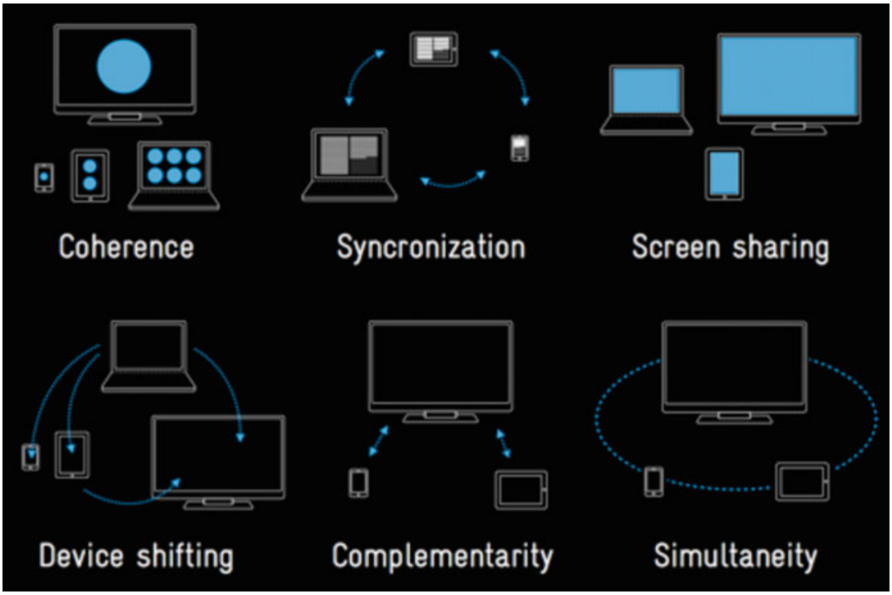


Fig. 9 Multiscreen patterns by Christophe Stoll (http://previous.precious-forever.com/2011/05/26/patterns-for-multiscreen-strategies/)

In order to consider the various ways that multiple screens can be used in conjunction with each other, the multiscreen patterns from Christophe Stoll are useful. He describes six patterns for multiscreen use, which are not mutually exclusive (Stoll 2011). The first pattern is "coherence," in which products or services should look and work coherently across devices, with features being optimized for each device or usage scenario. The "synchronization" pattern specifies that all devices should be in sync and be aware of and reflect changes happening in the other devices. "Screen sharing" refers to applications where multiple screens share the same source, and each screen shows either the same source or parts of it. With "device shifting" users can shift content from one device to another. The "complementarity" pattern can be applied to applications where devices complement each other, such as a second screen device that shows extra information about a television show. Finally, "simultaneity" refers to events happening at the same time, and the different devices can impact each other, like a voting app to play along with a quiz show (Fig. 9).

Summary

This section has discussed the different devices currently available which are being used either to watch television or video programming, or in conjunction with a television program, and how to take their properties into account when designing social applications for TV or online video. However, the technology landscape is changing rapidly, so let us have a peek look at how upcoming devices could change the TV and video experience.

In recent years, new screen-based devices are emerging which create new opportunities for possible interactions with TV or video content. One can imagine wearing Google Glass and getting personalized subtitles or being able to vote with an Apple Watch that notifies you with subtle tactile feedback when the next question comes up. As with the devices above, it will be necessary to take screen real estate, the personal character of the device, and the interaction modalities into account.

Although a screen might seem a necessity when designing for interactive television and online video, there are certain interactions that do not necessarily require a screen. Harboe et al. (2008), for example, used an Orb device to signal users with light and color codes when their friends or family at a remote location were watching TV, even when their own TV was off. Vatavu (2013) imagines a system where projectors are used to show content next to the television screen, so the main screen can stay uncluttered, and information is shown in the periphery. It is uncertain which types of devices will still be invented in the future, but many of them might hold interesting possibilities for interacting with TV and online video, potentially exploiting many more interaction modalities.

The next section will look at the way social interaction can be facilitated, a crucial aspect of many social TV and video applications.

Social Interaction

Another essential aspect of social television and online media services is the social interaction style that viewers use to interact with each other. The most relevant styles include but are not limited to: text, audio, video, nonverbal communication, and objects.

Very early socially aware media systems tried to replicate existing chat environments, by offering text communication capabilities while watching media (Abreu et al. 2002). While such services may work when consuming media in desktop computers or laptops (e.g., Zync) or when broadcasting and congregating using social networks, they are not appropriate for directly communicating when “leaning back” in the living room in front of the television set. For this reason, the research community focused as well on audio-based solutions that tried to replicate the living room metaphor with people conversing about a television program (Coppens et al. 2005; Metcalf et al. 2008). There has been some debate over whether text chat would be better than audio chat and in which situations. Based on a lab experiment, Geerts (2006) argued that voice chat is considered more

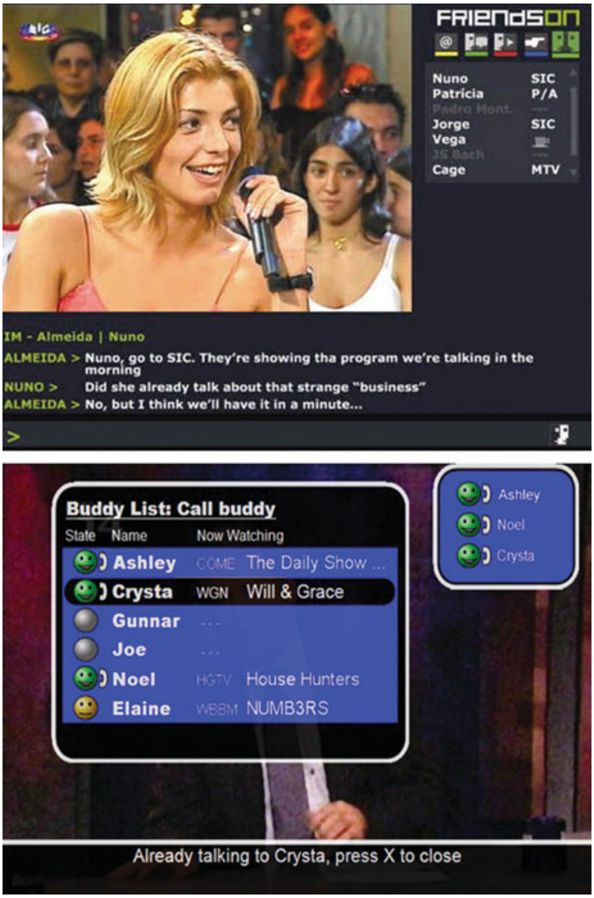


Fig. 10 Direct communication between viewers at different locations can be supported using text and audio

natural and direct and makes it easier to keep attention to the program, while text chat is more preferred by users having more experience with chatting on computers. On the other hand, Tullio et al. (2008) held a field trial for a week with a system that included both voice and text chat and concluded that text was used more often than voice communication, although both were usually used in combination. This last statement shows that it is best to offer multiple modes of communication, so users can choose which mode to use depending on the situation (Fig. 10).

Apart from text and audio, video can be used as the social interaction style. The main problem of this approach is that the visual attention of the viewer has to be shared between two different video feeds (see previous section). While this does not have to be a problem in itself, it is recommended only for certain specific scenarios and situations. Other less traditional interaction styles offered by existing services include nonverbal cues, like gestures that show the level of interest of the viewers during the show, and objects, like when directly sharing media content with others (videos or parts of videos). An interesting example of using gestures can be found in

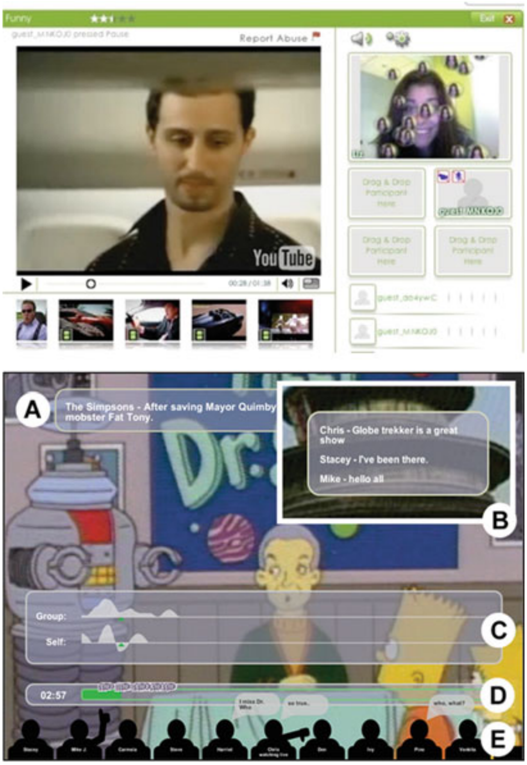


Fig. 11 Other social interaction styles include video and nonverbal communication, such as gestures or emoticons

Vatavu (2015). Using a representation of viewers called "audience silhouettes," he recreates the look and feel of Mystery Science Theater 3000 where members of the audience are seen at the bottom of a movie screen. (See e.g. *Mystery Science Theater 3000*: *25th Anniversary Edition*. (2013). Shout Factory.) Vatavu's audience silhouettes are captured live from viewers' movements, which they can use to communicate with each other through nonverbal means. A similar approach is taken in the CollaboraTV system (Nathan et al. 2008), where the characters at the bottom of the screen have more limited interaction capabilities and can use smileys or a limited set of hand gestures. Interestingly, the latter system also includes text interaction, synchronously as well as asynchronously (see section "Sync") (Fig. 11).

In summary, social television and online video applications should optimally support the communication process between users. They should support different levels of communication depending on the environment where the viewers are located, the content that is watched, and the intention of the users. Such support can be provided using a variety of social interaction styles such as text, audio, video, nonverbal communication, or objects.

This section discussed the different alternatives that can be offered to users to interact between them while watching television and online video content. The next section will focus on the level of awareness of others that services can provide.

Awareness

When designing social applications, awareness of other users in the system as well as what they are doing is crucial to enable the possibility to interact with each other in various ways as discussed in the previous section. In this section we will discuss three levels of awareness to include in social TV and online video applications: presence, identity, and activity awareness.

Presence

The lowest level of awareness is presence awareness. This means that it is possible to see that there are other people present but without further information about who they are or what they are doing. This type of awareness is ideal for large groups of people where the aim is not to have people to interact but rather that viewers are aware of other viewers.

A good example is the Zeebox app (now rebranded as Beamly) that shows the popularity of a TV channel by indicating how large the audience is at any given time. This information can be useful in deciding which show to watch, based on popularity (Fig. 12).

Identity

At the next level of awareness, not only presence is visible but also the identity of the people that are using the application. This is not so useful for large groups, as the information would quickly become too much to be shown on a screen, but is more relevant for displaying information about smaller groups where the users know each other or where interaction is stimulated such as on forums or chat groups.

An example of identity awareness, also clearly showing the difference with presence awareness, can be seen when you are logged into the Zeebox app that then shows which friends of a specific user are watching which show. This type of information can be much more useful than the general popularity of a television show, as people might prefer to watch content that is popular with their friends (Fig. 13).

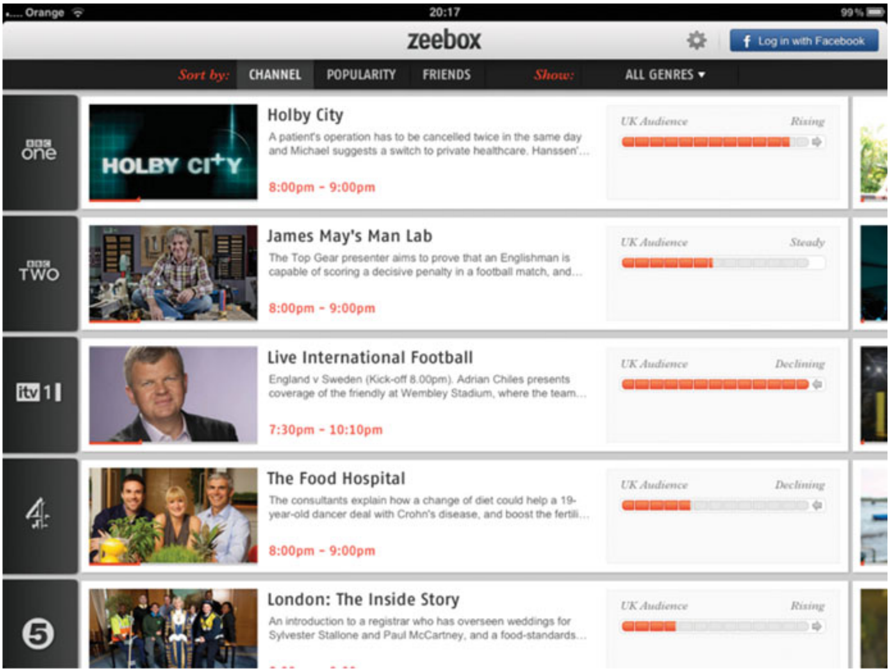


Fig. 12 Presence awareness in the Zeebox app

Activity

Finally, the highest level of awareness is where it is not only possible to see presence and identity but also the activities of other users. These activities can be varied, from simply showing if people are available to chat or not to visualizations that show actual movements of other users (Vatavu 2015).

Especially in systems where active communication or even collaborations are stimulated between users, this type of awareness is important for supporting interaction and making people aware if it is appropriate to contact someone or not or to react to specific actions of other users (Fig. 14).

Summary

As discussed in this section, there are different levels of awareness offering viewers a sense of other people that are online, who they are and what they are doing. This highly facilitates social interaction, as it indicates the possibility to interact with

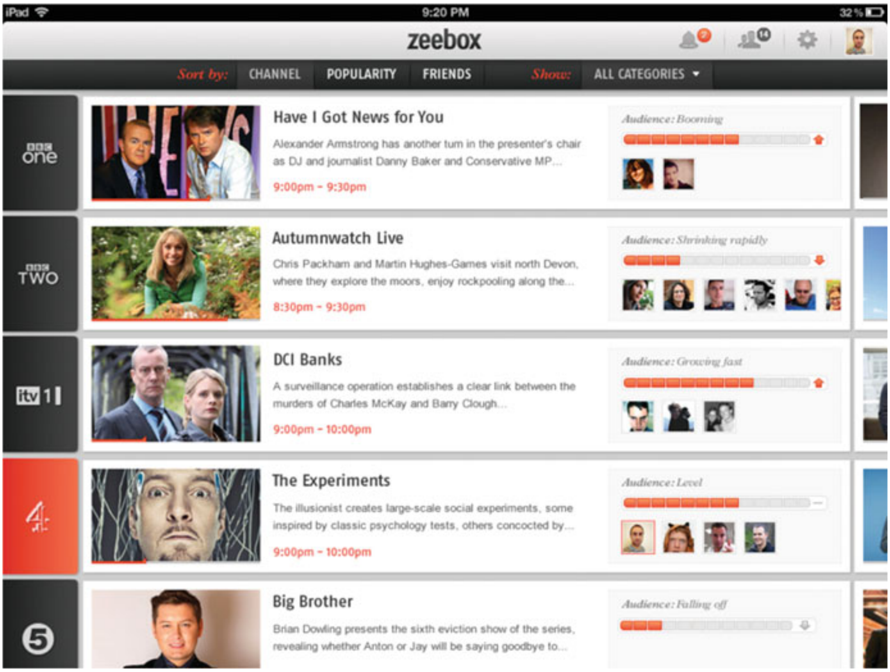


Fig. 13 Identity awareness in the Zeebox app

Fig. 14 Activity awareness in Motorola’s Social TV

each other. In the next section we will discuss how synchronization has an impact on the design of social TV and online video systems.

Sync

Social communication between audiences or individuals consuming media can happen either synchronously or asynchronously. In some cases the viewers expect to be able to interact with each other in real time, like in broadcasted sport events, quiz shows, and elimination shows (e.g., Big Brother). In all those cases, viewers' participation, like shouting after a goal of your favorite team, is an intrinsic part of the media experience. One core research problem for these cases is to achieve acceptable synchronization between the viewers conversing, given that most probably they are using different networks and devices (Kooij et al. 2014; Mekuria et al. 2012). In other cases, synchronicity is not that important. There are some services, for example, that allow commenting when watching media. These comments and impressions are then linked to the media content as extra material, which can be experienced by others as an overlay of the television content or in their secondary screens. These cases enable social communication between viewers but are less strict regarding the availability of viewers.

Media synchronization is a current and relevant challenge for the research community that requires multidisciplinary work (Stokking et al. 2015; Boronat et al. 2012). Apart from the more user experience factors (e.g., acceptable thresholds), novel technical solutions and standards are needed (Rainer et al. 2015; Montagud et al. 2013, 2014; Vaishnavi et al. 2011) (Fig. 15).

In the past, some applications offered strangers the option to directly communicate with each other on a channel where the content was not the same and was not synchronized (e.g., Joost's "channel chat") (Geerts and De Grooff 2009a). The non-synchronized content made it however difficult to find common ground with strangers to talk about the same things while watching, so it is doubtful if this combination would lead to successful communication. One design guideline is to discourage situations of direct communication between people watching different content or non-synchronized content. Another design guideline is to always provide the right level of synchronicity between the viewers. Such level of synchronization would be dependent on a number of factors such as type of content and interaction style (Geerts et al. 2008). For example, viewers will be more engaged in

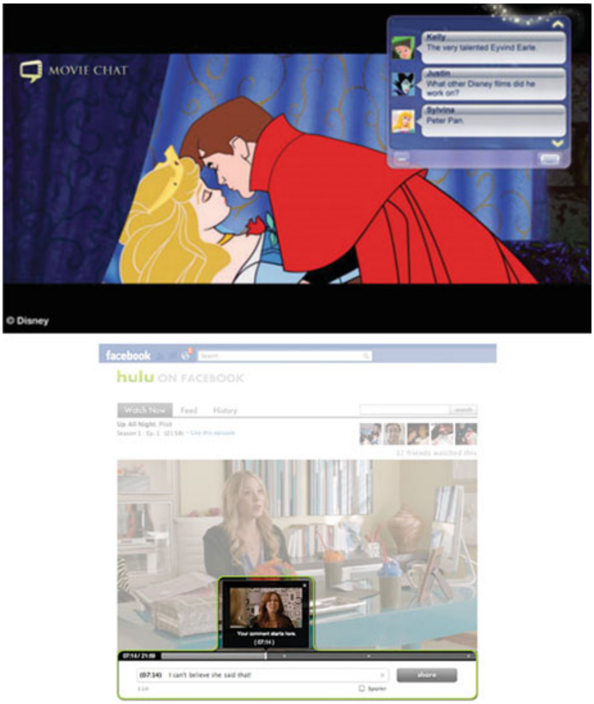


Fig. 15 Depending on the application, and its purpose, synchronous or asynchronous communication can be provided

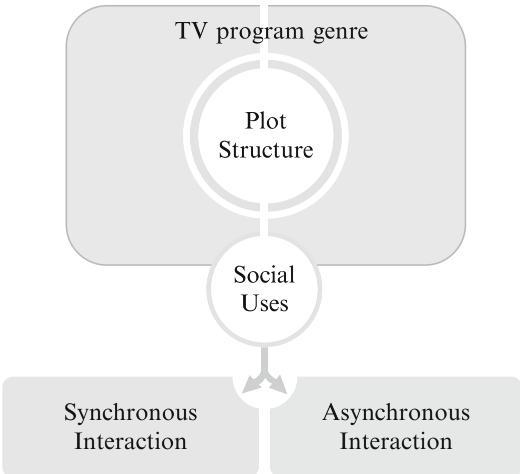


Fig. 16 The type of synchronization is dependent on the media content (and its genre) and on the purpose of the application

conversations with others for some types of content than for others, depending on the narrative level and the plot structure. Moreover, depending on the interaction style (audio or text) the level of synchronization will significantly vary (Geerts et al. 2011) (Fig. 16).

The next section will focus on social reach, the last of the salient features of social television and online video applications. The social reach refers to the relationships of the users involved in social communication while watching content.

Social Reach

The last part of the framework for social television and online video focuses on the types of relationships between users of an application. Five types of relationships are identified that have an impact on how to design social applications for TV or online video: household, family, friends, acquaintances, and strangers. The definition of each of these relationships is sometimes stretched a bit, as will become clear in this section, but the defining factor is that with each type of relationship the reach of the application is increased. Therefore the term social reach is used to indicate how the relationship type impacts the scope of the applications.

Household

Household refers to several people that have a close or intimate relationship and typically watch television or video content collocated, i.e., in the same room. Examples are couples or nuclear families with children, but also a group of friends watching together can be considered a household in the framework.

When designing social applications for a household, you need to take into account the social dynamics within the same room and provide applications that support this. Some examples of many possible instances of social dynamics are parental rules for television use, informal agreements over who controls the remote control, people walking in or leaving the room while watching, or different preferences for which content to watch. At the most basic level, this means making sure that any interactivity that you provide does not disturb other members of the household. If a group of friends is watching TV and all of a sudden a chat box with a private message pops up on TV, this can cause embarrassing situations (see Regan and Todd (2004) for a good example of this). If you want to be more actively stimulating social interaction within a household, then you can create applications that involve the different members by, for example, letting them play against each other or providing the option to vote individually and then see a common score (Leenheer et al. 2015).

Family

Family refers to people that have a close relation but do not live under the same roof. Apart from family, also close friends could be part of this group.

The fact that they do not live under the same roof and thus do not watch TV or online video collocated, like with the household category, means that different types of interactions need to be considered. Instead of looking at ways to support social dynamics in the same room, you need to consider how to enable remote communication, for example, through text or voice chat, but even video communication is possible. Since the relationship is quite close and typically the group is rather small, direct communication is one mode of interaction that can be implemented, but other activity types like being pitched against each other in games are also good ways to stimulate social interaction between family members and good friends.

Friends

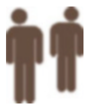
Friends are considered as those people that have a friendly relationship with each other and have one or more things in common like work or hobbies but are not that close with each other. Typically, the category “friends” is much bigger than the category “family” (including close friends), which means that social TV or online video applications for this category need to provide interaction modes or activities that are suitable for larger groups of people.

Direct communication is less convenient with larger groups, so social interaction will be more indirect. Examples are social electronic programming guides (EPGs) that indicate which of your friends have watched which shows, social recommendation systems, or aggregated scores of groups of friends that are playing along with a quiz show.

Acquaintances

Acquaintances are even less close than friends and are people that have a weak relationship with each other and do not have much in common. These are typically relationships where people do not want to share too many details or information with each other or are merely useful for instrumental reasons. Direct communication or interaction is most of the times not useful here, but a person may choose to follow one or more acquaintances if their tastes match and see which television show, movie, or video they recommend. However, Dezfuli et al. showed that the program genre has an influence on the willingness of people to directly interact with each other. For example, acquaintances that share a similar hobby or interest would like to communicate while watching a quiz or documentary about that shared interest (Dezfuli et al. 2011).

Strangers



Finally, strangers are people that you do not have a relationship with and that you may encounter by chance. While it might seem strange to develop social TV or online video applications for strangers, in real life there are occasions where you can watch TV together with strangers such as in a sports bar. The fact that other people are present, even if you do not know them, can enhance the atmosphere and create a feeling of sociability.

If you translate this to online social TV applications, it is worthwhile to visualize the presence of other people and their activity to create a feeling of being in a group of people watching a sports match, where the cheers to support the different teams can be seen or heard. This way, watching online TV or video with strangers can become a sociable experience (Weisz et al. 2007).

Summary

In this section different types of relationships were discussed with their impact on the design of social television and online video applications. When designing for a household, family, friends, acquaintances, or strangers, different features and design decisions have to be taken into account. As shown by Dezfuli et al. (2011), and similar to earlier discussions related to synchronization or devices, the program genre again plays an important role in the type of interaction that users prefer with different type of relationships. As a whole, this chapter shows the complex interrelations between the different aspects which social TV and online video applications constitute of and which the framework presented here tries to address.

The next section will conclude this chapter with a look at some current developments and the future of social TV and online video.

Conclusion

In recent years we have witnessed a radical transformation of the way we find and experience media content. This chapter argues that socialization around the content is one of the most relevant aspects of such transformation. For this reason this chapter provided a comprehensive, but approachable, framework for understanding the new media landscape and its novel social characteristics. The framework is composed of a number of atomic categories (activity, interactive device, social interaction, sync, awareness, and social reach) and helps in easily and consistently classifying social television and online media services by their most salient features.

Nevertheless, the evolution of online video and television is not restricted to their social aspects (Obrist et al. 2015a). We can distinguish four broad areas of research beyond sociability: content creation, media recommendations, device ecosystem, and feedback (Fig. 17).

Firstly, developing valued content is key, where truly interactive and transmedia experiences maximize the potential of growing consumer trends (Nandakumar 2014; Eversman et al. 2015; Silva et al. 2015). The mix of broadcast content with Internet content offers promising possibilities for this, with emerging standards such as HbbTV (Hybrid Broadcast Broadband Television) providing frameworks that enable these links for producers, for example, through synchronization protocols. This creates exciting opportunities for enriching TV programs with content from the Web but also to enable completely new media experiences (Nixon et al. 2013; Zorrilla et al. 2015).

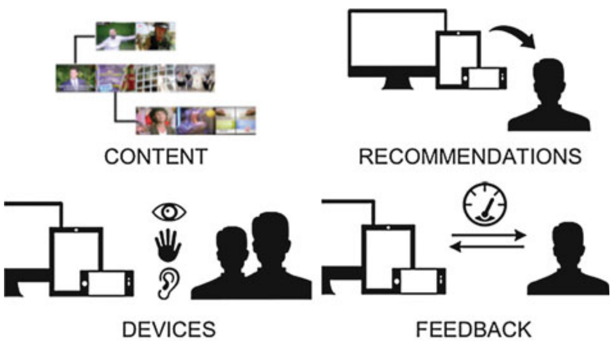


Fig. 17 The upcoming key challenges on the field of online media and television include content creation and storytelling, content recommendation and data science, the seamless usage of a heterogeneous set of devices for consuming content, and real-time monitoring of feedback

Secondly, understanding the media consumer is vital. A popular research topic in other domains and user recommendations are a prime area of research which raise specific challenges for broadcast viewing (Vanattenhoven and Geerts 2015a). Group recommendations can build on television as a social medium, and need further research, as well as methods for gathering accurate data to feed recommendation algorithms. Here we can learn from the advancements in data science.

A third area of research focuses on the complex ecosystem of devices people use to consume and to interact with media content. The wide range of devices available in many households is radically changing. It is important to gain insight into which devices people are using while watching TV and for which purposes, so this can be taken into account when creating programs and applications for TV or online video (McGill et al. 2014, 2015; Geerts et al. 2014). In contrast, we can design new media experiences that take into account these multiple devices and create more immersive and novel multisensory experiences, delivering richer and more meaningful content (Chambel et al. 2014; Obrist et al. 2014; Obrist 2015).

Finally, gathering audience feedback and insight into the impact of television and online video to consumers is vital. New technologies allow for more detailed and varied feedback (e.g., social media) from diverse audiences (Karahasanović and Heim 2015; Holz et al. 2015; Chaney et al. 2014). However, measuring who is doing what and how consumption is occurring is still a big challenge, given the complexities of multi-device and social viewing habits.

Daily we are witnessing new developments about how media content is found and experienced. Deep transformations have occurred since the days of the professionally crafted, linear, noninteractive movies and television programs. This chapter tried to shed some light on the implications, mostly focusing on the social and interactivity perspectives. Still, there are many unexplored challenges ahead in this young and fascinating field, such as the provision of truly connected experiences (Cesar et al. 2014a). One invaluable resource, for scholars and practitioners, for better understanding future developments is the yearly ACM International Conference on Online Video and Television (ACM TVX).

Recommended Reading

J. Abreu, P. Almeida, V. Branco, 2BeOn: interactive television supporting interpersonal communication. *Proceedings of the Eurographics Workshop on Multimedia* (2002), pp. 199–208

L. Barkhuus, Television on the internet: new practices, new viewers. *Extended Abstracts on Human Factors in Computing Systems* (ACM CHI) (2009), pp. 2479–2488

R. Bernhaupt, M. Obrist, A. Weiss, E. Beck, M. Tscheligi, Trends in the living room and beyond. *Proceedings of EuroITV* (2007), pp. 146–155

R. Bernhaupt, D. Wilfinger, A. Weiss, M. Tscheligi, An ethnographic study on recommendations in the living room: implications for design of iTV recommender systems. *Proceedings of EuroITV* (2008), pp. 92–101

F. Boronat, M. Montagud, H. Stokking, O. Niamut, The need of inter-destination synchronization for emerging social interactive multimedia applications. IEEE. Commun. Mag. **50**(11), 150–158 (2012)

P. Cesar, D. Geerts, Understanding social TV: a survey. *Proceedings of the Networked and Electronic Media Summit* (NEM Summit) (2011)

P. Cesar, D. C. Bulterman, A. J. Jansen, Usages of the secondary screen in an interactive television environment: control, enrich, share, and transfer television content. *Proceedings of EuroITV* (2008), pp. 168–177

P. Cesar, D.C.A. Bulterman, J. Jansen, D. Geerts, H. Knoche, W. Seager, Fragment, tag, enrich, and send: enhancing the social sharing of videos. ACM Trans. Multimed. Comput. Commun. Appl. **5**(3), 19 (2009)

P. Cesar, R. Kaiser, M.F. Ursu, Toward connected shared experiences. IEEE. Comput. **47**(7), 86–89 (2014a)

P. Cesar, M. Obrist, E.F. Churchill, T. Bartindale, ACM international conference on interactive experiences for television and online video (ACM TVX 2014). IEEE. MultiMedia. **21**(4), 112–c3 (2014b)

T. Chambel, P. Viana, V. M. Bove, S. Strover, G. Thomas, ImmersiveMe'14: 2nd ACM international workshop on immersive media experiences. *Proceedings of the ACM International Conference on Multimedia* (ACM MM) (2014), pp. 1255–1256

A.J.B. Chaney, M. Gartrell, J.M. Hofman, J. Guiver, N. Koenigstein, P. Kohli, U. Paquet, A large-scale exploration of group viewing patterns. *Proceedings of the ACM international conference on Interactive experiences for TV and online video* (ACM TVX) (2014), pp. 31–38

T. Coppens, F. Vanparijs, K. Handekyn, AmigoTV: a social TV experience through triple-play convergence. *Alcatel-Lucent white paper* (2005)

N. Dezfuli, M. Khalilbeigi, M. Mühlhäuser, D. Geerts, A study on interpersonal relationships for social interactive television. In *Proceedings of EuroITV* (2011), pp. 21–24

N. Dezfuli, M. Khalilbeigi, J. Huber, M. Özkorkmaz, M. Mühlhäuser, PalmRC: leveraging the palm surface as an imaginary eyes-free television remote control. Behav. Inform. Technol. **33** (8), 829–843 (2014)

N. Ducheneaut, R.J. Moore, L. Oehlberg, J.D. Thornton, E. Nickell, SocialTV: designing for distributed, social television viewing. Int. J. Hum. Comput. Interact. **24**(2), 136–154 (2008)

D. Eversman, T. Major, M. Tople, L. Schaffer, J. Murray, United universe: a second screen transmedia experience. *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video* (ACM TVX) (2015), pp. 173–178

M. Gawlinski, *Interactive Television Production* (Focal Press, Oxford, 2003)

D. Geerts, Comparing voice chat and text chat in a communication tool for interactive television. *Proceedings of the Nordic Conference on Human-computer Interaction* (2006), pp. 461–464

D. Geerts, P. Cesar, Interaction design for online video and television. *Extended Abstracts of the SIGCHI Conference on Human Factors in Computing Systems* (ACM CHI) (2015), pp. 2463–2464

D. Geerts, D. De Grooff (sup.), *Sociability Heuristics for Interactive TV. Supporting the Social Uses of Television*. (PhD) (2009a)

D. Geerts, D. De Grooff, Supporting the social uses of television: sociability heuristics for social TV. *Proceedings of the International Conference on Human Factors in Computing Systems* (ACM CHI) (2009b), pp. 595–604

D. Geerts, P. Cesar, D. Bulterman, The implications of program genres for the design of social television systems. *Proceedings of the international Conference on Designing interactive User Experiences For TV and Video* (2008), pp. 71–80

D. Geerts, I. Vaishnavi, R. Mekuria, O. van Deventer, P. Cesar, Are we in sync?: Synchronization requirements for watching online video together. *Proceedings of the International Conference on Human Factors in Computing Systems* (ACM CHI) (2011), pp. 311–314

D. Geerts, R. Leenheer, D. De Grooff, J. Negenman, S. Heijstraten, In front of and behind the second screen: viewer and producer perspectives on a companion app. *Proceedings of the ACM international conference on Interactive experiences for TV and online video* (ACM TVX) (2014), pp. 95–102

V. Hansen, Designing for interactive television v 1.0. BBCi & interactive tv programmes (BBC, 2006). Retrieved from http://www.bbc.co.uk/guidelines/futuremedia/desed/itv/itv_design_v1_2006.pdf

G. Harboe, In search of social television. in *Social Interactive Television: Immersive Experiences and Perspectives*, eds. by P. Cesar, D. Geerts, K. Chorianopoulos (IGI Global, 2009)

G. Harboe, C.J. Metcalf, F. Bentley, J. Tullio, N. Massey, G. Romano, Ambient social Tv: drawing people into a shared experience. *Proceedings of the International Conference on Human Factors in Computing Systems* (ACM CHI) (2008), pp. 1–10

C. Hesselman et al., Sharing enriched multimedia experiences across heterogeneous network infrastructures. Commun. Mag. **48**(6), 54–65 (2010)

C. Holz, F. Bentley, K. Church, M. Patel, "I'm just on my phone and they're watching TV": quantifying mobile device use while watching television. *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video* (TVX) (2015), pp. 93–102

A. Karahasanović, J. Heim, Understanding the behaviour of online TV users. Pers. Ubiquit. Comput. **19**, s 839–s 852 (2015)

W.J. Kooij, H.M. Stokking, R. van Brandenburg, P.-T. de Boer, Playout delay of TV signals: measurement system design, validation and results. *Proceedings of the ACM international conference on Interactive experiences for TV and online video* (ACM TVX) (2014), pp. 23–30

R. Leenheer, D. Geerts, J. Vanattenhoven, Learning lessons for second screen from board games. *Proceedings of the ACM international conference on Interactive experiences for TV and online video* (ACM TVX) (2015), pp. 143–148

R. Mekuria, P. Cesar, D.C.A. Bulterman, Digital TV: the effect of delay when watching football. *Proceedings of EuroITV* (2012), pp. 71–74

C. Metcalf, G. Harboe, J. Tullio, N. Massey, G. Romano, E.M. Huang, F. Bentley, Examining presence and lightweight messaging in a social television experience. ACM Trans. Multimed. Comput. Commun. Appl. **4**(4), 27 (2008)

M. McGill, J. Williamson, S.A. Brewster, How to lose friends & alienate people: sharing control of a single-user TV system. *Proceedings of ACM international conference on Interactive experiences for TV and online video* (ACM TVX) (2014), pp. 147–154

M. McGill, J. Williamson, S.A. Brewster, A review of collocated multi-user TV. Pers. Ubiquit. Comput. **19**, 743–759 (2015)

M. Montagud, F. Boronat, H. Stokking, Early event-driven (EED) RTCP feedback for rapid IDMS. *Proceedings of the ACM International Conference on Multimedia* (ACM MM) (2013), pp. 323–332

M. Montagud, F. Boronat, H. Stokking, P. Cesar, Design, development and assessment of control schemes for IDMS in a standardized RTCP-based solution. Comput. Netw. J. **70**(9), 240–259 (2014)

A. Nandakumar, J. Murray, Companion apps for long arc TV series: supporting new viewers in complex storyworlds with tightly synchronized context-sensitive annotations. *Proceedings of the ACM international conference on Interactive experiences for TV and online video* (ACM TVX) (2014), pp. 3–10

M. Nathan, C. Harrison, S. Yarosh, L. Terveen, L. Stead, B. Amento, CollaboraTV: making television viewing social again. *Proceeding of the international Conference on Designing interactive User Experiences For TV and Video* (2008), pp. 85–94

T. Neate, M. Jones, M. Evans, Mediating Attention for Second Screen Companion Content. *Proceedings of the International Conference on Human Factors in Computing Systems* (ACM CHI) (2015), pp. 3103–3106

Nielsen, *Shifts in Viewing: The Cross-Platform Report Q2 2014* (Sept 2014)

Nielsen, *Screen Wars: The Battle for Eye Space in a Tv-Everywhere World* (March 2015)

L. Nixon, M. Bauer, C. Bara, Connected media experiences: Web based interactive video using linked data. *Proceedings of the international conference on World Wide Web (companion)* (2013), pp. 309–312

M. Obrist, Special issue on personal interactive (TV) environments. Multimedia Systems **15**(3), 125–126 (2009)
M. Obrist, Multi-sensory media experiences. *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video* (ACM TVX) (2015), pp. 221–221
M. Obrist, A. Tuch, K. Hornbæk, Opportunities for odor: experiences with smell and implications for technology. *Proceedings of the International Conference on Human Factors in Computing Systems* (ACM CHI) (2014), pp. 2843–2852
M. Obrist, P. Cesar, D. Geerts, T. Bartindale, E.F. Churchill, Online video and interactive TV experiences. *Interactions* **22**(5), 32–37 (2015a)
M. Obrist, P. Cesar, S. Basapur, Forward to the theme issue on interactive experiences for television and online video. Pers. Ubiquit. Comput. **19**, 858 (2015b)
B. Rainer, S. Petscharnig, C. Timmerer, Merge and forward: self-organized inter-destination multimedia synchronization. *Proceedings of the ACM Multimedia Systems Conference* (ACM MMsys) (2015), pp. 77–80
T. Regan, I. Todd, *Media Center Buddies: Instant Messaging Around a Media Center – Microsoft Research. TechReport MSR-TR-2004-47* (2004), Retrieved 28 Jul 2015, from http://research.microsoft.com/apps/pubs/default.aspx?id=70063
A. Schade, *Video Usability* (2014), Retrieved 27 Jul 2015, from http://www.nngroup.com/articles/video-usability/
D.A. Shamma, Y. Liu, Zync with me: synchronized sharing of video through instant messaging, in *Social Interactive Television: Immersive Experiences and Perspectives,* eds. by P. Cesar, D. Geerts, K. Chorianopoulos. (IGI Global, 2009)
D.A. Shamma, L. Kennedy, E.F. Churchill, Tweet the debates: understanding community annotation of uncollected sources. *Proceedings of the SIGMM Workshop on Social Media* (2009), pp. 3–10
P. Silva, Y. Amer, W. Tsikerdanos, J. Shedd, I. Restrepo, J. Murray, A game of thrones companion: orienting viewers to complex storyworlds via synchronized visualizations. *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video* (ACM TCX) (2015), pp. 167–172
H. Stokking, P. Cesar, F. Boronat, M. Montagud, Media synchronization workshop. *ACM International Conference on Interactive Experiences for Television and Online Video* (ACM TVX) (2015), pp. 215–219
C. Stoll, Multiscreen patterns|precious, strategic design & visual language (2011). Retrieved from http://previous.precious-forever.com/2011/05/26/patterns-for-multiscreen-strategies/
C. Tichi, *Electronic Hearth: Creating an American Television Culture*, 1st edn. (Oxford University Press, New York, 1991). Hardback edition
J. Tullio, G. Harboe, N. Massey, Investigating the use of voice and text chat in a social television system. *Proceedings of EuroITV* (2008), pp. 163–167
M.F. Ursu et al., Interactive TV narratives: opportunities, progress, and challenges. ACM Trans. Multimed. Comput. Commun. Appl. **4**(4), 25 (2008)
I. Vaishnavi, P. Cesar, D. Bulterman, O. Friedrich, S. Gunkel, D. Geerts, From IPTV to synchronous shared experiences challenges in design: distributed media synchronization. Signal Process. Image Commun. **26**(7), 370–377 (2011)
J. Vanattenhoven, D. Geerts, Contextual aspects of typical viewing situations: a new perspective for recommending television and video content. Pers. Ubiquit. Comput. **19**, 761–779 (2015a)
J. Vanattenhoven, D. Geerts, Broadcast, video-on-demand, and other ways to watch television content: a household perspective. *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video* (ACM TVX) (2015b), pp. 73–82
J. Vanattenhoven, D. Geerts, D. De Grooff, Television experience insights from HbbTV. *2nd International Workshop on Interactive Content Consumption* at TVX 2014 (2014)
R.-D. Vatavu, There's a world outside your TV: exploring interactions beyond the physical TV screen. *Proceedings of EuroITV* (2013), pp. 143–152

R.-D. Vatavu, Audience silhouettes: peripheral awareness of synchronous audience kinesics for social television. *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video* (ACM TVX) (2015), pp. 13–22

R.-D. Vatavu, I.-A. Zaiti, Leap gestures for TV: insights from an elicitation study. *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video* (ACM TVX) (2014), pp. 131–138

R. Velt, S. Benford, S. Reeves, M. Evans, M. Glancy, P. Stenton, Towards an extended festival viewing experience. *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video* (ACM TVX) (2015), pp. 53–62

J.D. Weisz, S. Kiesler, H. Zhang, Y. Ren, R.E. Kraut, J.A. Konstan, Watching together: integrating text chat with video. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (ACM CHI) (2007), pp. 877–886

D. Williams, M.F. Ursu, P. Cesar, K. Bergström, I. Kegel, J. Meenowa, An emergent role for TV in social communication. *Proceedings of EuroITV* (2009), pp. 19–28

M. Zorrilla, N. Borch, F. Daoust, A. Erk, J. Flórez, A. Lafuente, A web-based distributed architecture for multi-device adaptation in media applications. Pers. Ubiquit. Comput. **19**, 803–820 (2015)

The Interactive TV Experience: Where We Came From and Where We Are Going

45

Michael J. Darnell

Contents

M.J. Darnell (✉)
User Experience Center America – Visual Displays, Samsung Research America, Mountain View, CA, USA
e-mail: mike.darnell@samsung.com

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_52

Abstract

Major events in the history of TV, especially in the USA, are reviewed from the perspective of the TV viewer's experience. Interactivity with TV is considered among these events. General TV experience themes that emerge from these historical events are discussed in this chapter as to their relationship to people's motivations for watching TV. The emergent themes include improvements in TV picture quality and increases in the amount of TV program choices available to the TV viewer. Additional emergent themes include increased ease and convenience of controlling the TV and increased information about the TV program. Finally, there is the theme of enabling the TV viewer to participate interactively in the TV experience. These themes are used as a basis to envision the future of TV experiences.

Keywords

DVR • EPG • Interactive TV • iTV • OTT • History of television • Picture quality • PVR • Remote control • User experience • VOD

Introduction

Interactive TV (iTV) is the approach to TV where the TV viewer is able to communicate with and influence the outcome of a TV program rather than simply view it. Notable reviews have been written on iTV (Carey 1996; Jensen 2008), particularly providing the perspective of technology and business. These reviews have described iTV as having a history of partial successes at best. The blame is placed largely on inadequate technology or unprofitable business models. The perspective taken in this chapter is a bit different, emphasizing the point of view of people's "user experience" with TV. User experience is defined as a person's overall experience with the TV, including how well it meets their needs and how easy and pleasing it is to use. A driving question is: How have the developments in TV in general and the interaction with the TV in particular affected and been affected by the TV viewers' experience?

A seminal study investigated people's primary motivations for watching TV (Lee and Lee 1995). The motives for watching TV can be summarized as follows:

1. To have a familiar routine
2. To relax, escape, and improve one's mood

3. To be informed and learn new things
4. To learn from and about other people
5. To socialize
6. To experience excitement, suspense, and an engrossing different world

This chapter aims to understand if the developments in TV that fulfilled these primary motivations were more successful than those that did not. If so, these primary motivations might serve as a good basis for understanding the past and predicting the future of TV. Human behavior is slow to change compared to technology. Thus, people's motivations for watching TV in the early days of TV may still be true today.

What Is Interactive TV?

Modern "smart TVs" have features similar to Internet-connected computers. People can browse the Internet, play games, access social media, and read and send email on their TV. Since most TVs do not have efficient input devices for these activities, most people do these activities only to a limited extent on the TV in part because the activities are more difficult to do on a TV than on a computer (Cesar et al. 2008b). Although these activities are interactive and are done on the TV screen, they are not generally what is meant by iTV.

The term iTV is used for TV applications that present additional information on the TV that is synchronized with TV programs (Bachmayer et al. 2009). For example, the TV viewer may be able to browse through football stats for a sports match on TV they are currently watching. This additional information and the broadcasted sports match share the TV screen at the same time. Interactive TV might also refer to apps that present alternate camera feeds and/or audio feeds for a game, giving the viewer the option for a more personalized experience (Zorik et al. 2013). Other iTV apps may present the Twitter feeds from people that are watching the same live game.

Interactive TV also encompasses apps by which people can "vote" or provide comments through their TV that influence the outcome or content of a live TV program (Quico 2003). Interactive TV includes applications by which people at home can play along with game shows on live TV. These activities are characterized by a "return path," by which people at home are able to influence the outcome of the TV program as well as to simply select from among options that the TV program offers.

These particular forms of iTV have never really caught on in the USA, although communicating with a TV program through other means (i.e., text messaging, Internet, or phone) has been extremely popular in certain instances. For example, *American Idol*, one of the most popular TV shows in the US television history, was a singing competition where viewers voted for performers on the show using the Internet or their phone.

Finally, there is interactive storytelling in which people's input while watching a show is used to dynamically affect the way the story unfolds (Murray 1997). This is perhaps the most advanced form of iTV. Like storytelling, game playing is also an ancient activity. A modern variation, video gaming, is played on the TV. Although people play games for entertainment and they are interactive, video games are not considered part of the iTV experience in this chapter.

More recently, second screen devices like laptops, tablets, and smartphones have been used to host the interactive part of iTV reserving the TV screen for the TV program (Geerts et al. 2014). One advantage of this arrangement is that the iTV app does not occlude the TV screen or cause the TV program to be reduced in size. Another advantage is that it is easier to interact with a second screen device than with a TV.

Where We Came From: A Brief History of TV in the USA

The following is a summary of the history of major TV developments including iTV in the USA from the perspective of the TV viewer. TV in the USA is the primary region covered in this chapter as the author's personal experience and expertise pertain to the TV landscape in the USA.

1940s: TV Explodes into Popularity

In the USA, the history of TV experience starts in 1940 after TV became commercially available (Genova 2001). There were only about 2000 TVs in the USA. Sixty percent were in New York City and another 30 % were in New Jersey. At that time, New York was the center of the new TV industry. Many of the early TV shows were adapted from radio shows. There were also ice hockey, basketball, and baseball matches on TV (Fisher 2000).

During the war years (1940–1945), TV production was halted in the USA. After the war ended, the US manufacturing switched from wartime products to consumer products, and television sales exploded in the USA going from 178,000 TVs sold in 1947 to 7,500,000 in 1950. Most people were buying their first TV.

1950s: Black and White TV

Early TV

In the early 1950s, there were two main controls (knobs) on TVs: volume and channel (see Fig. 1). Volume was a continuous control. The channel control had 12 detents. There were up to 12 broadcast television channels that could be tuned on a TV set (channels 2–13). There was also a power control. Sometimes power was combined with the volume control and sometimes it was a separate knob. A typical

Fig. 1 Early television with volume and channel controls

TV had a 12-inch diagonal black and white screen. Remote controls were not widely available.

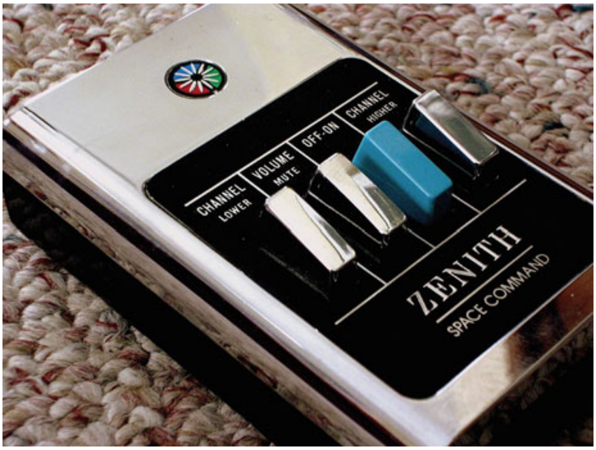
The typical TV viewer experience would be something like this: The person would walk over to the TV and turn the power knob. Then they would use the channel knob to cycle through the channels to the desired channel. In those days, there was no random access to a channel. After tuning, they would adjust the volume with the volume knob and then walk over to the sofa and sit down. This activity would be repeated when the show ended or when something better came on a different channel. A printed schedule of which programs were on at which times was locally published. Since there were usually only a few channels available and TV was only broadcast for part of the day, the schedules were brief. Different regions in the USA had different publications. A national schedule called *TV Guide* was published in 1953.

TV watching was largely a passive entertainment experience, often compared to sitting around a campfire listening to someone telling an interesting story.

The family would sit in front of the TV in their living room, choose a channel, and watch it together, relaxing and being entertained (see Fig. 2). Since there were few channels (Miller 1999) and TV programs were broadcast on a regular schedule, people adjusted their schedule around the broadcast schedule of their favorite shows.

Fig. 2 Family watching television in the 1950s

Fig. 3 Early Zenith wireless remote (Courtesy of Todd Ehlers, CC BY-SA 2.0)



Remote Controls Provide Convenience

In 1956, Zenith introduced the first commercial wireless remote controls (Luplow and Taylor 2012). Although remote controls were available before, they were not very popular, in part, because of the wires running from the remote to the TV, across the living room floor. The wireless remotes typically had four buttons to control TV power on/off, volume mute, channel up, and channel down (Fig. 3).

The user experience would be something like this: The person sits down on their sofa, leans over, grabs the remote from the coffee table, and presses the power (off/on) button. Leaning back on the sofa, with the remote control in hand, they press the channel up or channel down button repeatedly to cycle to the desired channel.

The wireless remote increased the frequency of interaction with the TV, because it made interaction much easier, frequently eliminating the tedious step of having to get up and walk over to the TV. One of those “new” interactions was “muting” the TV sound with the volume mute button. This was promoted as a way to avoid having to listen to “annoying” commercials. The avoidance of advertising is a theme that has persisted throughout the history of TV in the USA where advertisements were embedded within TV programs rather than between them. Although the wireless remotes initially were only available on high-end TVs, they predicted the future of TV interaction.

1960s: Color Makes TV More Realistic

Although color TV was first available in 1954, with the first color broadcast of the Tournament of Roses Parade, it was not until 1965 that most prime-time programs were color, and color TVs began selling in high volume. By 1973, most US households owned a color TV (McDonough 2004).

Today, a primary reason people give for preferring one TV over another is better picture quality. Good picture quality enhances the entertainment and ultimately the viewing experience. For most people, compared to black and white TV, color TV has better picture quality.

1970s: VCRs and Interactive TV Trials

VCRs Introduce TV Consumers to Time Shifting

Videocassette recorders (VCRs) began to achieve mass-market success in 1976 (Pearson 2004). The VCR was a device that allowed a broadcasted TV program to be recorded on videotape and then watched at a more convenient time. The most popular videocassette format was Video Home System (VHS). For most people, this was their first experience with the concepts of time shifting and controlling video (Fig. 4).

The typical TV viewer experience of setting up a scheduled recording would be something like this: After finding the broadcast schedule of the TV program in the newspaper or *TV Guide*, the viewer would turn on the TV, tune it to channel 4, turn on the VCR, put a videocassette in the VCR, and rewind it to make sure the tape was at the beginning. Then, the viewer would press a “menu” button on the VCR to display the programming menu on the TV. On the programming menu, the TV viewer would set up the recording by specifying the day, the start time, the duration, and the channel number. Finally, the TV viewer would have to remember to turn off the VCR. Failing to turn off the VCR would result in the program not being recorded!

Fig. 4 VHS videocassette

Most people did not actually record much with the VCR because it was relatively difficult and the resulting video quality was frequently not as good as the original broadcast. The "blinking 12:00" on the front of the VCR, indicating that the clock on the VCR was not set, became an iconic symbol of the difficulty of programming a VCR in particular and of using new consumer electronic devices in general. The VCR was mainly used to play prerecorded tapes, which people could purchase.

In addition to time shifting, the VCR introduced consumers to rewinding and fast-forwarding video. Most consumers were already familiar with rewind and fast-forward on audiotape, but now it could be done with videotape. This allowed the viewer to rewind the recording if something was missed or to re-watch a scene. It also allowed content and advertisements to be skipped over.

Probably the biggest benefit of the VCR was the advance in convenience with time shifting. VCRs allowed consumers to watch recorded programs on their own schedule.

Interactive TV: The QUBE Cable TV Deployment

In 1977, the Warner Cable *QUBE* system was deployed in Columbus, Ohio. QUBE was a limited deployment iTV system. The system had a special multibutton controller and cable box connected to the TV in consumer's homes. Besides regular TV channels, QUBE pioneered pay-per-view movie channels on which first-run movies would repeat over and over on a channel, so the wait to watch a movie from the beginning was relatively short (Lachenbruch 1977).

QUBE also had interactive TV programs, where the consumer could not only watch the programs but also "interact" with them. One program, *Talent Search*, allowed people to watch a contestant's performance and then vote on it using the controller. The people's votes would determine the outcome of the talent competition. *Screen Test* was a game show about the motion picture industry where the viewers at home could play along, using the controller, with contestants on the show. On the show *How Do You Like Your Eggs*, contestants would guess how the home audience responded to various questions (Forth 1999).

The *QUBE* system failed to become commercially sustainable. However, parts of the system lived on to success. Some of the channels developed for *QUBE* lived on such as *MTV* and *Nickelodeon*. Also, pay-per-view became commercially successful for many future TV services. However, the interactive programming was met with limited participation.

Although being able to play along with game shows is novel and fun, it may not match one of people's fundamental motivations for watching TV: relaxation. Playing games is a very popular form of entertainment, but it is not the relaxing passive form associated with storytelling. One wonders what percentage of entertainment is active and what is passive for the average person. Perhaps active engagement can never be expected to be the principal reason people watch TV.

1980s: Cable TV, Teletext, Universal Remote Controls, and Video Rentals

Cable TV Brings More Choices and Improves the Picture Quality

Cable TV became popular because people could receive more channels with cable than they could receive with an antenna, and they could receive the channels with consistently good reception. For many people, cable boxes allowed them to tune more than just 12 channels without having to buy a new TV. By this time, there were up to 28 national channels available.

Infrared Remote Controls Provide Convenience for the Masses

In 1980, the first "modern" remote control, using infrared (IR) technology, was deployed for a Viewstar cable TV box (Beschloss 2013). This allowed more buttons to be added to the remote control. Thereafter, IR remotes came into widespread use. As was described earlier, the remote control changed the way most people interacted with the TV. They no longer had to get off the sofa to change channels. They could easily channel *surf* or *zap*. They also had random access to channels via a number pad on the remote, which was much more efficient, especially as the number of channels increased. So instead of having to go through intervening channels in a sequence to get from one channel to the next, they could go directly to the desired channel by typing the channel number on the remote control's keypad. Advertising was introduced within programs because it was found that people would be more likely to remain on the channel if the advertisement was during a program rather than between programs. The added convenience of a remote control contributed to the more passive, relaxing nature of watching TV.

Teletext: Accessing Textual Information Through the TV

Teletext is a character-based interactive system for delivering general information to TVs. It was widely used in Europe and in the UK but is in use today only to a limited extent and is being phased out. The "color" buttons on most TV remote controls were used to interact with teletext.

Teletext provided weather, news, sports scores, and other information of general interest. People could navigate through a hierarchy of character-based information on their TV screen using their remote control. Teletext was more like using the TV as an information device (e.g., computer) rather than an interactive TV (Merriman 2014).

Although teletext provided information to be informed and learn new things, a primary motivation for watching TV, teletext did not provide the entertainment experience associated with informational TV programs. Teletext never really caught on in the USA. Consumers were unwilling to pay enough for teletext service to make a viable business (Carey 1996). In the UK and some European countries, governments backed the provision of teletext, so it did not have to be profitable.

Multiple Remote Controls Crowd the Coffee Table

By 1985, the US households typically had more than one remote control sitting on their coffee table and they needed to use multiple remotes to watch TV. These included a TV remote, a VCR remote, and, for some people, a cable box remote.

For most consumers, having to use only a single remote to watch TV is preferred to having to use multiple remote controls, because it simplifies the experience. By 1985, Philips made a "universal remote control" commercially available (Roebuck 2012). The universal remote was able to control multiple devices with a single remote.

In the following years, most cable box remotes became at least partially universal in that they could turn on the TV and control the TV volume as well as control the cable box. Also, some TV remotes were able to control cable boxes as well as the TV. However, the process of setting up the remote to control one's specific cable box or TV was complicated enough that most people did not do it. It was just easier to continue to use multiple remote controls.

Video Rentals Bring Choice and Time Shifting

In 1986, video stores became very popular as a way to rent movies to bring home and watch on TV. There were approximately 15,000 in the USA. A consumer could walk into the video store and see racks upon racks of VHS videocassette boxes. The VHS box showed all the pertinent information regarding the movie as well as attractive box art. Most people were drawn to the new releases, movies that just finished their run in movie theaters. New releases would generally be located around the periphery of the store arranged in alphabetical order by title. The older titles were arranged on the interior of the store, sometimes by genre. Consumers would walk around the store, look at the VHS boxes, and pick up the boxes of the videos they wanted to rent. At the checkout counter, the store clerk would take the boxes and the payment and retrieve the tapes from behind the counter, put them in protective boxes, and hand them to the consumer to take home. The rentals were relatively inexpensive: $3 or $4 per tape. The tapes had to be returned in the next day or two and had to be rewound, otherwise there was an extra charge. This became a part of many people's weekend ritual. Playing these prerecorded videocassettes became the primary use of VCRs. Initially, video stores rented only VHS tapes, but 10 years later switched over to digital video discs (DVD), which are described in a later section of this chapter (Phillips and Ferdman 2013).

For most consumers, VCRs and VHS tapes from the video store were their first exposure to video on demand. With video on demand, the consumer could watch the video from the beginning on their own schedule, and they could take control of the playback experience through start, stop, pause, rewind, and fast-forward functions. None of these experiences were available watching broadcast TV. The consumer had to watch the program when it was broadcasted or they would miss it. The consumer had little control.

1990s: Digital TV, Internet TV, DVDs, and Surround Sound

Digital Cable and Satellite TV: More Channels and Content Choices

In 1994 and 1995, TV was in the process of being switched over from analog to digital. By switching to digital, the cable networks were able to compress ten digital channels into the same bandwidth used by one analog channel. This made it possible to offer ten times as many channels to consumers, many of which could be pay-per-view.

Other technological advancements made it possible for individual consumers to receive satellite TV signals directly from satellites with a relatively small and inexpensive satellite dish, about 2 ft (0.6 m) in diameter, installed on the roof or side of their house. The programming offered by the DIRECTV and the Dish Network satellite TV companies was essentially the same as that offered by cable TV; thus, competition arose among TV services.

The digital TV services supplied households with a set top box (e.g., cable or satellite box) to provide the means to decode the digital TV signals into a form that could be displayed on the TV. This also enabled TV services to centrally control which channels could be viewed, depending on the “channel package” for which the individual consumer paid. As a result, the consumer user experience shifted mainly to the set top box’s on-screen user experience and remote control. The set top box remote control became the primary remote control. The TV remote control became used mainly for turning the TV on and off, controlling the audio volume, and switching TV inputs. In the following years, the set top box remote controls became capable of controlling the TV and other devices. Because you could use the set top box remote to power the TV on and off and control the volume, this made TV remotes largely unused.

When cable TV only had 28 channels, it was relatively easy to keep track of the TV schedule through a printed schedule. With the hundreds of channels available through digital cable, and the different TV services available in the same regions, keeping track of the TV schedule became more difficult to do with a printed schedule. Because cable and satellite boxes were essentially digital computers, it became possible to offer an on-screen “guide” (EPG) that was navigable by the user with the remote control (see Fig. 5).

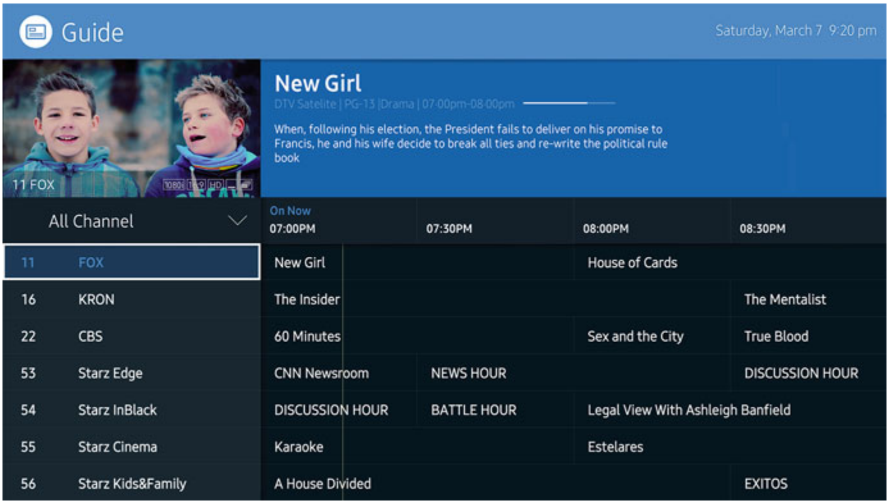


Fig. 5 Interactive TV "grid" guide (EPG)

The guide was a grid showing the shows that were on each channel at the current time and for several days into the future. The guide could also be used to tune a channel by using the remote control to navigate to and select the title of the program the TV viewer wanted to watch. Before digital TV, the channel up and channel down buttons on the remote, which changed channels serially, were the most widely used method of browsing live TV. Over time, the guide became the primary method of browsing live TV (Darnell et al. 2007). This is because there were too many channels to go through serially, one by one, and the channel changing speed of digital TV was too slow (e.g., ~2 s).

Although most people in the USA had a "pay TV" service with a digital cable box or satellite box, about 10 % used over-the-air (OTA) antennas. Several years after cable TV and satellite TV became digital, OTA signals began to be broadcasted in digital rather than analog format. This gave OTA TV better picture quality and more channels.

WebTV: Internet on TV and Web-Based iTV

In 1996, the Internet was exploding in popularity. Computers were still expensive for home use averaging $1500. *WebTV* was a relatively inexpensive ($100) set top box that could be connected to any TV and used to browse the Internet and send and receive email (Swisher 2013). WebTV came with a remote control, but most WebTV users purchased a wireless keyboard. Typing characters using an on-screen keyboard and remote control's up, down, left, and right arrow buttons was too tedious for email.

Later models of WebTV, called *WebTV Plus*, incorporated a TV tuner, an electronic program guide (EPG), and iTV apps. An analog cable or antenna could be connected to the *WebTV Plus*, and the *WebTV Plus* could be used to tune channels through its EPG. The WebTV EPG could also tune channels on a cable box or satellite box through an IR blaster, a small device connected to the WebTV which sent out IR control signals to the cable or satellite box.

The most popular iTV applications included Wheel of Fortune and Jeopardy. When the viewer tuned to one of these game shows that supported interactivity, an "i" icon would appear on the screen. The viewer could select the "i," and a web application would overlay a small area of the TV screen and allow the user to play along with the game show. Although these applications were popular in concept, they were not widely used, in part, because with dial-up modem speeds, the interactive app frequently lagged behind the game show, resulting in a poor interactive experience.

The TV features of *WebTV Plus* were not widely used because they were largely unnecessary for basic TV watching. With OTA antenna or analog cable, there generally were not that many channels to motivate the usage of the WebTV's EPG. Cable and satellite boxes, having lots of channels, included their own built-in, better integrated, and more responsive EPGs.

People generally take the path of least resistance when watching TV, minimizing their expenditure of energy. This is compatible with relaxation, a primary motivation for watching TV (Lee and Lee 1995).

DVDs Overtake VHS Videocassettes Bringing Better Picture Quality to Removable Media

In 1997, DVDs outsold VHS videocassettes for the first time (Farivar 2006). Besides having a smaller form factor, the video quality of DVD, 480 lines, was much better than VHS (~240 lines). All else being equal, a video image composed of more lines has higher resolution and better perceived quality than an image composed of fewer lines. The higher resolution of DVD was a significant improvement in picture quality. In addition, unlike VHS videocassettes, DVDs did not have to be rewound after viewing, saving the consumer time and effort.

Home Surround Sound Brings Higher Audio Quality to the TV Experience

At about the same time that DVDs became popular, home surround sound systems also became popular. An illustrative surround sound system, Dolby 5.1, included an audio-video (A/V) receiver, five speakers, and a subwoofer. The speakers were arranged to surround the TV viewer: one to the left and right of the TV and one centered above or below the TV and two behind the viewer, one left and one right. The placement of the subwoofer was not critical since it emitted bass frequencies, which are difficult to localize.

These systems were frequently attached to DVD players but sometimes also to the cable box or satellite box. Surround sound systems typically had a separate remote control to control the power, volume, and source for the A/V receiver. The advantage of surround sound was that the audio quality was much higher.

2000s: Flat Panel TVs, DVR, VOD, HDTV, OTT, YouTube, and Netflix

DVR/PVR Brings Time Shifting and Control Over Viewing Advertisements

In 1999, the digital video recorder (DVR), also known as the personal video recorder (PVR), was introduced commercially by three companies: TiVo, ReplayTV, and EchoStar. A digital video recorder is a device or feature with a hard disk drive that allowed the TV viewer to record TV programs for later playback. It also allowed the viewer to set up recurring recordings – to automatically record all upcoming broadcasts of a TV program.

Interestingly, a DVR is similar in concept to a VCR but is much easier to use and has better picture quality. This similarity probably made DVR a very slow-to-adopt technology. People may have reasoned: I already have a VCR and do not use it to record shows at all, so why would I need an easier-to-use recording device? However, once people tried DVR, they got hooked and could not live without it! After 15 years, about half of the households in the USA have DVR (Friedman 2013). Today, DVRs have enough storage for hundreds of shows. Some people do not watch live TV anymore; they record nearly everything and watch the recordings. This allows them to have an on-demand experience and to skip over the advertisements. It is interesting that, for DVR users, the most common way to interact with the TV is by skipping advertisements during DVR playback. This is a routine behavior and occurs on the average of every 8 min while watching recordings (Darnell et al. 2007).

Later developments in DVR allowed recordings to be shared between multiple TVs in the home.

Video On Demand Is the Ultimate Time-Shifting Feature

In 2003, video on demand (VOD) began displacing video rentals. Video on demand presents a hierarchically categorized collection of movies and TV programs which the TV viewer can browse through anytime on their TV screen with their remote control and watch immediately from the beginning (Stelter and Chozick 2013). Initially, video on demand presented recently released movies to view for a fee but later incorporated “free” movies and recent episodes of popular TV shows.

The experiential benefit of VOD over its predecessor, pay-per-view (PPV), is that the viewer could truly start watching the movie from the beginning at *any* time as well as pause, rewind, and fast-forward without the need for DVR. In contrast, PPV broadcasted the same movie repeatedly on a live TV channel, sometimes on several channels with offset start times, so the user only had to wait for 15–30 min to watch the movie from the beginning.

Video on demand essentially provided the same movies and TV programs available for rent at the video store but was much more convenient, because the person did not have to leave their home to procure and return the videos. For pay movies, there were generally no ads; however, for “free” programs, ads were frequency included. Sometimes, fast-forward was disabled during playback so people would be unable to skip the ads.

HDTV Provides an Increase in Picture Quality

In 2005, high-definition TV (HDTV) became popular. High-definition TV increased the vertical resolution of the TV picture from 480 lines to 720 lines and later to 1080 lines and changed the shape of the TV to 16×9 rather than 4×3. This was a noticeable improvement in picture quality and was particularly noticeable in sports programming. HDTV was the next step in improving the picture quality and thus enhancing the entertainment experience.

YouTube Provides a Platform for New Kinds of Content

YouTube was founded in 2005 (Fitzpatrick 2010). YouTube allowed individuals to easily share their own self-created videos with anyone on the Internet. Digital video cameras were widely available at this time. YouTube provided an opportunity for people to experiment and create new forms of video entertainment for a potentially large audience. The service was initially available on computers and was not available as an app on TVs until about 5 years later. YouTube is primarily a search-based experience. The person searches for videos on a particular topic (e.g., Taylor Swift) by typing in a keyword. Because of the difficulty of specifying text on a TV, and its incompatibility with relaxation, it is unclear whether this form of entertainment will ever become a primary way for people to watch TV.

Netflix Shows the Potential of a Video On Demand World

In 2005, Netflix was mailing out 1 million DVDs every day to their subscribers in the USA. For a relatively low subscription fee, people could create a prioritized list of their desired DVDs on the Netflix website, and several of the DVDs from their list would arrive at their home in a day or two. People could take as long as they wanted to watch the DVDs and return them in the mail, postage-free. When they were returned, new DVDs from their list would be mailed out. This convenience contributed to the demise of video stores.

A few years later, Netflix became the largest streaming TV service (Aurletta 2014). Netflix had a large collection of commercial-free movies and TV shows and was able to offer a subscription for $8 per month to watch as much as the subscriber wanted. Netflix also began producing their own exclusive, award-winning TV series and has become available in many countries across the world.

Flat Panel TVs Bring Larger Screens into People's Homes

The first television technology used commercially was the cathode ray tube (CRT). CRT TVs were deeper than they were wide and so occupied a lot of space. They were also very heavy and limited in diagonal screen size to about 35 in. (89 cm). In the late 1990s, liquid crystal display (LCD) technology was introduced, allowing TVs to be much wider for a given depth. These were known as "flat panel" TVs. By 2006, flat panel TVs were outselling other TVs (Sayer 2006) and there were multiple flat panel technologies (e.g., plasma, LED). For a given picture size, the weight and overall size of a flat panel TV were far less than competing technologies. As a result, people bought TVs with larger screens, resulting in a better TV picture.

OTT: The Ultimate in Time Shifting and Place Shifting

In 2007, over-the-top (OTT) TV became available. With OTT, TV programs and movies could be streamed over the Internet to a standard web browser, typically on demand (e.g., Netflix, Hulu). This provided a supplemental or alternative TV source. For some people, known as "cord cutters," because they canceled their cable or satellite TV service, it became their primary source of TV (Waniata 2015). Initially, OTT was only available on computers. To watch the content on TV, the computer needed to be connected to the TV. This is something relatively few people did and it typically required a high-definition television (HDTV). In addition to having a wide variety of popular TV shows and movies available on demand and in the case of Netflix, commercial-free, the browsing experience, characterized by box art and movie posters, was a much more enjoyable user experience than the text-based experience available on TVs at that time.

Streaming Media Players Make It Easy to Get OTT Content onto the TV

Over the next few years, streaming media players (SMPs) such as Apple TV and Roku became available (Towle 2014). These were devices that could be connected to a TV, similar to the way a cable or satellite TV set top box was connected, to stream OTT TV shows and movies directly to the TV, without the need for a computer. Like cable and satellite boxes, SMPs provided their own on-screen user experience and remote control.

The SMP on-screen user experiences were similar to the web browser style of user experience characterized by box art and movie posters. Browsing through box art for movies and TV shows was much more informative and enjoyable than browsing through text-only lists of titles, typically found on the TV services' video on demand systems of the era.

The SMP remote controls were simplified relative to typical TV services' remote controls. Since all programs on the streaming media players were on demand, there was no need for buttons on the remote control associated with navigating live TV such as number buttons, channel buttons, or a guide button. This resulted in a remote with many fewer buttons having a simpler and easy-to-use experience.

As an alternative to SMPs, popular game consoles (i.e., Xbox, PlayStation, Wii) also provided apps to stream OTT TV shows and movies to the TV. In the USA, game consoles continue to be the most popular means by which people stream OTT content to their TVs (Hoelzel 2014).

2010s: Smart TVs, 3D, Voice and Gesture Commands, Mobile Devices, and Ultra HD

Smart TVs Make OTT Access Even More Convenient

In 2010, TV manufacturers (e.g., Samsung, LG, Sony, etc.) began creating "smart TVs." These TVs had built-in web browsers and built-in apps (e.g., Netflix, Hulu Plus) that allowed the TV to stream OTT TV shows and movies without an extra device connected to the TV (Donovan 2012). From a user experience point of view, this simplified the user experience by eliminating another device (i.e., an SMP) connected to the TV. It also eliminated the need for another remote control since the TV's remote control could be used to access the OTT content.

3D Brings Greater Realism but with Negative Side Effects

In 2011, TV manufacturers began marketing 3D TVs. 3D content could be viewed on the TV by wearing special glasses. Although there had been very successful 3D movies in theaters (e.g., Avatar), for the most part, 3D on TV was not commercially successful (Bell 2013). Part of the problem was a lack of 3D content, but part of the problem was that the glasses were awkward and uncomfortable for many people to wear. Also, for many people, there is an increase in eyestrain and nausea as a result of viewing 3D (Obrist et al. 2012). By 2014, 3D TVs were being downplayed by the manufacturers.

Voice and Gesture Commands Strive to Provide More Convenient TV Control

Voice command technology for controlling TVs became widely available with the Xbox Kinect and Samsung and LG Smart TVs. The TV viewer could control the TV by voice rather than using a remote control (Katzmaier 2012). Voice commands were generally limited to navigating and selecting items visible on the TV screen. This limitation helped the TV viewer know what they could say, as well as helping the voice system recognize what the TV viewer said in a particular context.

TV viewers at home typically did not use these features often, and, when they did, their use was limited to a few iconic commands (e.g., "Xbox pause!," "Hi TV," "volume down!"). Early voice command systems were not very accurate, causing people to have to repeat commands routinely. In many situations, using voice commands took more energy to use than the remote control. Voice commands for TV were more of a secondary, novel way to control the TV than a primary way.

Voice command systems also allowed people to search for TV shows by voice (Perry 2014). This was a more potentially efficient use of voice because people were not limited to referring to the small number of items visible on the screen. Also, searching for TV shows by typing text on the TV using a remote control was a particularly tedious experience. However, this potential advantage for voice search was tempered by the fact that searching for TV shows by specific criteria such as title or actor is not a particularly frequent activity.

Xbox Kinect (White 2011) and Samsung Smart TV (Katzmaier 2012) also implemented the ability to use hand gestures to control the TV. A triggering gesture, such as a wave, was used to put the Xbox (TV) into a mode to map the viewer's hand movements to cursor movements on the screen. It was thought gestures might be easier than using a remote control because using gestures eliminated the time to find and pick up the remote and gestures acted directly on the TV screen rather indirectly through a remote control. Research showed that the gestures were not used often in part because it took more energy to make hand gestures than to press buttons on a remote control (Cabral et al. 2005). Also, people generally keep their remote control very close by, so reaching for it is generally not too difficult. People watching TV are usually relaxing and do not want to expend unnecessary energy. Gestures may be more appropriate for gaming and other active entertainment than for TV viewing.

Watching TV Content on a Mobile Device Can Be More Convenient

Apps on mobile devices (smartphones and tablets) enabled viewing of TV content on the mobile device (Nielsen 2013). Most apps provided streaming video on demand content (e.g., Netflix, Hulu Plus, ABC, ESPN). Cable and satellite TV providers also created apps (e.g., XFINITY TV) that allowed their customers to view live TV, DVR, and video on demand content on mobile devices. From a user experience perspective, being able to view TV on a mobile device is convenient; however, most people prefer viewing TV content on a TV when they have that option.

Watching TV on a mobile device is a continuation of a trend to have more TVs in the home. In the USA, the average household has 2–3 TVs. This enables people in the household to watch different shows at the same time in different rooms allowing individuals to watch what they want, rather than having to compromise with other members of the household.

Using a Mobile Device as a Remote Control

Some mobile apps were created for touchscreen mobile devices (e.g., smartphones and tablets) to simulate the TV remote control. These apps usually have the same button sets as traditional remote controls but typically have different button layouts to fit the form factor of the mobile device. On the positive side, the virtual buttons on a mobile app allow more descriptive and personalized button labels, as well as being able to use the context to organize buttons. For instance, when the TV viewer is watching live TV, the channel number buttons could be displayed on the touchscreen, but when using the DVD player, the DVD remote control's buttons could be displayed.

On the negative side, it is easier to use a physical button on traditional remote control because you can orient your fingers on it by feel. This is not possible on the touchscreen, which requires more effort to keep your finger on a particular button while looking at the TV (Moskovciak 2013). Also, with physical buttons, you can more easily orient by feel to commonly used nearby buttons, whereas to do so on a touchscreen generally requires looking at it. There are many very common remote control activities that are typically done by feel while looking at the TV (e.g., using the directional pad, using fast-forward and rewind, channel surfing, changing the volume).

Using a Mobile Device to Browse and Search for TV Content

Some apps on mobile touchscreen devices present a complete user interface for the TV. For instance, instead of accessing a guide (EPG) on the TV screen and scrolling through it with the remote control, one can access a guide on the mobile device screen and scroll through it using touchscreen gestures. Choosing a program in the mobile guide can play the program on the TV screen (e.g., TiVo, Dish, XFINITY).

This type of mobile app has the potential of providing a better interactive experience than the TV because it is potentially easier to scroll through lists (e.g., of TV shows or movies) on a mobile device's touchscreen than on a TV using a remote control. Similarly, to search for a TV program or movie, it is easier to enter text on the touchscreen than on a TV's on-screen keyboard and remote control. Once a TV program or movie is found, the person can either watch it on the device or on the TV (e.g., Netflix with Chromecast).

On the other hand, it is easier to *access* the guide on a TV using a remote control than on a general-purpose mobile device like a smartphone or tablet. For example, on a mobile device you might have to first unlock the device, then find the app, and open it before accessing the guide. On a typical remote, you simply have to press the guide button. This decreased ease of access on the mobile device could override the advantage of navigating using the mobile device.

TV Companion Apps Complement TV Programs

Some mobile device apps provide information synchronized with a live TV show (Cesar et al. 2008a). These are known as "companion apps" or "second screen apps." As the TV show plays on the TV, the companion app on the mobile device presents relevant information and options related to the current state of the show (Vascellaro 2011). For example, while watching a baseball game, a companion app could present stats for the current batter. Twitter and Facebook could also be used as companion apps because many people comment on popular live TV programs while watching them.

In 2013 and 2014, companion apps were also available on Samsung Smart TVs. When watching a live TV program, an app could be invoked to supplement the show. For example, a live Twitter feed from people watching and commenting on the same show could be viewed on the TV screen alongside the TV program. These features fulfill some of people's primary motivations for watching TV, such as helping the TV viewer to be more informed, learn new things, and socialize.

Ultra HD Takes Picture Quality to the Next Level

In 2014, TV manufacturers were increasing the production of ultra HD and 4 K TVs (Tretbar 2014). These TVs have about four times the resolution of HD TVs with 2160 lines versus 1080 lines in HDTV. For very large TVs (e.g., 80 in/200 cm), this increased resolution is dramatic. For smaller TVs (e.g., 46 in/117 cm), the benefit of the increased resolution is more subtle but still apparent especially in the text that appears on the TV.

What Themes Emerge from the History of TV Experiences?

This section discusses major recurring themes that emerge from the overview provided in the previous section of this chapter.

The major recurring themes include:

- Making the TV picture quality better
- Making more TV program choices available to the viewer while watching TV
- Making the control and interaction with the TV more convenient and easier
- Providing additional information to supplement TV programs
- Enabling the TV viewer to participate interactively in the TV experience (i.e., iTV)

A summary for each of the themes is provided below pointing out their relevance in relation to creating, enhancing, and further shaping TV experiences.

Improving Picture Quality Enables a Better TV Experience

Improved picture quality has occurred throughout the history of TV starting with color TV, then cable TV, digital TV, DVD, HDTV, flat panel TV, 3D TV, and most recently ultra HD and 4 K. Picture quality is one of the most important factors in the development of TV and consequently the TV experience. Although experiencing good picture quality is not one of the fundamental motivations for watching TV, it contributes to the overall aesthetics of the TV experience and may help facilitate some of the primary motivations for watching TV (Lombard and Ditton 1997). Such primary motivations include "escape" and "experiencing an engrossing different world."

Making More TV Program Choices Available Enables a Better TV Experience

Making more TV program choices available to the TV viewer when they watch TV is also a major theme that has occurred throughout the history of the TV user experience starting with the VCR and cable TV, and then video rentals, digital cable/satellite TV, DVR time shifting, video on demand (VOD), and over-the-top (OTT) TV. Having more choices as to what to watch during a given TV viewing session helps the TV viewer to better fulfill some of the primary motivations for watching TV. These include "escape," mood improvement, being informed, learning new things, and experiencing suspense and excitement.

Making TV Control and Interaction Easier Contributes to a Relaxing Experience

TV watching is largely a lean-back experience and relaxation is a primary motivation for watching TV (Lee and Lee 1995). Being able to control the TV with minimal effort is compatible with relaxation. Many of the innovations in the history of TV experience have centered on making it easier to control the TV. These include the wireless remote control, IR remotes, universal remotes, voice and gesture control, and control of TV through mobile devices.

Providing Additional Information Through the TV Facilitates Learning

Another recurring theme in the history of the TV experience has been providing information through the TV. This is exemplified by digital TV program info, WebTV, teletext, and companion apps. Metadata about TV programs including the program title, synopsis, and actors is becoming standard on digital cable and satellite TV. In addition, apps available on smart TVs (e.g., Samsung Sports Live) and smartphones (e.g., The Walking Dead Story Sync) provide information to enhance live TV programs.

Efforts to provide general information through the TV, including teletext and web browsing, have also been a recurring theme, supporting one of the primary motivations for watching TV: to be informed and learn new things. However, this use of TV has been met by limited success in part because it is relatively difficult to interact with a TV compared to other more specialized devices like laptops, tablets, and smartphones. In addition, although TV viewers prefer a large screen for video consumption, they may not prefer one for the consumption of non-video information.

Enabling the TV Viewer to Participate Interactively in the TV Program

Another theme to emerge in the history of the TV experience is interactive TV (iTV). There have been many small interactive TV deployments exemplified by QUBE, WebTV's interactive game show apps, and various apps providing information synchronized with TV programs. A recent example is second screen companion apps, where the interactive part of the experience is on a laptop, tablet, or smartphone, and the TV is reserved to view the program.

Interactive TV has had limited success. Although it is compelling to some people, it has never become widespread. This may be due in part to the content or technological limitations. However, it may simply be that for most people TV watching is a passive experience more compatible with relaxation, which is a primary motivation for watching TV.

Summary

At the beginning of this chapter, it was hypothesized that the innovations and resulting experiences throughout the history of TV that were related to primary motivations for watching TV would be more successful than those that were not. This hypothesis was only partially supported. Two of the themes, better picture quality and more content choices, were not directly related to the primary motivations of people to watch TV. Two other themes, making control of the TV easier and providing supplementary information, were both related to several primary reasons people watch TV: relaxation and learning, respectively. Finally, the theme of actively participating in a TV program (iTV) was not a primary motivation for watching TV. To conclude, successful innovations in the TV experience came from multiple sources. Some fulfilled people's needs and primary motivations (e.g., relaxation, being informed, and learning), but others were more technology driven. For example, picture quality improvements could be done because they were technologically possible and improved the TV experience, not because they fulfilled people's basic needs for watching TV.

Where Are We Going: The Future of TV Experiences

This section predicts the future of TV experiences based on the major recurring themes that emerged from the history of TV experiences discussed in the previous section of this chapter.

Better Picture Quality

Since improvements in picture quality have been a recurring theme in the history of the TV experience, improvements in picture quality are predicted to continue into the future. In 5 years, in 2020, most TVs will be ultra HD 4 K and the newest TVs may even be 8 K TVs (6,320 lines). Resolution, a primary enabler of improved picture quality, will probably continue to increase in the future until the average person simply cannot see the additional detail in their living room because of the limits of the human vision.

Larger TV Screens and Multiple Screens

The average size of TV screens will continue to get larger as the overall TV size, weight, and cost decrease for a given screen size. Screen size will ultimately be limited by people's living room scale and furniture placement, which largely determines the TV viewing distance today.

Watching video on laptops, tablets, and smartphones will continue to grow but will only offset TV screens in situations where these other devices are more convenient. Convenience being equal, people will continue to prefer viewing on larger TV screens.

More Video Content and More On-Demand Content

Increasing the amount of video programming choices available during a TV viewing session is expected to continue in the future. This includes both the total amount of choices and also the immediate availability of the TV viewer's routinely watched shows. Even though, on average, people watch a relatively small number of programs, they value having a large number of programs from which to choose. Most of the growth will not occur in the number of scheduled/live TV channels available but in the number of on-demand programs available. Viewing programs on demand is much more convenient than viewing them on a schedule, because once they become available you can watch them whenever you want.

The increase in on-demand programs has been a recent theme in the history of TV experience. In 10 years, in 2025, it is expected that the majority of TV watching will be on demand. An exception to this will be sporting events and other time-critical events such as contests, election results, etc. These events will continue to be broadcast live, in part, because the indeterminacy of the outcome adds to the enjoyment of the experience (Vosgerau et al. 2005).

Less Live TV and Less Use of the Grid Guide

The primary way that people navigate live TV today is by using an interactive grid guide (EPG). The grid guide displays a matrix of channels by time slots. In the future, when less time is spent watching scheduled/live TV channels, and more time is spent watching on-demand programs, the grid guide will become largely unnecessary. People will continue want to know when new episodes of the shows they watch will be available, but will not need a grid guide for that purpose. The strength of a grid guide is showing all of the available TV choices during a given time slot, but with on-demand, everything is available for weeks. The only programs where the time element is critical will be sports matches and the few other programs people want to watch live. A grid guide will not be necessary for that purpose since the viewer is at most interested in a only few live shows at any one time.

Less DVR and Fewer Advertisements

DVR will likely continue to exist as long as it gives people more control than on-demand programs over skipping advertisements. If on-demand programs allowed the viewer to skip the advertisements similar to DVR, DVR would probably disappear quickly, because DVR takes more effort to set up and maintain recordings in comparison to on-demand. Also, the on-demand experience of having complete seasons, organized by episode to watch anytime, is not present with DVR. The feature of DVR that will continue to exist for people is the list of programs that they watch routinely. This list will be composed mostly of on-demand programs rather than DVR programs.

Simplified Remote Controls

When nearly all content is on-demand, the simplification of the remote control will become widespread, similar to the Apple TV, Roku, and Amazon remote controls of today. There will be no need for number buttons, channel buttons, a guide button, or other buttons that support scheduled/live TV channels.

Aggregated On-Demand Browsing Experience

Today, on-demand programs are provided by multiple content providers through their individual apps, each with a different experience (e.g., Netflix, Amazon, Hulu, Vudu, HBO GO, XFINITY, etc.). To the extent that people will watch programs from multiple content providers on a daily basis, having to maintain a separate list of TV programs for each provider, having to remember which provider provides a particular program, and having to switch between different provider's experiences may be a less than optimal situation. One way of addressing these issues would be to have a system that aggregates the TV programs and movies from multiple content providers into a single experience, similar to the way the cable TV systems today aggregate channels from many different providers into a single experience.

More Enjoyable Browsing Experience

The browsing experience of choosing movies and TV programs to watch will continue to be enhanced with more photos, videos, graphics, and information. This richer browsing experience will become a more enjoyable part of the TV experience.

Lean-Forward TV Experiences Will Continue to Be Limited

Although TVs will continue to have computer-like functions, because TVs are essentially computers, people will continue to use these functions only to a limited extent. TV will continue to be primarily a lean-back, relaxing experience. Informational content will need to be viewed as part of a lean-back experience to be successful.

Interactive TV apps that provide information synchronized with TV programs or allow the TV viewer to actively affect the TV program will be more successful on second screens than on the TV. This is because it is easier to interact on a second screen device than on the TV. These applications will be more successful with younger people who are used to multitasking.

Conclusion

The developments in the history of television including interactive TV were reviewed from the perspective of TV viewer's experiences. Five recurring themes emerged from this review:

- Making the TV picture quality better
- Making more TV program choices available to the viewer while watching TV
- Making the control and interaction with the TV more convenient and easier
- Providing additional information to supplement the TV program
- Enabling the TV viewer to participate interactively in the TV experience (i.e., iTV)

These themes were compared with people's primary motivations for watching TV. The primary motivations included:

- To have a familiar routine
- To relax, escape, and improve one's mood
- To be informed and learn new things
- To learn from and about other people
- To socialize
- To experience excitement, suspense, and an engrossing different world

It was hypothesized that themes that supported primary motivations would be more successful than those that did not and thus could be used as a basis to predict the future of TV. This hypothesis was not really supported.

Some of the themes, for example, making the control of the TV more convenient and easier, mapped well into relaxation, a primary motivation for watching TV. In contrast, some of the themes, for example, making the TV picture quality better, did not map directly into a primary motivation for watching TV but was instead more technology driven. The picture quality could be improved by increasing the resolution because the state of the art made this cost effective.

Some of the themes came from directly addressing user needs as reflected in the primary motivations for watching TV, whereas other themes seemed to arise from technological opportunities.

Thus, the future of TV user experience was predicted based on the five recurring themes from the history of TV, rather than on people's primary motivations for watching TV. The TV experience is currently undergoing rapid changes, so it will be truly interesting to see what it will be like in 10 years.

References

K. Aurletta, Outside the box: Netflix and the future of television (2014), http://www.newyorker.com/tech/elements/object-of-interest-remote-control. Accessed 23 July 2015

S. Bachmayer, A. Lugmayr, G. Kotsis, New social & collaborative interactive TV program formats. Proceeding MoMM '09 Proceedings of the 7th international conference on advances in mobile computing and multimedia (ACM, New York, 2009), pp. 121–128. http://dl.acm.org/citation.cfm?id=1821776

D. Bell, 4 Reasons the 3D TV movement is already dead (2013), http://www.cracked.com/quick-fixes/4-reasons-3d-tv-movement-already-dead/. Accessed 4 May 2015

S. Beschloss, Object of interest: remote control. The New Yorker (2013), http://www.newyorker.com/tech/elements/object-of-interest-remote-control. Accessed 30 May 2015

M. Cabral, C. Morimoto, M. Zuffo, On the usability of gesture interfaces in virtual reality environments. LCIHC'05 Proceedings of the 2005 Latin American conference on human-computer interaction (ACM, New York, 2005), pp. 100–108. http://dl.acm.org/citation.cfm?id=1111370

J. Carey, Winky dink to stargazer: five decades of interactive television (1996), http://www.columbia.edu/cu/business/courses/download/B9201-XX/carey/history_of_interactive_tv.pdf. Accessed 2 May 2015

P. Cesar, D.C. Bulterman, A.J Jansen, Usages of the secondary screen in an interactive television environment: control, enrich, share, and transfer television content. in *Proceedings of the 6th European Conference on Changing Television Environments* (Springer, Berlin/Heidelberg, 2008a), pp. 168–177

P. Cesar, K. Chorianopoulos, J. Jensen, Social television and user interaction. Comput. Entertain. **6** (1), 4 (2008b)

M. Darnell, How do people really interact with TV?: naturalistic observations of digital TV and digital video recorder users. Comput. Entertain. **5**(2), Article No. 10 (2007)

J. Donovan, Smart TVs: how do they work? (2012), http://techcrunch.com/2012/01/13/smart-tvs-now-and-next/. Accessed 31 May 2015

C. Farivar, DVD players finally overtake VCRs in the US (2006), http://www.engadget.com/2006/12/21/dvd-players-finally-overtake-vcrs-in-us/. Accessed 30 May 2015

D. Fisher, Chronomedia 1940 (2000), http://www.terramedia.co.uk/Chronomedia/years/1940.htm. Accessed 2 May 2015

L. Fitzpatrick, A brief history of YouTube (2010), http://content.time.com/time/magazine/article/0,9171,1990787,00.html. Accessed 3 May 2015

F. Forth, QUBE Columbus interactive TV warner cable (1999), http://www.qube-tv.com. Accessed 15 May 2015

W. Friedman, As streaming rises, DVR penetration slows (2013), http://www.mediapost.com/publications/article/215069/as-streaming-tv-rises-dvr-penetration-slows.html. Accessed 4 May 2015

D. Geerts, R. Leenheer, D. De Grooff, J. Negenman, S. Heijstraten, In front of and behind the second screen: viewer and producer perspectives on a companion app. Proceeding TVX'14 (2014), pp. 95–102. http://dl.acm.org/citation.cfm?id=2602312

T. Genova, Television history: the first 75 years (2001), http://www.tvhistory.tv/facts-stats.htm. Accessed 2 May 2015.

M. Hoelzel, Video game consoles are still the most popular streaming devices for TVs (2014), http://www.businessinsider.com/video-game-consoles-are-still-the-most-popular-streaming-devices-for-tvs-2014-6. Accessed 31 May 2015

J. Jensen,. in *Interactive Television – A Brief Media History*, ed. by M. Tscheligi, M. Obrist, A. Lugmayr. Changing Television Environments: 6th European Conference, EuroITV 2008, Salzburg, Austria, 3–4 July 2008, Proceedings, (Springer, Berlin/Heildelberg, 2008), pp. 1–10

D. Katzmaier, Samsung smart interaction: hands-on with voice and gesture control (2012), http://www.cnet.com/news/samsung-smart-interaction-hands-on-with-voice-and-gesture-control/. Accessed 31 May 2015

D. Lachenbruch, Will It Play In Columbus? TV Guide Magazine (1977), 24 Dec 1977, pp. 2–5

B. Lee, R. Lee, How and why people watch TV: implications for the future of interactive television. J. Advert. Res. **35**, 9–18 (1995)

M. Lombard, T. Ditton, At the heart of it all: the concept of presence. J. Comput. Mediat. Commun. (1997). doi:10.1111/j.1083-6101.1997.tb00072.x/full

W. Luplow, J. Taylor, Channel surfing redux: a brief history of the TV remote control and a tribute to its coinventors. IEEE Consum. Electron. Mag. **1**(4), 24–29 (2012)

S. McDonough, Color TV turns 50. Associated Press (2004), http://old.chronicle.augusta.com/stories/2004/03/25/liv_409907.shtml. Accessed 30 May 2015

C. Merriman, Teletext is 40 years old. The Inquirer (2014), http://www.theinquirer.net/inquirer/news/2372556/teletext-is-40-years-old

J. Miller, History of American broadcasting: US TV Stations as of 1950 (1999), http://jeff560.tripod.com/1950tv.html. Accessed 2 May 2015

M. Moskovciak, Forget Smartphone remotes: here is why real buttons win (2013), http://www.cnet.com/news/forget-smartphone-remotes-heres-why-real-buttons-win/. Accessed 3 May 2015

J. Murray, *Hamlet on the Holodeck: The Future of Narrative in Cyberspace* (The Free Press, New York, 1997)

Nielsen, Free to move between screens: the cross platform report (2013), http://www.nielsen.com/us/en/insights/reports/2013/a-look-across-media-the-cross-platform-report-q3-2013.html. Accessed 31 May 2015

M. Obrist, D. Wurhofer, T. Meneweger, T. Grill, M. Tscheligi, Viewing experience of 3DTV: an exploration of the feeling of sickness and presence in a shopping mall (2012), http://ac.els-cdn.com/S1875952112000055/1-s2.0-S1875952112000055-main.pdf?_tid=196a6654-07bf-11e5-b85d-00000aacb35e&acdnat=1433095485_3c79ea8192a0f8ef369230ff5fc88f96. Accessed 31 May 2015

S. Pearson, Changes to video recorders and VCR technology over the last 50 years (2004), http://www.thepeoplehistory.com/vcr.html. Accessed 15 May 2015

T. Perry, CES 2014 trends: new remotes and interfaces to make smart TVs actually usable (2014), http://spectrum.ieee.org/tech-talk/consumer-electronics/audiovideo/ces-2014-trends-getting-smart-tv-under-control. Accessed 3 May 2015

M. Phillips, R. Ferdman, A brief, illustrated history of Blockbuster which is closing the last of its US stores (2013), http://qz.com/144372/a-brief-illustrated-history-of-blockbuster-which-is-closing-the-last-of-its-us-stores/. Accessed 2 May 2015

C. Quico, Are communication services the killer applications for Interactive TV? or I left my wife because I am in love with the TV set. *Proceedings of the 1st European conference on interactive television: from viewers to actors* (2003), https://www.google.com/url?sa = t&rct = j&q = &esrc = s&source = web&cd = 1&ved = 0CB8QFjAAahUKEwjR3rTMyLjHAhXEOj4KHbcjABA&url = http%3A%2F%2Fwww.springer.com%2Fcda%2Fcontent%2Fdocument%2Fcda_downloaddocument%2F9781848822740-c2.pdf%3FSGWID%3D0-0-45-730300-p173868329&ei = ikHWVdHxBMT1-AG3x4CAAQ&usg = AFQjCN-GOVznq81SEg9IaWgZEXRLLROgVrA

K. Roebuck, *3D Television* (Emereo PTY Limited, Brisbane, Australia, 2012), p. 910

P. Sayer, LCDs outsell CRTs in Europe (2006), http://www.macworld.com/article/1054094/lcdsales.html. Accessed 1 June 2015

B. Stelter, A. Chozick, Viewers start to embrace television on demand (2013), http://www.nytimes.com/2013/05/21/business/media/video-on-demand-viewing-is-gaining-popularity.html?_r = 0. Accessed 3 May 2015

K. Swisher, Microsoft quietly shuts down MSN TV, once known as WebTV. AllThingsD (2013), http://allthingsd.com/20130706/microsoft-quietly-shuts-down-msn-tv-once-known-as-webtv/. Accessed 2 May 2015

A. Towle, Streaming media increases in popularity (2014), http://guardianlv.com/2014/07/streaming-media-increases-in-popularity/. Accessed 3 May 2014

A. Tretbar, Study projects 4K will expand like HD, only faster (2014), http://www.digitaltrends.com/home-theater/4k-will-follow-path-similar-to-hd-growth-pattern-but-faster/. Accessed 3 May 2015

J. Vascellaro, Help for watching TV (2011), http://www.wsj.com/articles/SB10001424052970204226204576601072612523158. Accessed 4 May 2015

J. Vosgerau, K. Wertenbroch, Z. Carmon, Indeterminacy and live television (2005), http://www.insead.edu/facultyresearch/research/doc.cfm?did = 1651

R. Waniata, Cord cutting 101: how to quit cable for online streaming video (2015), http://www.digitaltrends.com/home-theater/how-to-quit-cable-for-online-streaming-video/. Accessed 31 May 2015

C. White, Control netflix movies using gestures or voice with kinect (2011), http://mashable.com/2011/04/14/kinect-netflix/. Accessed 31 May 2015

G. Zorik, L. Barkhuus, A. Engsrom, E. Onnevall, Panoramic video: design challenges and implications for content interaction. **EuroITV '13**: proceedings of the 11th European conference on interactive TV and video (ACM, 2013)

Part XI

Social and Ethical Issues

Addiction and Entertainment Products 46

Mark D. Griffiths and Halley M. Pontes

Contents

Abstract

Technological forms of entertainment have become increasingly popular among both adults and adolescents. However, there have been a growing number of reports about excessive use of entertainment technology and potentially addictive use (e.g., to video games, mobile phones, the Internet, etc.). The present chapter briefly overviews addiction to these entertainment products (i.e., "technological addictions") by defining addiction and arguing that technological addictions are a type of behavioral addiction. The chapter also reviews the empirical literature concerning online gaming addiction and social networking addiction, as well as examining the differences between Internet gaming addiction and Internet addiction. The chapter demonstrates that issues encountered by contemporary researchers and clinicians regarding the assessment of online addictions appear complex and include several factors. It is concluded that there is a clear need to distinguish between addictions *to* the Internet and addictions *on* the Internet. Gambling or gaming addicts who chooses to engage in online gambling and gaming are not Internet addicts – the Internet is just the

M.D. Griffiths (✉) • H.M. Pontes
International Gaming Research Unit, Psychology Division, Nottingham Trent University, Nottingham, UK
e-mail: mark.griffiths@ntu.ac.u; halley.pontes2013@my.ntu.ac.uk

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_10

place where they conduct their chosen (addictive) behavior. Based on empirical research, it is evident that excessive entertainment technology use appears to be at least potentially addictive. Further research is needed on whether activities such as video game addiction and Internet addictions such as social networking addiction are distinct clinical entities.

Keywords
Gaming addiction • Online gaming addiction • Facebook addiction • Social networking addiction • Behavioral addiction • Internet addiction • Online addictions • Video game addiction • Internet gaming disorder

Introduction

The popularity of technology as an entertainment phenomenon has become ever increasing in the lives of both adults and adolescents. Coupled with this, there have been a growing number of reports in the media about excessive use of entertainment technology (e.g., video games, mobile phones, the Internet, etc.). Although the concept of technological addictions appears to have its supporters in the media, there is much skepticism among the academic community – not least among those working in the field of addiction research. It is not hard to understand such attitudes. For many, the concept of technological addictions seems farfetched particularly if their concepts and definitions of addiction involve the taking of psychoactive drugs. Despite the predominance of drug-based definitions of addiction, there is now a growing movement which views a number of behaviors as potentially addictive including many behaviors which do not involve the ingestion of a psychoactive drug (e.g., gambling, computer game playing, exercise, sex, Internet use) (Griffiths 2005). Such diversity has led to new all encompassing definitions of what constitutes addictive behavior.

Defining Addiction: Technological Addiction as a Behavioral Addiction

It has been consistently argued that behaviors such as problematic gambling and problematic video game playing are no different from (say) alcoholism or heroin addiction in terms of the core components of addiction (Griffiths 2005). If it can be shown that a behavior such as excessive gambling can be a bona fide addiction, then there is a precedent that any behavior which can provide continuous rewards in the absence of a psychoactive substance can be potentially addictive (i.e., a behavior as opposed to a chemical addiction). Such a precedent "opens the floodgates" for other excessive behaviors to be theoretically considered as potential addictions (e.g., the Internet and videogame use) (Griffiths 2005).

For many years, it has been alleged that pathologies exist among excessive Internet users and video game players. For instance, as early as 1983, Soper and Miller claimed "video game addiction" was like any other addiction and consisted of a compulsive behavioral involvement, a lack of interest in other activities, association mainly with other addicts, and physical and mental symptoms when attempting to stop the behavior (e.g., the shakes). Griffiths (1998) and Young (1998) argued a similar case for excessive Internet users. Such addictions have been termed "technological addictions" (Griffiths 1995) and have been operationally defined as nonchemical (behavioral) addictions that involve excessive and problematic human-machine interaction. Technological addictions can either be passive (e.g., television) or active (e.g., video games) and usually contain inducing and reinforcing features that may contribute to the promotion of addictive tendencies (Griffiths 1995). Technological addictions can thus be viewed as a subset of behavioral addictions and feature core components of addiction first outlined by Brown (1993) and modified by Griffiths (2005), i.e., salience, mood modification, tolerance, withdrawal, conflict, and relapse.

Research into the area of technological and online addictions is underpinned by three fundamental questions: (1) What is addiction? (2) Do technological and online addictions actually exist? (3) If technological and online addictions exist, what are people actually addicted to? The first question continues to be a much-debated question both among psychologists within the field of addiction research as well as those working in other disciplines. For many years, the first author has operationally defined addictive behavior as any behavior that features all the core components of addiction (Griffiths 1995, 2005). It is argued that any behavior (e.g., video game playing, social networking, mobile phone use) that fulfills these six criteria is therefore operationally defined as an addiction. In the case of a technological addiction it would be:

- *Salience*: This occurs when some kind of technology use (e.g., video game playing, Internet use, mobile phone use) becomes the most important activity in the person's life and dominates their thinking (e.g., preoccupations and cognitive distortions), feelings (e.g., cravings), and behavior (e.g., deterioration of socialized behavior). For instance, even if the person is not actually playing on a video game they will be thinking about the next time that they will be.
- *Mood modification*: This refers to the subjective experiences that people report as a consequence of engaging in their chosen technological behavior and can be seen as a coping strategy (i.e., they experience an arousing "buzz" or a "high" or paradoxically tranquilizing feel of "escape" or "numbing").
- *Tolerance*: This is the process whereby increasing amounts of time engaged in a technological behavior is required to achieve the former mood-modifiying effects. This basically means that for someone engaged in Internet use or video game playing, they gradually build up the amount of the time they spend engaged in the behavior even though time, in and of itself, does not necessarily translate into addiction.

- *Withdrawal symptoms*: These are the unpleasant feeling states and/or physical effects that occur when the technological behavior is discontinued or suddenly reduced (e.g., the shakes, moodiness, irritability).
- *Conflict*: This refers to the conflicts between the technology user and those around them (e.g., interpersonal conflict), conflicts with other activities (e.g., job, schoolwork, social life, hobbies, and interests), or conflicts from within the individual themselves (e.g., intrapsychic conflict and/or subjective feelings of loss of control) which are concerned with spending too much time engaged in activities such as Internet use or video game play.
- *Relapse*: This is the tendency for repeated reversions to earlier patterns of technology use to recur and for even the most extreme patterns typical of the height of excessive technology use to be quickly restored after periods of abstinence or control.

Having operationally defined addiction, it is the present authors' belief that technological and online addictions exist but that it affects only a small minority of users. There appear to be many people who use technology excessively but are not addicted as measured by these (or any other) criteria (Griffiths 2010). The third question is perhaps the most interesting and the most important when it comes to researching in this field. What are people actually addicted to when they use the Internet, mobile phone, or play video games excessively? Is it the interactive medium of playing, the aspects of its specific style (e.g., an anonymous and disinhibiting activity), or the specific types of games (e.g., aggressive games, strategy games, etc.)? This has led to much debate among those of us working in this field. Research being carried out into Internet addiction may lead to insights about other technological addictions such as video game addiction and mobile phone addiction (and vice versa). For instance, Young (1999) has claimed that Internet addiction is a broad term covering a wide variety of behaviors and impulse control problems. This is categorized by five specific subtypes:

- *Cybersexual addiction*: Compulsive use of adult websites for cybersex and cyberporn
- *Cyber-relationship addiction*: over-involvement in online relationships
- *Net compulsions*: Obsessive online gambling, shopping, or day trading
- *Information overload*: Compulsive web surfing or database searches
- *Computer addiction*: Obsessive computer game playing (e.g., *Doom*, *Myst*, *Solitaire,* etc.)

In reply to Young, Griffiths (2000a) argued that many of these excessive Internet users are not "Internet addicts" but just use the Internet excessively as a medium to fuel other addictions. Put very simply, a gambling addict or a video game addict who engages in their chosen behavior online is not addicted to the Internet. The Internet is just the place where they engage in the behavior. However, in contrast to this, there are case study reports of individuals that appear to be addicted to the Internet itself (e.g., Young 1998; Griffiths 1998, 2000b). These are usually people (and very often adolescents in their late teenage years) that use Internet chat rooms or play fantasy role playing games – activities that they would not engage in except on the Internet itself. These individuals to some extent are engaged in text-based virtual realities and take on other social personas and social identities as a way of making users feel good about themselves (Griffiths 2000b). In these cases, the Internet may provide an alternative reality to the user and allow them feelings of immersion and anonymity that may lead to an altered state of consciousness. This in itself may be highly psychologically and/or physiologically rewarding. Obviously for those playing online computer games (theoretically a combination of both Internet use and video game play), these speculations may provide insights into the potentially addictive nature of video games for those playing in this medium.

The cognitive-behavioral model of pathological Internet use [PIU] (Davis 2001) was developed to describe the set of symptoms related to PIU and its etiology. This model was very influential in the early 2000s because it was the first to clearly distinguish between specific pathological Internet use [SPIU] and generalized pathological Internet use [GPIU]. As noted by Davis, SPIU can be broadly defined as a type of Internet addiction where people are dependent on a specific function of the Internet (e.g., gaming), whereas GPIU relates to a general, multidimensional overuse of the Internet (e.g., use of many different online applications).

One of the salient features of this model is the emphasis on the importance and role of maladaptive cognitions in the development and maintenance of PIU rather than focusing on the well-documented behavioral factors (e.g., withdrawal, tolerance) associated with PIU. In order to explain the nature of the cognitive theory of PIU, Davis introduced the concepts of distal and proximal contributory causes of PIU. Distal causes include preexisting psychopathology (e.g., depression, social anxiety, substance dependence) and behavioral reinforcement (provided by the Internet itself throughout the experience of new functions and situational cues which contribute to conditioned responses). Proximal causes involve maladaptive cognitions that are seen as a sufficient condition that can lead to both GPIU and SPIU and also cause the set of symptoms associated with PIU. Another important proximal cause that contributes to the causal pathway of GPIU is related to the social context of the individual (e.g., lack of social support, social isolation). As hypothesized by Davis, GPIU involves spending abnormal amounts of time on the Internet, either wasting time with no directive purpose or spending excessive amounts of time in chat rooms. Moreover, procrastination is also assumed to play an important role in both the development and maintenance of GPIU.

In this model, symptoms of PIU primarily derive from maladaptive cognitions. These symptoms relate more to cognitive symptoms and as such may include obsessive thoughts about the Internet, diminished impulse control, inability to cease Internet use, and the generalized feeling that the Internet is the only place where individuals feel good about themselves. Other symptoms may include thinking about the Internet while offline, anticipating future time online, decreasing interest for other activities or hobbies, and social isolation.

More recently, Brand et al. (2014) further developed Davis' model by taking into account important neuropsychological mechanisms and control processes mediated by executive functions and prefrontal cortical areas. Therefore, this model attempts to explain and understand the development and maintenance of both generalized Internet addiction [GIA] and specific Internet addiction [SIA] (e.g., Internet gaming disorder). According to Brand et al., it is important to distinguish between functional Internet use, GIA, and SIA. While functional Internet use encompasses the use of the Internet as a tool for dealing with personal needs and goals in everyday life in a healthy way, both GIA and SIA may serve different purposes in the context of addiction. Furthermore, in the development and maintenance of GIA, the user has some needs and goals that can be satisfied using certain Internet applications. It is assumed that psychopathological symptoms (e.g., depression, anxiety) are predisposing factors for developing GIA. Moreover, social cognitions (e.g., perceived social isolation, lack of offline support) are also assumed to be related to GIA.

Accordingly, particular emphasis is given to Internet use expectancies as it may involve anticipations of how the Internet can be helpful for distracting individuals from thinking about their problems and/or escaping from reality. Such activity can also be used to enhance positive mood states and/or minimize negative mood states. These expectancies may also interact with the user's general coping style and self-regulation capacities. Therefore, when going online, the user receives reinforcement in terms of dysfunctional coping strategies with negative feelings or problems in everyday life. While Internet use expectancies are positively reinforced, given the strong reinforcement character of certain Internet applications, the cognitive control concerning the Internet use becomes more effortful. This should particularly be the case if Internet-related cues interfere with executive processes.

In the development and maintenance of SIA, it is argued by Brand et al. (2014) that psychopathological symptoms are also particularly involved in this type of Internet addiction. Therefore, it is hypothesized that specific person's predispositions increase the probability that an individual receives gratification from the use of certain applications and overuses these applications again. In this framework, it is postulated that the expectancy that such Internet applications can satisfy certain desires increases the likelihood that these applications will be used frequently and that the individual begins to lose of control over the use of such applications. Consequently, gratification is experienced, and therefore the use of such applications and also the specific Internet use expectancies and the coping style are reinforced positively. Another assumption of this perspective for understanding SIA is that the more general psychopathological tendencies (e.g., depression, social anxiety) are negatively reinforced due to the fact that additional specific Internet applications can be used to distract from problems in the real life or to avoid negative feelings, such as loneliness or social isolation. Contrary to the theory proposed by Davis where the model of PIU has been put to test by developing a theory-driven instrument to assess Internet addiction, the theoretical framework put forth by Brand et al. still remains to be tested empirically.

To date, there have been many types of excessive and/or problematic activity that have been conceptualized as a technological and/or online addiction. This has included television addiction, mobile phone addiction, video game addiction, Internet addiction, social networking addiction, online auction addiction, online sex addiction, and online gambling addiction. Due to space restraints, a review of all of these different types of entertainment addictions is beyond the scope of this chapter. Therefore, the rest of this chapter briefly overviews the empirical research on two entertainment addictions – online gaming addiction and social networking addiction.

A Brief Overview of Online Gaming Addiction

Following the release of the first commercial video games in the early 1970s, it took until the 1980s for the first reports of video game addiction to appear in the psychological and psychiatric literature (e.g., Egli and Meyers 1984; Keepers 1990; Shotton 1989; Soper and Miller 1983). However, these studies were somewhat observational, anecdotal, and/or case studies, primarily based on samples of teenage males, mostly based on a particular type of video game using a particular medium (i.e., "pay-to-play" arcade video games). The 1990s saw a small but significant increase of research into video game addiction with all of these studies being carried out in the UK and on adolescents typically surveying children in school settings (e.g., Griffiths and Hunt 1998; Phillips et al. 1995). In contrast to the early 1980s studies, these studies mainly examined non-arcade video game playing (i.e., home console games, handheld games, PC gaming). However, all of these studies were self-report surveys, relatively small scale, and the main problem was that all of them assessed video game addiction using adapted versions of the DSM-III-R or DSM-IV criteria for pathological gambling. Although there are clearly many similarities between gambling and video gaming, they are different behaviors, and specific video game screening instruments should have been developed. Based on further analysis of the adapted DSM criteria used, these studies were later criticized as being more likely to be assessing video game preoccupation rather than video game addiction (Charlton 2002).

The 2000s saw a substantial growth in the number of studies on video game addiction particularly as gaming expanded into the new online medium where games could be played as part of a gaming community (i.e., massively multiplayer online role playing games [MMORPGs] such as *World of Warcraft* and *Everquest*). According to a number of systematic reviews (e.g., Kuss and Griffiths 2012), approximately sixty studies were published on gaming addiction between 2000 and 2010, and a vast majority of these examined MMORPG addiction and was not limited to the study of adolescent males. Furthermore, many of these studies collected their data online, and a significant minority of studies examined various other aspects of video game addiction using nonself-report methodologies. These include studies using polysomnographic measures and visual and verbal memory tests (Dworak et al. 2007), medical examinations including the patient's history, and physical, radiologic, intraoperative, and pathologic findings (Cultrara and Har-El 2002), functional magnetic resonance imaging (Han et al. 2010; Hoeft et al. 2008; Ko et al. 2009), electroencephalography (Thalemann et al. 2007), and genotyping (Han et al. 2007).

When looking at estimated prevalence rates of problematic video gaming, they range from 1.7 % (e.g., Rehbein et al. 2010) to 8–10 % among general samples (e.g., Gentile et al. 2011). Prevalence rates among video game players were in some cases much higher (e.g., Grusser et al. 2007; Hussain et al. 2012). These studies also indicate that, in general, males are significantly more likely than females to report problems relating to their gaming. According to King and colleagues (2012, 2013), the differences in methods of assessing game-based problems may partly account for differences in prevalence rates. Furthermore, many studies fail to assess prior problems (i.e., lifetime prevalence). King et al. (2012, 2013) also noted that some studies do not consider subclinical cases (i.e., meeting some but not all criteria for problematic use), and the presence of comorbid psychopathology is not routinely assessed.

From a substantive perspective, there are some generalizations that can be made with regard to the demographic characteristics of gamers and problem gamers. The literature, to date, suggests that adolescent males and young male adults appear to be at greater risk of experiencing problematic video game play. However, the course and severity of these problems are not well known, and the finding that this group is more at risk may be a consequence of sampling bias as well as the fact that this group plays video games more frequently than other sociodemographic groups (Griffiths et al. 2012). It has also been suggested that university students may be especially vulnerable to developing problematic video gaming. Reasons for this include their flexible tuition and study hours, ready access to high-speed broadband on a 24/7 basis, and multiple stressors associated with adjusting to new social obligations and/or living out of home for the first time (King et al. 2012).

Irrespective of whether problematic video game play can be classed as an addiction, there is now a relatively large number of studies all indicating that excessive video game play can lead to a wide variety of negative psychosocial consequences for a minority of affected individuals. These were summarized by Griffiths et al. (2012) and include sacrificing work, education, hobbies, socializing, time with partner/family, and sleep, increased stress, an absence of real-life relationships, lower psychosocial well-being and loneliness, poorer social skills, decreased academic achievement, increased inattention, aggressive/oppositional behavior and hostility, maladaptive coping, decreases in verbal memory performance, maladaptive cognitions, and suicidal ideation. This list of potential psychosocial consequences clearly indicates that excessive gaming is an issue irrespective of whether it is an addiction. In addition to the reported negative psychosocial consequences, Griffiths et al. (2012) also noted there are many reported health and medical consequences that may result from excessive video game playing. These include epileptic seizures, auditory hallucinations, enuresis, encoprisis, obesity, wrist pain, neck pain, tenosynovitis, blisters, calluses, sore tendons, numbness of fingers, hand-arm vibration syndrome, sleep abnormalities, psychosomatic challenges, and repetitive strain injuries.

A number of studies have examined the role of different personality factors, comorbidity factors, and biological factors and their association with gaming addiction. In relation to personality traits, a review by Griffiths et al. (2012) reported that gaming addiction had been shown to have association with neuroticism, aggression and hostility, avoidant and schizoid interpersonal tendencies, loneliness and introversion, social inhibition, boredom inclination, sensation seeking, diminished agreeableness, diminished self-control and narcissistic personality traits, low self-esteem, state and trait anxiety, and low emotional intelligence. It is hard to assess the etiological significance of these associations with gaming addiction as they may not be unique to the disorder. Further research is therefore needed. Research has also shown gaming addiction to be associated with a variety of comorbid disorders. This includes attention deficit hyperactivity disorder, symptoms of generalized anxiety disorder, panic disorder, depression, social phobia, school phobia, and various psychosomatic symptoms (Griffiths et al. 2012).

Through use of fMRI, biological research has shown that gaming addicts show similar neural processes and increased activity in brain areas associated with substance-related addictions and other behavioral addictions, such as pathological gambling (e.g., significant activation in the left occipital lobe, parahippocampal gyrus, dorsolateral prefrontal cortex, nucleus accumbens, right orbitofrontal cortex, bilateral anterior cingulate, medial frontal cortex, and the caudate nucleus) (Han et al. 2010; Hoeft et al. 2008; Ko et al. 2009). It has also been reported that gaming addicts (like substance addicts) have a higher prevalence of two specific polymorphisms of the dopaminergic system (i.e., Taq1A1 allele of the dopamine D2 receptor and the Val158Met in the Catecholamine-O-Methyltransferase) (Han et al. 2009) which suggests that among some players there might be some genetic predisposition to develop video game addiction.

But what is it that makes gaming so addictive for the small minority? It is the authors' view that, addiction is all about constant reinforcement, or put more simply, being constantly rewarded while playing the game. Gaming rewards can be physiological (e.g., feeling "high" or getting a "buzz" while gamers are playing or beating their personal high scores), psychological (e.g., the feeling gamers get when they have complete control in a specific situation or know their strategic play helped them win), social (e.g., being congratulated by fellow gamers when doing something well in the game), and, in some cases, financial (e.g., winning a gaming tournament). Most of these rewards are – at least to some extent – unpredictable. Not knowing when the next reward will come keeps some players in the game. In short, they carry on gaming even though they may not have received an immediate reward. They simply hope that another reward is "just around the corner" and keep on playing.

As a consequence of this upsurge in research over the last decade, the Substance Use Disorder Work Group (SUDWG) recommendeded that the DSM-5 (American Psychiatric Association 2013) include a sub-type of "Internet gaming disorder" [IGD] in section 3, the Appendix, ("Emerging Measures and Models"). The main text includes all the mental disorders that have been clinically identified and that have robust empirical support. Section 3 of the DSM is outside of the main text and features disorders that have some empirical evidence but that still need further empirical validation before appearing in the main text. According to Petry and O'Brien (2013), IGD will not be included as a separate mental disorder in the main text of the DSM until the (i) defining features of IGD have been identified, (ii) reliability and validity of specific IGD criteria have been obtained cross-culturally, (iii) prevalence rates have been determined in representative epidemiological samples across the world, and (iv) etiology and associated biological features have been evaluated.

One of the key reasons that IGD was not included in the main text of the DSM-5 was that the SUDWG concluded that no standard diagnostic criteria were used to assess gaming addiction across most studies. A recent review of instruments assessing problematic, pathological, and/or addictive gaming by King and colleagues (2013) reported that 18 different screening instruments had been developed and that these had been used in 63 quantitative studies comprising 58,415 participants. This comprehensive review identified both strengths and weaknesses of these instruments. The main strengths of the instrumentation included: (i) the brevity and ease of scoring, (ii) excellent psychometric properties such as convergent validity and internal consistency, and (iii) robust data that may aid the development of standardized norms for adolescent populations. However, the main weaknesses identified in the instrumentation included: (i) core addiction indicators being inconsistent across studies, (ii) a general lack of any temporal dimension, (iii) inconsistent cut-off scores relating to clinical status, (iv) poor and/or inadequate inter-rater reliability and predictive validity, and (v) inconsistent and/or dimensionality. It has also been noted by a number of authors that the criteria for IGD assessment tools are theoretically based on a variety of different potentially problematic activities including substance use disorders, pathological gambling, and/or other behavioral addiction criteria. There are also issues surrounding the settings in which diagnostic screens are used as those used in clinical practice settings may require a different emphasis than those used in epidemiological, experimental, and neurobiological research settings.

The fact that IGD was included in Section 3 of the DSM-5 appears to have been well received by researchers and clinicians in the gaming addiction field (and by those individuals that have sought treatment for such disorders and had their experiences psychiatrically validated as this may help them feel less stigmatized). Irrespective of approach or model, the components and dimensions that comprise addiction outlined at the start of this chapter are very similar to the IGD criteria in Section 3 of the DSM-5. For instance, the six addiction components directly map onto the nine proposed criteria for IGD (of which five or more need to be endorsed and resulting in clinically significant impairment). More specifically: (1) *preoccupation with Internet games* [salience]; (2) *withdrawal symptoms when Internet gaming is taken away* [withdrawal]; (3) *the need to spend increasing amounts of time engaged in Internet gaming* [tolerance]; (4) *unsuccessful attempts to control participation in Internet gaming* [relapse/loss of control]; (5) *loss of interest in hobbies and entertainment as a result of, and with the exception of, Internet gaming* [conflict]; (6) *continued excessive use of Internet games despite knowledge of* *psychosocial problems* [conflict]; (7) *deception of family members, therapists, or others regarding the amount of Internet gaming* [conflict]; (8) *use of the Internet gaming to escape or relieve a negative mood* [mood modification]; and (9) *loss of a significant relationship, job, or educational or career opportunity because of participation in Internet games* [conflict].

However, as suggested by several authors (e.g., Griffiths et al. 2014), for IGD to be included in the section on "substance-related and addictive disorders," the gaming addiction field must unite and start using the same assessment measures so that comparisons can be made across different demographic groups and different cultures. For epidemiological purposes, Koronczai and colleagues (2011) asserted that the most appropriate measures in assessing problematic online use (including Internet gaming) should meet six requirements. Such an instrument should have: (i) brevity (to make surveys as short as possible and help overcome question fatigue), (ii) comprehensiveness (to examine all core aspects of IGD as possible), (iii) reliability and validity across age groups (e.g., adolescents vs. adults), (iv) reliability and validity across data collection methods (e.g., online, face-to-face interview, paper and pencil), (v) cross-cultural reliability and validity, and (vi) clinical validation. It was also noted that an ideal assessment instrument should serve as the basis for defining adequate cut-off scores in terms of both specificity and sensitivity. To fulfill all these requirements, future research should adjust the currently used assessment tools to the newly accepted DSM-5 criteria and take much more efforts to reach and study clinical samples, which is an unequivocal shortcoming of gaming research.

A Brief Overview of Social Networking Addiction

Social networking sites (SNSs) are virtual communities where users can create individual public profiles, interact with real-life friends, and meet other people based on shared interests. SNS usage patterns from both consumer research and empirical research indicate that overall, regular SNS use has increased substantially over the last few years (Kuss and Griffiths 2011). SNSs are predominantly used for social purposes, mostly related to the maintenance of established offline networks, relative to individual ones (Kuss and Griffiths 2011). However, recent evidence suggests that individuals may feel compelled to maintain their online social networks in a way that may, in some circumstances, lead to using SNSs excessively.

Based on the relatively sparse literature to date, it would appear that in some individuals, SNS addiction incorporates the experience of the "classic" addiction symptoms, namely, mood modification (i.e., engagement in SNSs leads to a favorable change in emotional states), salience (i.e., behavioral, cognitive, and emotional preoccupation with the SNS usage), tolerance (i.e., ever increasing use of SNSs over time), withdrawal symptoms (i.e., experiencing unpleasant physical and emotional symptoms when SNS use is restricted or stopped), conflict (i.e., interpersonal and intrapsychic problems ensue because of SNS usage), and relapse (i.e., addicts quickly revert back to their excessive SNS usage after an abstinence period).

It is generally accepted that a combination of biological, psychological, and social factors contributes to the etiology of addictions that may also hold true for SNS addiction (Griffiths et al. 2014). From this it follows that SNS addiction shares a common underlying etiological framework with other substance-related and behavioral addictions. However, due to the fact that the engagement in SNSs is different in terms of the actual expression of (Internet) addiction (i.e., pathological use of SNSs rather than other Internet applications), the phenomenon may be worthy of individual consideration, particularly when considering the potentially detrimental effects of both substance-related and behavioral addictions on individuals who experience a variety of negative consequences because of their addiction.

According to a recent review (Griffiths et al. 2014), the twenty or so empirical studies examining SNS addiction fall into one of four types: (i) self-perception studies of social networking addiction, (ii) studies of social networking addiction utilizing a social networking addiction scale, (iii) studies examining the relationship between social networking and other online addictions, and (iv) studies examining social networking addiction and interpersonal relationships. The review noted that all the studies suffered from a variety of methodological limitations. Many of the studies attempted to assess SNS addiction, but mere assessment of addiction tendencies does not suffice to demarcate real pathology. Most of the study samples were generally small, specific, self-selected, convenient, and skewed with regards to young adults and female gender (e.g., Andraessen et al. 2012; Cabral 2011; Cam and Isbulan 2012; Elphinston and Noller 2011; Koc and Gulyagci 2013; Turel and Serenko 2012). This may have led to the very high addiction prevalence rates (up to 34 %) reported in some studies as individuals from these sociodemographic groups are likely to be more heavy social networking users. Consequently, empirical studies need to ensure that they are assessing addiction rather than excessive use and/or preoccupation.

It was recently noted that for many researchers, Facebook addiction has become almost synonymous with social networking addiction (Griffiths 2012b). However, Facebook is just one of many websites where social networking can take place. Most of the scales that have been developed have specifically examined excessive Facebook use such as the Bergen Facebook Addiction Scale (Andraessen et al. 2012), the Facebook Addiction Scale (Cam and Isbulan 2012), and the Facebook Intrusion Questionnaire (Elphinston and Noller 2011), i.e., addiction to one particular commercial company's service (i.e., Facebook) rather than the whole activity itself (i.e., social networking). The real issue here concerns what people are actually addicted to and what the new Facebook addiction tools are measuring.

For instance, Facebook users can play games like *Farmville*, can gamble on games like poker, can watch videos and films, and can engage in activities such as swapping photos or constantly updating their profile and/or messaging friends on the minutiae of their life (Griffiths 2012b; Kuss and Griffiths 2011). Therefore, "Facebook addiction" is not synonymous with "social networking addiction" – they are two fundamentally different things as Facebook has become a specific website where many different online activities can take place – and may serve different purposes to various users. What this suggests is that the field needs a psychometrically validated scale that specifically assesses “social networking addiction” rather than Facebook use. In the aforementioned scales, social networking as an activity is not mentioned, therefore the scale does not differentiate between someone potentially addicted to *Farmville* or someone potentially addicted to constantly messaging Facebook friends.

To explain the formation of SNS addiction, Turel and Serenko (2012) summarized three overarching theoretical perspectives that may not be mutually exclusive:

- *Cognitive-behavioral model*: This model emphasizes that “abnormal” social networking arises from maladaptive cognitions and is amplified by various environmental factors and eventually leads to compulsive and/or addictive social networking.
- *Social skill model*: This model emphasizes that “abnormal” social networking arises because people lack self-presentational skills and prefer virtual communication to face-to-face interactions, and it eventually leads to compulsive and/or addictive use of social networking.
- *Socio-cognitive model*: This model emphasizes that “abnormal” social networking arises due to the expectation of positive outcomes, combined with Internet self-efficacy, and deficient Internet self-regulation eventually leads to compulsive and/or addictive social networking behavior.

Based on these three models, Xu and Tan (2012) suggest that the transition from normal to problematic social networking use occurs when social networking is viewed by the individual as an important (or even exclusive) mechanism to relieve stress, loneliness, or depression. They contend that those who frequently engage in social networking are poor at socializing in real life. For these people, social media use provides continuous rewards (e.g., self-efficacy, satisfaction), and they end up engaging in the activity more and more, eventually leading to many problems (e.g., ignoring real-life relationships, work/educational conflicts, etc.). The resulting problems may then exacerbate individuals’ undesirable moods. This then leads such individuals to engage in the social networking behavior even more as a way of relieving dysphoric mood states. Consequently, when social network users repeat this cyclical pattern of relieving undesirable moods with social media use, the level of psychological dependency on social networking increases.

Whether social networking addiction exists is debatable depending upon the definition of addiction used, but there is clearly emerging evidence that a minority of social network users experience addiction-like symptoms as a consequence of their excessive use (Griffiths et al. 2014). Studies endorsing only a few potential addiction criteria are not sufficient for establishing clinically significant addiction status. Similarly, significant impairment and negative consequences that discriminate addiction from mere abuse have (to date) generally not been assessed in published studies. Thus, future studies have great potential in addressing the emergent phenomenon of SNS addiction by means of applying better methodological designs, including more representative samples, and using more reliable and valid addiction scales so that current gaps in empirical knowledge can be filled.

In addition, specific attention needs to be paid to selecting larger samples that are representative of a broader population in order to increase the respective study's external validity. The generalizability of results is essential in order to demarcate populations at risk for developing addiction to SNSs. Similarly, it appears necessary to conduct further psychophysiological studies in order to assess the phenomenon from a biological perspective. Furthermore, clear-cut and validated addiction criteria need to be assessed. It is insufficient to limit studies into addiction to assessing just a few criteria. The differentiation of pathology from high frequency and problematic usage necessitates adopting frameworks that have been established by the international classification manuals. Moreover, in light of clinical evidence and practice, it appears essential to pay attention to the significant impairment that SNS addicts experience in a variety of life domains as a consequence of their abusive and/or addictive behaviors.

Similarly, the results of data based on self-reports are not sufficient for diagnosis because research suggests that they may be inaccurate (Griffiths et al. 2014). Conceivably, self-reports may be supplemented with structured clinical interviews and further case study evidence as well as supplementary reports from the users' significant others. Research into social networking addiction is needed specifically in relation to clinical applicability and criteria for diagnosis. Furthermore, research is needed to examine gender differences as there appears to be a higher prevalence of problems among females (as opposed to other problematic online behaviors such as gaming addiction which is more prevalent among males; Kuss and Griffiths 2012). Such observations strengthen the rationale for a clear-cut social networking addiction classification rather than an umbrella term of Internet addiction.

Internet Addiction and Internet Gaming Addiction

As can be seen from earlier in the chapter, research into online addictions has greatly increased. Alongside this, there have been scholarly debates about whether Internet addiction really exists. Some may argue that because Internet use does not involve the ingestion of a psychoactive substance, then it should not be considered a genuine addictive behavior. However, the latest (fifth) edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) (American Psychiatric Association 2013) reclassified "gambling disorder" as an addiction disorder rather than as a disorder of impulse control. The implications of this reclassification are potentially far reaching. The most significant implication is that if an activity that does not involve the consumption of drugs (i.e., gambling) can be a genuine addiction accepted by the psychiatric and medical community, there is no theoretical reason why other problematic and habitual behaviors (e.g., shopping, work, exercise, sex, video gaming, etc.) cannot be classed as a bona fide addiction.

There have also been debates among scholars that consider excessive problematic Internet use to be a genuine addiction as to whether those in the field should study generalized Internet addiction (the totality of all online activities) and/or specific addictions on the Internet such as Internet gambling, Internet gaming, and

Internet sex (Griffiths 2012a; Griffiths and Szabo 2014). As noted in an earlier section of this chapter, Griffiths (2000a) has constantly argued that there is a fundamental difference between addictions *on* the Internet and addictions *to* the Internet. More specifically, it has been argued that the overwhelming majority of individuals that are allegedly addicted to the Internet are not Internet addicts but were individuals that use the medium of the Internet as a vehicle for other addictions (Griffiths 2000a).

Prior to the publication of the latest DSM-5 (American Psychiatric Association 2013), there had also been debates as to whether Internet addiction should be introduced into the text as a separate disorder (Petry and O'Brien 2013). As noted earlier, the DSM-5 now includes a sub-type of problematic Internet use (i.e., Internet gaming disorder [IGD]) in Section 3 ("Emerging Measures and Models"). However, far from clarifying the debates surrounding generalized versus specific Internet use disorders, the section of the DSM-5 discussing IGD noted that:

> There are no well-researched subtypes for Internet gaming disorder to date. Internet gaming disorder most often involves specific Internet games, but it could involve non-Internet computerized games as well, although these have been less researched. It is likely that preferred games will vary over time as new games are developed and popularized, and it is unclear if behaviors and consequence associated with Internet gaming disorder vary by game type. . .Internet gaming disorder has significant public health importance, and additional research may eventually lead to evidence that Internet gaming disorder (also commonly referred to as Internet use disorder, Internet addiction, or gaming addiction) has merit as an independent disorder. (p. 796)

Two immediate problematic issues arise from these assertions. Firstly, IGD is clearly seen as synonymous with Internet addiction as the text claims that Internet addiction and Internet use disorder are simply other names for IGD. Secondly – and somewhat confusingly – it is asserted that IGD (which is by definition Internet based) can also include offline gaming disorders.

With regards to the first assertion, Internet addiction and Internet gaming addiction are not the same, and recent empirical research clearly shows that to be the case. For instance, Király and colleagues (2014) examined the interrelationship and the overlap between Internet addiction disorder and IGD in terms of (among other variables) gender, time spent using the Internet and/or online gaming, and preferred online activities. They collected their data from a nationally representative sample of over 2,000 Hungarian adolescents. They found that IGD was much more strongly associated with being male and that Internet addiction was positively associated with online chatting, online gaming, and social networking, while IGD was only associated with online gaming. The authors argued that IGD appears to be a conceptually different behavior than Internet addiction and that their data supported the notion that Internet addiction and IGD are separate nosological entities. A further complicating factor is that many researchers have used the Internet Addiction Test (Young 1998) to assess online gaming addiction. This may have been one of the reasons as to why the DSM-5 asserted that IGD and Internet addiction are the same disorder.

The second assertion that IGD can include offline video gaming is both baffling and confusing. Some researchers consider video games as the starting point for examining the characteristics of gaming disorder (Griffiths et al. 2012), while others consider the Internet as the main platform that unites different addictive Internet activities, including online games (Young 1998, 1999). For instance, it has been argued that although all addictions have particular and idiosyncratic characteristics, they share more commonalities than differences (i.e., salience, mood modification, tolerance, withdrawal symptoms, conflict, and relapse) and likely reflects a common etiology of addictive behavior (Griffiths 2005). For the present authors, IGD is clearly a sub-type of video game addiction. Similarly, Porter and colleagues (2010) do not differentiate between problematic video game use and problematic online game use. They conceptualized problematic video game use as excessive use of one or more video games resulting in a preoccupation with and a loss of control over playing video games and various negative psychosocial and/or physical consequences. For Young (1999) – and as noted earlier – "cyber-relationship addictions," "cyber-sexual addictions," "net compulsions" (gambling, day trading), and "information overload" are all Internet addictions. However, many would argue that these – if they are addictions – are addictions *on* the Internet, not *to* it.

However, recent studies (Demetrovics et al. 2012; Kim and Kim 2010) have made an effort to integrate both approaches. For instance, Kim and Kim (2010) claim that neither the first nor the second approach adequately captures the unique features of MMORPGs and argue an integrated approach is a necessity. More specifically they argue that: "*Internet users are no more addicted to the Internet than alcoholics are addicted to bottles*" (p. 389). The Internet is just a channel through which individuals may access whatever content they want (e.g., gambling, shopping, chatting, sex). On the other hand, online games differ from traditional standalone games, such as offline video games, in important aspects such as the social dimension or the role-playing dimension that allow interaction with other real players. Consequently, it could be argued that IGD can either be viewed as a specific type of video game addiction, as a variant of Internet addiction, or as an independent diagnosis. However, the idea that IGD can include offline gaming disorders does little for clarity or conceptualization.

Finally, it is also worth mentioning that there are some problematic online behaviors that could be called Internet addictions as they can only take place online. The most obvious activity that fulfills this criterion is social networking as it is a "pure" online activity and does not and cannot take place offline (Griffiths et al. 2014). Other activities such as gambling, gaming, and shopping can still be engaged in offline (as gamblers can go to a gambling venue, gamers can play a standalone console game, shoppers can go to a retail outlet). However, those engaged in social networking would not (if unable to access the Internet) walk into a big room of people and start chatting to them all. However, even if social networking addiction is a genuine Internet addiction, social networking itself is still a specific online application and could still be considered an addiction on the internet, rather than to it.

Based on empirical evidence, Internet gaming addiction (or any of the alternate names used to describe problematic gaming) is not the same as Internet addiction. The gaming studies' field needs conceptual clarity, but as demonstrated, the DSM-5 itself is both misleading and misguided when it comes to the issue of Internet gaming disorder.

Conclusion

This chapter has demonstrated that research into entertainment addictions is becoming an increasingly studied phenomenon. The amount and the quality of research in the gaming addiction field have progressed much over the last decade but are still in its infancy compared to other more established behavioral addictions, such as pathological gambling. There have also been several noticeable trends that can be drawn from reviewing the literature on gaming addiction:

- There has been a significant increase in empirical research decade by decade since the early 1980s.
- There has been a noticeable (and arguably strategic) shift in researching the mode of video game play. In the 1980s, research mainly concerned "pay-to-play" arcade video games. In the 1990s, research mainly concerned standalone (offline) video games played at home on consoles, PCs, or handheld devices. In the 2000s, research mainly concerned online massively multiplayer video games.
- There has been a noticeable shift in how data are collected. Up until the early 2000s, data about video game behavior was typically collected face-to-face, whereas contemporary studies collect data online, strategically targeting online forums where gamers are known to (virtually) congregate. These samples are typically self-selecting and (by default) unrepresentative of the general population. Therefore, generalization is almost always one of the methodological shortcomings of this data collection approach.
- Survey study sample sizes have generally increased. In the 1980s and 1990s, sample sizes were typically in the low hundreds. In the 2000s, sample sizes in their thousands – even if unrepresentative – are not uncommon.
- There has been a diversification in the way data are collected including experiments, physiological investigations, secondary analysis of existing data (such as that collected from online forums), and behavioral tracking studies.
- There has been increased research on adult (i.e., non-child and non-adolescent) samples reflecting the fact that the demographics of gaming have changed.
- There has been increasing sophistication in relation to issues concerning assessment and measurement of problematic video game play and video game addiction. In the last few years, instruments have been developed that have more robust psychometric properties in terms of reliability and validity. However, there are still some concerns as many of the most widely used screening instruments were adapted from adult screens and much of the video game literature has examined children and adolescents. King and colleagues (2013)

> assert that to enable future advances in the development and testing of interventions for video game-related problems, there must be some consensus among clinicians and researchers as to the precise classification of these problems.

Clearly more research is needed on whether activities such as video game addiction and Internet addictions (e.g., social networking addiction) are distinct clinical entities. From a research standpoint, it is evident that excessive entertainment technology use appears to be at least potentially addictive. With respect to video games, there is also a need for a general taxonomy of video games as it could be the case that particular types of games are more addictive than others. Devising such reliable taxonomy of video games may prove challenging to researchers due to the ever evolving nature of games and also the fact that nowadays, the boundaries of specific game genres are blurred because very often one game can feature several sub-genres of games within itself all at once (e.g., Grand Theft Auto, which may be classed as a racing and/or as a first-person shooter game simultaneously). Another major problem is that video games can be played in lots of different ways including handheld consoles, on a personal computer, home video game consoles, on arcade machines, on the Internet, and on other portable devices (e.g., mobile phones, i-Pods). It may be the case that some of these media for playing games (such as in an arcade or on the Internet) may be more addictive because of other factors salient to that medium (e.g., disinhibition on the Internet). Therefore, future research needs to distinguish between excessive play in different media.

There is also the question of developmental effects, i.e., does video game playing or social networking have the same effect regardless of age? It could well be the case that playing video games and social networking have a more pronounced addictive effect in young children but less of an effect (if any) once they have reached adulthood. There is also the social context of playing, i.e., does playing in groups or individually, with or against each other, affect potential addictiveness of games in any way? All of these questions still need further empirical investigation.

It does appear that excessive entertainment technology use can have potentially damaging effects upon a minority of individuals who display compulsive and addictive behavior, and who will do anything possible to “feed their addiction.” Using these individuals in research would help identify the roots and causes of addictive playing and the impact of such behavior on family and school life. It would be clinically useful to illustrate problem cases even following them longitudinally and recording developmental features of technological addictions. This would help determine the variables that are salient in the acquisition, development, and maintenance of such behaviors.

There is no doubt that technology usage among the general population will continue to increase over the next few years and that if social pathologies (including video game addiction and social networking addiction) do exist then this is certainly an area for development that should be of interest and concern to all those involved in the addiction research field. Until there is an established body of literature on the psychological, sociological, and physiological effects of technological addiction, directions for education, prevention, intervention, and treatment will remain limited in scope.

The issues encountered by contemporary researchers and clinicians regarding the assessment of online addictions appear complex and include several factors. Firstly, it has been noted how historically the use of inconsistent heterogeneous and non-consensual nomenclatures to describe what appears to be the same phenomenon (i.e., gaming addiction) has influenced the development of a varied number of definitions and frameworks for understanding and assessing what appears to represent the same disorder. Secondly, despite being important at some point, these definitions and frameworks largely contributed to the "boom" in the development of several psychometric tools for assessing gaming addiction irrespective of their viability. Thirdly, as outlined by reviews on the assessment of most used psychometric tools, these tools have a wide range of problems.

Some of these conceptual problems found in the literature regarding the assessment of gaming addiction are important because without conceptual clarity and empirical support for treatment efficacy, it is also premature to offer clinical guidelines for the treatment of generalized Internet addiction or gaming addiction. Furthermore, because some of the early conceptualizations where online gaming addiction was seen as a sub-type of Internet addiction, a trend for historically assessing online gaming addiction using generalized Internet addiction tools has been established and translated by a substantial number of studies (particularly those focusing on the neuropsychology of gaming addiction) using this method. This in turn, has contributed to some of the methodological problems in the literature regarding the assessment and understanding of gaming addiction.

There is clearly a need to distinguish between addictions *to* the Internet and addictions *on* the Internet. As noted throughout this chapter, gambling addicts who choose to engage in online gambling, as well as a computer game addicts who play online are not Internet addicts – the Internet is just the place where they conduct their chosen (addictive) behavior. These people display addictions *on* the Internet. However, there is also the observation that some behaviors engaged on the Internet (e.g., cybersex, cyberstalking, etc.) may be behaviors that the person would only carry out on the Internet because the medium is anonymous, non-face-to-face, and disinhibiting (Griffiths 2012).

In contrast, it is also acknowledged that there are some case studies that seem to report an addiction to the Internet itself (e.g., Young 1998; Griffiths 2000b). Most of these individuals use functions of the Internet that are not available in any other medium, such as chat rooms or various role playing games. These are people addicted *to* the Internet. However, despite these differences, there seem to be some common findings, most notably reports of the negative consequences of excessive technology use (neglect of work and social life, relationship breakdowns, loss of control, etc.), which are comparable to those experienced with other, more established addictions. Based on the empirical research to date, technological addictions (and thus addictions to entertainment) do exist. However, how prevalent they are remains an area that is highly debatable and still needs to progress further both conceptually and methodologically.

References

American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders – Text Revision*, 5th edn. (American Psychiatric Association, Washington, DC, 2013)

C.S. Andraessen, T. Tosheim, G.S. BrunBerg, S. Pallesen, Development of a Facebook addiction scale. Psychol. Rep. **110**, 501–517 (2012)

M. Brand, K.S. Young, C. Laier, Prefrontal control and internet addiction: a theoretical model and review of neuropsychological and neuroimaging findings. Front. Hum. Neurosci. **8**, 375 (2014)

R.I.F. Brown, Some contributions of the study of gambling to the study of other addictions, in *Gambling Behavior and Problem Gambling*, ed. by W.R. Eadington, J.A. Cornelius (University of Nevada Press, Reno, 1993), pp. 241–272

J. Cabral, Is generation Y addicted to social media? Elon J. Undergrad. Res. Commun. **2**(1), 5–13 (2011)

E. Cam, O. Isbulan, A new addiction for teacher candidates: social networks. Turkish Online J. Educ. Technol. **11**(3), 14–19 (2012)

J.P. Charlton, A factor analytic investigation of computer 'addiction' and engagement. Br. J. Psychol. **93**, 329–344 (2002)

A. Cultrara, G. Har-El, Hyperactivity-induced suprahyoid muscular hypertrophy secondary to excessive video game play: a case report. J. Oral Maxillofac. Surg. **60**, 326–327 (2002)

R.A. Davis, A cognitive-behavioral model of pathological internet use. Comput. Hum. Behav. **17**, 187–195 (2001)

Z. Demetrovics, R. Urbán, K. Nagygyörgy, J. Farkas, M.D. Griffiths, O. Pápay, G. Kokonyei, K. Felvinczi, A. Oláh, The development of the Problematic Online Gaming Questionnaire (POGQ). PLoS ONE **7**(5), e36417 (2012). doi:10.1371/journal.pone.0036417

M. Dworak, T. Schierl, T. Bruns, H.K. Struder, Impact of singular excessive computer game and television exposure on sleep patterns and memory performance of school-aged children. Pediatrics **120**, 978–985 (2007)

E.A. Egli, L.S. Meyers, The role of video game playing in adolescent life: is there a reason to be concerned? Bull. Psychon. Soc. **22**, 309–312 (1984)

R.A. Elphinston, P. Noller, Time to face it! Facebook intrusion and the implications for romantic jealousy and relationship satisfaction. Cyberpsychol. Behav. Soc. Netw. **14**, 631–635 (2011)

D.A. Gentile, H. Choo, A. Liau, T. Sim, D. Li, D. Fung, A. Khoo, Pathological video game use among youths: a two-year longitudinal study. Pediatrics **127**, 319–329 (2011)

M.D. Griffiths, Technological addictions. Clin. Psychol. Forum **76**, 14–19 (1995)

M.D. Griffiths, Internet addiction: does it really exist? in *Psychology and the Internet: Intrapersonal, Interpersonal and Transpersonal Applications*, ed. by J. Gackenbach (Academic, New York, 1998), pp. 61–75

M.D. Griffiths, Internet addiction – time to be taken seriously? Addict. Res. **8**, 413–418 (2000a)

M.D. Griffiths, Does internet and computer "addiction" exist? Some case study evidence. Cyber Psychol. Behav. **3**, 211–218 (2000b)

M.D. Griffiths, A "components" model of addiction within a biopsychosocial framework. J. Subst. Use **10**, 191–197 (2005)

M.D. Griffiths, The role of context in online gaming excess and addiction: some case study evidence. Int. J. Ment. Heal. Addict. **8**, 119–125 (2010)

M.D. Griffiths, Internet sex addiction: a review of empirical research. Addict. Res. Theory **20**, 111–124 (2012a)

M.D. Griffiths, Facebook addiction: concerns, criticisms and recommendations. Psychol. Rep. **110**(2), 518–520 (2012b)

M.D. Griffiths, N. Hunt, Dependence on computer games by adolescents. Psychol. Rep. **82**, 475–480 (1998)

M.D. Griffiths, D.J. Kuss, Z. Demetrovics, Social networking addiction: an overview of preliminary findings, in *Behavioral Addictions: Criteria, Evidence and Treatment*, ed. by K. Rosenberg, L. Feder (Elsevier, New York, 2014), pp. 119–141

M.D. Griffiths, A. Szabo, Is excessive online usage a function of medium or activity? An empirical pilot study. J. Behav. Addict. **3**, 74–77 (2014)
M.D. Griffiths, D.J. Kuss, D.L. King, Video game addiction: past, present and future. Curr. Psychiatr. Rev. **8**, 308–318 (2012)
M.D. Griffiths, D.L. King, Z. Demetrovics, DSM-5 internet gaming disorder needs a unified approach to assessment. Neuropsychiatry **4**(1), 1–4 (2014). doi:10.2217/npy.13.82
S.M. Grüsser, R. Thalemann, M.D. Griffiths, Excessive computer game playing: evidence for addiction and aggression? Cyber Psychol. Behav. **10**, 290–292 (2007)
D.H. Han, Y.S. Lee, C. Na, J.Y. Ahn, U.S. Chung, M.A. Daniels et al., The effect of methylphenidate on internet video game play in children with attention-deficit/hyperactivity disorder. Compr. Psychiatry **50**, 251–256 (2009)
D.H. Han, J.W. Hwang, P.F. Renshaw, Bupropion sustained release treatment decreases craving for video games and cue-induced brain activity in patients with internet video game addiction. Exp. Clin. Psychopharmacol. **18**, 297–304 (2010)
F. Hoeft, C.L. Watson, S.R. Kesler, K.E. Bettinger, A.L. Reiss, Gender differences in the mesocorticolimbic system during computer game-play. J. Psychiatr. Res. **42**, 253–258 (2008)
Z. Hussain, M.D. Griffiths, T. Baguley, Online gaming addiction: classification, prediction and associated risk factors. Addict. Res. Theory **20**, 359–371 (2012)
G.A. Keepers, Pathologicical preoccupation with video games. J. Am. Acad. Child Adolesc. Psychiatry **29**, 49–50 (1990)
D.L. King, P.H. Delfabbro, M.D. Griffiths, Clinical interventions for technology-based problems: excessive internet and video game use. J. Cogn. Psychother. Int. Q. **26**, 43–56 (2012)
D.L. King, M.C. Haagsma, P.H. Delfabbro, M.S. Gradisar, M.D. Griffiths, Toward a consensus definition of pathological video-gaming: a systematic review of psychometric assessment tools. Clin. Psychol. Rev. **33**, 331–342 (2013)
M.G. Kim, J. Kim, Cross-validation of reliability, convergent and discriminant validity for the problematic online game use scale. Comput. Hum. Behav. **26**, 389–398 (2010)
O. Király, M.D. Griffiths, R. Urbán, J. Farkas, G. Kökönyei, Z. Elekes, D. Domokos Tamás, Z. Demetrovics, Problematic internet use and problematic online gaming are not the same: findings from a large nationally representative adolescent sample. Cyberpsychol. Behav. Soc. Netw. **17**, 749–754 (2014)
C.H. Ko, G.C. Liu, S.M. Hsiao, J.Y. Yen, M.J. Yang, W.C. Lin et al., Brain activities associated with gaming urge of online gaming addiction. J. Psychiatr. Res. **43**, 739–747 (2009)
M. Koc, S. Gulyagci, Facebook addiction among Turkish college students: the role of psychological health, demographic, and usage characteristics. Cyberpsychol. Behav. Soc. Netw. **16**, 279–284 (2013)
B. Koronczai, R. Urban, G. Kokonyei, B. Paksi, K. Papp, B. Kun, Z. Demetrovics, Confirmation of the three-factor model of problematic internet use on off-line adolescent and adult samples. Cyberpsychol. Behav. Soc. Netw. **14**, 657–664 (2011)
D.J. Kuss, M.D. Griffiths, Online social networking and addiction: a literature review of empirical research. Int. J. Environ. Public Health **8**, 3528–3552 (2011)
D.J. Kuss, M.D. Griffiths, Online gaming addiction: a systematic review. Int. J. Ment. Heal. Addict. **10**, 278–296 (2012)
N.M. Petry, C.P. O'Brien, Internet gaming disorder and the DSM-5. Addiction **108**, 1186–1187 (2013)
C.A. Phillips, S. Rolls, A. Rouse, M.D. Griffiths, Home video game playing in schoolchildren: a study of incidence and patterns of play. J. Adolesc. **18**, 687–691 (1995)
G. Porter, V. Starcevic, D. Berle, P. Fenech, Recognizing problem video game use. Aust. N. Z. J. Psychiatry **44**, 120–128 (2010)
F. Rehbein, M. Kleimann, T. Mossle, Prevalence and risk factors of video game dependency in adolescence: results of a German nationwide survey. Cyberpsychol. Behav. Soc. Netw. **13**, 269–277 (2010)

M. Shotton, *Computer Addiction?: A Study of Computer Dependency* (Taylor and Francis, London, 1989)
W.B. Soper, M.J. Miller, Junk time junkies: an emerging addiction among students. Sch. Couns. **31**, 40–43 (1983)
R. Thalemann, K. Wölfling, S.M. Grüsser, Specific cue reactivity on computer game-related cues in excessive gamers. Behav. Neurosci. **12**, 614–618 (2007)
O. Turel, A. Serenko, The benefits and dangers of enjoyment with social networking websites. Eur. J. Inf. Syst. **21**, 512–528 (2012)
H. Xu, B.C.Y. Tan, Why do I keep checking Facebook: effects of message characteristics on the formation of social network services addiction (2012), http://elibrary.aisnet.org/Default.aspx?url=http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1216&context=icis2012
K. Young, *Caught in the Net: How to Recognize the Signs of Internet Addiction and a Winning Strategy for Recovery* (Wiley, New York, 1998)
K.S. Young, Caught in the net: how to recognize the signs of Internet addiction - and a winning strategy for recovering. New York: John Wiley & Sons, Inc. (1998)
K. Young, Internet addiction: evaluation and treatment. Stud. Br. Med. J. **7**, 351–352 (1999)

Applied Entertainment: Positive Uses of Entertainment Media

47

Jeffrey H. Goldstein

Contents

Abstract

This chapter considers the role of entertainment media in education, health, and quality of life. Because of its potential to affect our well-being, entertainment can be seen as a public health issue. When freely chosen, entertainment can produce desired states such as relaxation or arousal and can induce the range of human emotions that enrich daily life. The emotional and social satisfactions provided by entertainment are supplemented by their impact on executive functioning and health.

Entertainment serves the range of “uses and gratifications” familiar to media students (cognitive, social, emotional/physiological).

J.H. Goldstein (✉)
Institute for Cultural Inquiry, Utrecht University, Utrecht, The Netherlands
e-mail: J.Goldstein@uu.nl

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_9

Among the cognitive benefits of entertainment media are the maintenance or improvement of problem solving and enhanced perceptual skills. Listening to music or watching television can produce positive cognitive effects. Music, in addition to its mood management function, also affects brain development, language, and cognitive development.

One undeniable feature of play is fun. Positive emotions, including humor, contribute to a sense of well-being and health.

Video gaming can be beneficial for brain development and functioning. The positive effects of video gaming may also prove relevant in therapeutic interventions targeting psychiatric disorders, particularly depression. Studies of the noninstitutionalized elderly suggest that digital games can speed reaction time and may positively influence executive function and have social and emotional benefits. Exergames are a substitute for physical exercise when outdoor play is not feasible.

If entertainment is a public health issue, it is largely in the area of mental health that it has its greatest impact. Enjoying music, a film, a video game, or a You Tube video can improve mood, strengthen friendships, and increase competence.

Digital entertainment media have been used in basic scientific research. Games can teach STEM subjects efficiently by reaching a large audience.

Keywords
Entertainment • Video games • Film • Television • Music • Humor • Health • Education

Introduction

Entertainment is a public health issue, and this chapter seeks to illustrate this assertion with a selection of recent studies (see Spencer 2006). The emphasis is on the application of ordinary "off-the-shelf" entertainment – film, television, music, and video games – for the enlightenment, health, and well-being of its audience, participants, or users. Not included are media produced expressly to teach, such as "serious games" (see the *International Journal of Serious Games*) (http://journal.seriousgamessociety.org/index.php?journal=IJSG), or new technologies, such as robots and virtual reality, designed to achieve practical goals, such as household chores. The primary focus here is on media produced or used primarily to entertain or amuse rather than enlighten, where learning and healthful outcomes occur as a by-product of play and entertainment.

The global entertainment industry – film, television, video games, and recorded music – was estimated to be worth $1.6 trillion in 2012, making it one of the world's largest industries. The average person spends nearly 12 h/day using screen-based media, half of which is for entertainment (http://www.market-truth.com/screen-time-hours-day-spend-front-screens/).

In 2004, Rauterberg reviewed the literature on positive effects of entertainment technology, examining the content and context of use in education, prosocial behavior, and therapy. This chapter continues this selective review of the uses of entertainment. The studies reviewed are intended to suggest the wealth of possibilities for the uses and applications of entertainment media, even though for some of the topics discussed in this chapter there are also studies with null or negative results.

Studies in the areas of education, health and medicine, mental health and personal well-being, and science are reviewed. The underlying theoretical thread that runs through these discussions is loosely referred to as "uses and gratifications theory" and includes Zillmann's mood management theory. In this view, the selection of leisure activities, including media content, is based on an individual's expectation of what that selection will deliver in terms of satisfactions. According to mood management theory, the selection and consumption of specific media content can serve to alter and regulate prevailing physiological and mood states. To the extent that entertainment media are used to enhance positive emotions or to cope with negative ones, they contribute to health and well-being.

Music, films, television series, and video games are all used to teach, heal, and amuse. Business schools use episodes of the TV series *The Office* in lessons on management styles and office politics. Medical schools rely increasingly on video games as training methods, diagnostic tools, and therapeutic interventions. Entertainment can promote social bonds and a sense of community, as happens on a massive scale during World Cup football matches.

Learning and Education

Entertainment varies in the degree to which information and instruction are built into the product. Many games have a dual intention, sometimes referred to as "gamification" or "edutainment." They mix play and instruction in a hybrid that can be referred to as "guided play." For instance, games like Pixelberry's *High School Story*, developed in partnership with the National Eating Disorders Association, can be tailored to address issues such as eating disorders and obesity.

When media have a specific educational goal, we can speak of playful learning, a combination of free play and guided play (Fisher et al. 2011). What is learned during play? And does this transfer to situations outside of play? The capacity model developed by Fisch (2000) to account for transfer of learning from television is applicable to other media, including video games and film. Transfer is most likely when four conditions are met: The audience must understand the content, must create an abstract mental representation of that content, remember the content and see its relation to the new problem or setting, and must apply the remembered content to the new problem.

Play and Learning

Play has been defined as any activity freely chosen, intrinsically motivated, and personally directed. It has no goal other than itself. The cognitive processes involved in play are similar to those involved in all learning: motivation, meaning, repetition, self-regulation, and higher-order information processing.

Neuroscientist Jaak Panksepp found that play stimulates production of a protein, "brain-derived neurotrophic factor," in the amygdala and the prefrontal cortex, areas responsible for organizing, monitoring, and planning for the future (see Panksepp 2010). In one study, 2 h a day of play with objects produced changes in the brain weight and efficiency of experimental animals (Rosenzweig and Bennett 1972).

Play-based education has been popular at least since Friedrich Froebel (1782–1852) invented the Kindergarten and Froebel toys to develop specific skills. *Sesame Street* is a modern manifestation of this notion (see Fisch 2004). For example, playing with blocks promotes language development. In one study, children aged one and a half to two and a half who were provided with sets of molded plastic building bricks with which to play had significantly higher language scores 6 months later, compared with a control group (Christakis et al. 2007). Sets of plastic building bricks were distributed to a random group of families with children aged 1½–2½ years who were registered at a pediatrics clinic. The parents also received two newsletters with suggestions of things that they and their child could do with the blocks (such as sorting blocks by color and seeing how big a stack they could make). Data were gathered from 140 families. Most of the children who received blocks reported playing with them. Receiving blocks was associated with significantly higher language scores in a sample of middle- and low-income children 6 months later. The researchers conclude that playing with blocks can lead to improved language development in children from middle- and low-income families.

Why does block play promote language? As children manipulate objects, they begin to understand more about their qualities and relations. Older children begin to make up stories or scripts for these objects, which underlie further understanding of them. One theory that may explain associations between early exposure and subsequent cognitive and linguistic outcomes is based on the development of mental schemes that Vygotsky referred to as "scaffolding." Mental schemes are internal models of the world that a child uses to understand and master the environment. They are the precursors of thought and language. "Through play, that is, unstructured manipulation of objects, the child begins to develop a mental picture of and cognitive categories about the objects around him or her. These mental schemes underlie an understanding of object permanence, the development of memory, and the roots of impulse control and language" (Christakis et al. 2007).

In a study of *Baby Einstein* videos, Zimmerman, Christakis, and Meltzoff (2007) reported that exposure to *Baby Einstein* videos was negatively associated with language development. Ferguson and Donnellan (2014) replicated their study. Caregivers of 392 children aged 6–16 months and 358 children aged 17–27 months reported on media exposure, language development, and control variables related to child/parent interaction and demographic characteristics. Results indicated that exposure to baby videos could be construed as positive, neutral, or negative depending upon the statistical analysis. The effect size estimates were generally negligible across analyses. Exposure to educational programming tended to be positively related to language development. Infants exposed to no media actually had lower levels of language development compared to infants with some exposure. Given these results, the Zimmerman et al. (2007) data set does not permit strong inferences about a connection between exposure to media and language development in young children.

Young pre-readers or early readers, aged 3–5, may benefit from electronic toys that provide spoken instructions and reading help. Book readers allow the child to touch a page or a word to hear letters, words, or entire stories read aloud. Computers, smart toys, and other new media have the potential to facilitate language development if parents mediate the experience, for example, labeling objects, asking questions, repeating dialogue, and describing the content (Hirsh-Pasek et al. 2015; Weber 2006).

Video Games and Learning

In their everyday lives, children come across barcode scanners in supermarkets, mobile telephones, portable computers, cash dispensers, and parking lot ticket machines, all of which confront them with modern technology. Electronic toys and digital games are the child's first hands-on introduction to this world, giving children an opportunity to learn about technology as well as with technology.

Video games can be powerful brain-training tools that can improve such cognitive skills as visual attention, concentration, navigation, multitasking, and task switching, all while simultaneously increasing speed and accuracy. Games have taught reading, vocabulary, and math to elementary school pupils and health care and safe sexual practices to adolescents (Kato 2010).

Until around age 2 years, flat screens do not hold much interest to a child, and the child's limited fine motor control makes using computers and electronic game devices difficult. Preschoolers, aged 3–5, increasingly realize their ability to influence events on a screen and use and pretend using cell phones and computers. From ages 6 to 12, the full plethora of contemporary media becomes available.

Manipulating blocks on a computer screen, as in *Tetris*, may promote the same abilities as manipulating real blocks. Computerized blocks were devised by Itoh et al. (2005) using a simple interface that requires no computer expertise. Shapes in the physical environment are then represented on a screen, enabling users to interact with cyberspace via these physical objects. First, using a set of computerized blocks, children construct a shape with which they want to play in cyberspace. The computer automatically recognizes the constructed structure in real time and then retrieves 3D virtual models closely matching the constructed structure. Children then play the virtual model's multimedia contents. Children can play in cyberspace while manipulating the constructed object in their hands. Hybrid video games-interactive toys, such as Disney's *Infinity 2*, are a current manifestation of this technology.

Digital games seem to have the same basic functions as other media in serving adolescents' mood management, stimulation seeking, social relationships, and self-presentation. The challenge and mastery that accompany play and games, involvement, and opportunities for social contact, both during and after play online and on the playground, are all powerful attractors. Motivations for playing video games vary with age and developmental stage.

An overview of the benefits of playing video games was published in 2014 by Granic, Lobel, and Engels. They note that 97 % of children and adolescents in the USA play video games for at least 1 h per day. Potential benefits are noted in four domains: cognitive, motivational, emotional, and social. For example, playing video games, including first-person shooter games, may boost children's learning, health, and social skills. Playing video games can strengthen a range of cognitive skills such as spatial navigation, reasoning, memory, and perception. A 2013 meta-analysis found that playing violent video games was just as successful as academic courses in improving a player's capacity to think about objects in three dimensions. This has implications for education and career development, given the importance of spatial skills for achievement in science, technology, engineering, and mathematics. Playing video games may help children develop problem-solving skills. The more adolescents reported playing strategic video games, such as role-playing games, the more they improved in problem solving and school grades the following year (Granic et al. 2014).

Johnson (2005) argues that video games and television dramas confer valuable cognitive skills on players and viewers. Popular culture is increasingly complex and intellectually challenging. For example, instead of having a single plot, today's best TV shows require us to follow as many as a dozen plot threads in a single episode. Scripts contain numerous allusions to previous episodes and external stories, novels, and events. Close attention is required to understand what is going on, and this effort makes us smarter, Johnson argues.

Video games are particularly good at training strategic thinking and problem solving. Computer games incorporate many learning principles by putting learners in the role of decision-maker, confronting them through progressively harder challenges, and engaging the player in experimenting with different ways of learning and thinking. Well-produced simulation games encourage visualization, experimentation, and creativity in developing strategy.

Children learn to use these media largely through trial and error – through exploration, experimentation and play, and collaboration with others, both in face-to-face and virtual forms. Electronic game playing is a "multi-literate" activity: it often involves interpreting complex three-dimensional visual environments, reading both on-screen and off-screen texts (such as games magazines and websites), and processing auditory information. In computer games, success derives from the acquisition of skills and knowledge. Young people have to learn to "read" subtle nuances, often on the basis of minimal cues. They have to learn the rules and etiquette of online communication and to shift quickly between genres or language registers. Provided they are sensible about divulging personal information, chat rooms and social networking sites provide young people with a safe arena for rehearsing and exploring aspects of identity and personal relationships that may not be available elsewhere (Jansz 2006). Simulation games enable activities otherwise too costly or too dangerous, difficult, or impractical to implement in the classroom.

Gaming expertise is linked to executive functioning, self-monitoring, pattern recognition, problem solving, decision-making, qualitative thinking, and superior short-term and long-term memory (Shaffer 2008). As all toys do, electronic toys and digital games keep players on task for a longer period of time, giving whatever benefits they may offer a greater probability of being realized.

Designing games and writing apps are part of digital literacy and entail new approaches to information and representation (Owston et al. 2009).

Cognition and Perception

Reading a book, watching television drama series, listening to or playing music, and playing video games have all been linked to language development and mental facility. For instance, playing video games promotes a variety of visual skills (tracking, object rotation) and abstract thinking (Spence and Fong 2010). These effects extend beyond the game to other situations. Electronic games have been among the most successful means for reducing the typically reported sex differences in spatial abilities (Dye et al. 2009).

Approximately 500,000 students graduate annually in the USA with an undergraduate degree in science, technology, engineering, or mathematics (Mayo 2009). These numbers pale in comparison to the reach of a single computer video game, such as *World of Warcraft*, with over 12 million subscribers. According to Mayo, video games can yield a 7–40 % positive learning increase over a lecture. Learning is enhanced because the player has control over the game. Game-based learning often requires the formation of hypotheses, experimentation, and discovery, the very bases of science. Game-based learning has the potential to deliver science and math education to millions of users simultaneously.

Many critics contend that children are no longer able to engage in authentic, spontaneous play and that the narratives, symbols, and scenarios of their pretend play have been taken over by the media, depriving children of the opportunity to develop their imagination and autonomy. The concern is that technology will displace more desirable activities, like outdoor play, reading, or socializing with friends, and there is some support for this view (Valkenburg and Peter 2009). Yet much research suggests that youth are far from being passive recipients of media. In their play and games, children and adolescents actively appropriate cultural commodities, making their own discriminations and judgments, while combining and reworking them in myriad ways. Researchers have explored whether children who play with electronic toys play less creatively or imaginatively than children who play with more traditional toys – blocks and dolls, paper, and crayons – that do not involve electronics. Although there is insufficient research to provide a clear answer, the concerns do not seem to be justified (Bergen 2004). In one experiment, 6–8-year-old children displayed the same reasoning skills and performed similarly (required the same number of moves for the solution) whether a task was in the form of a board game or a computer game (see Goldstein 2011).

A review by Russ and Dillon (2011) found no decrease in children's imagination or creativity in the period 1985–2008. They found that imagination in play significantly increased over time. There was no evidence of change in organization of a story or in overall expression of affect in play. Among 12-year-olds, video game playing was strongly related to creativity (Jackson et al. 2012). Research on television and children's imaginative play finds that the content, but not the quantity, of fantasy is affected by program content. Contemporary play objects, by virtue of their electronic functions and affordances, invite exploration and discovery. According to Resnick (2006) "today's technology can open new opportunities for children to playfully explore, experiment, design, and invent."

Jackson (2012) examined the positive effects of video game playing. In her research on the relationship between video game playing and cognitive outcomes, she found that 12-year-old children who played video games more were more creative than those who played them less.

Götz et al. (2005) conducted a multinational study of the fantasy worlds of 8–10-year-old children in Germany, Israel, South Korea, and the USA. Children build upon a wealth of information gathered from a wide range of sources, including their own personal experience and media, and freely interweave them to create rich fantasy backdrops for play. Some children stay fairly close to the original media script in their fantasy play. "This raises the question of whether such media texts inhibit children's imagination, so that there is less originality and more imitation in the fantasy. According to our study, this is the exception rather than the rule... Contrary to popular belief, children make sophisticated use of these mediated worlds. They mix and match settings and specific objects within them in ways that facilitate their own fantasy worlds and allow them to best experience their wishes in these worlds. They highlight and expand on those aspects of the original media worlds that are particularly attractive to them and adapt or erase those that hinder or are not relevant to the wished for experience. One might say they play the role of editor..." (Götz et al. 2005, pp. 197–199).

A review by Latham et al. (2013) found that among the demonstrated skills and benefits of video games are as follows:

- Hand-eye coordination and reaction time
- Spatial visualization
- Visual attention
- Visual search strategies
- Task switching
- Visual perception and the use of visual information
- Use of video games as a cognitive intervention, for example, with the elderly

Meta-analyses by Powers et al. (2013) investigated how video games impact information-processing skills (auditory processing, executive functions, motor skills, spatial imagery, and visual processing). Quasi-experimental studies (72 studies) compare habitual gamers with controls; true experiments (46 studies) use commercial video games in training. "Video games led to improved information processing in both the quasi-experimental studies, and the true experiments. Whereas the quasi-experimental studies yielded small to large effect sizes across domains, the true experiments yielded negligible effects for executive functions, which contrasted with the small to medium effect sizes in other domains."

Hutchinson and Stocks (2013) examined the effects of action video game play on performance for encoding global motion. There were no significant differences between video game player (VGP) and non-video game player (NVGP) performance for either translational or rotational motion. For radial motion, VGPs were significantly better than NVGPs at discriminating contracting, but not expanding, elements. The authors postulate that VGPs are selectively sensitive to radial contractions because, while this type of optic flow information is abundant in action video games, it is not encountered in abundance in the real world. In their research, 16 action video game players (VGPs) (mean age, 22 years) and 16 non-video game players (NVGPs) (mean age, 21 years) took part. All participants were male. Participants watched a screen showing 400 white dots moving on a gray background. While some of the dots moved randomly, a certain number in each test moved in a coherent pattern – either translational (up or down), radial (contraction or expansion), or rotational (anticlockwise or clockwise). The fewer consistently moving dots it took for the participant to correctly identify the overall pattern, the higher their motion perception. But in all cases apart from radial contraction, there was no significant difference between gamers and nongamers. "In summary, VGPs and NVGPs exhibited equivalent sensitivity to global motion, with the exception of radial contractions, for which VGP performance was superior." The skill tested in this study was akin to the visual field when walking backwards. This is a common movement in action video games, as when a player is exploring an area or dodging enemies.

Some game genres are better suited to particular kinds of learning. For example, playing action video games, like *Grand Theft Auto*, *Call of Duty*, or *Battlefield*, alters a range of visual skills. Dye et al. (2009) established changes in different aspects of visual attention in habitual video game players as compared to nonplayers. Nonplayers trained on an action video game show marked improvement from their pretraining abilities. Table 1, from Prensky (2005), further suggests that learning content be matched to game genre.

Evidence to date suggests that both brief and extensive exposure to video game play can result in a broad range of enhancements to various cognitive and perceptual faculties that generalize beyond the original context. Despite promise, video

Table 1 Types of learning from games

Content	Examples	Games
Facts	Laws, formulae, specs	Memory
Skills	Interviewing, teaching, project leader	RPGs, adventure, detective
Judgment	Management, ethics	RPGs, multiplayer, strategy
Reasoning	Tactical thinking	Puzzles
Systems	Health care, complex systems	Simulation games

Adapted from Prensky (2005)

game research is host to a number of methodological issues that require addressing before progress can be made in this area. Future work is required to identify the mechanism that allows the act of video game play to generate such a broad range of generalized enhancements (Latham et al. 2013).

Music, Cognition, and Intelligence

As with playing video games, listening to, or playing, music involves multiple senses and multiple regions of the brain. Music is processed by different areas of the brain working closely together to make sense of things such as melody, harmony, and rhythm. The ability to recognize patterns is necessary to appreciate music; it is also a key component of intelligence. So we should expect a relationship between musical comprehension and some forms of intelligence, including spatial reasoning, memory, language, math comprehension, and creativity (Weinberger 1998).

"Brain plasticity" is the term for the capacity of the human brain to alter in response to the environment. Hearing music affects brain plasticity and the way a child's brain develops. The first sound a developing infant hears is the rhythmic heartbeat of its mother, a sort of prenatal music. During fetal development, hearing will improve, and by 36 weeks, a baby can memorize music that it hears frequently, often to the extent that the music will be recognized after birth (Winkler et al. 2009).

Exposure to music prenatally appears to help postnatal coordination and physical skills. In one study, babies in utero were exposed to recordings of different musical components that increased in complexity during the pregnancy. In total, each baby was exposed to between 50 and 90 h of music. The babies exposed to music seemed to make faster progress in some areas compared to babies who had not been exposed to music. Their pre-speech became evident sooner, and there were noticeable differences in hand-eye coordination, visual tracking, facial mirroring, general motor coordination, and the ability to hold a bottle with both hands (Lafuente et al. 1997).

Children from preschoolers through adolescents who receive from 6 months to 2 years of piano lessons show improved cognitive skills, spatial skills, and vocabulary compared to control groups that did not receive music lessons (Piro and Ortiz 2009). Music participation, both inside and outside of school, is associated with measures of academic achievement among children and adolescents. In one study, listening to a Mozart piano sonata increased spatial-temporal reasoning (Rauscher et al. 1997).

Adult musicians' brains show clear differences from those of nonmusicians, particularly in areas relating to listening, language, and the connection between the two sides of the brain (Hyde et al. 2009).

In sum, music, in addition to its mood management function, also affects brain development, language, and cognitive development.

Health

One undeniable feature of play is fun. Enjoyment is the main reason for playing. Positive emotions contribute significantly to a sense of well-being and health and improve the quality of life for children and adults.

Play has long been used in child clinical psychology and pediatrics as a diagnostic tool and therapeutic remedy for physical and emotional trauma. Entertainment media are used increasingly in education, health care, and scientific research. Games, apps, and infomercials are often used to support public health campaigns. Health messages have been imbedded in soap operas in the USA to discourage drunk driving and to reduce cigarette consumption, while in Latin America televised dramas, *telenovelas*, have helped to reduce the birth rate and sexually transmitted disease (Kennedy et al. 2004).

Games and Health

As early as 1991, Kinder could write that video games "have considerable educational and therapeutic value for a diverse range of groups – including adolescents, athletes, would be pilots, the elderly in old age homes, cancer patients undergoing chemotherapy, stroke victims, quadriplegics, and young children suffering from palsy, brain damage, and Down's syndrome" (p. 112).

Video games are increasingly used with both children and adults in a growing variety of health-care settings to impart information, promote adherence to exercise and medical regimens, and for training physicians. Games have aided in the management of weight loss and dieting, as well as chronic diseases such as diabetes and asthma (Kato 2010; Lieberman 2009). Games for mobile phones and tablets have been used to increase physical exercise, improve diet adherence, and increase awareness about health and diet (e.g., Byrne et al. 2012).

Brain Development and Functioning

Video gaming can be beneficial for the brain. Brain regions can be specifically trained. Video gaming causes increases in the brain regions responsible for spatial orientation, memory formation, and strategic planning as well as fine motor skills. The positive effects of video gaming may also prove relevant in therapeutic interventions targeting psychiatric disorders.

Playing *Super Mario* leads to an increase in nerve cells in regions of the brain involved in spatial navigation, memory formation, strategic planning, and fine motor skills of the hands (the right hippocampus, right prefrontal cortex, and the cerebellum). Researchers at the Max Planck Institute in Berlin asked adults to play the video game *Super Mario 64* over a period of 2 months for 30 min a day. A control group did not play video games. Brain volume was quantified using magnetic resonance imaging (MRI). In comparison to the control group, the video gaming group showed increases of gray matter, in which the cell bodies of the nerve cells of the brain are situated. These plasticity effects were observed in the right hippocampus, right prefrontal cortex, and the cerebellum. These brain regions are involved in functions such as spatial navigation, memory formation, strategic planning, and fine motor skills of the hands. The more these changes were pronounced, the more desire the participants reported to play the video game. "While previous studies have shown differences in brain structure of video gamers, the present study can demonstrate the direct causal link between video gaming and a volumetric brain increase. This proves that specific brain regions can be trained by means of video games," according to the authors. Video games could be therapeutically useful for patients with mental disorders in which brain regions are altered or reduced in size, e.g., schizophrenia, posttraumatic stress disorder, or neurodegenerative diseases such as Alzheimer's dementia (Kühn and Gallinat 2014).

In a second study, Kühn et al. (2014) examined cerebral plasticity during 2 months of training with *Super Mario 64*. Forty-eight men and women, average age 24, played on a Nintendo DS console for at least 30 min/day. The researchers found "significant gray matter (GM) increase in right hippocampal formation (HC), right dorsolateral prefrontal cortex (DLPFC) and bilateral cerebellum in the training group. The HC increase correlated with changes in . . . navigation strategy." Kühn et al. note that video game training augments gray matter in brain areas crucial for spatial navigation, strategic planning, working memory, and motor performance, as well as changes in navigation strategy. "Video game training could therefore be used to counteract known risk factors for mental disease such as smaller hippocampus and prefrontal cortex volume in, for example, post-traumatic stress disorder, schizophrenia and neurodegenerative disease."

Brain training games appear to improve executive functions in the elderly and to transfer to situations outside the game. Nouchi et al. (2012) recruited 32 elderly volunteers through an advertisement in the local newspaper and randomly assigned them to either of two game groups, *Brain Age* or *Tetris*. All participants were nongamers who reported playing less than 1 h of video games per week over the past 2 years. Participants in both the *Brain Age* and the *Tetris* groups played their game for about 15 min per day, at least 5 days per week, for 4 weeks. Each group played for a total of about 20 days. Measures of four cognitive functions were made before and after training: global cognitive status, executive functions, attention, and processing speed. Results showed that the effects of the brain-training game were transferred to executive functions and to processing speed. However, the brain-training game showed no transfer effect on global cognitive status or attention. Results showed that playing *Brain Age* for 4 weeks could lead to improve cognitive functions (executive functions and processing speed) in the elderly. This suggests a possibility by which the elderly could improve executive functions and processing speed in short-term training. Long-term effects and relevance for everyday functioning remain uncertain as yet.

Exergaming

Exergames are video games that require gross motor activity, thereby combining gaming with physical exertion. Exergames are a substitute for physical exercise when outdoor play is not feasible, as in inclement weather or in various institutional settings (Peng et al. 2011). According to a study by Graves et al. (2010), playing new-generation active computer games uses significantly more energy than playing sedentary computer games but not as much energy as playing the sport itself. The energy used when playing active Wii Sports games was not of sufficient intensity to contribute toward the recommended daily amount of exercise in children.

As video game play becomes increasingly recognized as a cognitively, and sometimes physically, demanding activity, the dramatic increase in the amount of time devoted to gaming in today's culture draws attention to a variety of other potentially related issues. For example, with increased rates of chronic obesity, especially in the USA, along with concerns that exercise in the form of physical play has been replaced by sedentary gaming, there is great interest in determining whether mimetic "exergames" such as Wii Fit might be beneficial to health and fitness. In a recent review of 38 randomized control studies of video game training on health-related outcomes, Primack et al. (2012) reported benefits associated with video game play for psychological therapy as well as physical therapy and physical activity outcomes but concluded that poor study quality was a concern.

If exergames are enjoyable to players, they are likely to stay on task longer, and this may ultimately provide a modicum of exercise. Beyond energy expenditure, exergames can have effects on perceptions of confidence, control, and engagement. In a study of young adults aged 18–35, exergames that are more enjoyable produce more energy expenditure (Lyons et al. 2014).

Therapy and Rehabilitation

The potential use of video games in rehabilitation with respect to their behavioral, physiological, and motivational effects was reviewed by Lohse et al. (2013). They focus on motor learning through commercial video games and the potential to increase patient engagement with therapy. This is important because patient nonadherence with therapy is a major barrier to rehabilitation. The authors integrate game design, motor learning, and rehabilitation science to provide criteria by which therapists can help patients choose appropriate games for rehabilitation. "Research suggests that video games are beneficial for cognitive and motor skill learning in both rehabilitation science and experimental studies with healthy subjects. Physiological data suggest that gameplay can induce neuroplastic reorganization that leads to long-term retention and transfer of skill. There is interdisciplinary evidence suggesting that key factors in game design, including choice, reward, and goals, lead to increased motivation and engagement. We maintain that video game play could be an effective supplement to traditional therapy. Motion controllers can be used to practice rehabilitation-relevant movements, and well-designed game mechanics can augment patient engagement and motivation in rehabilitation. . .."

NASA developed video games that use biofeedback to train pilots to stay alert during long flights and calm during emergencies. Signals from sensors attached to the player and body are fed through a signal-processing unit to a video game joystick. As the player's brain waves approach an optimal, stress-free pattern, the joystick becomes easier to control. The technology is commercially available and is used to treat symptoms arising from brain injuries, attention deficit hyperactivity disorder, and learning disabilities. The system allows off-the-shelf video games, such as racing games, to be controlled through brain wave activity. Video games can be controlled using many forms of biofeedback, including galvanic skin response (GSR), heart rate, and temperature (Mason et al. 2004). “Mind-operated devices” can be complex. For example, a person could learn to control the lighting in a room, compose music, or move a cursor on a computer monitor by controlling his or her physiology.

Commercially available interactive game platforms like the Nintendo Wii (Wii) and PlayStation (PS) are increasingly used in the rehabilitation of a wide variety of patients. Such games have gained popularity in burn rehabilitation because they encourage range of motion while distracting from pain. Although interactive video games were not originally designed for rehabilitation purposes but rather for entertainment, they nevertheless may help achieve rehabilitative goals. In a study by Parry et al. (2012), upper extremity motion of 24 normal children was measured using 3D motion analysis during play with the two types of interactive video games most commonly described for use after burn: Wii and PS Eye Toy. Active range of motion for shoulder flexion and abduction during play with both PS and Wii was within functional range, thus supporting the idea that interactive video games offer activities with therapeutic potential to improve range of motion. PS resulted in higher demands and longer duration of upper extremity motion than Wii and therefore may be the preferred tool when upper extremity range of motion and muscular endurance are the goals of rehabilitation. When choosing a suitable interactive video game for application in rehabilitation, the user’s impairment together with the therapeutic attributes of the interactive video game should be considered to optimize outcome.

Entertainment Media and the Elderly

For the elderly, the main cognitive abilities that change over time are perception, attention, memory, and executive functioning. Studies of the noninstitutionalized elderly suggest that digital games can speed reaction time, and may positively influence executive function, and have social and emotional benefits.

Several proposals have been made to build upon the use and integration of video game technology in rehabilitation and training for older adults. In particular, the use of video games for health requires a system for selecting games that are appropriate for the physical, cognitive, and social requirements of older adults or those living with a disability (Marston and Smith 2012).

Kueider et al. (2012) conducted a systematic review to examine the efficacy of computer-based cognitive interventions for cognitively healthy older adults. Studies were included if they met the following criteria: average sample age of at least 55 years at time of training; participants did not have Alzheimer’s disease or cognitive impairment; and the study measured cognitive outcomes as a result of training. Studies published between 1984 and 2011 were classified into three groups by the type of computerized program used: classic cognitive training tasks, neuropsychological software, and video games. All three methods were effective in improving cognitive abilities in cognitively normal older adults. Most studies reported older adults did not need to be technologically savvy in order to successfully complete or benefit from training. Computerized training is an effective, less labor-intensive alternative to more traditional paper-and-pencil cognitive training approaches. "Based on the evidence reviewed, video games appear to be an effective means of enhancing reaction time, processing speed, executive function, and global cognition in older adults."

An important goal of cognitive training is to slow or reverse age-related declines in cognitive abilities. Basak and colleagues (2008) trained older adults in a real-time strategy video game for 23.5 h in an effort to improve their executive functions. A battery of cognitive tasks, including tasks of executive control and visuospatial skills, was assessed before, during, and after video game training. The trainees improved significantly more than the control participants in executive control functions, such as task switching, working memory, visual short-term memory, and reasoning.

Multitasking behavior has become ubiquitous in today's technologically dense world, and substantial evidence has accrued regarding multitasking difficulties and cognitive control deficits in our aging population. Anguera et al. (2013) show that multitasking performance, as assessed with a custom-designed three-dimensional video game (*NeuroRacer*), exhibits a linear age-related decline from 20 to 79 years of age. By playing an adaptive version of *NeuroRacer* in multitasking mode, older adults (60–85 years old) reduced multitasking costs compared to control groups, attaining levels beyond those achieved by untrained 20-year-old participants, with gains persisting for 6 months. Furthermore, age-related deficits in neural signatures of cognitive control, as measured with electroencephalography, were remediated by multitasking training. Critically, this training resulted in performance benefits that extended to untrained cognitive control abilities (enhanced sustained attention and working memory), with an increase in midline frontal theta power predicting the training-induced boost in sustained attention and preservation of multitasking improvement 6 months later. These findings highlight the robust plasticity of the prefrontal cognitive control system in the aging brain.

Exergame training based on physically simulated sport could have cognitive benefits for older adults. If exergame play has the cognitive benefits of conventional physical activity and also has the intrinsic attractiveness of video games, then it might be an effective way to induce desirable lifestyle changes in older adults. Maillot et al. (2014) developed an active video game training program using a pretest – posttest design comparing an experimental group (24 $\times$ 1 h of training) with a control group without treatment. Participants completed a battery of neuropsychological tests, assessing executive control, visuospatial functions, and processing speed, to measure the cognitive impact of the program. They were also given a battery of functional fitness tests to measure the physical impact of the program. The trainees improved significantly more than the control participants in measures of physical function and cognitive measures of executive control and processing speed but not on visuospatial measures. Engagement in physically simulated sport games yielded benefits to cognitive and physical skills that are directly involved in functional abilities older adults need in everyday living.

Mental Health

If entertainment is a public health issue, it is largely in the area of mental health that it has its most direct impact. Enjoying music, a film, a video game, or a You Tube video can improve mood, strengthen friendships, and increase competence.

Simple casual games that are easy to access and can be played quickly, such as *Angry Birds* or *Candy Crush*, can improve players' moods, promote relaxation, and ward off anxiety. Even if video games only make people happier, this is itself a benefit.

Video games may be effective tools to learn resilience in the face of failure. By learning to cope with ongoing failures in games, children build emotional resilience they can rely upon in their everyday lives (Granic et al. 2014).

Online gaming spaces can be socially accommodating environments for socially inhibited individuals, such as the socially inept, socially anxious, or shy. In his review of research in 2004, Rauterberg concluded that "socially anxious and lonely people find more honest and intimate human relationships with others on the Internet than in the real world, and they tend to successfully integrate these online relationships into their offline lives" (Rauterberg 2004, p. 20). Recent studies corroborate Rauterberg's conclusion (Kowert et al. 2014; Shen and Williams 2011; Trepte et al. 2012). Kowert et al. (2014) evaluated the relationship between gaming-related friendships and shyness. Among a representative sample of German game players, emotionally sensitive (shy) players use online gaming spaces differently from their less emotionally sensitive counterparts and report differences in their in-game friendship networks. This suggests that online games hold the potential to be socially advantageous for shy individuals by allowing them to overcome their traditional social difficulties and generate new friendships as well as strengthen old ones.

Video games help combat soldiers experience fewer trauma-related nightmares (Gackenbach et al. 2013), and casual games can alleviate symptoms of depression (Russoniello et al. 2013) and anxiety (Fish et al. 2014). Granic et al. (2014) and Marston and Smith (2012) recommend that teams of psychologists, clinicians, and game designers work together to develop approaches to mental health care that integrate video game playing with traditional therapy.

Traditional forms of entertainment, including films and books, have been used as an adjunct in psychotherapy (Wedding and Niemiec 2014) and to promote healthy development (Niemiec and Wedding 2008). In *the Film Club* (2007), David Gilmour describes how he educated his adolescent son with the help of classic films.

Whitbourne et al. (2013) document why adults of different ages play casual video games, what benefits they perceive, and how regularly they play. They compared the online survey responses of 10,308 adults ranging from 18 to 80 years of age to questions regarding PopCap's free online game, *Bejeweled Blitz*. All respondents cited playing against friends as their main reason for playing. However, there were differences by age in the second most frequently cited reason. "Middle-aged adults cited stress relief, and older adults reported that they seek the game's challenges. As a result of playing CVGs, younger adults noted that they felt sharper and experienced improved memory; older adults were more likely to feel that their visuospatial skills and response time benefited. Adults aged 60 and older had heavier patterns of game play than did adults under the age of 60 years. . .. They believe that they derive a number of benefits that, for older adults, appear to offset declines in age-sensitive cognitive functions."

Depression

Depression is estimated to affect more than 300 million people worldwide. Although there has been some success in treatment of this illness with pharmaceuticals and cognitive behavior therapy, these are often costly or unavailable. Games and entertainment can be successful in dealing with depression.

Li et al. (2014) reviewed the existing literature on game-based digital interventions for depression through a meta-analysis of randomized controlled trials (RCTs). Nineteen studies were included in the review. Four types of game interventions – psychoeducation and training, virtual reality exposure therapy, exercising, and entertainment – were identified. The meta-analysis revealed a moderate effect size of the game interventions for depression therapy at posttreatment. Interventions based on psychoeducation and training had a smaller effect than those based on the other forms, and self-help interventions yielded better outcomes than supported interventions. The review and meta-analysis support the effectiveness of game-based digital interventions for depression.

Internet use in general is associated with less depression among retired older adults in the USA (Cotten et al. 2014). The reduction in depression was especially pronounced among those who live alone.

In a series of experiments, Fish et al. (2014) and Russoniello et al. (2013) found that a prescribed regimen of playing casual video games significantly reduced symptoms of clinical depression and anxiety, as measured with the Patient Health Questionnaire-9. They recommend that clinicians consider these low-cost video games as a possible intervention to address psychological and somatic symptoms associated with depression or anxiety.

Humor and Laughter

Humor and laughter are closely related, but they are not synonymous. Humor is the underlying cognitive and emotional process that frequently, but not necessarily, leads to laughter. Both have been studied in relation to physical and mental health. Humor/laughter can be means of coping with stress, easing pain, warding off illness, heightening a sense of well-being, and increasing social support.

Laughter is analgesic, helping to raise the pain threshold while lowering blood pressure (Stuber et al. 2007). Laughter has been found to reduce anger, anxiety, depression, and stress, increase pain threshold, reduce risk of myocardial infarction, improve lung function, and reduce blood glucose levels (Ferner and Aronson 2013). A number of studies have reported increases in immunity following laughter-inducing stimuli (Martin 2007). Research by Kimata (2001, 2004) found that watching a humorous film can reduce allergic reactions in individuals with allergies and can reduce the symptoms of bronchial asthma.

A 2011 study by Dunbar et al. noted that the sense of well-being associated with laughter might result from the release of endorphins. They conducted six experiments in both the laboratory (watching videos) and naturalistic contexts (watching stage performances), using change in pain threshold as an assay for endorphin release. The results show that pain thresholds were significantly higher after laughter than in the control condition. This pain-tolerance effect was due to laughter itself and not simply due to a change in positive affect. Laughter, through an endorphin-mediated opiate effect, may play a crucial role in social bonding.

Humor is an index of psychological well-being (Herzog and Strevey 2008). Laughter of depressed patients differs from that of healthy individuals and has the potential to serve as a diagnostic tool for depression and other neuropsychiatric pathologies (Navarro et al. 2014).

Comedy is nearly always atop the list of most popular films and television programs. Given the mood management, pain thresholds, and health sustaining properties of laughter, one could argue that we are amusing ourselves to health.

Science and Research

In addition to the many areas of human endeavor affected by entertainment, one must add science. Films, television programs, video games, and digital toys have been used in scientific research as measurement and training instruments (McMahon et al. 2011). Video games and digital toys have been used to study and develop perception, cognition, motivation, and social relationships. Cell phones, iPads, digital cameras, global positioning systems, robotics, and eye-tracking technology have all been used to study play (Alexander et al. 2009, Campoli et al. 2011). Handheld portable game devices have much to recommend them as psychological tests. The equipment is robust, inexpensive, small, light, and portable. Jones (1984) used two electronic toys, *Simon Says* and *Split Second*, to measure short-term memory, reaction time, and pattern recognition during a mountaineering expedition.

Weinmann et al. (2013) examined the potential of entertainment education for promoting engagement with a science issue, genetic engineering. What features of entertainment media experience increase viewers' perceived knowledge about an issue? A survey was conducted with 103 participants, who were shown a movie about genetic engineering beforehand. Results suggested that the greater the viewers' transportation, identification, and enjoyment, the more they felt informed. Viewers' degree of transportation proved to be sufficient to explain their perceived knowledge.

To help solve a long-standing problem in molecular biology (the crystal structure of M-PMV retroviral protease), Khatib and colleagues (2011) "challenged players of the protein folding game Foldit to produce accurate models of the protein. Remarkably, Foldit players were able to generate models of sufficient quality for successful molecular replacement and subsequent structure determination. The refined structure provides new insights for the design of antiretroviral drugs."

Computer games are being studied for their possible role as entertainment, measurement, and remedial training methods for manned spaceflight (e.g., Voynarovskaya et al. 2010).

Video games are both an expressive medium and a persuasive medium; they represent how real and imagined systems work, and they invite players to interact with those systems and form judgments about them. Persuasive games are used increasingly in advertising and politics (Bogost 2007; Nightingale 2014).

Political entertainment, such as *The Daily Show with Jon Stewart* and *The Colbert Report*, has stretched the boundaries of political reporting, combining humor with investigative journalism. As yet, there is little research on how such television content affects political views (Holbert 2014).

Conclusion

Entertainment is highly varied, from more or less passive to extremely active and from rote repetition to cerebral and creative. As such, its tentacles reach nearly every aspect of life, touching our emotions, teaching new ideas and skills, and affecting our social relationships and our health. The research presented here is designed to suggest areas where entertainment has played an integral role in education, health, and the quality of life.

Recommended Reading

G.M. Alexander, T. Wilcox, R. Woods, Sex differences in infants' visual interest in toys. Arch. Sex. Behav. **38**(3), 427–433 (2009)

J.A. Anguera, J. Boccanfuso et al., Video game training enhances cognitive control in older adults. Nature **501**, 97–105 (2013). doi:10.1038/nature12486

C. Basak et al., Can training in a real-time strategy video game attenuate cognitive decline in older adults? Psychol. Aging **23**, 765–777 (2008)

D. Bergen, Preschool children's play with "talking" and "nontalking" rescue heroes: effects of technology-enhanced figures on the types and themes of play, in *Toys, Games, and Media*, ed. by J. Goldstein, D. Buckingham, G. Brougére (Lawrence Erlbaum Associates, Mahwah, NJ, 2004), pp. 295–206

I. Bogost, *Persuasive Games: The Expressive Power of Videogames* (MIT Press, Cambridge, MA, 2007)

S. Byrne et al., Caring for mobile phone-based virtual pets can influence youth eating behaviors. J. Child. Media **6**(1), 83–99 (2012). doi:10.1080/17482798.2011.633410

D. Campoli et al., Instrumented toys for assessing spatial cognition in infants. Front. Mech. Eng. **6**(1), 82–88 (2011). doi:10.1007/s11465-011-0208-0

D.A. Christakis, F.J. Zimmerman, M. Garrison, Effect of block play on language acquisition and attention in toddlers: a pilot randomized controlled trial. Arch. Pediatr. Adolesc. Med. **161**(10), 967–971 (2007)

S.R. Cotten et al., Internet use and depression among retired older adults in the United States: a longitudinal analysis. J. Gerontol. Ser. B Psychol. Sci. Soc. Sci. (2014). doi:10.1093/geronb/gbu018

F.S. Din, J. Calao, The effects of playing educational videogames on kindergarten achievement. Child Study J. **31**, 95–102 (2001)

R.M. Dunbar et al., Social laughter is correlated with an elevated pain threshold. Proc. R. Soc. Biol. (2011). doi:10.1098/rspb.2011.1373

M.W.G. Dye, C.S. Green, D. Bavelier, Increasing speed of processing with action videogames. Curr. Dir. Psychol. Sci. **18**(6), 321–326 (2009). doi:10.1111/j.1467-8721.2009.01660.x

C.J. Ferguson, M.B. Donnellan, Is the association between children's baby video viewing and poor language development robust? A reanalysis of Zimmerman, Christakis, and Meltzoff (2007). Dev. Psychol. **50**, 129–137 (2014). doi:10.1037/a0033628

R.E. Ferner, J.K. Aronson, Laughter and MIRTH (methodical investigation of risibility, therapeutic and harmful): a narrative synthesis. Br. Med. J. **347**, f7274 (2013). doi:10.1136/bmj.f7274

S. Fisch, A capacity model of children's comprehension of educational content on television. Media Psychology **2**, 63–92 (2000)

S. Fisch, *Children's Learning from Educational Television: Sesame Street and Beyond* (Routledge, New York, 2004)

M.T. Fish, C.V. Russoniello, K. O'Brien, The efficacy of prescribed casual videogame play in reducing symptoms of anxiety: a randomized controlled study. Games Health J. **3**(5), 291–295 (2014). doi:10.1089/g4h.2013.0092

K. Fisher et al., Playing around in school: implications for learning and educational policy, in *Oxford Handbook of the Development of Play*, ed. by A.D. Pellegrini (Oxford University Press, New York, 2011), pp. 341–362

A. Fuchslocher, J. Niesenhaus, N. Kramer, Serious games for health: an empirical study of the game 'Balance' for teenagers with diabetes mellitus. Entertain. Comput. **2**, 97–101 (2011)

J. Gackenbach et al., Video game play as nightmare protection: a replication and extension. Dreaming **23**(2), 97–111 (2013). doi:10.1037/a0032455

D. Gilmour, *The Film Club: A True Story of a Father and Son* (Dundurn, Toronto, 2007)

J.H. Goldstein, Play and technology, in *Oxford Handbook of the Development of Play*, ed. by A.D. Pellegrini (Oxford University Press, New York, 2011), pp. 322–340. ISBN 13: 978-0195393002

M. Götz et al., *Media and the Make-Believe Worlds of Children: When Harry Potter Meets Poké mon in Disneyland* (Lawrence Erlbaum Associates, London, 2005)

I. Granic, A. Lobel, R.C.M.E. Engels, The benefits of playing video games. Am. Psychol. **69**, 66–78 (2014). doi:10.1037/a0034857

L.E.F. Graves et al., The physiological cost and enjoyment of Wii Fit in adolescents, young adults, and older adults. J. Phys. Act. Health **7**(3), 393–401 (2010)

T.R. Herzog, S.J. Strevey, Contact with nature, sense of humor, and psychological well-being. Environ. Behav. **40**, 747–776 (2008). doi:10.1177/0013916507308524

K. Hirsh-Pasek et al., *A Mandate for Playful Learning in Preschool: Presenting the Evidence* (Oxford University Press, New York, 2006)

K. Hirsh-Pasek et al., Putting education in 'educational' apps: lessons from the science of learning. Psychol. Sci. **16**, 3–34 (2015). doi:10.1177/152910061556971

R.L. Holbert, Strike while the iron is hot: seizing on recent advancements to propel forward the study of political entertainment media. Mass Commun. Soc. **17**, 303–306 (2014). doi:10.1080/15205436.2014.917041

C.V. Hutchinson, R. Stocks, Selectively enhanced motion perception in core video gamers. Perception **42**, 675–677 (2013). doi:10.1068/p7411

K.L. Hyde et al., Musical training shapes structural brain development. J. Neurosci. **29**, 3019–3025 (2009)

Y. Itoh et al., A computerized interactive toy: TSU.MI.KI. Entertain. Comput. ICEC **2005**, 507–510 (2005)

L.A. Jackson, The upside of videogame playing. Games Health J.: Res. Dev. Clin. Appl. **1**(6), 452–455 (2012). doi:10.1089/g4h.2012.0064

L.A. Jackson et al., Information technology use and creativity: findings from the children and technology project. Comput. Hum. Behav. **28**(2), 370–376 (2012). doi:10.1016/j.chb.2011.10.006

J. Jansz, The emotional appeal of violent video games for adolescent males. Commun. Theory **15**, 219–241 (2006)

S. Johnson, *Everything Bad Is Good for You: How Today's Popular Culture Is Actually Making Us Smarter* (Riverhead, New York, 2005)

M.B. Jones, Videogames as psychological tests. Simul. Gaming **15**, 131–157 (1984)

P.M. Kato, Videogames in health care: closing the gap. Rev. Gen. Psychol. **14**(2), 113–121 (2010)

M.G. Kennedy et al., Increases in calls to the CDC national STD and AIDS hotline following AIDS-related episodes in a soap opera. J. Commun. **54**, 287–301 (2004)

F. Khatib et al., Crystal structure of a monomeric retroviral protease solved by protein folding game players. Nat. Struct. Mol. Biol. **18**, 1175–117 (2011). doi:10.1038/nsmb.2119

H. Kimata, Differential effects of laughter on allergen-specific immunoglobulin and neurotrophin levels in tears. Percept. Mot. Skills **98**(3), 901–908 (2001)

H. Kimata, Effect of viewing a humorous vs. non-humorous film on bronchial responsiveness in patients with bronchial asthma. Physiol. Behav. **81**, 681–684 (2004)

M. Kinder, *Playing with Power* (University of California Press, Berkeley, 1991)

R. Kowert, E. Domahidi, T. Quandt, The relationship between online video game involvement and gaming-related friendships among emotionally sensitive individuals. Cyberpsychol. Behav. Soc. Netw. **17**(7), 447–453 (2014). doi:10.1089/cyber.2013.0656

A.M. Kueider et al., Computerized cognitive training with older adults: a systematic review. PLoS One **7**(7), e40588 (2012). doi:10.1371/journal.pone.0040588

S. Kühn, J. Gallinat, Amount of lifetime video gaming is positively associated with entorhinal, hippocampal and occipital volume. Mol. Psychiatry **19**, 842–847 (2014). doi:10.1038/mp.2013.100

S. Kühn et al., Playing Super Mario induces structural brain plasticity: gray matter changes resulting from training with a commercial video game. Mol. Psychiatry **19**, 265–271 (2014). doi:10.1038/mp.2013.120

M.J. Lafuente et al., Effects of the Firstart method of prenatal stimulation on psychomotor development: the first six months. Pre-Perinatal Psychol. J. **11**, 151 (1997)

A.J. Latham, L.M. Patston, L.J. Tipper, The virtual brain: 30 years of video-game play and cognitive abilities. Front. Psychol. **4**, 629 (2013). doi:10.3389/fpsyg.2013.00629

J. Li, Y.-L. Theng, S. Foo, Game-based digital interventions for depression therapy: a systematic review and meta-analysis. Cyberpsychology, Behav. Soc. Netw. **17**(8), 519–527 (2014). doi:10.1089/cyber.2013.0481

D.A. Lieberman, Designing serious games for learning and health in informal and formal settings, in *Serious Games: Mechanisms and Effects*, ed. by U. Ritterfield, M. Cody, P. Vorderer (Routledge, New York, 2009), pp. 117–130

K. Lohse et al., Video games and rehabilitation: using design principles to enhance engagement in physical therapy. J. Neurol. Phys. Ther. JNPT **37**, 166–175 (2013). doi:10.1097/NPT.0000000000000017

E.J. Lyons et al., Engagement, enjoyment, and energy expenditure during active video game play. Health Psychol. **33**, 1174–181 (2014). doi:10.1037/a0031947.

P. Maillot, A. Perrot, A. Hartley, Effects of interactive physical activity video game training on physical and cognitive function in older adults. Psychol. Aging **29**, 589–600 (2014). doi:10.1037/a0026268

R.A. Mar, K. Oatley, The function of fiction is the abstraction and simulation of social experience. Perspect. Psychol. Sci. **3**, 173–192 (2008)

H.R. Marston, S.T. Smith, Interactive videogame technologies to support independence in the elderly: a narrative review. Games Health J.: Res. Dev. Clin. Appl. **1**, 139–152 (2012). doi:10.1089/g4h.2011.0008

R.A. Martin, *The Psychology of Humor* (Elsevier, Amsterdam, 2007)

S.G. Mason et al., Real-time control of a videogame with a direct brain-computer interface. J. Clin. Neurophysiol. **21**, 404–408 (2004)

M.J. Mayo, Videogames: a route to large-scale STEM education? Science **323**, 79–82 (2009)

R.P. McMahon et al., Considerations for the use of commercial video games in controlled experiments. Entertain. Comput. **2**, 3–9 (2011). doi:10.1016/j.entcom.2011.03.002
J. Navarro et al., Validation of laughter for diagnosis and evaluation of depression. J. Affect. Disord. **160**, 43–49 (2014)
R. Niemiec, D. Wedding, *Positive Psychology at the Movies: Using Films to Build Virtues and Character Strengths* (Hogrefe, Boston 2008)
M. Nightingale, Behind the scenes: working with Hollywood to make positive social change, in *Media and the Well-Being of Children and Adolescents*, ed. by A.B. Jordan, D. Romer (Oxford University Press, New York, 2014)
R. Nouchi et al., Brain training game improves executive functions and processing speed in the elderly: a randomized controlled trial. PLoS One **7**, r29676 (2012). www.plosone.org
R. Owston et al., Computer game development as a literacy activity. Comput. Educ. **53**, 977–989 (2009)
J. Panksepp, Science of the brain as a gateway to understanding play. An interview with Jaak Panksepp. Am. J. Play **2**, 245–277 (2010)
J. Panksepp, C. Trevarthen, The neuroscience of emotion in music, in *Communicative Musicality: Exploring the Basis of Human Companionship*, ed. by S. Malloch (Oxford University Press, New York, 2009)
A. Parente, R. Parente, Mind-operated devices: mental control of a computer using biofeedback. CyberPsychol. Behav. **9**(1), 1–4 (2006). http://www.liebertonline.com/doi/abs/10.1089/cpb.2006.9.1
S. Parry et al., Commercially available interactive videogames in burn rehabilitation: therapeutic potential. Burns **38**, 493–500 (2012). doi:10.1016/j.burns.2012.02.010
W. Peng, J.-H. Lin, J. Crouse, Is playing exergames really exercising? A meta-analysis of energy expenditure in active videogames. Cyberpsychol. Behav. Soc. Netw. (2011). doi:10.1089/cyber.2010.0578
J.M. Piro, C. Ortiz, The effect of piano lessons on the vocabulary and verbal sequencing skills of primary grade students. Psychol. Music **37**, 325–347 (2009)
A.T. Pope, E.H. Bogart, Extended attention span training system: video game neurotherapy for attention deficit disorder. Child Study J. **26**(1), 39–50 (1996)
K.L. Powers et al., Effects of video-game play on information processing: a meta-analytic investigation. Psychon. Bull. Rev. **20**, 1055–1079 (2013). doi:10.3758/s13423-013-0418-z
M. Prensky, *Digital Game-Based Learning* (McGraw-Hill, New York, 2005)
B.A. Primack et al., Role of video games in improving health-related outcomes: a systematic review. Am. J. Prev. Med. **42**(6), 630–8 (2012). doi:10.1016/j.amepre.2012.02.023
F.H. Rauscher et al., Music training causes long-term enhancement of preschool childrens' spatial-temporal reasoning. *Neurological Research*, **19**, 2–8 (1997)
M. Rauterberg, Positive effects of entertainment technology on human behavior, in *Building the Information Society*, ed. by R. Jacquart (IFIP, Kluwer Academic, Toulouse, 2004), pp. 51–58
M. Resnick, Computer as paintbrush: technology, play, and the creative society, in *Play = Learning: How Play Motivates and Enhances Children's Cognitive and Social-Emotional Growth*, ed. by D.G. Singer, R.M. Golinkoff, K. Hirsh-Pasek (Oxford University Press, New York, 2006)
M.R. Rosenzweig, E.L. Bennett, Cerebral changes in rats exposed individually to an enriched environment. J. Comp. Physiol. Psychol. **80**, 304–313 (1972)
S.W. Russ, J.A. Dillon, Changes in children's pretend play over two decades. Creat. Res. J. **23**(4), 330–338 (2011). doi:10.1080/10400419.2011.621824
C.V. Russoniello, M. Fish, K. O'Brien, The efficacy of casual videogame play in reducing clinical depression: a randomized controlled study. Games Health J. **2**, 341–346 (2013). doi:10.1089/g4h.2013.0010
D.W. Shaffer, *How Computer Games Help Children Learn* (Palgrave Macmillan, New York, 2008)
C. Shen, D. Williams, Unpacking time online: connecting Internet and massively multiplayer online game use with psychosocial well-being. Commun. Res. **38**, 123–149 (2011)

I. Spence, J. Fong, Videogames and spatial cognition. Rev. Gen. Psychol. **14**(2), 92–104 (2010)
M. Spencer, *Two Aspirins and a Comedy: How Television Can Enhance Health and Society* (Paradigm, Boulder, 2006)
M. Stuber et al., Laughter, humor and pain perception in children: a pilot study. eCAM (2007). doi:10.1093/ecam/nem097
S. Trepte, L. Reinecke, K. Juechems, The social side of gaming: how playing online computer games creates online and offline social support. Comput. Hum. Behav. **28**, 832–839 (2012). doi:10.1016/j.chb.2011.12.003
P.M. Valkenburg, J. Peter, Social consequences of the internet for adolescents. Curr. Dir. Psychol. Sci. **18**, 1–5 (2009)
N. Voynarovskaya et al., Nonverbal behavior observation: collaborative gaming method for prediction of conflicts during long-term missions, ed. by H.S. Yang. ICEC 2010, LNCS, vol. 6243 (2010), pp. 103–114
D.S. Weber, Media use by infants and toddlers: a potential for play, in *Play = Learning: How Play Motivates and Enhances Children's Cognitive and Social-Emotional Growth*, ed. by D.G. Singer, R.M. Golinkoff, K. Hirsh-Pasek (Oxford University Press, New York, 2006)
D. Wedding, R.M. Niemiec, *Movies and Mental Illness. Using Films to Understand Psychopathology* (Hogrefe, Ashland, 2014)
N. Weinberger, Brain, behavior, biology, and music: some research findings and their implications for educational policy. Arts Educ. Policy Rev. **99**, 28–36 (1998)
C. Weinmann et al., Approaching science by watching TV: what do entertainment programs contribute to viewers' competence in genetic engineering? Educ. Media Int. **50**(3), 149–161 (2013). doi:10.1080/09523987.2013.839152
S. Whitbourne, S. Ellenberg, K. Akimoto, Reasons for playing casual video games and perceived benefits among adults 18 to 80 years old. Cyberpsychol. Behav. Soc. Netw. **16**(12), 892–897 (2013). doi:10.1089/cyber.2012.0705
I. Winkler et al., Newborn infants detect the beat in music. PNAS Proc. Natl. Acad. Sci. U.S.A. **106**(7), 2468–2471 (2009). doi:10.1073/pnas.0809035106
F.J. Zimmerman, D.A. Christakis, A.N. Meltzoff, Associations between media viewing and language development in children under age two years. J. Pediatr. **151**, 364–368 (2007). doi:10.1016/j.jpeds.2007.04.071

Introduction to the Ethics of New and Emerging Science and Technology

48

Tjalling Swierstra

Contents

Abstract

This chapter reconstructs the patterns of moral argumentation typical for debates between enthusiastic proponents and skeptic opponents about the (expected) impacts of technologies. These patterns frame how issues are viewed and constitute a tool kit on which discussants draw. Section "General Definitions"

T. Swierstra (✉)
Department of Philosophy, Faculty of Arts and Social Sciences, Maastricht University, Maastricht, MD, The Netherlands
e-mail: t.swierstra@maastrichtuniversity.nl

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_33

briefly explains key concepts like morality and different types of ethics and offers a reflection on the relation between morality, ethics, and technology. Section "Visions of Technological Development" discusses two perspectives on technological development that play an important role in the preliminary debate whether ethical assessment of technology makes any sense in the first place: technological determinism and voluntarism. Section "Meta-Ethics" introduces the reader to two meta-ethical arguments regarding the dynamic relation between ethics and technology. The first concerns the likelihood that the NEST will eventually confront us with problems. Here optimists and skeptics or pessimists oppose one another. The second meta-ethical issue is whether morality must adapt to technology or vice versa. Section "Normative Ethics" provides an overview of the normative ethical arguments in favor of a certain proposed technology and juxtaposes them to the arguments used by more skeptical opponents. The arguments are grouped according whether they regard consequences, principles, justice, or conceptions of the good life. In the concluding section "To Conclude," I point out some ways in which real-life deliberations and discussions about NEST benefit from knowing these arguments and argumentative patterns.

Keywords
Emerging technologies • Ethics • Public deliberation • Public controversies

Introduction

When lecturing on the ethics of technology, I like to invite the audience to engage in a thought experiment. How would our lives look like without technology. People look around, and then the air conditioner comes to a grinding halt. The next moment walls come tumbling down, and the roof disappears into thin air. We flinch as we find ourselves on the ground, with the chairs gone. We have a moment to realize how gray our neighbor actually is, and how pale her face, before we realize we have to cover our own naked bodies too. Fortunately, contact lenses and glasses have disappeared. And as our tongue starts exploring the sudden cavities in our mouth, a man collapses as his pacemaker gives out. It is hard to call for help, without a phone.

The point is, of course, that modern lives have become utterly entangled with the technological artifacts and systems we created. Indeed, we live in a human-made world: a techno-tope rather than in a biotope. In a sense, this has always been the case. We date back the origins of the human species as far back as – well – the earliest artifacts we manage to find. However, technological artifacts have become much more ubiquitous since the so-called Scientific Revolution of the seventeenth and eighteenth centuries and more particularly since the Industrial Revolution of the nineteenth and twentieth centuries. How we relate to the natural world around us, to our fellow human beings, and even to ourselves, all these relations are co-shaped by the dynamics of a rapidly evolving technology.

Some people are happier than others about this, but regardless of one's optimistic or pessimistic stance, no one can avoid the question: under what conditions is which technology conducive to human (and nonhuman) flourishing? How to design, develop, use, and distribute technology? This is the vast and rapidly expanding domain of *technology ethics* (Franssen 2009). Increasingly, these ethical questions are posed and explored in public debates. Topics include nuclear energy, fracking, climate change, genetic engineering, nanoparticles, production of life-saving drugs for poor countries, the impact of web surfing, privacy, mobility and migration, robots, and so on and so forth. Actually, in technological cultures, it is much harder to think of controversies in which technology does *not* play a major role, than of controversies where it does.

The aim of this chapter is not to tell the reader what is good and bad technology or how it should (not) be used. Instead, the aim is to introduce the reader to recurring argumentative patterns in this debate. Although technologies change and diverge, debates about new and emerging science and technologies (abbreviated as NEST) show remarkable structural similarities. Indeed, NEST-ethical debates more or less follow a shared grammar. It is this grammar that the following chapter seeks to identify and elucidate (This chapter builds upon: Swierstra and Rip (2007).). If citizens, technology developers, and policy makers learn to understand and apply this grammar, this will help them to think and discuss about the ethical aspects of new and emerging technologies.

Section "General Definitions" briefly explains some key concepts and offers a reflection on the relation between morality, ethics, and technology. Section "Visions of Technological Development" discusses two perspectives on technological development that play an important role in the preliminary debate whether ethical assessment of technology makes any sense in the first place. Section "Meta-Ethics" introduces the reader to some meta-ethical arguments regarding the dynamic relation between ethics and technology. Section "Normative Ethics" provides an overview of the normative ethical arguments in favor of a certain proposed technology and juxtaposes them to the arguments used by more skeptical opponents. I am confident that the reader will recognize most if not all of these general arguments. In the concluding Section "To Conclude," I point out some ways in which real-life deliberations and discussions about NEST benefit from knowing these arguments and argumentative patterns.

General Definitions

Morality and Ethics

The concepts "morality" and "ethics" are often used interchangeably. For our goal, however, it is useful to distinguish the two. Here, morality refers to a special category of values and norms that guide us in our ordinary lives. "Values" indicate what we think is important; "norms" prescribe the conduct that will help realize these values in practice. Values typically are formulated in the form of substantives;

norms in the form of short sentences that start with "You should" or "You should not." For example, trust is a value, whereas that you should stick to your promises is the associated norm. Norms articulate the "how" of values; values articulate the "why" of norms. Not all norms are *moral* norms, nor are all values *moral* values. We reserve the adjective "moral" for values and norms that carry a lot of weight, either because they formulate fundamental rules of social intercourse or because they articulate what it means to flourish as a human being (Sayer 2011). These fundamental norms typically take the form of obligations and prohibitions. Their special status is manifest in the fact that moral rules are accompanied by praise and blame, rewards and punishments, to motivate people to live according to these norms and values. Take note: this is a formal (empty) definition of "moral." The definition does not tell us what norms and values are moral. People can and do disagree about whether a particular rule is moral or not. For instance, (female) chastity is held to be a central moral value in some cultures, whereas in others it is considered as a private lifestyle choice.

An interesting observation about morality is that it largely exists in the form of routines that are so obvious that in the normal course of our lives, we are barely aware of their impact on our thinking, feeling, and acting (Dewey 1994; Keulartz et al. 2004). It would be quite disturbing if you would first consider to kill an annoying colleague, only to then decide after ample reflection that (unfortunately) this would be immoral. The idea simply should not have entered your conscious thinking in the first place (Williams 1985). So naturally comes the moral taboo on murder to us that we obey automatically. As a result, the most obvious and powerful moral rules and values tend to be the least visible and articulate in daily practice.

Ethics, by contrast, refers to explicit reflecting on and discussing morality. A precondition for ethics is that moral norms or values have become explicit. The main reason they do so is because they have for some reason become problematic. At such a moment, morality stops being self-evident, commonsensical, and invisible and gets articulated into explicit topics of reflection and conversation. In such a situation, "cold" morality turns into "hot" ethics: invisible, solid, moral routines become fluid in ethics, so they – if necessary – can be readjusted to the new situation.

Under what circumstances do morals stop being self-evident? First, a moral norm becomes visible when it is violated or disobeyed. Breaking a rule is the surest way to make one aware that the rule existed in the first place. Second, and related to the first reason, our moral routines get shaken up into visibility when others contest them. If we meet people with different (sub)cultural backgrounds, we not only realize that they have different norms and values, but we also become more consciously aware of our own – at least to the extent that they differ from the ones held high by the other person. A third way moral norms or values lose their self-evidence is when they conflict. Most people value honesty and trustworthiness. But what if a friend confides in you that he betrays his wife, and then his wife later calls to inquire whether you know something? In a moral dilemma, we are forced to choose which norm to obey or what value to put first, and that forces us to articulate the values we want to, but now cannot, pursue.

Technology, Morality, and Ethics

Technological societies embody a fourth way in which morality is turned into ethics. Moral routines thrive in a stable environment because there they receive constant confirmation. But as modern, technological societies are highly dynamic, they do not provide for such a stable environment. Technological change "heats up" morality into ethics. When asked around 1995 whether they would like to have a mobile phone, many people responded negative because they did not like the idea that one would be available everywhere and every time. Twenty years later emotions flair and accusations fly if someone does not respond quickly enough. Twenty years ago, phoning someone was considered to be non-problematic. Nowadays many prefer to text, because phoning is considered to be intrusive (Turkle 2010).

A more futuristic example: presently everyone opposes doping in sports as that gives unfair advantages. The only factor that is allowed to determine who wins and loses is how good one is. But no one finds it problematic that, apart from training and willpower, all kinds of "natural" biological differences between athletes impact their performance. What if advances in biotechnology would allow us to neutralize these "natural differences"? Would we then still deem doping unfair? Or would we see it as a moral obligation to help the "biologically disadvantaged," because only by doping them the competition can really be considered "fair"?

In short, new technologies have the potential to *destabilize* parts of our tacit, implicit morality and thus turn them into topics for explicit ethical reflection, debate, and struggle.

Rule Ethics and Life Ethics: Consequences, Principles, Justice, and the Good Life

Ethical discussions are about good and bad. We engage in such discussions not for fun, but because it is unclear what to do, for example, because we have to choose between two evils or because applying existing norms results in counterintuitive consequences. In such situations people broadly use two kinds of ethical considerations: *rule-ethical* arguments and *good life* arguments – sometimes abbreviated as the *right* or the *good* (Rawls 1988).

The rule-ethical approach is almost automatically chosen when interests conflict. To avoid that the strongest wins, ethics aims for solutions based on impartial rules, acceptable to all involved parties. Reasonable people, is the idea, voluntarily conform to conclusions drawn on such grounds. The question is then of course: how do you find or design such rules? There exist – in general terms – three different answers to that rule-ethical question.

The first answer comes in the form of a meta-rule that states: Choose the practical alternative *with the best consequences for the largest group*. This type of rule ethics is called *consequentialist*. To determine what the "best" consequences are, consequentialist ethics usually refers to "pain" and "happiness," as no one likes pain and everyone aims for happiness. Ethics then is about maximizing the happiness (or minimizing the pain) of as many stakeholders as possible. Consequentialism thus predominantly prioritizes the interests of the collective, the common interest, over the interests of the individual.

The second answer comes in the form of a meta-rule that states: Choose the practical alternative *that meets a fundamental moral principle*. Such principles are supposed to capture the fundamental rules of, and preconditions for, successful social interaction. Often such rules come in the form of prohibitions, duties, and rights. There are things you simply do not do because they are intrinsically wrong, regardless of the consequences. Even though one would have the chance to ultimately save thousands, we do not test new drugs on prisoners. As this example, illustrates, moral principles tend to defend the rights of the individual against the interests of the collective. Therefore moral principles under normal circumstances outweigh consequences (although that is a rule of thumb that allows for exceptions; see section "Normative Ethics").

A third answer to the ethical question "what moral rules should we obey to solve our conflicts in a reasonable way" is geared toward a special, but crucial, subcategory of ethical issues: how to *fairly* or *equitably* distribute a scarce good, that is, a good for which demand exceeds supply. This is the problem of *distributive justice*. Which distribution is just depends on the criterion we think is adequate: equality, merit, need, and chance are the most common candidates for such criteria. I will return to these in section "Normative Ethics."

In brief, rule ethics includes discussions about consequences, principles, and (distributive) justice. In modern, pluralist and liberal, societies, public debate is conducted mainly on the basis of this kind of rule-ethical arguments. The reason for this is that rule ethics is considered as sufficiently objective and impartial to allow for a reasonable consensus in many cases. Initially, however, ethics was much broader. For the ancient Greeks, for example, ethics in the first place revolved around the questions: how to live a good (admirable) life? How to be a good person? Religions and ancient wisdom teachings too mostly revolve around this question: how to be a good Christian, Jew, Muslim, Hindu, Buddhist, Taoist, or Confucian? Since the religious wars of the sixteenth century that marked the beginning of "modernity" in the West, religion, together with this "good life ethics," has been progressively banished from the public sphere. In modern, liberal and pluralist, societies, public ethics tends to focus on a thin ethics of "traffic rules" that allow peaceful coexistence of parties that would otherwise be fighting. What constitutes a good life is delegated to the individual's personal discretion and thus to the private sphere (Swierstra 2002). In a democracy, each citizen should be free to pursue her or his conception of the good life – of course on the condition that she/he does not harm others in doing so (Mill 1989). However, this public-private split is less sharp than often suggested. For instance, discussions about politicians, euthanasia, or – less dramatic – the influence of computer games on the character of the players usually contain more or less explicit references to ideas about what constitutes a good life or a good person (Sandel 2010).

Normative Ethics, Meta-ethics, and Descriptive Ethics

Finally, a last conceptual distinction is relevant for the ethics of new and emerging science and technology: meta-ethics versus normative ethics. Normative ethics analyzes ethical problems and attempts to provide a reasoned solution. Cloning, is that allowed? Is human enhancement a good idea? Should genetic tests be promoted, prohibited, or released? Should violent computer games be outlawed?

Meta-ethics refers to the philosophical reflection on the methods and foundations of normative ethics. It asks, for instance, whether ethical debates can be rational, and if so, in what sense of "rational." Or it examines whether ultimately not all ethical arguments can be traced back to consequences or rather to principles. Meta-ethics is primarily fodder for philosophers but not exclusively. Anyone who in a discussion groans that "values are ultimately merely subjective, aren't they, so let's talk about something else" engages in meta-ethics. We will see in the section "Meta-Ethics" that some meta-ethical arguments play an important role in NEST-ethical discussions.

Descriptive ethics, finally, is geared toward describing existing moralities or – as in the following sections – existing ways to discuss ethical questions. Descriptive ethics is thus a form of sociology or ethnography. Whereas the findings of normative ethics are to be evaluated in terms of "right" and "wrong," the empirical claims of descriptive ethics ask for an evaluation in terms of "true" or "false." So, the ideas and argumentative patterns explicated in the following sections are a form of descriptive ethics. This means that you as a reader are invited to check my inventory of recurring arguments and argumentative patterns not in terms of whether you agree with them but in terms of whether the inventory is realistic and complete.

Visions of Technological Development

In the following sections, I introduce the reader to recurring arguments and argumentative patterns in a NEST-ethical debate. I reconstruct this debate as one between technology optimists and pessimists. But I want to stress that nowadays undiluted technology optimism and pessimism have become equally rare. Even the staunchest optimist no longer denies that technological development created major problems (even if she/he holds that the solution to these problems is more and better technology). And even the blackest pessimist no longer denies that we inhabit an irreversibly technological world (even if she/he holds that we need more social and less technological solutions). Nor blindly embracing nor blindly rejecting technology (with a capital T) is a realistic option nowadays.

The first issue usually addressed in NEST-ethical discussions is not (meta- or normative) ethical but factual: can technology development be influenced? (Bijker et al. 1987 (2012); Smits and Marx 1994) If we lack sufficient grip on scientific and technological developments, an ethical assessment makes no sense. Some (even partial) control over technology is a precondition for any NEST-ethics.

To the question whether technology can be steered, roughly two negative answers exist. The first answer is descriptive: you *cannot* steer, even if you try. The second negative answer is normative: you *should* not steer, because the costs of doing so outweigh the possible benefits. The affirmative response similarly has a descriptive and a normative form. On the one hand, historical and sociological evidence suggests that social actors do guide the course of technology development (descriptive). On the other hand, one could argue that this societal influence is not sufficiently subjected to adequate (usually understood as: democratic) control (normative).

Determinism: Descriptive

The position that technological development cannot be guided is known as "technological determinism." This determinism is justified in at least three different ways. Technology development is first presented as autonomous, necessary, and solely determined by the laws of nature. Technological development is like a train that has no option but to follow the rails (even if the direction of those rails can only be established retrospectively). Steering is not an option; drive! Although this variant of technological determinism has long been dominant, it seems to have lost much of its force. This may have to do with the fact that technology development is becoming increasingly expensive. As a result it becomes more difficult to disregard that technological developments depend on societal factors like corporate and/or political funding.

Over the last decades a second justification for technological determinism has gained ground. The (international, global) competition has taken over the role of the unstoppable force driving the development of technology. If we do not develop this technology, then our competitors will. In an open market all we can do is play along and try to compete successfully.

So-called "technological path dependencies" provide a third argument why influencing the course of technological development is futile. A technology is not like an intangible idea, which can be simply refuted and replaced. On the contrary, technology possesses a material robustness and it is embedded in a techno-social network. As a consequence, a technology cannot be easily reviewed or overturned. Even if we now think electric cars are a better idea than gasoline cars, we have erected a physical infrastructure around the petrol car consisting of engines, petrol stations, refineries, and providers. Furthermore, we have become attached to the sound of the explosion engine and to the way it behaves. These factors make it hard to change tracks. The technological choices of the past limit our technological options for the future. We are forced to work in the old mold.

Determinism: Normative

Whereas the preceding arguments were predominantly descriptive, the next ones are mainly normative: even if we *could* control technological development, we *should* not attempt to because such control can only be achieved at the expense of other fundamental values.

The general version of this normative justification of technological determinism is that scientific and technical development are inseparable from social progress. Whoever attempts to interrupt or steer technological development automatically jeopardizes social progress.

A more specific, economics version of this argument holds that in a free market economy, technology cannot be regulated. Manufacturers and consumers should be free to supply and purchase what they choose to – at least as long as they do not harm others. State interference can only result in market inefficiencies and other woes.

In another version of this argument, academic freedom plays the crucial role. Only the academic community can and must decide which scientific and technological research is worthwhile. Society has given universities a mandate to investigate freely, regardless of political, religious, economic, and ideological interferences. Paradoxically, it is precisely because of this unfettered freedom that universities can serve society. Society should not steer scientific and technological progress; it only decides whether scientific and technological findings will be applied and how. The reason society is allowed this latter type of decision is that these are not based on facts (what is the case) but on values (what should be the case), and here scientific and technological experts have no say.

Voluntarism: Descriptive and Normative

Opposing this (descriptive and normative) determinism is what can be referred to as "voluntarism." Drawing on historical and sociological research, voluntarists claim that social factors constantly influence technological development. Scientific and technological research is done by people, and so by definition people exercise influence (Collins and Pinch 1998; Sismondo 2004). The question is not one of determinism but of politics: who is pulling the strings, and who should do so ideally? (Winner 1980; Morozov 2014)

In public discussions on NEST, often a certain distrust is detectable. Behind the joyful expectations around a NEST, some suspect the "spin" of interest groups that try to "sell" their technologies. How to be sure that it is not large industries, such as the pharmaceutical industry, the food industry, or the oil companies, who decide what is examined or produced, hiding their influence behind the veil of technological determinism?

Even many who believe that business and government are not doing so badly, concur with greater democratic control over the course of technological development. This plea usually translates first into the requirement of transparency: it should at least be clear who decides what and on what grounds? Why do we invest in this and not rather in that? A minimal democratic standard is that it must be possible to hold agents accountable for their choices. This is a powerful incentive for technology developers to make sure that their products and production are safe, healthy, and sustainable. A more radical democratic requirement is that citizens should be in a position to think and talk in advance about the technology that will eventually help shape their lives. This requirement would shift the discussion from "how to avoid harmful technologies" to the more positive, aspirational, question "among all the possible technologies that we can spend money and energy on, which ones do we hold to be the most worthwhile?" (Von Schomberg 2013; Stilgoe et al. 2013)

This question whether a particular NEST is desirable or not usually only gets posed *after* the technology has been introduced (in)to society. First, we find ourselves invited to marvel at the cloned sheep Dolly, smart energy meters, and Google glasses, and only afterwards there is room for the ethical questions. Of course, we do not want to waste time worrying about technological fantasies that may never materialize. But retrospective assessing has a major drawback: often so much has been invested in the NEST, so many stakeholders have rallied around it, and it has become so intertwined with other technologies that the genie cannot be put back into the bottle. "Retrospective ethics" quickly degenerates into commenting on a fait accompli. For "prospective ethics" the problem is the reverse. "Upstream" much is still unclear. How feasible are the technological expectations? And thinking about the social consequences of a NEST, how can we be sure that society will at that time not have changed drastically anyway? (Martin 2010) Prospective ethics is unavoidably speculative, and the chance to reach agreement under such conditions seems nil.

This problem is known as the knowledge-control dilemma of Collingridge who formulated it in 1980 (Collingridge 1980). When there is still something to steer, we lack the necessary knowledge to do so; by the time we have that knowledge, the technology has already "solidified" and become socially embedded. Advocates of a NEST can mobilize this dilemma to avoid ethical debate: "It is now too early for such a debate. First wait until further research has distinguished fact from fiction." Only to then later point out, when the technology has materialized and we are finally able to distinguish fact from fiction, "that now unfortunately it is no longer realistic to try to turn things back." This rhetoric works because it refers to a real dilemma.

Skeptics, however, have ways to respond. Firstly, they argue that ethics is not "added" to scientific and technological research, as if these would be value free. Most research is motivated by the desire, hope, and expectation that it will help achieve wonderful things: less disease, less hunger, more wealth, and more justice. (Such positive expectations also serve to mobilize financial and/or political support for the research.) In this sense, research is guided by ethical considerations from the outset. And if early expectations regarding NEST are speculative, this applies as much to hopes as it does to fears.

Similarly, skeptics object to the argument that when a technology has been developed, it is too late for ethics. In the first place, one can never pinpoint a precise moment when a technique is indeed "finished." Artifacts are intermediate stages in a continuous development with several generations. For example, it is impossible to say at what stage is your phone "ready." Secondly, even in a late state, an artifact can be improved at on the basis of ethical concerns. Not only are, for example, modern cars more sustainable than previous generations, but moral concerns also played a major incentive behind research geared at producing stem cells without having to harvest them from embryos.

Meta-ethics

In NEST-ethical debates two meta-ethical issues play a central role. The first concerns the likelihood that the NEST will eventually confront us with problems that require ethical deliberation. Optimists assess that chance as negligible; their opponents are more pessimistic. The second meta-ethical issue is whether morality must adapt to technology or vice versa. Here we find fundamentalists and relativists opposing one another.

Trust in Technology

NEST-ethical discussions are superfluous if scientific and technological progress always results in social progress. This optimistic stance is based on the so-called linear model of technology development. According to this model there runs a straight line from basic research through applied research, product development and dissemination in society, to social progress.

Technology optimists present themselves self-consciously as prophets of a new age and downplay social and ecological problems: these problems will be solved by technological progress itself. Pessimists by contrast point out that their technology always causes unintended and unforeseen problems and that we should not proceed as long as we do not yet know how big those problems will be and whether we have indeed solutions for them. The so-called precautionary principle dictates that the burden of proof lies with the optimists who think it is safe to proceed with a certain technological development.

Optimism and pessimism are based on conflicting images of technology. Optimists tend to see technology as a neutral instrument. Technology simply provides us with new possibilities; it is up to us to make good and wise use of them. If someone kills another person using a hammer, one does not blame the hammer or its designer, only its user. Or, as the motto of the National Rifle Association has it, " guns don't kill people; people kill people." Pessimists by contrast tend to stress that technology is neither passive nor neutral (Ihde 1993). It is not neutral because it incorporates specific values, such as the desire to maximize efficiency or the desire to control (natural and social) reality. Technology is not passive either. In Europe we do not hand out guns, as we think that guns in a certain sense *make* killers. For the pessimist, technology is therefore never to be trusted.

It is important, however, to stress that the naked fact that technologies are active and value laden is in itself not sufficient to become a pessimist: we can also try to build moral values into the technology – e.g., sustainability or compassion – or use the technology to "nudge" people to do the right thing, for example, using speed bumps to motivate drivers to slow when passing a school (Akrich 1992; Latour

1992; Thaler and Sunstein 2008). Similarly, instrumentalism doesn't necessarily imply optimism. If one has a bleak view of humankind, one can be very pessimist about how technologies will be used.

Interaction Technique and Morality

When a new technology is introduced, commonly its revolutionary character is stressed. In part this reflects the pride and enthusiasm of the scientists and technology developers involved, but the hype also serves to generate attention and to mobilize financial, political, and policy support (Borup et al. 2006). When in 2000 the human genome was presented, British Prime Minister Blair spoke of "a revolution in medical science, which will prove to be much more important than the discovery of antibiotics in the last century." His US colleague President Clinton drew a comparison with the discoveries of Galileo and the splitting of the atom: "Today we are learning the language in which God created life."

But interestingly enough, although with regard to science and technology revolution is stressed, the public is simultaneously assured that with regard to our morality, it will be business as usual. The revolutionary technology, so it is said, will only help to realize our existing, unproblematic goals. All that changes is that we can do the things we want to do more effectively. Modernity is exceptional in its embrace of scientific and technological change, whereas in more traditional cultures new knowledge or new technology is often seen as threatening. But with these traditional cultures, most moderns still share a deep-rooted conservatism regarding morality. Only a few people are willing to face the idea that not only our facts and our artifacts but also our values are "provisional" and subject to change.

The question is whether moral and technological change can really be separated that easily (Jasanoff 2004). Is it not rather logical to assume, skeptics will not fail to emphasize, that revolutionary technologies will destabilize our moral routines too? (Swierstra et al. 2010) Four arguments characterize this part of the NEST-ethical discussion.

The Argument of Precedence

Technology advocates can try to ease the fear that the NEST is at odds with the accepted morality by downplaying its "novelty." They can point to non-controversial precursors of the controversial technique in history or in nature, thus demonstrating that morally speaking there is nothing new under the sun. Is genetic modification something new and scary? Of course not; we have been doing that since the beginning of agriculture and animal husbandry, only slower and less efficiently. Is cloning a revolution? No, nature does it herself; we call those clones "twins." Is nanotechnology a miracle? No, beer brewing is also based on manipulation of nature at the nanolevel. Does human enhancement constitute a revolution? Of course not, it is essentially the same as sending your child to school.

Obviously not everybody will be convinced by this "argument of precedence." Opponents will identify morally relevant differences between the claimed precedent and the NEST under discussion. To them, there are substantial differences between twins and clones (twins are the same age, clones not) or between animal breeding and its genetic manipulation (one uses chance and proceeds slowly; the other is goal-oriented and fast) that make that the NEST is not automatically as acceptable as its claimed "precedent."

The Slippery Slope Argument

A special way to debunk the argument of precedence is through warning that the new technology will put us on a "slippery slope" (Van der Burg 1991). In this argument, the conflict between the proposed technology and the existing morality gets relocated to the future. Expect this argument to be used in cases where the new technology at first sight seems rather innocuous or beneficial. For example, is it not great that with a simple genetic test we can avoid that a severely disabled child in constant pain is born? But, so opponents will object, where lies the limit? Does this technology ultimately not lead to a Brave New World with no place for "differently abled" fellow citizens?

Or suppose we develop a gene therapy to make criminals less aggressive. That seems noble, but will we not degenerate into a society where those in power will manipulate our minds?

Of course, technology optimists are not convinced that such a slippery slope even exists. For them, the slippery slope argument is a form of determinism that disregards that at every step of the way, we can always decide to stop and even retrace our steps.

The Habituation Argument Versus the Argument of Moral Decline

Creating a reassuring continuity with the past or with nature is not always a promising argumentation strategy. And some technology optimists refuse to bow to what they see as moral conservatism. They will openly admit that the NEST is at odds with current morality. To then assertively add: all the better! Since the invention of fire and the wheel, technological innovations have always met with initial resistance and moral panics. But this resistance always fades out after a while. When the first train huffed and puffed between Amsterdam and Haarlem, academics warned that such inhuman speed would bring women to miscarriage and would spoil the milk in the udders of the shocked cows. And the first test tube baby was considered either as a monster or a miracle, but it is now hard to find a school class without an IVF child. In brief, people get used to the new technology; given some time, morality slavishly adapts to the new technical reality (Haldane 1924).

Opponents of this view cannot of course deny that morality in the past has often coevolved with the technology and that most people indeed tend to acquiesce in this. But this fact, of course, does not justify the normative conclusion that those moral adaptations should be welcomed. Maybe one day we will produce people on an assembly line – as described in Aldous Huxley's *Brave New World* (1932) – and maybe we will even come to accept this as normal. But that does not mean that such a technology therefore would also be good, anymore than that the majority of the Germans in Nazi-Germany being anti-Semite would make anti-Semitism morally okay! It just means that large groups of people can go morally astray. As there is moral progress, there is also moral decline. These voices in the debate oppose the moral relativism they perceive behind the habituation argument.

Normative Ethics

The arguments explored in the previous section circumvent the direct normative question whether a certain NEST is morally desirable or not. In the following section I describe patterns in normative ethical discussions about NEST that deal with this question. I distinguish the arguments depending on whether they are related to consequences, principles, justice, or the good life (see section "Visions of Technological Development"). Obviously, arguments of technology proponents call forth matching counterarguments by skeptics and vice versa.

Consequences

Technology is not always designed with a specific purpose in mind, such as to provide for a specific social need (demand pull). It is also common for researchers to give their curiosity free rein, to be especially interested in a proof of principle, or to stumble almost by accident on something for which the marketing department subsequently finds a purpose – if necessary by creating new needs (technology push). Even if there was from the outset an intended purpose, that is often not the (sole) purpose that is finally realized. The first steam machine was designed to pump water from the mine, and it took a while before someone had the bright idea to put wheels under the device. As soon as a technology exists, engineers and users start to invent new applications that often overshadow the original intentions behind the technology.

That being said, it is still true that a NEST is usually presented to the outside world as an instrument to realize specific (desirable) consequences. As mentioned earlier, hardcore technology optimism is no longer as strong as it once was. Not only is it now clear to everyone that technology also has drawbacks – such as for the environment – but scientific research and technological development have also become increasingly expensive and large scale. Technological research now easily spans several nations, and costs quickly run in the millions if not billions. As a result, society – or a corporation – is no longer willing to write a blank check for scientists and technologists. These have to mobilize financial, political, and public support for their projects by justifying their projects in terms of their utility. To do this they typically apply consequentialist ethical arguments, in the form of expectations, hopes, and promises: if you invest now, tomorrow you will reap the benefits (cure for cancer, solution to hunger, peace through better communication, etc). These promises and expectations are rarely recognized as ethical arguments, but they are of course.

Opponents of a NEST also apply such consequentialist arguments, only now these take the form of doubts or fears rather than hopes. Whether the arguments are in favor of a NEST or against it, they can be challenged in four ways:

(a) Consequentialist arguments take the form of (positive or negative) expectations about the future. Their speculative character means that their *plausibility* can always be questioned. And that becomes easier if the promised future lies further in the distance and if more factors play a role. If the intended effects of a particular NEST are assessed to be improbable, the argument for investing in it weakens accordingly. (Or vice versa: if the projected risks are found to be small, there is less reason for precaution.)
(b) Even if expectations are deemed plausible, this does not end the discussion. After all, maybe we possess *a superior alternative* to the proposed NEST. For example, in the debate about nuclear energy, some argue for wind and solar energy, even if they accept the technical feasibility of nuclear energy. Or some argue that we do not need genetically modified rice to alleviate hunger in the Third World, if only we would distribute wealth more fairly or help to install democratic ways of governance.
(c) Consequentialist arguments are also contested by pointing out unintended and undesirable *side effects* (Tenner 1996). Through trial and error we have learned that technologies always do more than what they were developed for. Technology Assessment (TA) was created in the seventies to explore such side effects of NEST in advance. Adverse environmental impacts are the best-known example of such unintended and undesirable consequences, but one can also think of the students who suddenly discovered Ritalin as a means to increase the concentration during examinations. If the intended effects are outweighed by the unintended and undesirable side effects, this can turn the scales on the NEST in question.
(d) Finally, a consequentialist argument can be criticized on the grounds that the hoped-for result is actually *not as desirable as it is presented to be*. For example, so-called trans-humanists are enthusiastic advocates of various techniques that promise to physically and mentally "enhance" people. But what exactly do we mean by "enhanced"? Is it really progress if we grow taller, if we do not age, or if pills help us to concentrate longer so that we can work longer days? (Sharon 2014)

I now turn to argumentation patterns that revolve around (the interpretation and application of) ethical principles.

Rights, Duties, and Responsibilities

The principle ethical (or "deontological") part of morality exists in the form of (moral, not necessarily legal!) prohibitions, rights, obligations, and responsibilities. Rights and prohibitions/duties/responsibilities correspond with each other: if X has a certain right, it means that Y has a corresponding prohibition/duty/responsibility to ensure that that right is respected. It is useful to distinguish positive rights (claim rights) from negative rights (freedoms). In the first case the other party has to do something; in the second case they must abstain from doing something.

An important principle ethical argument in favor of a NEST is that it will help to fulfill a duty that is based on important positive rights, such as the right to good health care or to information or to a life without hunger or terrorism. In this case we have the moral obligation, or responsibility, to develop this (medical) technology. In NEST-ethical discussions, negative rights are also frequently mobilized: if I want to develop a particular technology and I do not harm anyone in doing so, others may not prevent this. This is the much mobilized moral principle of free choice, for instance, “I am not forcing you to use preimplantation genetic diagnosis, so you should not deny me my right to use it.”

Moral principles carry much weight, and they cannot – as in the case with consequentialist arguments – be undercut by questioning their plausibility. This is because principles apply regardless of the consequences. This does not mean they cannot be disputed. The weakness of general principles is that there often exists a large gap between the principle and the concrete problems to which they are applied. Principles always need to be interpreted and applied wisely. This gap leaves room for doubt, which typically takes four forms:

(a) The principle mobilized by advocates of a NEST can be *outweighed by another*, conflicting moral principle. For example, electronic patient records may indeed lead to better care and thus help reduce human suffering (which is a moral duty), but does that gain indeed overrule the increased risk of privacy violations (privacy is a moral right)? Or does the right to “own” one’s body materials (like body tissue containing genetic information) carry as much weight as the right of patients who can be helped with this material?
(b) Opponents can also argue that the principle invoked by the proponents is indeed important but does not apply in the case of this particular technology. For example, “infertile couples indeed have a right to enhance their children, e.g., by sending them to school, but this right does not extend to germ line intervention.”
(c) An opponent can object that although proponents indeed appeal to a crucial moral principle and although that principle does indeed apply to the NEST in question, it still does not justify it. For instance, advocates of human enhancement justify this technology by appealing to individual freedom: if people want to make use of these techniques, they have got the moral right to do so. Opponents, however, object that human enhancement is actually incompatible with individual freedom, because in a competitive society everyone will eventually be forced to enhance herself/himself.
(d) A last way to cast doubt on a deontological justification (or refutation) of a NEST is by appealing to consequentialist ethics. There are (rare) cases when the damage to the collective is so huge that it can be justified to restrict the moral rights of the individual. In open societies this is not done lightly, but even

these have laws that suspend civil liberties in times of national danger. Discussions about NEST rarely touch on national danger, but there are recurring references to the so-called tragedy of the commons. In specific situations respecting individual rights causes collective disaster. Many people think, for example, that it is unwise to grant prospective parents the moral right to determine the sex of their children, as that would (does) lead to a huge imbalance in the sex ratio, with devastating consequences for society as a whole. What is smart for one can be dumb for all. In such a situation consequentialism can indeed sometimes trump deontology.

Justice

A third part of NEST-ethical discussions concerns distributive justice. This issue presupposes some form of scarcity as only then the ethical problem of distribution manifests itself. In the case of NEST, scarcity occurs at two stages. In the development stage it has to be decided on which scientific research and technological development to spend scarce resources like money, time, and energy. After a technology has been introduced to society, new questions arise regarding the distribution of its benefits and costs. The answer to the first question depends partly on a satisfactory answer to the second question. Few think it is ethically okay to spend scarce resources if the technology will eventually only benefit a privileged group while allocating the costs to the poor.

In the case of a new technology, benefits and costs should be shared justly. But what is "just"? In practice, four criteria vie with each other. Scarce goods are distributed on the basis of equality, merit, need, or chance. In some situations we prefer one criterion, in other situations another. For example, education is mostly distributed on the basis of equality; piece rates are paid based on performance; health care is given on the basis of need; and if someone wins a lottery, we simply congratulate her or him.

In NEST-ethical discussions we see mostly the first three criteria at work. In the general rhetorics, a technology is usually supposed to benefit everyone more or less equally. For example, for many it would be unfair if the rich could enhance themselves and the poor not. A recurring motif in NEST-ethical discussions is therefore how to avoid a gap between technological haves and have-nots. The merit criterion is particularly prominent in discussions about intellectual property rights: it is only fair that those who invested in, and took risks for, developing a NEST also reap the (first, financial) benefits. The need criterion is evident in arguments like it is perverse to use biotechnology to help obese Westerners rather than help the millions of poor who suffer from malaria.

Even if discussants agree that technology should profit everyone more or less equally, there is room for (political) disagreement about how to achieve this desired outcome. In the discussion regarding technological haves and have-nots, we encounter two opposing views. The first view defends the "trickle down effect": it is inevitable that first only the wealthy benefit from a new technology, but thanks to this technological avant-garde, eventually prices will drop so that ultimately everyone profits. The wealthy elite paves the way for the poor masses. This optimism is challenged by skeptics who argue that in many cases the distance between the elite and the large mass did not get smaller at all. Even if the poor do profit, the rich profit much more. The position of the poor may improve in absolute terms, but in relative terms they are now worse off. The only solution, in the eyes of these critics, is that the government intervenes on behalf of the poor and powerless, for example, by forcing the pharmaceutical industry to give priority to drugs aimed at diseases common in poor countries.

The Good Life

Rule-ethical considerations are widely accepted as legitimate contributions to public opinion (which does not mean that everyone will agree with all considerations). Good life considerations are more controversial. The discussion rules of a liberal, pluralist, society determine that there is little public patience with such considerations. Received opinion dictates that everyone should be free to live their lives as they see fit, as long as they do not get in other people's way. The question of the good life, including the sometimes religiously motivated answers, usually gets dismissed as a private concern. The assumption is that the question of what it means for a human being to live a good life does not allow for rational, objective, and impartial debate. This impression is strengthened by the fact that good life ethics typically takes the form of stories: myths, stories from a Holy Book, fairy tales, fables, novels, films, urban legends, etc. This kind of "narrative argument" is often considered to be less compelling than arguments that point at consequences, principles, or justice.

But the proclaimed privatization of good life ethics is theory more than practice. In NEST-ethical discussions we constantly encounter visions of the good life, based on deep-rooted beliefs about what it means to be human and about our place in the cosmos. Robots are already helping to ease the burden of caring for our loved ones, but to what extent do we really want to be "freed" from that burden? Is care not an integral element of the good life? Does playing violent computer games make us more violent prone and less empathic, less good persons? Does the pervasive availability of Internet porn change our experience of sex and intimacy for the better or worse? In relation to NEST two more abstract good life ethical issues are particularly relevant. The first issue is how we should deal with boundaries. The second issue is to what extent people should use science and technology to exert control over reality.

Boundaries

Boundaries basically allow for two types of reaction: either you respect them or you try to transgress them.

Advocates of a NEST typically defend the second option. Their patron saint is Prometheus, the Titan who ignored the express prohibition by Zeus and gave fire (technology) to humankind. Here, boundaries get dismissed as *frontiers*: temporary limitations that somehow demand to be overstepped. Nothing in the world is as it has to be; in principle, everything is a candidate for change and improvement. Technology changes reality into an object of our choice. It is up to us moderns to decide about the shape of the world. History is filled with examples of heroes who pushed the boundaries of human knowledge and skill because they were untroubled by taboos or apparent impossibilities. In the words of American philosopher Richard Dworkin (2000): “Playing God is indeed playing with fire. But that is what we mortals have done since Prometheus, the patron saint of dangerous discoveries. We play with fire and take the consequences, because the alternative is cowardice in the face of the unknown.” As is evident from this quote, this is a fairly masculine discourse, in which doubters get easily dismissed as “sissies.”

But boundaries and limits also have defenders:

(a) Some stress the importance of religiously sanctioned boundaries: “It is not good for us to play *God*.” We all know what happened when Adam and Eve ate the apple and when their descendants built the Tower of Babel. For instance, in the debate on genetic modification some appeal to the idea of a God-willed creation possessing an intrinsic order that demands our respect. Although this religious argument is familiar, in its pure form it is actually rare in Western democracies. One reason for this is that the argument carries little weight in a secular society. In addition, it is essentially an authority argument, and that has not much tracking in a democratic society. Finally, it is not so clear what God actually wants from us. Some even argue that God, as He created us in His image, *wants* us to be creators like Him. For these reasons, the argument is usually accompanied by an auxiliary explanation of why God has good reasons to set these boundaries, such as the inviolability of all life (opposing genetic selection) or that death is exactly the thing that gives meaning to our lives (opposing research on immortality).

(b) Closely related to the previous argument is the appeal to *nature*. The claim is that nature has its own intrinsic value and order, which commands our respect. Many people want a “natural” life, back to “nature”, and so on. It is not a coincidence that in these discussions there are constant references to Frankenstein’s monster, an “unnatural” mixture of live and dead material, an impossible combination of creature and artifact. Like the appeal to God’s will, the appeal to nature is controversial. As early as the eighteenth century, the English philosopher David Hume protested against the “naturalistic fallacy.” That something is the case (factual) can never serve as an argument that it should be the case (normative). Nature is often terrible, and few are willing to stand naked and refrain from unnatural products like clothes, penicillin, and computers. Some go so far to argue that humans are “naturally” unnatural. Even the idea that there is such a thing as a natural order is increasingly under attack. According to modern biology, nature is subject to permanent change and chance and the result of mindless tinkering rather than of a master plan. Nature herself appears to be the first to transgress natural boundaries.

Interestingly, the debate is not settled with these objections. It proves to be surprisingly hard to genuinely bid farewell to the ideal of "naturalness." Think about how food is packaged and marketed or about how patients may refuse to have "chemicals in their body," but also about a futuristic technology as tissue engineering that is "sold" in terms of "using the natural capacities of the body to heal itself." It is as yet unclear what basic (moral) intuitions may resonate in this persistent appeal to nature. It is possible that the desire to respect natural limits serves as an important counterweight to the tendency to approach nature in purely instrumental and objectifying terms. I will return to this motif below.

(c) Less controversial is the argument that points to our *cognitive* limitations. There are boundaries to what we can know. Not only do we not know everything we would like to know, how big certain risks are exactly, but we do not even know what we do not know – technologies always surprise us by suddenly failing us, as Icarus found out when he flew too close to the sun, or by having unexpected effects. Although science and technology promise mastery and control, they can also unleash forces we are unable to control. This fear underlies the ancient myth of Pandora's box – which was opened out of curiosity and then contained all the plagues of humanity – and the more recent fable of the sorcerer's apprentice, who called forth forces he then did not know how to control.

This admittance of our cognitive limits is reflected in policy in the form of the previously mentioned precautionary principle, formulated in the 1970s by the German philosopher Hans Jonas (1973). This principle states that we should refrain from developing new technologies with unknown hazards for humans and the environment. Or more precisely, in such cases it is up to the advocates of the technology to make it plausible that these dangers do not exist or are manageable.

Again, many disagree. NEST-advocates emphasize that concerns about possible hazards are largely based on emotion and speculation. Instead, they insist on the need for "sound science": worries about new technology only deserve attention when they are based on irrefutable scientific evidence. The debate on global climate offers ample examples of both positions.

Fifty years ago, the German-American philosopher Günther Anders (1956) added a new dimension to this awareness of our cognitive limitations: not only are we not in a position to imagine what our technologies will do, but it is equally uncertain whether our moral imagination can keep up with what we can do. In earlier times, it was common that people could imagine more than they could actually do. According to Anders, in modern times, the reverse is true: we are now able to do – or destroy – more than we are able to imagine. We cannot really imagine the suffering inflicted by the A-bombs thrown on Hiroshima and Nagasaki. One or two deaths can be conceived, but not a hundred thousand deaths. And exactly because our moral imagination fails, it becomes easier to throw the bomb. Similar concerns can be found in contemporary debates about, for example, research on aging: does our moral imagination suffice to realistically imagine a world without death?

(d) Finally, there are boundaries placed on us by *tradition*. Precisely because of its dynamic nature, technology is at odds with inherited ways of doing, seeing, and appreciating. The past is littered with traditional communities that have disintegrated as a result of the advent of science, technology, and modernity. According to the founding-father of conservatism, the eighteenth century politician and philosopher Edmund Burke, tradition forms a reservoir of practical wisdom accumulated by humankind through much trial and error. But for such traditional life and worldviews, the "disenchanted" modern world of science and technology has no patience. Therefore, science and technology, or so say the skeptics, pave the way for a cold and inhospitable materialism, selfishness, and nihilism.

Control

It was the mission of the enlightenment to subject nature. As moderns, we are all to a larger or lesser degree shaped by that ambition. Technology promises to emancipate us from our dependence on the mercy of the world around us by submitting and controlling it. That is technology's essential promise. And who would opt for a return to a pre-technological world? In that world our lives would indeed be miserable, cold, fragile, and short.

But our appreciation of "technology" of course does not imply that we should be happy with all technologies or that we would not be allowed to think that certain forms of technological mastery go too far. In fact, our attitude toward technological mastery is ambivalent. As we saw above, some people want to subject technological progress to religious, natural, cognitive, or moral boundaries. But in the NEST-ethical debate, there are also positions that question this desire to control itself. This happens in three ways: the desire to control is wrong because it produces perverse effects, because it is futile and self-refuting, and because it jeopardizes other important values. (I take these three arguments from Hirschman (1991).)

A first motif in NEST-ethical discussions about technological control and mastery is that it is self-refuting. In a perverse turnaround, the technology that promises to free us ends up enslaving us. Just as the master could never trust his slave, so moderns cannot trust their technologies. Do we control the machine or does the machine control us? How many hours a day are we spending in obeying the dictates of technology – slaving at the conveyor belt, answering our mails, etc. Technology promised us control over the world, so the critics say, but has now apparently turned into something uncontrollable itself. Even technology enthusiasts agree with this (an opinion we discussed above under the heading of technological determinism) (Ellul and Merton 1964).

A second motif is that technological control will ultimately prove to be futile. Technology promises to satisfy our needs and so to take away our discomfort, our frustration, our feeling of lack, and our suffering. But, such is the gist of this criticism, this endeavor underestimates the adaptive nature of human needs. Thanks to technology we live more comfortably than kings in the Middle Ages. But are we, with all this technology, really happier today than before, so the critics ask rhetorically? Do our smartphones really make us feel more fulfilled than our parents were? Have all these time-saving devices really resulted in more leisure time for all? Are our needs today better satisfied than before, or have they coevolved with technological development, so that on balance nothing was won and technological progress was futile?

Thirdly, the desire to control can jeopardize other important values. According to the critics, technology has a tendency to define all problems in technical terms, which then of course require a technical solution. If you go around with a hammer in your hand, inevitably you will end up looking for nails. But some problems are not technical, but, for example, social or cultural. Can we really solve environmental problems by applying smarter technologies alone, or do we not – at least also – need a change in mentality and behavior? This narrowing of the focus is referred to as a "technical fix." (To which technology enthusiasts of course counter that their opponents suffer from a social fix, expecting too much from frail human consciences.)

Technological control marginalizes other, social, ways to approach problems and thus jeopardizes important values incorporated in those alternative solutions, e.g., solidarity, cooperation, democracy, etc. But there may be a deeper sense in which technological control jeopardizes core values. When we control nature, we simply take what we want. From this perspective, the technological drive to control is little else than a form of violation. The pollution and depletion of mother earth is but the foreseeable outcome of this attitude.

This control attitude not only does injustice to nature, the victim, but also to ourselves, the violators. The American philosopher Albert Borgmann (1984) points to the impoverishment of our existence as a result of technology, which offers us everything on a silver platter. Modern consumers are barely aware of the fact that their house is warm, that there is food on their table, and that they move from A to B. Because they delegate all the work to technology, these feats have become devoid of existential meaning. Technology reduces us to passive consumers of existence, rather than people who live life to the fullest – including the imperfections, the suffering, and all the hard work inherent in real life.

Whoever subjects his environment, or other beings, no longer has a *relationship* with that environment, with those other beings, in the true sense of the word. For a genuine relationship requires reciprocity: the other person must possess some form of robustness, thanks to which she/he can appear as a real "other" rather than as a passive extension of ourselves. Much as we sometimes wish that our partner would meet our desires more perfectly, there are very few people who would opt for a custom-made love-bot. Too much control leads to boredom and a feeling of loneliness. Most people want to "receive" children, not "order" them. They want to love the child with all his or her "given" quirks, instead of designing an ideal child (Sandel 2009). We may have a deeply seated desire for being in control, but we have an equally deeply seated need for reality to resist our touch. This existential need underlies our ill-articulated fear of a resistance-free world, which, though no longer carrying the risk of suffering, loss, humiliation, and defeat, is also unable to provide warmth, gratitude, satisfaction, depth, meaning, and surprise.

The most famous example of this abhorrence, evoked by perfect technological control over humans and the world, is *Brave New World* (1932). In this novel, Aldous Huxley evokes a world where everyone is happy, cheerful, well fed, and healthy. And yet, or because of this, it is a deeply uncanny world. The novel expresses something that cannot easily be expressed in terms of the prevailing rule-ethical language of pain, happiness, rights, and justice. Its narrative confronts us with the metaphysical loneliness of humans who ultimately only meet their mirror image as the world obediently bends to our wishes and desires and thus becomes ephemeral. Although we avoid suffering and misfortune as much as possible, we simultaneously recognize that both are inextricably linked to a meaningful and truly human existence. Technology's promise of control thus both enables us to live human lives and jeopardizes it.

To Conclude

In the previous sections, the reader was introduced to arguments and counterarguments that dominate discussions on new and emerging science and technology. Why is this useful? To be clear, the goal is not to suggest that these discussions are superfluous because it is essentially always the same discussion. Not only do we have to judge over and over again how plausible hopes and fears are in particular cases, we also have to weigh rights against each other and evaluate which moral principles apply to concrete cases and how to apply them wisely. Similar with the arguments pertaining to distributive justice and the good life, what criterion to apply in which situation and to what degree to trust on the trickle down effects, these are deeply political questions that will need to be debated over and over. And although there are marked differences in how people assess boundaries and technological control, most people are sufficiently ambivalent on these issues that they keep reflecting on them.

So, the aim of the overview is not to write a computer program that will do the deliberation for us by ticking off the boxes. The aim is rather the opposite. NEST-ethical debates can be enriched if participants know what considerations have proven relevant in previous discussions. NEST-ethics may not provide any answers, but it can help us in asking the right questions.

Recommended Reading

M. Akrich, The description of technical objects, in *Shaping Technology/Building Society*, ed. by W.E. Bijker, J. Law (MIT Press, Cambridge, MA, 1992), pp. 205–224

G. Anders, *Die Antiquiertheit des Menschen* (CH Beck, München, 1956)

W.E. Bijker, T.P. Hughes, T.J. Pinch (eds.), *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology* (MIT Press, Cambridge, MA, 2012) (originally published in 1987)

A. Borgmann, *Technology and the Character of Contemporary Life: A Philosophical Inquiry* (University of Chicago Press, Chicago, 1984)

M. Borup, N. Brown, K. Konrad, H. van Lente, The sociology of expectations in science and technology. Technol. Anal. Strat. Manage. **18**(3-4), 285–298 (2006)

D. Collingridge, *The Social Control of Technology* (Pinter, London, 1980), pp. 128–133

H. Collins, T. Pinch, *The Golem at Large: What You Should Know about Technology* (Cambridge University Press, Cambridge, 1998)

J. Dewey, in *The Moral Writings of John Dewey*, ed. by J. Gouinlock (Prometheus, Amherst, NY, 1994) (revised edition)

R. Dworkin, *Sovereign Virtue* (Harvard University Press, Cambridge, MA, 2000), p. 446

J. Ellul, R.K. Merton, *The Technological Society* (Vintage books, New York, 1964)

M. Franssen, G.-J. Lokhorst, I. Van de Poel, Philosophy of Technology, in: *Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/entries/technology/#DevEthTec (2009, revised 2013)

J.B.S. Haldane, *Daedalus or Science and the Future* (EP Dutton, New York, 1924)

A.O. Hirschman, *The rhetoric of reaction* (Harvard University Press, Cambridge, MA, 1991)

D. Ihde, *Postphenomenology: Essays in the Postmodern Context* (Northwestern University Press, Evanston, Ill, 1993)

S. Jasanoff (ed.), *States of knowledge: The co-production of science and social order* (Routledge, New York, 2004)

H. Jonas, Technology and responsibility: Reflections on the new tasks of ethics. Soc. Res. **40**(1), 31–54 (1973)

J. Keulartz, M. Schermer, M. Korthals, T. Swierstra, Ethics in a technological culture: A programmatic proposal for a pragmatist approach. Sci. Technol. Hum. Value. **29**(1), 3–29 (2004)

B. Latour, Where are the missing masses? The sociology of a few mundane artefacts, in *Shaping Technology/Building Society: Studies in Sociotechnical Change*, ed. by W. Bijker, J. Law (MIT Press, Cambridge, MA, 1992), pp. 225–258

B.R. Martin, The origins of the concept of ‘foresight’ in science and technology: an insider’s perspective. Technol. Forecast. Soc. Change **77**(9), 1438–1447 (2010)

J.S. Mill, in *On Liberty and Other Writings*, ed. by C. Stefan (Cambridge University Press, Cambridge, 1989)

E. Morozov, *To Save Everything, Click Here: The Folly of Technological Solutionism* (PublicAffairs, New York, 2014)

J. Rawls, The priority of right and ideas of the good. Philos Public Affairs **17**, 251–276 (1988)

M.J. Sandel, *The Case Against Perfection* (Harvard University Press, Cambridge, MA, 2009)

M.J. Sandel, *Justice: what’s the right thing to do?* (Macmillan, New York, 2010)

A. Sayer, *Why Things Matter to People: Social Science, Values and Ethical Life* (Cambridge University Press, Cambridge, UK, 2011)

T. Sharon, *Human Nature in an Age of Biotechnology. The Case for Mediated Posthumanism* (Springer, Dordrecht, 2014)

S. Sismondo, *An Introduction to Science and Technology Studies* (Blackwell Publishers, London, 2004)

M.R. Smith, L. Marx, *Does Technology Drive History?: The Dilemma of Technological Determinism* (MIT Press, Cambridge, MA, 1994)

J. Stilgoe, R. Owen, P. Macnaghten, Developing a framework for responsible innovation. Res. Policy **42**(9), 1568–1580 (2013)

T. Swierstra, Moral vocabularies and public debate. The cases of cloning and new reproductive technologies, in *Pragmatist ethics for a technological culture*, ed. by J. Keulartz, M. Korthals, M. Schermer, T. Swierstra (Kluwer Academic Publishers, Deventer, 2002)

T. Swierstra, A. Rip, Nano-ethics as NEST-ethics: Patterns of moral argumentation about new and emerging science and technology. Nanoethics **1**, 3–20 (2007)

T. Swierstra, H. van de Bovenkamp, M. Trappenburg, Forging a fit between technology and morality The Dutch debate on organ transplants. Tech. Soc. **32**(1), 55–64 (2010)

E. Tenner, *Why things bite back: Technology and the revenge of unintended consequences* (Knopf, New York, 1996)

R.H. Thaler, C.R. Sunstein, *Nudge: Improving Decisions about Health, Wealth, and Happiness* (Yale University Press, New Haven, CT, 2008)
S. Turkle, *Alone Together. Why We Expect More from Technology and Less from Another* (Basic Books, New York, 2010)
W. Van der Burg, The slippery slope argument. Ethics **102**, 42–65 (1991)
R. Von Schomberg, A vision of responsible research and innovation, in *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society* (Wiley, London, 2013), pp. 51–74
B. Williams, *Ethics and the Limits of Philosophy* (Fontana Press/Collins, London, 1985)
L. Winner, Do artifacts have politics? Daedalus **109**(1), 121–136 (1980)

Negative Effects of Video Game Play

49

Christopher L. Groves and Craig A. Anderson

Contents

Abstract

Video game play has become a ubiquitous form of entertainment in modern society. As a result, interest has accrued from parents, educators, policy makers, and scientists alike regarding the potential effects of this relatively new media. The current chapter has several goals. The first is to describe the research findings regarding the most heavily studied topic in video game research: the effects of violent video game content on aggressive affect, cognition, and behavior. The second goal is to describe the psychological processes that give rise to aggressive (and nonaggressive) negative outcomes of video game play. These psychological processes are described by the General Aggression Model

C.L. Groves (✉) • C.A. Anderson
Center for the Study of Violence, Department of Psychology, Iowa State University, Ames, IA, USA
e-mail: cgroves@iastate.edu; caa@iastate.edu

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_13

(GAM) and the domain-specific theories which GAM incorporates. These "smaller" theories include (but are not limited to) script theory, attribution and decision-making, cognitive neoassociation theory, learning theories, and desensitization. Several other negative outcomes of video game play are also described which include risk taking, attention problems, impulsivity, reduced helping, stereotyping, and video game addiction. Some discussion focuses on specific types of game content (e.g., game mechanics and themes) to the outcomes observed among game players. Lastly, special attention is paid to explaining that the processes that give rise to negative effects are often the same processes that give rise to positive effects and that the notion that games should be considered either "good" or "bad" is much too simplistic. It is our hope that this chapter serves to provide a clear understanding of the negative effects of video game play but also underscores that games also have tremendous potential for positive outcomes.

Keywords
Video game • Aggression • Attention problems • Stereotyping • General aggression model

The advent of computerized technology has transformed the entertainment industry. Tablets, phones, home computers, and game consoles have provided unprecedented access to movies, television shows, and video games. No shift in human entertainment is as marked as the advent of video games. Indeed, this relatively new form of entertainment has exploded in popularity since its beginnings in the 1980s. In a 2009 survey, 88 % of American youth between the ages of 8 and 18 reported playing video games at least occasionally with an average time spent per week at 13.2 h (Gentile). The amount of time spent on video games continues to increase, though comparable data are scarce. Interestingly, increases in video game time have not led to decreases in television time.

The number of hours that children and adolescents spend playing video games has led policy makers, parents, teachers, and researchers alike to question the potential negative and positive effects of this relatively new medium. Mirroring the concerns expressed in the early years of television use, violent content present within video games has sparked both research and debate. Although the current chapter's focus is on the negative effects of video game play, we emphatically note that video games are not inherently "bad." Theoretically, there are likely a variety of cognitive and social benefits of video game play, and many such positive effects have received empirical support (e.g., Prot et al. 2014a). Well-designed video games are excellent teachers that are highly motivating, engaging, and responsive to the player's skills, but the lessons drawn from video games can lead to both positive and negative outcomes.

In fact, some of the processes that give rise to known harmful effects are precisely the same processes that produce some of the positive outcomes. There is perhaps no clearer example of this phenomenon than the contrast between the effects of prosocial (Prot et al. 2014b) and violent games (Anderson et al. 2010).

As will be discussed in more detail throughout this chapter, comprehensive reviews of the empirical research literature show that violent video game play, in which players frequently kill or harm other characters, serves as a causal risk factor for aggressive behavior (Anderson 2004; Anderson et al. 2004, 2010; Anderson and Bushman 2001; Ferguson 2007a, b; Greitemeyer and Mügge 2014; Sherry 2001). Importantly, these effects are seen even among the analyses conducted by both critics and proponents of violent media effects (Boxer et al. 2015) and persist despite numerous tests of alternative explanations. Similarly, prosocial video games, in which players help other game characters in nonviolent ways, produce increases in empathy (Prot et al. 2014b) and helping behavior following game play (Gentile et al. 2009).

The academic focus on these effects reflects an understanding that individuals who consume media do not do so passively. Movies, television shows, and many games include rich stories with themes, lessons, and portrayals which impact consumers in a lasting manner. With this in mind, it is not surprising that research has focused heavily on major content themes within media, particularly video games over the past 20 years. This chapter focuses on the negative outcomes of violent video game play for two reasons. First, and foremost, it is the topic requested of us by the book editors. Second, it is by far the most heavily researched topic of studies on video game effects. There are growing research literatures on positive uses of a wide variety of types of video games, ranging from educational uses, training in numerous industries, and adjunct physical and psychological health therapies. But, those fascinating literatures lie outside the purview of this chapter.

Violent content is present in a large portion of mass media (Linder and Gentile 2009; Smith et al. 1998; Wilson et al. 1997, 1998; Thompson and Haninger 2001; Thompson et al. 2006; Yokota and Thompson 2000). It has been the focus of research for decades, beginning with film and (later) televised violence, but has experienced renewed interest with the advent of video game play. The hypothesized effects of screen violence on increased aggressive behavior have been confirmed by major scientific organizations and by a number of large-scale meta-analyses (American Academy of Pediatrics 2009; American Psychological Association 2005; Anderson 2003; Anderson et al. 2010; Bushman and Huesmann 2006; Hearold 1986; International Society for Research on Aggression 2012; Paik and Comstock 1994; Wood et al. 1991). Special focus will be placed on explicating the psychological processes at work that give rise to violent content effects. Following this discussion, we will touch upon a variety of other negative effects of video game play, and some focus will be placed on the video game mechanics (characteristics of the games) that are the primary drivers of these more recently discovered effects.

Aggression in Media Effects Research

Aggression is defined as “any behavior directed toward another individual that is carried out with the proximate (immediate) intent to cause harm. In addition, the perpetrator must believe that the behavior will harm the target, and that the target is motivated to avoid the behavior" (Anderson and Bushman 2002, pp. 28). Because it is both unethical and impractical to induce violent behavior in the laboratory, most research on violent video game effects uses milder forms of aggressive behavior as the outcome variable of interest.

Aggressive behavior is measured using a variety of laboratory tools. A common measure is one of the many different versions of Taylor's Competitive Reaction Time Task (e.g., Anderson and Dill 2000; Anderson et al. 2004, 2007; Anderson 2003; Bartholowet al. 2005; Engelhardt et al. 2011; Katori 2001; Konijn et al. 2007). This task involves asking a research participant to compete against an ostensible other participant in reaction time trials. Each trial involves the participant clicking a box in the center of the screen once it turns from green to red. Their goal is to "beat" their opponent's click. Prior to each trial, the participant (and ostensibly, their partner) selects a noise volume and duration. If the participant wins the reaction time trial, they administer the noise blast to their opponent. These volume levels typically range from 60 dB (the sound of a normal conversation) to 105 dB (the sound of a lawn mower) and can last from 0.5 to 5 full seconds.

Another common tool used for the measure of aggression is with the use of the hot sauce paradigm (e.g., Barlett et al. 2009; Yang et al. 2014). This method involves asking a participant to partake in a food-tasting task in which other participants in the same study are asked to eat a variety of foods. The participant is given information indicating that an ostensible other participant does not like spicy food. They are then asked to select the amount of hot sauce this person must consume in order to earn their research credit. Aggressive behavior is coded by measuring the amount of hot sauce assigned to this other participant.

Although these serve as two common methods for measuring aggression in laboratory studies, there are many others. Also, it is important to note that the establishment of a phenomenon is not based on a mere few aggressive indices. Strong conclusions regarding the effect of a given stimuli (e.g., video games) rest upon the ability for the effect of that stimulus to generalize across a variety of measures that tap into the same conceptual construct (i.e., behavior intended to harm another person). For this reason, research regarding the effects of violent video games uses a number of additional measures and has established relationships between violent video game play and verbal aggression (Krcmar and Farrar 2009), delivering painful electric shocks to another person (Sakamoto et al. 2001), children's peer ratings of aggressiveness, getting into arguments with teachers, getting into physical fights (Anderson et al. 2007; Möller and Krahé 2009), children's aggressive behavior observed during free play (Silvern and Williamson 1987), rating someone as less deserving of financial support (Cicchirillo and Chory-Assad 2005), making another person hold his or her hand in painfully cold ice water for an extended period of time (Ballard and Lineberger 1999), and interfering with another person's ability to win a prize (e.g., Saleem et al. 2012). Correlational and longitudinal studies have additionally used more extreme measures of aggression and found violent video game effects on carrying weapons on school property (e.g., Denniston et al. 2011; Ybarra et al. 2014), trait aggressiveness (e.g., Anderson et al. 2008; Krahe et al. 2012; Möller & Krahé 2009; Shibuya et al. 2004) violent juvenile delinquency (e.g., DeLisi et al. 2013; Hoph et al. 2008), and violent behavior (e.g., Anderson and Dill 2000; Anderson et al. 2007; Graber et al. 2006).

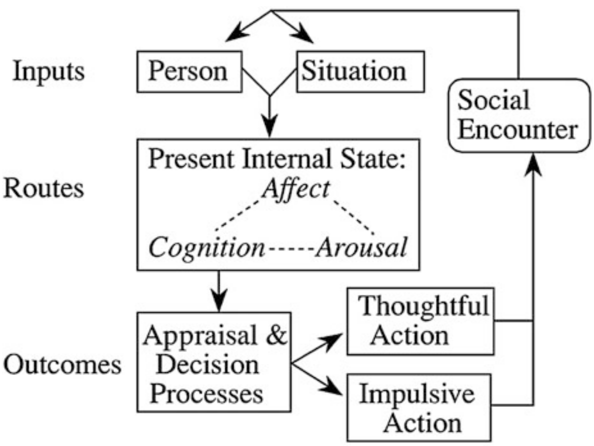
Theories of Violent Media Effects

A major purpose of psychological research on media effects is to establish whether or not a hypothesized relationship exists between engaging with specific forms of content (e.g., violent content) and outcomes of individuals (e.g., aggressive behavior). An important additional goal, however, is to develop a better understanding of *how* these effects occur. Numerous theories have contributed to our understanding of violent content effects on aggressive behavior. Each of these theories is incorporated into a larger framework called the General Aggression Model (GAM; Anderson and Bushman 2002). First, we will describe GAM before explaining each of these more domain-specific theories in turn.

The General Aggression Model is a framework that includes social, biological, and cognitive factors to help explain aggressive behavior, thoughts, and affect (emotions). Within GAM, a distinction is made between short- and long-term processes. First, we begin with the short-term processes as described by the single-cycle episode.

The single-cycle episode (Fig. 1) represents the psychological processes at work within any single given situation. As with all behavior (not just aggression), there are two forms of input that influence the behavior during a given social encounter: personological and situational variables. The personological factor includes all of the qualities of the individual that are carried across situations. These variables include all of the biological and enduring personality traits possessed by an individual including (but certainly not limited to) gender, the presence of testosterone, genetic predispositions toward aggressions, aggressive personality, agreeableness, and the like. Personological factors also include enduring beliefs (e.g., beliefs

Fig. 1 Single-cycle episode within GAM (Note. Reprinted from Anderson and Bushman 2002)



about the appropriateness of using aggression to resolve conflicts), perceptual and cognitive styles (e.g., the tendency to perceive intentional harm or threats in everyday social interactions), and attitudes. The situational factors include all characteristics of the social encounter that influence aggressive tendencies, either increasing them (e.g., receiving an insult) or decreasing them (e.g., being complimented). These factors include (but again, are certainly not limited to) the presence of provocation, exposure to violent content, temperature of the room, and aggression-inhibiting factors such as the presence of parental figures or police officers, as well as being in a setting in which aggression is socially penalized (e.g., church, a courtroom).

These inputs are not simply additive in that each personological and situational input simply adds to (risk factor) or subtracts from (protective factor) the propensity toward aggression. Many of these inputs interact with each other and these interactions can be quite complex. For example, when an individual is in a situation in which it is unclear whether a second party is provoking the individual (situational factor), the tendency to interpret this ambiguous situation as either a real provocation or merely an accidental harm depends on the individual's trait-like tendency to interpret ambiguous social information in hostile or friendly terms (hostile attribution bias, a personological factor discussed in more detail later in this chapter).

These two forms of input (situational and personological) then influence internal states of the individual. These internal states include affect, cognition, and arousal. Importantly, these internal states are highly interactive. That is, aggression-related cognitions influence arousal and affect, and vice versa. For example, when insulted, an individual may think that the provocation was unjustified (cognition) which influences affect (anger) and arousal (e.g., heart rate). Another important consideration is that the inputs described by GAM may affect one or more of these internal states and they do not necessarily all activate simultaneously. Again, within these two portions (and other portions as well) of the single-cycle episode, a number of interactions are possible. GAM does not directly elucidate the nature of these interactions but does provide indications of which variables may potentially interact.

In the third stage, the internal state leads to decision-making and appraisal processes. In this stage, the individual develops an immediate appraisal of how to behaviorally respond to the given situation. Like the previous stages, a variety of complex processes are at work during this stage that are elucidated more fully in other psychological literatures (e.g., spontaneous inference processes, Krull 1993; Krull and Dill 1996; explanation processes, Anderson et al. 1996). When making a decision to act, individuals produce an immediate appraisal of how to react to the situation. This may lead directly to an action (an impulsive action) or may lead to a reappraisal process. In this case, the individual may reappraise their decision and decide upon an alternative action. Reappraisal processes are engaged when the individual possesses the resources including cognitive resources (such as short-term memory) and time to reappraise the decision. Two additional criteria must also be met before a reappraisal is engaged. They must find the outcome of the action to be important and unsatisfying. When a predicted outcome of the behavior is satisfying, appraisal processes terminate and the behavior is engaged. The outcome may be

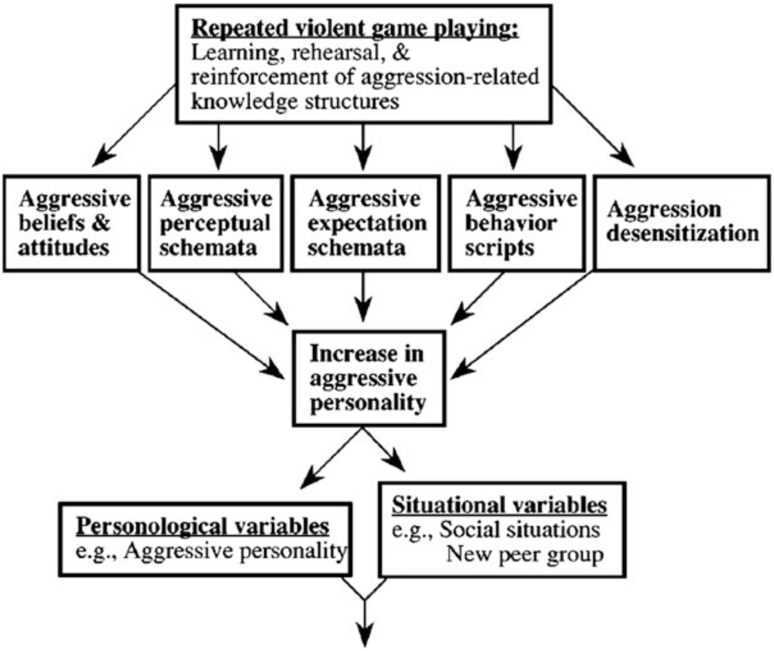


Fig. 2 Long-term processes represented in GAM (Note. Reprinted from Anderson and Bushman 2002)

unsatisfying but considered unimportant, in which case, the behavior is engaged as it would be a waste of cognitive resources to continue the reappraisal process.

The outcome of a reappraisal is not always nonaggressive. When considering the previously mentioned example in which someone is verbally insulted, the individual may initially decide to respond by returning the verbal insult. They may, however, find that the outcome of this response is important and unsatisfying after considering the unjustified nature of the original insult, the damage to one's social image, or the history of provocations received from the target individual. These examples are essentially cognitive appraisals that can additionally influence affect and arousal. This illustrates the interactive relationship that the reappraisal process has with the individual's present internal state. They may therefore decide to physically hit their provocateur which may be seen as a more justified response which can also deter future retaliation or insults.

Once the behavior has been selected and enacted, it feeds into the social situation and affects the next social encounter. The events derived from the behavior are then incorporated into the situational input factor and a new cycle begins. Each single-cycle episode can be seen as a learning trial in which the individual utilizes, develops, and reinforces knowledge structures associated with the episode. This leads to long-term changes in the individual's personality factor and influences future single-cycle episodes (see Fig. 2). These knowledge structures and related processes that lead to long-term changes in personality are discussed next.

Knowledge Structures and Priming Effects

A crucial theoretical backbone of GAM is the consideration of knowledge structures in understanding aggressive behavior. Cognitive neoassociation theory (Berkowitz 1990, 1993) was developed in an attempt to expand upon the frustration-aggression hypothesis which posits that aggression results from frustration which is a reaction to having one's goals thwarted. Cognitive neoassociation theory states that aggression results from the experience of aversive events. These aversive events directly produce negative affect, which, in turn, activate fight-or-flight tendencies. Importantly, this theory incorporates a knowledge structure approach to assist in understanding why aggressive cues increase aggression. Aggressive cues are simply stimuli that are related to aggression. The theory rests upon the understanding that long-term memory exists as a network of related concepts and the relationships between concepts vary in strength (Collins and Loftus 1975; Fig. 3). For example, the word "gun" is more closely related to the word "murder" than the word "orange." When a given concept is activated (primed), a spreading of activation occurs in which concepts that are closely related to the primed concept are also activated. This conceptual structure includes thoughts, affect, and behavioral tendencies which are all activated when presented with a cue. When viewing images of violence, aggression-related concepts are activated in memory, and the mind is effectively primed to operate using these knowledge structures.

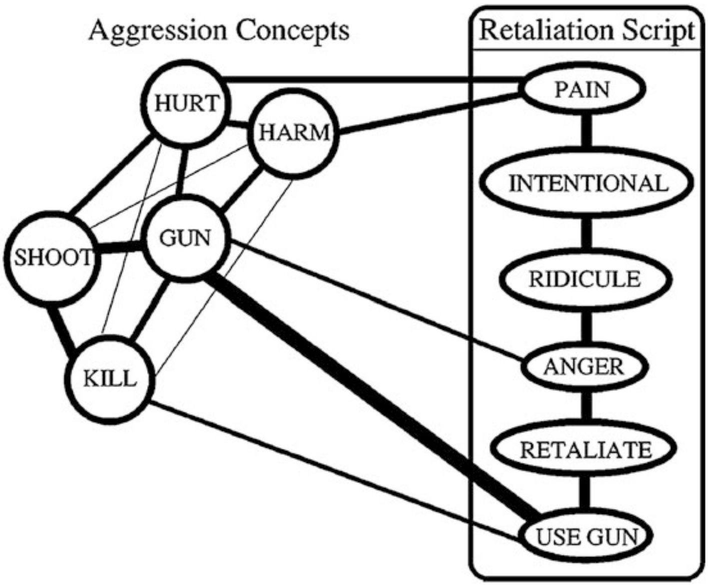


Fig. 3 Example organization of knowledge structures (Note. Reprinted from Anderson and Bushman 2002)

This priming effect is among the most influential and heavily studied phenomena across multiple domains of psychology. To demonstrate the generalizability of this process, some nonaggressive examples will be presented first. In a famous series of studies by Bargh et al. (1996), participants were randomly assigned to complete one of three versions of a sentence scramble test in which participants were tasked with unscrambling a small set of words to form a sentence. In one version, half of the sentences included words that were conceptually related to the concept of "rudeness" (e.g., disturb, intrude, interrupt, impolite). In another version, these terms were replaced with words related to the concept of "politeness" (e.g., respect, considerate, yield, courteous). A third group of participants unscrambled sentences containing neutral words, unrelated to either rudeness or politeness. After completing this task, the participant approached the experimenter to receive their next task. The experimenter, however, speaks to a confederate for an extended period of time in which she/he does not acknowledge the participant. Of primary interest was whether participants interrupted the experimenter in order to request their next task. As predicted, participants who unscrambled sentences containing words related to "rudeness" were more likely to interrupt the experimenter compared to participants primed with either neutral or polite words. In a subsequent experiment, participants primed with "elderly"-related words walked more slowly when leaving an experiment. In the words of Bargh et al. (1996), "Thinking has the function of preparing the body for action. . ." (p. 232), and this process is highly automatic and difficult to control without direct conscious awareness and effort.

As noted above, priming effects rely on existing associations in memory (e.g., the concept of elderly is closely related to the notion of slow, methodical walking). One method for examining these associations is with a tool called the Implicit Association Task. In this task, concepts are presented on a screen in which participants are asked to respond to conceptual pairings. Words or images from the conceptual categories are presented on the center of the computer screen. Participants are instructed to press a key on either the left or right side of the keyboard based on which category is presented. For example, participants may be asked to press the "i" key if either word related to "good" or words related to "candy" are presented and press the "e" key if words related to "bad" or words related to "vegetable" are presented. Participants are presented with rapid sequences of these images, and they are instructed to categorize them as quickly as possible. These categories are then swapped so the "i" key is pressed when words related to "bad" or words related to "candy" are presented and the "e" key is pressed when words related to "good" and words related to "vegetable" are displayed. When an association between concepts is strong, reaction times are faster when the two categories share the same key press. In this case, participants' reaction times are likely to be faster when the concepts of "good" and "candy" are paired when compared to the concepts of "bad" and "candy" are paired. In other words, because these concepts are more closely related to each other, faster reaction times occur when they are paired.

This method was used in research by Klimmt et al. (2010) in which participants were randomly assigned to play a video game as either a soldier or a race car driver.

They then completed a version of the IAT in which they were asked to pair words related to the self (e.g., "I," "me") and words related to the character they played as (e.g., "brave," "speed"). They found that reaction times were faster when participants paired self-related words with the character they played as, indicating that game play enhanced the associations participants had with qualities of the character they played as and their self-concept. In other words, video game play shifted participants' self-concept to be more in line with the character they had played. Thus, those who played a video game as a soldier temporarily had a self-concept more closely linked to "brave," whereas those who had played a video game as a race car driver had their self-concept temporarily more closely linked with "speed."

In a similar study, Bluemke et al. (2010) randomly assigned participants to play either a violent game in which participants shot at other characters in the game or played a nonviolent game in which all aspects of the game were identical except that participants were asked to water flowers (prosocial) or tag shapes (neutral). Participants then completed an aggression-related IAT. Those who had just played a violent game produced faster response times to self-related words and aggression-related words among the violent game players. In essence, violent game play induced a shift toward more aggressive self-concept, and this shift can be seen as causal due to the experimental design of the study.

Other research has used different methods of testing these associations. One common method involves the use of the word-completion task. The critical difference between the IAT and the word-completion task is that the IAT measures associations between concepts, while the word-completion task measures the cognitive accessibility of a given concept. This task requires that participants fill in the blanks of incomplete words as fast as they are able. Critically, some of these words can be filled complete either aggressive or nonaggressive words. For example, the word fragment "explo_e" can be completed to form the word "explode" or "explore." Research using these tasks has observed that participants asked to play violent games are more likely to complete these fragments with aggressive words, relative to those asked to play nonviolent games (Carnagey and Anderson 2005; Barlett and Rodeheffer 2009). These studies and others illustrate an increase in aggressive cognitions (an umbrella term which encompasses aggressive thought accessibility) which appear to be stronger drivers of violent video game effects on aggression than other routes (e.g., affect, arousal; Anderson and Dill 2000; Barlett and Anderson 2013; Carnagey and Anderson 2005).

Thus far, this knowledge structure approach has been used to describe only short-term effects. There is, however, strong evidence that the long-lasting associations that individuals have between concepts have potent implications for aggressive behavior in the long term. Furthermore, one way that strong, long-lasting links between such concepts develop is with practice. Just as one way to memorize a telephone number or multiplication tables is to repeatedly rehearse them, so too do various concepts become more strongly linked in long-term memory with frequent rehearsal.

A well-established finding in research on aggression is that aggressive cues, such as guns, automatically increase the accessibility of aggressive thoughts through the processes described above (e.g., Anderson et al. 1998). Bartholow and colleagues

(2005) used this effect to test whether individuals with contrasting associations with specific types of guns would respond differently to such primes. To do this, they exposed hunters and nonhunters to images of either hunting rifles or assault rifles. They found that images of hunting rifles were more likely to increase aggressive thoughts and behaviors among nonhunters. For hunters, however, aggression increased when they viewed images of assault rifles but not hunting rifles. Presumably, this is because hunters have a very different life history with hunting guns, a history that associates such guns with family outings (e.g., learning to hunt with dad). This finding indicates that the presence of an aggression prime relies on the existing knowledge structure to guide thoughts and behavior. Because the hunters had previously held associations with hunting rifles that were not necessarily aggressive (i.e., hunting rifles are used for sport, entertainment, or fun, rather than killing), the presence of a hunting rifle image did not activate aggression-related knowledge structures to guide behavior.

This fascinating work nicely illustrates how personal histories of associative pairing (e.g., hunting rifles and fun or sport) influence the behavioral outcome of any situational prime. There is, however, evidence that the development of such associations cannot only be measured over time (e.g., Krahé and Möller 2010) but trained as well. In a study by Penton-Voak and colleagues (2013), high-risk youth were provided with a weeklong training program in which they used a computerized program that presented them with several pictures of faces that ranged from unambiguously happy through completely ambiguous to unambiguously angry. Following the presentation of each image, these youth were asked to state whether the person was angry or happy. After providing each rating, participants were provided with feedback indicating whether the person was actually angry or happy. The key manipulation in this experiment was whether participants were "corrected" by the program to indicate that the ambiguous faces were indeed happy or angry. For some participants, the ambiguous faces were identified by the program as angry, whereas for other participants, the program identified ambiguous faces as happy. For participants who received feedback indicating that the ambiguous faces were indeed happy, self-reported aggressive behavior and anger were lower immediately following the training; this training effect persisted for at least 2 weeks (at which point data collection ceased). Further, the ratings of staff members working with these youth also reflected this behavioral shift.

From the perspective of GAM and cognitive neoassociation theory, this training paradigm modified the associations between the ambiguous faces and the appraisals regarding the internal state of these individuals, in essence changing the person's knowledge structures involving perception of other people's emotional states. Each time an ambiguous face is paired with the confirmation that the individual is happy, the features of the face are associated with the appraisal that the person is happy. With repeated trials utilizing a variety of faces, the features of ambiguous faces in general become associated with this appraisal and long-term behavioral changes result. Because ambiguous faces are now more likely to engender more positive appraisals, it is less likely that these individuals will respond to similar faces with anger or provocation.

Collectively, these studies illustrate how aggression-related knowledge structures (or nonaggressive knowledge structures) can develop over time and help drive behavior. When playing a video game with violent content, aggression-related concepts are activated and help drive behavior in the short term (i.e., produce a priming effect). Further, as individuals engage in violent game play over extended periods of time (e.g., months, years), aggression-related knowledge structures are developed and elaborated upon – creating a strong network of aggression-related concepts that are more readily activated and are therefore more likely to drive behavior.

Script and Social Learning Theory

Script theory posits that individuals possess knowledge structures that are responsible for guiding behavior based on a given situational context (Huesmann 1988, 1986). In a classic example, many are very familiar with what one is supposed to do when dining at a restaurant. Individuals arrive at the restaurant and (often) wait to be seated; they view the menu and order drinks then food. They eat, then pay, and eventually leave. Scripts such as this one are reinforced by viewing socially acceptable and rewarded behavior, a process elucidated by social learning theory (Bandura 1977).

When individuals play violent video games, they repeatedly view aggressive behavior in a rewarding context. Heroes who slay enemies are rewarded with social praise by other characters in the game; they receive monetary rewards, new weapons and armor, and the like. Further, the consequences of aggressive behavior are frequently not present at all in violent games. In reality, aggressive and violent actions lead to pain, death, fear, collateral damage, and social consequences that are rarely portrayed in video games. There are no grieving family members on the streets of the Grand Theft Auto franchise following a player’s rampage.

These qualities of games make the act of aggression appear more attractive and less threatening than in reality. Players are better able to imagine themselves experiencing reward for aggressive actions and simultaneously do not fully appreciate the consequences of such actions. For example, violent television use is associated with aggressive fantasizing in boys (Viemerö and Paajanen 1992). Further, imagining oneself engaging in behavior increases the individual’s intentions to actually enact the behavior (Anderson and Godfrey 1987). In addition, individuals who imagine themselves as the character within a violent game are also more likely to behave aggressively (Konijn et al. 2007).

Aggressive Attitudes and Beliefs

Another major factor to consider when understanding aggressive behavior is the way in which individuals interpret the world, hold beliefs about aggressive actions, and process social information (Crick and Dodge 1994; Dodge 2010).

Two well-understood effects of violent video game play include the development of hostile attribution biases and increases in normative beliefs about aggression.

When we see others engage in a given behavior, we make attributions about why they engaged in the behavior. For many of these decisions, this process is quite simple. When we see another individual eating, it is likely because they are hungry (or perhaps bored). At other times, information is not available to make accurate attributions regarding behavior. Instead, we rely on heuristics or mental shortcuts that allow us to make quick decisions but spare us from excessive information collection or the production of numerous possible explanations. These heuristic decision-making processes consequently lead to errors and are often based on the knowledge structures that have been developed throughout a lifetime. For example, aggressive children tend to demonstrate a hostile attribution bias. This is a tendency to interpret ambiguous provocations (i.e., a behavior that could be interpreted as either benign or hostile) in hostile terms (Crick and Dodge 1994; Dodge 2010; Orobio de Castro et al. 2002). These individuals tend to perceive the behaviors of others as hostile, while most others would interpret the same behavior as accidental or otherwise. Violent video game play has been associated with this tendency in multiple longitudinal studies in which violent video game players demonstrated increases in this tendency (Anderson et al. 2007; Gentile et al. 2011; Möller and Krahe 2009). A common measure requires that participants be asked to read ambiguous story stems describing a social encounter depicting ambiguous behavior (e.g., being bumped in a hallway) before asking participants to explain the reason for the behavior (e.g., it was done on purpose, the hallway was crowded). In these studies, individuals who reported high violent game play were more likely to make hostile attributions regarding the ambiguous behavior. In one of these longitudinal studies (Anderson et al. 2007), high-frequency violent game play early in the school year predicted increases in hostile attribution biases, which, later in the school year, predicted physical aggression.

Other work has focused on the beliefs that individuals have regarding aggressive behavior – specifically their normative beliefs about aggression. For example, exposure to violent media (television and movies) increases individuals' beliefs that the world is dangerous (Bryant et al. 1981; Gerbner et al. 1982). This is an important factor when considering how aggressive behavior is influenced by beliefs regarding the appropriateness of such actions. In fact, research findings illustrate the relationship between violent video game play and pro-violence attitudes (Funk et al. 2004) and that this relationship, in turn, leads to increases in aggressive behavior (Möller and Krahé 2009).

Desensitization to Violence

Another route through which violent video game play increases aggression is by desensitizing individuals to violence. This is a reduced emotional and physiological response to viewing violence in real life. In one major study on the topic, Carnagey et al. (2007) randomly assigned participants to play a violent or nonviolent game for

20 min. They then asked all participants to view a 10-min-long videotape of real-life violence while measuring participants' heart rate and galvanic skin response (another measure of physiological arousal). They found that individuals who had played the violent game displayed less physiological arousal than their nonviolent game-playing counterparts.

This tendency is not only present in the short term. Bartholow et al. (2006) measured the neurological responses of participants while they viewed images of actual violence. They found that the amount of long-term exposure to violent media was associated with reduced neural responding in an area of the brain that has been associated with the aversive motivational system. Further, this reduced response predicted increases in aggressive behavior.

When we are faced with the decision to aggress, an important factor we consider is how aversive the act of aggression appears to us. For individuals who find the notion of acting aggressively to be unpleasant, nonaggressive alternatives are more likely to be selected (Bartholow et al. 2006; Engelhardt 2011). Further, desensitization also reduces empathy (Anderson et al. 2010; Funk et al. 2004) and reduces helping or prosocial behavior (Bushman and Anderson 2009).

Relational Aggression

In addition to the effects of violent media on typical physical and verbal forms of aggressive behavior, some relatively new work has uncovered similar effects on relational forms of aggression. Relational aggression is a behavior enacted with the intent of harming another by manipulating another's relationships. This form of aggression can be either direct, as when someone threatens to withdraw their friendship from the target, or indirect, as when spreading rumors about another person. Sometimes relational aggression is verbal, but it also can be nonverbal or even a non-behavior (e.g., intentionally not inviting a classmate to a party that includes most other classmates). Some content analyses have been conducted on this topic, finding that relational aggression is well represented in some forms of media such as reality TV shows (Coyne et al. 2010) and popular adolescent TV shows (Coyne and Archer 2004). Much of the early work in the violent media domain has focused on the presence of physical aggression, yet these content analyses suggest that a great deal of aggressive content has been overlooked by this older literature. Research in this domain possesses two closely related foci. The first is the effect of relationally aggressive media content on aggression-related outcomes (e.g., Coyne et al. 2011). The second is the effect of physically violent media content on relational aggression (e.g., Möller and Krahe 2009). Both of these cross-domain content effects have been found. For instance, Coyne et al. (2011) found that viewing relationally aggressive media was associated with increases in relational aggression between intimate partners (e.g., threatening to break up with one's romantic partner as a coercive tactic) and this effect occurred for both males and females. Möller and Krahe (2009) found that playing violent video games was related to increases in relational aggression in a cross-sectional analysis (both measures taken at the same time); however, this cross-domain relationship did not persist over time as violent game play did not predict relational aggression 30 months later.

Another interesting finding is the discovery of short-term cross-domain effects in experimental studies. In the work by Coyne et al. (2008), participants were randomly assigned to watch a relationally aggressive video clip, a physically aggressive video clip, or a nonaggressive video clip. The researchers found that viewing either aggressive clip led to increases in both relationally aggressive behavior and physically aggressive behavior, suggesting that a relatively "general" aggression-related knowledge network is activated by viewing either of these forms of aggression. Interestingly, these authors speculated that long-term effects may demonstrate more specificity, which is precisely what was found by Möller and Krahe (2009). Work in this domain is very promising as it provides the opportunity for researchers to examine any potential specificity in aggression-related knowledge structures, as well as factors that may foster or inhibit this type of specified knowledge structure development.

Nonaggressive Outcomes of Violent Game Play

As evidenced by some of the previous discussion, viewing violence in games does not solely affect aggressive behavior. Although studies of the effect of violent content on aggressiveness have made up the bulk of research attention, other outcomes of violent game play also have received some attention. Violent actions in games are often heavily associated with other game mechanics (e.g., high action) and themes (e.g., revenge) that lead to these harmful effects. Some of these effects appear to be related to violent video game play in particular while others may be solely due to high media consumption in general.

Prosocial Behavior

As mentioned previously, violent game play has been found to reduce prosocial behaviors (Anderson et al. 2010; Greitemeyer and Mügge 2014). In two studies on this effect, researchers examined the influence of violent media use on ecologically valid (i.e., applies well to real-world events) measures of helping behavior (Bushman and Anderson 2009). In the first study, participants were randomly assigned to play a violent or a nonviolent game. After playing the game, they were asked to complete a questionnaire. While filling out this questionnaire, a staged fight occurred outside of the laboratory. Individuals who had played the violent game were less likely to rate the fight as serious and were less likely to "hear" the fight in the first place, relative to those who had played a nonviolent game. In their second study, they asked a woman to stand outside of a movie theater and struggle in picking up her crutches. Of primary interest was whether the length of time it took for movie goers to help would vary as a function of the type of movie they watched.

As the researchers predicted, they found that it took longer for individuals to help if they had just watched a violent movie compared to a nonviolent movie. Importantly, when she struggled to pick up her crutches before the movie, violent and nonviolent movie goers helped her in an equivalent amount of time, suggesting that the violent movie viewing itself was responsible for the helping decrements, not characteristics of the people going to the movie. These findings of violent media effects on prosocial behavior are not unique to these two studies and have been demonstrated by several research groups (Anderson et al. 2010; Rothmund et al. 2011; Sheese and Graziano 2005).

Attention and Impulsivity

Attention problems have been an important area of focus in the research on negative outcomes of video game play. Briefly, it should be noted that there are several possible definitions of attention. Some scholars focus on what is termed as visual attention. One aspect of visual attention is the amount of visual space within one's field of view in which information can be processed – which is termed one's useful field of view. It, however, may be more accurate to say that this is a form of visual processing, rather than attention. Regardless, video game play has been causally associated with improvements in this visual task (Green and Bavelier 2006).

Another, more colloquial use of the term attention, refers to the ability to maintain attention on a singular target – especially a target that does not naturally capture attention from viewers (e.g., a school lecture). It is this form of attention that several studies have identified decrements resulting from high television viewing in childhood (e.g., Christakis et al. 2004, 2013; Landhuis et al. 2007; Levine and Waite 2000; Mistry et al. 2007). Others have linked this propensity to habitual video game play in particular (Gentile 2009; Bioulac et al. 2008; Mistry et al. 2007).

A theoretical explanation for these effects is dubbed the Excitement Hypothesis (Gentile et al. 2012). Video games (and screen media generally) are naturally exciting and stimulating. They possess a number of cues that naturally draw players' attention such as violence or sexualized imagery (Ivory 2006; Linder and Gentile 2009), as well as more basic attention-grabbing features such as sound effects, video editing, or flickering lighting (Kubey and Csikszentmihalyi 2002). This continuous attentional "grabbing" may increase individuals' threshold for stimulation required to draw attention, and, therefore, they may find more menial tasks such as listening to a teacher, parent, or employer more difficult. Alternatively, it may simply be the case that children who have attention problems are more attracted to video game play and other electronic media. In the work by Gentile et al. (2012), support was found for both of these accounts. In other words, it appears that video game play exacerbates attention problems (also seen by Swing et al. 2010), and individuals with attention problems are also attracted to video games.

Related to this issue, the effect of video game play on impulsivity has also demonstrated a bidirectional relationship (Gentile et al. 2012). Individuals who played violent games in particular were more likely to agree to statements such as "I do things without thinking" and "I act on the spur of the moment" (Swing and Anderson 2014). Further, this finding supported a unique route through which violent media use increased aggressive behavior (Swing and Anderson 2014).

Unfortunately, it may be difficult for game designers to create games that avoid developing problems such as these. Part of the attraction that individuals undoubtedly experience for video games is that they are exciting and grab their attention. It may be prudent for designers to develop games that require slow-paced thinking which may influence attention and impulsivity in a positive manner. Though, more research is needed in this area to ensure that this is the case.

Risk Taking

Related to the observed effects of violent game play on impulsivity, risk-taking behaviors are also seen to result from certain types of video game play. Much of the research on this topic has focused on the relationship between playing racing video games and risky driving behaviors. Theoretically, the processes that give rise to increases in risky driving behaviors from racing games are similar to the processes that are responsible for violent content effects on aggression. In one report by Fischer et al. (2009), four separate experiments were conducted with a focus on elucidating the processes underlying the racing game – risky driving effect. As expected by the researchers, and predicted by GAM, participants assigned to play a racing game exhibited increased inclinations for reckless driving, and this effect was partially fueled by changes in participants' perceptions of themselves as risky drivers. In other words, participants who were randomly assigned to play a racing video game were consequently more likely to view themselves as risky drivers and therefore displayed more risk-taking tendencies.

This tendency is not limited solely to brief exposures to racing games. In another study by Beullens et al. (2011), adolescents were studied over a period of 2 years. The researchers found that respondents who played racing video games were more likely to report increased risk-taking behavior including speeding and "fun riding" (taking risks while driving to make driving more entertaining). These effects also appeared to be mediated by changes in players' attitudes and intentions regarding these behaviors. In other words, racing game players tended to have more positive attitudes toward these risk-taking behaviors, and these attitudes predicted intentions to engage in the behavior, which, in turn, predicted actual speeding behaviors. Lastly, to rule out the possibility that these relationships could be explained by an attraction hypothesis (i.e., individuals susceptible to risky driving tend to play more racing games), the authors observed these effects even after statistically controlling for aggression and sensation seeking. This provides some evidence that racing video game play influences these attitudes and behaviors, not necessarily the other way around.

While these studies focused largely on the impact of risky driving games on driving behavior, other work elucidates the impact of more general risk-glorifying games on other risk-taking behaviors. In a 4-year longitudinal study by Hull et al. (2014), participants were asked to report the frequency with which they played mature-rated games and their engagement in a number of risk-taking behaviors including (but not limited to) smoking, sex without a condom, and binge drinking. The authors found that mature-rated video game play was associated with increases in all of the measured risk-taking behaviors. The authors also measured a number of mediating variables and found that the increases in risk-taking behaviors were at least partially fueled by the impact of mature-rated game play on increased sensation seeking, more positive attitudes toward deviant behavior, and affiliation with more delinquent peers.

One particularly interesting result of this study was that the impact of game play on risk-taking behaviors appeared to vary as a function of the specific games that participants played. Participants were asked to report whether they had played three specific games (Spider-Man 2, Manhunt, and Grand Theft Auto III). The authors were interested in understanding whether characteristics of the protagonist would have a measurable impact on risk-taking behaviors. They found that the game with a protagonist who largely engages in risk-taking behaviors with the intent of helping others (Spider-Man 2) was only weakly related to the players' real-world risk-taking behaviors, compared to games with protagonists who possess more deviant motives (Manhunt, Grand Theft Auto III). This makes some sense when considering the self-conceptual shifts that occur during game play discussed previously. Characters such as Spider-Man are likely not viewed as inherently "risky" individuals. They engage in risky behavior solely for the purpose of helping others. Deviant characters, on the other hand, are more likely viewed as inherently risky individuals as their risk-taking behaviors (e.g., theft, murder) are more often conducted in service of their own self-interests (e.g., fun, to obtain money, or infamy). As players engage in the roles of these characters over time, the self-concept is repeatedly paired with concepts related to the characters which then drive behavior. While not a primary purpose of the study, this finding is certainly interesting, and more work is needed to fully elucidate the processes hypothesized to be involved, especially for the long-term effects observed in this study.

Stereotyped Depictions

While depictions of stereotypes have a history in the television and movie domain, relatively less work has been done on this topic with regard to video games. Much of this work has focused on stereotypical portrayals of both men and women in video games. Female characters in video games are frequently depicted as sexualized, attractive, and weak, while males are more often portrayed as aggressive, muscular, and dominant (Beasley and Collins-Standley 2002; Dill and Thill 2007; Stermer and Burkley 2012). In the work by Dill and colleagues (2008), participants either viewed gender-stereotyped images of video game characters or viewed images of professional men and women. They then read a real-life vignette describing a sexual harassment incident in which a male college professor harassed a female student. They found that male participants found the sexual harassment to be more tolerable when they had viewed the sexualized images of video game characters. In the other work by Beck et al. (2012), participants were randomly assigned to watch a series of events within a video game in which violence and sexual exploitation against women are portrayed in the game Grand Theft Auto IV or watched a video of a baseball video game being played. They found that individuals viewing the sexually stereotyped game play were more likely to endorse rape myth attitudes among male participants.

Other work has focused on depictions of characters as racially stereotyped. Researchers Saleem and Anderson (2013) conducted research indicating that individuals playing a video game that depicted Arab characters as terrorists observed increased anti-Arab attitudes and perceptions of Arabs as aggressive compared to participants who played other games without such depictions, although they also observed increased anti-Arab attitudes among participants who played a game with Russian terrorists. This finding suggests that there exists a strong preexisting association between Arabs and terrorism, perhaps partially resulting from televised coverage of Arabs in a stereotyped or unbalanced manner.

Unlike some of the other negative effects of video game play, it is much easier for video game developers to produce games that do not exacerbate stereotyped beliefs of players. Violent content and high pace of action are often highly related in the entertainment value of video games, at least among the general population – making the avoidance of these mechanics difficult for many developers who wish to sell as many games as possible. Stereotyped depictions of women and racial groups are, at least in contrast, much easier to avoid including in games. The benefits of such game design decisions are not likely to reflect in a developer's checkbook, but will provide a measurable positive effect on individuals, and society at large.

Conclusion

We would like to end this chapter in a manner similar to the way in which it began: by pointing out that video games are not inherently bad and restating the veracity with which these conclusions are reported. While not the focus of the current chapter, there are a number of psychological and social benefits to video game play. The negative outcomes we have described in this chapter rely on the same basic psychological processes that are known to produce positive outcomes (e.g., the prosocial benefits of prosocial game play versus the aggressive problems resulting from violent play; Greitemeyer and Mugge 2014). Other benefits of certain types of games, such as improvements in the useful field of view, may partially be responsible for the decrements in other areas, such as proactive executive control. As an example, for children who are better able to process visual information in their periphery, they may be more easily distracted by attention-grabbing cues in the classroom (e.g., the child fidgeting nearby; Gentile et al. 2014). These examples illustrate the flexibility of psychological processes to give rise to a variety of outcomes and exemplify the complexity inherent in media effects research. Thus, the common question we hear from parents and policy makers – "Are video games good or bad for children and adolescent?" – is much too simplistic.

A subset of the research we have described here – particularly the effects of violent content on aggressive behavior – has been the subject of intense debate in the scientific literature (Anderson et al. 2013; Bushman and Huesmann 2014; Elson and Ferguson 2013a, b; Ferguson 2014; Groves et al. 2014; Krahe 2014; Warburton 2013). Critics of violent media effects research are a small but vocal group who do not agree with the vast majority of media effects scholars and pediatricians (Anderson et al. 2015; Bushman et al. 2015). Debate within the sciences is most often seen as a healthy component of the scientific process. However, when all alternative explanations have been empirically tested and disagreement persists despite consistent empirical support, such persistent denial of well-established findings stymies scientific progress, as well as misinforms and confuses parents, teachers, game players, and developers. It is our hope that the evidence described here helps to provide a clear picture of the negative effects of video game play and assists in informing consumers and developers alike.

There are many positive aspects of video games, as attested by other chapters in this volume and by some of our own research and published papers. Indeed, both the authors of this chapter are big fans of video games in general and are excited about the future of games and their potential for positive effects in a wide array of contexts.

Recommended Reading

American Academy of Pediatrics, Policy statement – media violence. Pediatrics **124**(5), 1495–1503 (2009)

American Psychological Association, APA calls for reduction of violence in interactive media used by children and adolescents (2005), Retrieved 20 June 2013 from http://www.apa.org/news/press/releases/2005/08/video-violence.aspx

C.A. Anderson, An update on the effects of playing violent video games. J. Adolesc. **27**, 113–122 (2004)

C.A. Anderson, B.J. Bushman, Effects of violent video games on aggressive behavior, aggressive cognition, aggressive affect, physiological arousal, and prosocial behavior: a meta-analytic review of the scientific literature. Psychol. Sci. **12**(5), 353–359 (2001)

C.A. Anderson, B.J. Bushman, Human aggression. Annu. Rev. Psychol. **53**, 27–51 (2002)

C.A. Anderson, K.E. Dill, Video games and aggressive thoughts, feelings, and behavior in the laboratory and in life. J. Pers. Soc. Psychol. **78**, 772–790 (2000)

C.A. Anderson, S.S. Godfrey, Thoughts about actions: the effects of specificity and availability of imagined behavioral scripts on expectations about oneself and others. Soc. Cogn. **5**(3), 238–258 (1987)

C.A. Anderson, Violent video games: Myths, facts, and unanswered questions. Psychological Science Agenda: Science Briefs, October, **16**(5), 1–3 (2003)

C.A. Anderson, D.S. Krull, B. Weiner, Explanations: processes and consequences, in *Social Psychology: Handbook of Basic Principles*, ed. by E.T. Higgins, A.W. Kruglanski (Guilford Press, New York, 1996), pp. 271–296

C.A. Anderson, A.J. Benjamin, B.D. Bartholow, Does the gun pull the trigger? Automatic priming effects of weapon pictures and weapon names. Psychol. Sci. **9**, 308–314 (1998)

C.A. Anderson, N.L. Carnagey, M. Flanagan, A.J. Benjamin, J. Eubanks, J.C. Valentine, Violent video games: specific effects of violent content on aggressive thoughts and behavior. Adv. Exp. Soc. Psychol. **36**, 199–249 (2004)

C.A. Anderson, D.A. Gentile, K.E. Buckley, *Violent Video Game Effects on Children and Adolescents: Theory, Research, and Public Policy* (Oxford University Press, New York, 2007)

C.A. Anderson, A. Sakamoto, D.A. Gentile, N. Ihori, A. Shibuya, S. Yukawa, M. Naito, K. Kobayashi, Longitudinal effects of violent video games on aggression in Japan and the United States. Pediatrics **122**(5), e1067–e1072 (2008). doi:10.1542/peds.2008-1425

C.A. Anderson, A. Shibuya, N. Ihori, E.L. Swing, B.J. Bushman, A. Sakamoto, H.R. Rothstein, M. Saleem, Violent video game effects on aggression, empathy, and prosocial behavior in eastern and western countries: a meta-analytic review. Psychol. Bull. **136**, 151–173 (2010)

C.A. Anderson, M. DeLisi, C.L. Groves, Subtracting from scientific knowledge about media effects. [Review of: adolescents, crime, and the media: a critical analysis by Christopher J. Ferguson]. PsycCRITIQUES **58**(51), Article 2 (2013). doi:10.1037/a0034788

C.A Anderson, L. Andrighetto, B.D. Bartholow, L. Begue, P. Boxer, J.F. Brockmyer, . . . W Warburton, Consensus on media violence effects: comment on Bushman, Gollwitzer, and Cruz. Psychol. Pop. Media Cult. **4**(3), 215–221 (2015)

P. Arriaga, F. Esteves, P. Carneiro, M.B. Monterio, Are the effects of unreal violent video games pronounced when playing with a virtual reality system? Aggress. Behav. **34**, 521–538 (2008)

M.E. Ballard, R. Lineberger, Video game violence and confederate gender: effects on reward and punishment given by college males. Sex Roles **41**, 541–558 (1999)

A. Bandura, *Social Learning Theory* (General Learning Press, New York, 1977)

J.A. Bargh, M. Chen, L. Burrows, Automaticity of social behavior: direct effects of trait construct and stereotype activation on action. J. Pers. Soc. Psychol. **71**(2), 230 (1996)

C.P. Barlett, C.A. Anderson, Examining media effects: the general aggression and general learning models, in *Media Effects/Media Psychology*, ed. by E. Scharrer (Blackwell-Wiley, San Francisco, 2013), pp. 1–20e

C.P. Barlett, C. Rodeheffer, Effects of realism on extended violent and nonviolent video game play on aggressive thoughts, feelings, and physiological arousal. Aggress. Behav. **35**(3), 213–224 (2009)

C.P. Barlett, O. Branch, C. Rodeheffer, R.J. Harris, How long do the short term video game effects last? Aggress. Behav. **35**, 225–236 (2009)

B.D. Bartholow, C.A. Anderson, N.L. Carnagey, A.J. Benjamin, Individual differences in knowledge structures and priming: the weapons priming effect in hunters and nonhunters. J. Exp. Soc. Psychol. **41**, 48–60 (2005a)

B.D. Bartholow, M.A. Sestir, M.D. Davis, Correlates and consequences of exposure to videogame violence: hostile personality, empathy, and aggressive behavior. Personal. Soc. Psychol. Bull. **31**, 1573–1586 (2005b)

B.D. Bartholow, B.J. Bushman, M.A. Sestir, Chronic violent video game exposure and desensitization: behavioral and event-related brain potential data. J. Exp. Soc. Psychol. **42**, 532–539 (2006)

B. Beasley, T.C. Collins-Standley, Shirts vs. skins: clothing as an indicator of gender stereotyping in video games. Mass Commun. Soc. **5**, 279–293 (2002)

V.S. Beck, S. Boys, C. Rose, E. Beck, Violence against women in video games: a prequel or sequel to rape myth acceptance?. J. Interpers. Violence. **27**(15), 3016–3031 (2012), 0886260512441078

L. Berkowitz, On the formation and regulation of anger and aggression: a cognitive neoassociationistic analysis. Am. Psychol. **45**, 494–503 (1990)

L. Berkowitz, Pain and aggression: some findings and implications. Motiv. Emot. **17**(3), 277–293 (1993)

K. Beullens, K. Roe, J. Van den Bulck, Excellent gamer, excellent driver? The impact of adolescents' video game playing on driving behavior: a two-wave panel study. Accid. Anal. Prev. **43**(1), 58–65 (2011)

S. Bioulac, L. Arfi, M.P. Bouvard, Attention deficit/hyperactivity disorder and video games: a comparative study of hyperactive and control children. Eur. Psychiatry **23**(2), 134–141 (2008)

M. Bluemke, M. Friedrich, J. Zumbach, The influence of violent and nonviolent computer games on implicit measures of aggressiveness. Aggress. Behav. **36**(1), 1–13 (2010)

P. Boxer, C.L. Groves, M. Docherty, Video games do indeed influence children and adolescents' aggression, prosocial behavior, and academic performance: a clearer reading of Ferguson. Perspect. Psychol. Sci. **10**(5), 671–673 (2015)

J. Bryant, R.A. Carveth, D. Brown, Television viewing and anxiety: an experimental examination. J. Commun. **31**(1), 106–119 (1981)

B.J. Bushman, C.A. Anderson, Comfortably numb desensitizing effects of violent media on helping others. Psychol. Sci. **20**(3), 273–277 (2009)

B.J. Bushman, L.R. Huesmann, Short-term and long-term effects of violent media on aggression in children and adults. Archives of Pediatrics & Adolescent Medicine, **160**(4), 348–352 (2006)

B.J. Bushman, L.R. Huesmann, Twenty-five years of research on violence in digital games and aggression revisited: a reply to Elson and Ferguson (2013). Eur. Psychol. **19**, 47–55 (2014)

B.J. Bushman, M. Gollwitzer, C. Cruz, There is broad consensus: media researchers agree that violent media increase aggression in children, and pediatricians and parents concur. Psychol. Pop. Media Cult. **4**(3), 200–214 (2015)

N.L. Carnagey, C.A. Anderson, The effects of reward and punishment in violent video games on aggressive affect, cognition, and behavior. Psychol. Sci. **16**(11), 882–889 (2005)

N.L. Carnagey, C.A. Anderson, B.J. Bushman, The effect of video game violence on physiological desensitization to real life violence. J. Exp. Soc. Psychol. **43**, 489–496 (2007)

D.A. Christakis, F.J. Zimmerman, D.L. DiGiuseppe, C.A. McCarty, Early television exposure and subsequent attentional problems in children. Pediatrics **113**(4), 708–713 (2004)

D.A. Christakis, M.M. Garrison, T. Herrenkohl, K. Haggerty, F.P. Rivara, C. Zhou, K. Liekweg, Modifying media content for preschool children: a randomized controlled trial. Pediatrics **131**, 431–438 (2013)

V. Cicchirillo, R.M. Chory-Assad, Effects of affective orientation and video game play on aggressive thoughts and behaviors. J. Broadcast. Electron. Media **49**, 435–449 (2005)

A.M. Collins, E.F. Loftus, A spreading activation theory of semantic processing. Psychol. Rev. **82**, 407–428 (1975)

S.M. Coyne, D.A. Nelson, F. Lawton, S. Haslam, L. Rooney, L. Titterington, L. Ogunlaja, The effects of viewing physical and relational aggression in the media: Evidence for a cross-over effect. J. Exp. Soc. Psychol. **44**(6), 1551–1554 (2008)

S.M. Coyne, J. Archer, Indirect aggression in the media: a content analysis of British television programs. Aggress. Behav. **30**(3), 254–271 (2004)

S.M. Coyne, S.L. Robinson, D.A. Nelson, Does reality backbite? Physical, verbal, and relational aggression in reality television programs. J. Broadcast. Electron. Media **54**(2), 282–298 (2010)

S.M. Coyne, D.A. Nelson, N. Graham-Kevan, E. Tew, K.N. Meng, J.A. Olsen, Media depictions of physical and relational aggression: connections with aggression in young adults' romantic relationships. Aggress. Behav. **37**(1), 56–62 (2011)

N.R. Crick, K.A. Dodge, A review and reformulation of social information-processing mechanisms in children's social adjustment. Psychol. Bull. **115**, 74–101 (1994)

M. DeLisi, M.G. Vaughn, D.A. Gentile, C.A. Anderson, J. Shook, Violent video games, delinquency, and youth violence: new evidence. Youth Violence Juvenile Justice **11**, 132–142 (2013). doi:10.1177/1541204012460874

M.M. Denniston, M.H. Swahn, M.F. Hertz, L.M. Romero, Associations between electronic media use and involvement in violence, alcohol and drug use among United States high school students. West. J. Emerg. Med. **12**(3), 310 (2011)

K.E. Dill, K.P. Thill, Video game characters and the socialization of gender roles: young people's perceptions mirror sexist media depictions. Sex Roles **57**, 851–865 (2007)

K.E. Dill, B.P. Brown, M.A. Collins, Effects of exposure to sex-stereotyped video game characters on tolerance of sexual harassment. J. Exp. Soc. Psychol. **44**(5), 1402–1408 (2008)

K.A. Dodge, Social information processing patterns as mediators of the interaction between genetic factors and life experiences in the development of aggressive behavior, in *Understanding and Reducing Aggression, Violence, and Their Consequences* (American Psychological Association, Washington, DC, 2010)
M. Elson, C.J. Ferguson, Twenty-five years of research on violence in digital games and aggression. Eur. Psychol. **19**(1), 33–46 (2013a)
M. Elson, C.J. Ferguson, Does doing media violence research make one aggressive? The ideological rigidity of social-cognitive theories of media violence and a response to Bushman and Huesmann (2013), Krahé (2013), and Warburton (2013). Eur. Psychol. **19**, 68–75 (2013b)
C.R. Engelhardt, B.D. Bartholow, G.T. Kerr, B.J. Bushman, This is your brain on violent video games: neural desensitization to violence predicts increased aggression following violent video game exposure. J. Exp. Soc. Psychol. **47**, 1033–1036 (2011a)
C.R. Engelhardt, B.D. Bartholow, J.S. Saults, Violent and nonviolent video games differentially affect physical aggression for individuals high vs. low in dispositional anger. Aggress. Behav. **37**(6), 539–546 (2011b)
C.J. Ferguson, The good, the bad and the ugly: a meta-analytic review of the positive and negative effects of violent video games. Psychiatry Q. **78**, 309–316 (2007a)
C.J. Ferguson, Evidence for publication bias in video game violence effects literature: a meta-analytic review. Aggress. Violent Behav. **12**, 470–482 (2007b)
C.J. Ferguson, When is a book review really something else? A response to the review of adolescents, crime and the media: a critical analysis. PsycCRITIQUES. **59**(10), (2014)
C.J. Ferguson, J. Kilburn, Much ado about nothing: the misestimation and over interpretation of violent video game effects in Eastern and Western nations ~ comment on Anderson et al. (2010). Psychol. Bull. **136**(2), 174–178 (2010)
P. Fischer, T. Greitemeyer, T. Morton, A. Kastenmüller, T. Postmes, D. Frey, . . . J. Odenwälder, The racing-game effect: why do video racing games increase risk-taking inclinations?. Pers. Soc. Psychol. Bull. **35**, 1395–1409 (2009)
J.B. Funk, H.B. Baldacci, T. Pasold, J. Baumgardner, Violence exposure in real-life, video games, television, movies, and the internet: is there desensitization? J. Adolesc. **27**, 23–39 (2004)
D.A. Gentile, Pathological video game use among youth 8 to 18: A national study. Psychol. Sci. **20**, 594–602 (2009)
D.A. Gentile, C.A. Anderson, S. Yukawa, M. Saleem, K.M. Lim, A. Shibuya et al., The effects of prosocial video games on prosocial behaviors: international evidence from correlational, longitudinal, and experimental studies. Personal. Soc. Psychol. Bull. **35**, 752–763 (2009)
D.A. Gentile, C.L. Groves, J.R. Gentile, The General Learning Model: Unveiling the Learning Potential in Video Games. In F. Blumberg (Ed.), *Learning by Playing: Video Gaming in Education*. New York, NY: Oxford University Press, 121–142 (2014)
D.A. Gentile, S.M. Coyne, D.A. Walsh, Media violence, physical aggression and relational aggression in school age children: a short-term longitudinal study. Aggress. Behav. **37**, 193–206 (2011)
D.A. Gentile, E.L. Swing, C.G. Lim, A. Khoo, Video game playing, attention problems, and impulsiveness: evidence of bidirectional causality. Psychol. Pop. Media Cult. **1**(1), 62–70 (2012)
G. Gerbner, L. Gross, M. Morgan, N. Signorielli, Charting the mainstreaming: television's contributions to political orientations. J. Commun. **32**, 100–127 (1982)
J.A. Graber, T. Nichols, S.D. Lynne, J. Brooks-Gunn, G.J. Botvin, A longitudinal examination of family, friend, and media influences on competent versus problem behaviors among urban minority youth. Appl. Dev. Sci. **10**, 75–85 (2006)
C.S. Green, D. Bavelier, Effect of action video games on the spatial distribution of visuospatial attention. J. Exp. Psychol. Hum. Percept. Perform. **32**(6), 1465 (2006)
T. Greitemeyer, D.O. Mügge, Video games do affect social outcomes: a meta-analytic review of the effects of violent and prosocial video game play. Personal. Soc. Psychol. Bull. **40**(5), 578–589 (2014)

C.L. Groves, C.A. Anderson, M. DeLisi, A response to Ferguson: more red herring. *PsycCRITIQUES*. **59**(10) article 9 (2014). doi:10.1037/a0036266
S. Hearold, A synthesis of 1043 effects of television on social behavior, in *Public Communication and Behavior*, ed. by G. Comstock, vol. 1 (Academic, New York, 1986), pp. 66–133
W.H. Hoph, G.L. Huber, R.H. Weib, Media violence and youth violence: a 2-year longitudinal study. J. Media Psychol. **20**, 79–96 (2008)
L.R. Huesmann, Psychological processes promoting the relation between exposure to media violence and aggressive behavior by the viewer. J. Soc. Issues **42**(3), 125–139 (1986)
R.L. Huesmann, An information processing model for the development of aggression. Aggress. Behav. **14**, 13–24 (1988)
J.G. Hull, T.J. Brunelle, A.T. Prescott, J.D. Sargent, A longitudinal study of risk-glorifying video games and behavioral deviance. J. Pers. Soc. Psychol. **107**(2), 300 (2014)
International Society for Research on Aggression, Report of the media violence commission. Aggress. Behav. **38**, 335–341 (2012)
J.D. Ivory, Still a man's game: gender representation in online reviews of video games. Mass Commun. Soc. **9**(1), 103–114 (2006)
T. Katori, Bouryokuteki bideogemu no kougeki sokushin kouka to sougosayousei [The effects of violent video games and interactivity on aggression]. in *Proceedings of the 42nd convention of the Japanese Society of Social Science,* pp. 602–603 (2001)
C. Klimmt, D. Hefner, P. Vorderer, C. Roth, C. Blake, Identification with video game characters as automatic shift of self-perceptions. Media Psychol. **13**(4), 323–338 (2010)
E.A. Konijn, M.N. Bijnank, B.J. Bushman, I wish I were a warrior: the role of wishful identification in the effects of violent video games on aggression in adolescent boys. Dev. Psychol. **43**, 1038–1044 (2007)
B. Krahé, I. Möller, Longitudinal effects of media violence on aggression and empathy among German adolescents. J. Appl. Dev. Psychol. **31**(5), 401–409 (2010)
B. Krahé, Restoring the spirit of fair play in the debate about violent video games: a comment on Elson and Ferguson (2013). Eur. Psychol. **19**, 56–59 (2014)
B. Krahe, R. Bushing, I. Moller, Media violence use and aggression among German adolescents: associations and trajectories of change in a three-wave longitudinal study. Psychol. Pop. Media Cult. **1**, 152–156 (2012)
M. Krcmar, K. Farrar, Retaliatory aggression and the effects of point of view and blood in violent video games. Mass Commun. Soc. **12**(1), 115–138 (2009)
D.S. Krull, Does the grist change the mill? The effect of the perceiver's inferential goal on the process of social inference. Personal. Soc. Psychol. Bull. **19**(3), 340–348 (1993)
D.S. Krull, J.C. Dill, On thinking first and responding fast: flexibility in social inference processes. Personal. Soc. Psychol. Bull. **22**(9), 949–959 (1996)
R. Kubey, M. Csikszentmihalyi, Television addiction. Sci. Am. **286**(2), 74–81 (2002)
C.E. Landhuis, R. Poulton, D. Welch, R.J. Hancox, Does childhood television viewing lead to attention problems in adolescence? Results from a prospective longitudinal study. Pediatrics, **120**(3), 532–537 (2007)
L.E. Levine, B.M. Waite, Television viewing and attentional abilities in fourth and fifth grade children. J. Appl. Dev. Psychol. **21**(6), 667–679 (2000)
J.R. Linder, D.A. Gentile, Is the television rating system valid? Indirect, verbal, and physical aggression in programs viewed by fifth grade girls and associations with behavior. J. Appl. Dev. Psychol. **30**(3), 286–297 (2009)
K.B. Mistry, C.S. Minkovitz, D.M. Strobino, D.L. Borzekowski, Children's television exposure and behavioral and social outcomes at 5.5 years: does timing of exposure matter? Pediatrics **120**, 762–769 (2007)
I. Möller, B. Krahe, Exposure to violent video games and aggression in German adolescents: a longitudinal analysis. Aggress. Behav. **35**, 75–89 (2009)
B. Orobio de Castro, J.W. Veerman, W. Koops, J.D. Bosch, H.J. Monshouwer, Hostile attribution of intent and aggressive behavior: a meta-analysis. Child Dev. **73**, 916–934 (2002)

H. Paik, G. Comstock, The effects of television violence on antisocial behavior: a meta-analysis1. Commun. Res. **21**(4), 516–546 (1994)
I.S. Penton-Voak, J. Thomas, S.H. Gage, M. McMurran, S. McDonald, M.R. Munafò, Increasing recognition of happiness in ambiguous facial expressions reduces anger and aggressive behavior. Psychol. Sci. **24**, 688–697 (2013)
S. Prot, C.A. Anderson, D.A. Gentile, S.C. Brown, E.L. Swing, The positive and negative effects of video game play, in *Children and Media*, ed. by A. Jordan, D. Romer (Oxford University Press, New York, 2014a), pp. 109–128
S. Prot, D.G. Gentile, C.A. Anderson, K. Suzuki, E. Swing, K.M. Lim, Y. Horiuchi, M. Jelic, B. Krahé, W. Liuqing, A. Liau, A. Khoo, P.D. Petrescu, A. Sakamoto, S. Tajima, R.A. Toma, W.A. Warburton, X. Zhang, C.P. Lam, Long-term relations between prosocial media use, empathy and prosocial behavior. Psychol. Sci. **25**, 358–368 (2014b)
T. Rothmund, M. Gollwitzer, C. Klimmt, Of virtual victims and victimized virtues: differential effects of experienced aggression in video games on social cooperation. Personal. Soc. Psychol. Bull. **37**, 107–119 (2011)
A. Sakamoto, M. Ozaki, R. Narushima, T. Mori, K. Sakamoto, M. Takahira et al., Terebigemu asobi ga ningen no bouryoku koudou ni oyobosu eikyo to sono katei: Joshidaigakusei ni taisuru 2-tsu no shakaishinrigakuteki jikken [The influence of video game play on human violence and its process: two social psychological experiments of female university students]. Stud. Simul. Gaming **11**(1), 28–39 (2001)
M. Saleem, C.A. Anderson, Arabs as terrorists: effects of stereotypes within violent contexts on attitudes, perceptions and affect. Psychol Violence **3**, 84–99 (2013)
M. Saleem, C.A. Anderson, D.A. Gentile, Effects of prosocial, neutral, and violent video games on children's helpful and hurtful behaviors. Aggress. Behav. **38**, 281–287 (2012). doi:10.1002/ab.21428
B.E. Sheese, W.G. Graziano, Deciding to defect: the effects of video-game violence on cooperative behavior. Psychol. Sci. **16**, 354–357 (2005)
J.L. Sherry, The effects of violent video games on aggression. Hum. Commun. Res. **27**(3), 409–431 (2001)
A. Shibuya, A. Sakamoto, N. Ihori, S. Yukawa, Media bouryoku heno sesshoku, sesshoku kankyo ga kougekisei ni oyobosu choukiteki eikyou: Shougakusei heno paneru chosa [The long-term effects of media violence and its situational variables on aggression: A panel study to elementary school children]. in *Proceeding of the 45th convention of the Japanese Society of Social Psychology*, pp. 248–249 (2004)
S.B. Silvern, P.A. Williamson, The effects of video game play on young children's aggression, fantasy, and prosocial behavior. J. Appl. Dev. Psychol. **8**(4), 453–462 (1987)
S.L. Smith, K. Lachlan, R. Tamborini, Popular video games: Quantifying the presentation of violence and its context. J. Broadcast. Electron. Media **47**(1), 58–76 (1998)
S.P. Stermer, M. Burkley, Xbox or SeXbox? An examination of sexualized content in video games. Soc. Personal. Psychol. Compass **6**(7), 525–535 (2012)
E.L. Swing, C.A. Anderson, The role of attention problems and impulsiveness in media violence effects on aggression. Aggress. Behav. **40**, 197–203 (2014). doi:10.1002/ab.21519
E.L. Swing, D.A. Gentile, C.A. Anderson, D.A. Walsh, Television and video game exposure and the development of attention problems. Pediatrics **126**(2), 214–221 (2010)
K.M. Thompson, K. Haninger, Violence in E-rated video games. J. Am. Med. Assoc. **286**, 591–598 (2001)
K.M. Thompson, K. Tepichin, K. Haninger, Content and ratings of mature-rated video games. Arch. Pediatr. Adolesc. Med. **160**, 402–410 (2006)
V. Viemerö, S. Paajanen, The role of fantasies and dreams in the TV viewing-aggression relationship. Aggress. Behav. **18**, 109–116 (1992)
W. Warburton, Apples, oranges, and the burden of proof – putting media violence findings into context: a comment on Elson and Ferguson (2013). Eur. Psychol. **19**, 60–67 (2013)

B.J. Wilson, D. Kunkel, D. Linz, W.J. Potter, E. Donnerstein, S.L. Smith, E. Blumenthal, M. Berry, Violence in television programming overall: University of California, Santa Barbara study, in *National Television Violence Study*, vol. 1 (Sage, Thousand Oaks, 1997), pp. 3–268

B.J. Wilson, D. Kunkel, D. Linz, W.J. Potter, E. Donnerstein, S.L. Smith, E. Blumenthal, M. Berry, Violence in television programming overall: University of California, Santa Barbara study, in *National Television Violence Study*, vol. 2 (Sage, Thousand Oaks, 1998), pp. 3–204

W. Wood, F.Y. Wong, J.G. Chachere, Effects of media violence on viewers' aggression in unconstrained social interaction. Psychol. Bull. **109**(3), 371 (1991)

G.S. Yang, L.R. Huesmann, B.J. Bushman, Effects of playing a violent video game as male versus female avatar on subsequent aggression in male and female players. Aggress. Behav. **40**, 537–541 (2014). doi:10.1002/ab.21551

M.L. Ybarra, L.R. Huesmann, J.D. Korchmaros, S.L. Reisner, Cross-sectional associations between violent video and computer game playing and weapon carrying in a national cohort of children. Aggress. Behav. **40**, 345–358 (2014). doi:10.1002/ab.21526

F. Yokota, K.M. Thompson, Violence in G-rated animated feature films. J. Am. Med. Assoc. **283**, 2716–2720 (2000)

Unconscious Emotions in Media Content 50

Huang-Ming Chang, Leonid Ivonin, and Matthias Rauterberg

Contents

H.-M. Chang (✉) • L. Ivonin • M. Rauterberg
Industrial Design, Eindhoven University of Technology, Eindhoven, The Netherlands
e-mail: Huang.M.Chang@gmail.com; Leonid.Ivonin@gmail.com; G.W.M.Rauterberg@tue.nl; ifip.tc14.chair@gmail.com

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4_41

Abstract

Stories and human society are inseparable. Before human history, our ancestors created myths as the oldest form of narratives. Due to the advances of digital media, stories are nowadays communicated in various ways. Yet the structure and the key elements of modern stories still remain the same. This chapter starts with a discussion about the relationship between stories and human society. It is curious that human beings keep being intrigued by stories of similar structures and scenes. A reasonable speculation is that these identical components of stories might induce emotional experiences at an unconscious level. Can emotions be unconscious? While the answer to this question is still open to debate, more and more empirical evidence suggests that humans might have unconscious emotions in certain conditions. In order to provide a holistic theoretical overview focusing on unconscious emotions, a review of contemporary theories, models, and measures of emotion is provided. This review illustrates an integrated overview of research into unconscious emotions and also points out potential challenges and promising approaches that could inform future research on unconscious emotions in media content.

Keywords

Media content • Unconscious emotion • Emotion theory • Emotion model • Emotion measurement • Review

Introduction

As part of human nature, people are always interested in exploring the world where they live in. Why does the sun come up from the east? Why are there four seasons? This world is mysterious, not only for our ancestors but also for people in the modern society. Before human beings were capable of justifying a hypothesis with objective evidence (i.e., the scientific revolution), our ancestors formed an imaginative worldview about how this world is built. Myths were thus created to serve the function of media. The relationship between myth and science is a subject as old as that of myth and science themselves. The modern view toward myths varies chronologically by the centuries. In the nineteenth century, myth and science were commonly considered to be opposite and incompatible. In the extreme, people at that time thought of defeating myths by scientific approaches as a sign of a modern society. It was believed that because of the advances of science, this world is becoming less mysterious, so modern people have to abandon myth. On the contrary, in the twentieth century, myth and science were usually taken to be compatible, and therefore myths are once again embraced by the modern society. There is no doubt that the quality of material life substantially relies on science and technology, but it needs to be noted that myths have been enriching the mental life of modern people in many ways.

Myth is the oldest form of narrative in human history. Since there are no historical records or scientific proof, it is difficult to justify if these mythical narratives are true facts or just merely imaginary stories made up by ancient people. Nevertheless, myths are still influencing modern society and still fascinating to people nowadays. Myths, as a primeval form of living reality, have been woven into people's lives in many different ways. It is usually linked with ritual. Together, they were used to regulate how ancient people were to behave toward nature and construct their basic view of how this world was made. Myths and rituals divide an individual human life into three phases: childhood, adolescence, and retirement. Each of these phases is characterized by corresponding motive complexes: love, power, and death (Rauterberg 2011). In general, there are two ways of interpreting myths: literal and symbolic (Dundes 1988). Literal interpretations look into factual or historical bases for myths, while symbolic interpretation prefers to regard myths as a code requiring some deciphering. In the early nineteenth century, myths were treated as nothing different from fables, inventions, and fictions. Thus, rationalists treated myths as delusions and pure imaginations. From a literal perspective, myths would be merely an unrealistic and unscientific version of history. The greatest value gained from myths would probably be anthropology, rituals, and religions, which mostly aim at finding a reasonable explanation or rational speculation for the reality of ancestral societies. Interestingly, it was found that ancient people tend to present symbols as an instance of association with reality, instead of looking for rational causes (Kirk 1975). It appears that myths were not meant to give a rational explanation toward the reality but to provide a symbolic representation that shows the natural tendency of how ancient people interpret the reality that they lived in.

The nature of myth is different from fictions that people read today. More than just storytelling, some scholars consider a myth as a way of making sense of the physical world (May 1991). A myth is a sacred narrative usually explaining how the world and humankind came to be in its present form. One of the foremost functions of myth is to establish models of behavior. Through its myth, a healthy society gives its people relief from guilt and anxiety and encourages humans to discover true values in their lives. Rather than merely imaginary stories, myth is a narrative resurrection of a primeval reality. "It is a living reality, believed to have once happened in primeval times, and continuing ever since to influence the world and human destinies" (Malinowski 1954). Indeed, it is surprising that even though the content of myths seems unrealistic and irrational, myths are still manifested in modern media content, such as artwork, music, and movies. Numerous ancient myths are being reedited as a modern version and remade into various forms of digital media. Although the content of the myths might change overtime, the skeleton of the stories still remains intact. The human society grows with myth, while myth manifests the thinking and behavioral pattern of the people within the society. This circular relationship has been continuously working along with human history, and, nowadays, many traces of myth still can be seen in the modern society.

The Unconscious

Some cultural rituals and customs can be traced back to historical events, but others could only count on myths, fairy tales, and legends. Although these fictional narratives cannot be proved to be true, they lie at the heart of people and form the knowledge that guides people how to see things and act in the society. Fictional narratives have been an important medium for the growth of culture and society. They provide a deep and immersive simulative experience of social interactions. Furthermore, it is believed that people's thoughts and behaviors are deeply and implicitly influenced by their social and cultural context. In recent years, psychologists start to investigate how people could adapt themselves into a given social situation. The results suggest that this adaption not necessarily be a conscious action but can be unconscious (Bargh and Morsella 2008). Indeed, many of our decisions and behaviors are irrational and driven by the unconscious. While most people think irrationality is harmful for decision-making, empirical studies have revealed the merits of utilizing the unconscious to achieve better results in a complex decision-making task (Dijksterhuis 2004). Sometimes people make decisions based on preferences instead of rational reasoning. In fact, it is an inborn capability of humans to look for better conditions in terms of physical and mental well-being. While a cognitive task is beyond the capability of rational thinking, the intuitive, unconscious, irrational thinking might take over and solve the problem with "gut feelings" (Gigerenzer 2007).

The interests in human unconscious were first revealed by Freud (1922). After being quiet for a long time, research on the unconscious regains the attention of the science community in recent years. Some psychologists strictly hold a conscious-centric model of the human mind, considering the unconscious as a low-level mental processing system that handles subliminal-strength stimulation from the environment, while the conscious processes still play the primary role in causing human judgment and behavior (Bargh and Morsella 2008). Nevertheless, other scientists uphold the priority of the unconscious and suggest that the unconscious processing is not limited at the low-level perception but can also handle high-level mental activities, such as decision-making, goal setting, attitude, and emotion regulation. The distinction between conscious reasoning and unconscious thinking is significant. According to Kahneman (2003), there are three cognitive systems of the human mind: perception, intuition, and reasoning. Perception and intuition form an experiential system that processes the incoming information in real time and generates corresponding reactions automatically. Only part of the incoming information would pass the filter of attention and enter into the rational system for logical reasoning.

In addition to the cognitive account for the unconscious, psychologists also take into account the impact from the social and cultural context. People who live in a common social context would automatically and implicitly develop collective intentionality – a kind of knowledge that is formed and shared within a group of people (Barrett 2012). This knowledge provides a basic reference for people to understand the external world and thus generate emotion, memory, and other mental content. When being part of a society, people tend to adapt themselves into the given social context, and this adaption might take place outside of conscious awareness. While it seems to be a valid assumption, it is still unclear how this collective intentionality is achieved within a large group of people, e.g., a society. Human communication is complicated and versatile. Besides the literal interpretation toward the use of languages, human beings tend to utilize a higher level of skill to communicate abstract notions. Some sociologists propose that all kinds of human communication are symbolic (Manis and Meltzer 1978). Their fundamental premise is that humans do not directly react to the ontological-existing reality but respond to their understanding of this reality. It is a natural tendency for humans to express abstract notions in an associative and symbolic manner. Symbolic meaning appears to be an important layer that lies between the human mind and the physical world. The meaning-making process is automatic, spontaneous, and irreversible. Emotion, as one of the essential psychological phenomena, allows people to respond to the meaning they perceived from the world and thus shapes their behaviors.

Media and Entertainment

While the pure science looks for precision and rationality, the application domain of entertainment aims to create emotionally intense, rich-in-meaning media content. Entertainment is part of people's daily lives. The history of entertainment can be traced back to the ancient times. It is not only because humans need to rest from work but also due to the desire for enriching the mental and spiritual life. The essence of entertainment is versatile. Rather than a means for merely having fun, entertainment plays the role in maintaining mental sustainability and facilitating social transformation. It can also be a kind of media that delivers cultural experiences at an unconscious level. With the advances of media technology, the modern form of entertainment offers tremendous potential for making positive effects on human behaviors in various ways. In this regard, it is needed to consider how entertainment gradually changes with the growth of human's society and how future entertainment technologies might in turn make impact to people's lives.

Since the 1980s, the emergence of computational technology and the advances of multimedia communication have considerably changed people's lives. Today, multimedia systems have widely spread and permeated to people's daily lives. Mobile technology and Internet enhance the accessibility of media content, and communication becomes richer in terms of content. Entertainment, as an important category of media content, becomes one of the driving forces that thrust the development of new types of multimedia systems. Media technology has broadened the bandwidth of communication for transmitting entertaining content to people by using multimodal channels. However, the content of the entertainment is equally important and worth studying. Myths, as the media content in ancient times, were spread through word of mouth and symbolically represented as artwork. These mythical stories were not lost along with the growth of human society and still are being manifested in various media forms, such as video games and movies. This curious phenomenon has demonstrated the universal human capacity to classify, to codify, and to communicate their experiences symbolically in different time scales.

There are two kinds of interpretations toward media content: rational and emotional. With regard to functional purposes, communication needs to be precise and logical, so the media content has to be descriptive and concrete in order to achieve a clear understanding of the enclosed message. On the contrary, media content for entertaining purposes particularly emphasizes the emotional aspect of the communication. While the rational interpretation of media content looks for simplicity, emotional interpretation of media content is abstract, dynamic, symbolic, and experiential. This categorization also resonates with the distinction between the two cognitive systems of the human mind. The emotional aspect of communication encompasses a considerable amount of information that would be handled by higher-dimension mental process: the unconscious. With these two interpretations, a pure, holistic experience thus can be delivered in a more comprehensive manner – a rich combination of communication channels that allows conscious and unconscious information flow freely.

Emotions in Media Content

In this chapter, it is intended to explore the relationship between two main subjects: emotion and media content. Media can be referred to any means of information communication. While the form of the media is progressing with the advances of technology, it is needed to concentrate on how media content becomes meaningful to the audience. Meaning, on the other hand, is becoming an important topic in psychological research on emotion. In the early years, it was assumed that emotion was directly caused by external stimuli in the environment, such as natural threats (snakes or earthquakes) cause the feeling of fear. Today, psychologists have developed more comprehensive explanations for the activation of emotion. Mainstream psychological theories suggest that the meaning of the given situation plays an important role in activating emotions (Arnold 1960). For example, natural threats are interpreted as a cause of death, and this interpretation thus activates the feeling of fear. In this regard, the symbolic meaning of media content plays a key role in bridging the gap between emotion and media content.

Narratives have been the primary media content for transmission of emotional experience since early times of the mankind. In essence, narratives are a kind of virtual reality, which pulls the reader into a fictional world. As it is mentioned earlier, fictional narratives usually reflect the worldview of the modern time. Different from history, which is meant for documenting the true fact, the function of narratives is to render the social atmosphere in an implicit manner and deliver rich emotional experiences along with the interaction among the characters. Most people tend to enjoy the mediated emotional experience rather than rationally analyze the logic of the content. Just as children love often-told bedtime stories, sometimes adults enjoy watching classic movies over and over again and never tire of them. Narratives can induce profoundly emotional experience. However, very few studies have addressed this topic within psychology probably due to the misconception that narratives are mere made-up illusions (Oatley 1999). Recent years have seen some researchers advocating the importance of narratives in psychology. They suggest that narratives operate as a means of simulating social and emotional experience (Oatley 1999; Mar and Oatley 2008). The evoked emotion may even influence the viewer's daily life after viewing. Although most of these studies particularly focus on literacy narratives, they still shed some light on narratives in various media types, such as paintings, comics, theater plays, and movies. If media are considered as an instrument for transmitting emotional experience, narratives can thus be defined as a design pattern of media content. In this regard, exploring the relationship between the pattern of narratives and its corresponding emotional responses appears to be a promising research direction.

Emotion is an important aspect of media experience, and it has been of great interest to the design community. Studies on emotional aspect of design usually adopt psychological theories of emotion as their theoretical basis. However, the mainstream psychological studies solely focus on utilitarian emotions (Scherer 2005). These types of emotions can be considered utilitarian in the sense of facilitating our adaptation to events that have important consequences for our well-being, such as fear triggers fleeing from danger. However, emotions in media content do not seem to have survival values for maintaining physical sustainability. This view appears to be limited and cannot explain emotional experiences in the context of media design, i.e., why people enjoy watching movies while movies do not bring any survival values? While it has been well researched that emotion has biological and social functions, other aspects of emotion remain mysterious and intriguing. Moreover, the question of whether emotions can be unconscious draws great attention in recent years. Psychologists hold various theoretical positions and use different measures to approach this hypothetical phenomenon. According to some accounts, emotion cannot, by definition, be unconscious (Izard 2009), while other accounts hold that emotions and other psychological phenomena can be unconscious (Kihlstrom 2008). For further justification of unconscious emotion, scientists are seeking experimental evidence to support their belief that emotion may be unconscious, at least under certain conditions. Today, the issue of unconscious emotion is still open to debate; it is difficult to judge whether the phenomenon exists or is just an anecdote. Notwithstanding the uncertainty surrounding them, the early findings have encouraged researchers to reassess existing conceptual and operational definitions of emotion and look for ways to accommodate the new evidence within their theoretical framework.

Theories, Models, and Measurement

Putting all emotion-related literature on a continuum between theory and practice, the main body of emotion research can be organized into three building blocks: *theories*, *models*, and *measurement*. In most cases, researchers confuse *theories* with *models* and refer to both of them as *theories*, but they take entirely different epistemological positions. Theories of emotion treat emotion as a natural phenomenon to be explained. Thus, they elaborate the ontology and functionality of emotions and focus on its conceptual definition. On the other hand, models of emotion are built to demonstrate the taxonomy of emotion by conceptualizing all kinds of emotions in a meaningful way. These models of emotion decompose emotions into meaningful and manipulatable parameters so that they can be used for practical purposes, such as affective computing. In order to provide evidence relevant to hypotheses derived from specific theories and models, various measures of emotion have been developed. Operationalizing the concept of emotion in this way has allowed researchers to move beyond biology and philosophy and assert the credibility of psychology as a science. However, the measurement of emotion is challenging because it is difficult to ensure that "the measure does index the emotion it is assumed to be related to." The validity of measurement is often a concern in psychology. Biased measurement methods significantly influence the results of experiments, and the interpretation of data can vary depending on the theoretical approach adopted.

When considering a new hypothesis such as the existence of unconscious emotion, researchers need to pay attention to the epistemic tradition within which theory, model, and measures are applied in order to decide whether the evidence is sufficient to support or reject the hypothesis. It is required to include at least two of the building blocks to fulfill the construct validity of research. Current empirical evidence for the existence of unconscious emotion is derived from research based on different assumptions and is insufficient to answer all the objections that have been raised. In order to understand clearly how the empirical data should be interpreted in terms of the various current theoretical perspectives, it is necessary to review carefully the various theoretical arguments put forward by proponents of the concept of unconscious emotion.

A Unified Structure

In order to review various theories of emotion, it is necessary to develop a unified descriptive structure to represent the process of emotion activation (see Fig. 1). The basic process by which an emotion emerges can be broken down into three stages: elicitation, emotional state, and reaction (Rottenberg et al. 2007). Usually the term "emotion" refers only to the second stage, which is the emotional state of the individual at any given moment. However, the cause of the emotion (i.e., elicitation) and how this emotion influences the individual (i.e., the individual's reaction

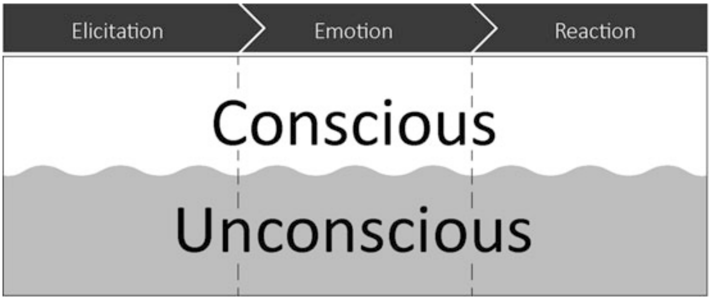


Fig. 1 The conceptual structure of emotion on which the review is based

to the emotion) are equally important; if these components of emotion are neglected, the discussion becomes purely philosophical and ceases to be grounded in empirical data. The elicitation of emotion can take almost any form; any information that comes from the external world may elicit an emotion. The concept is not restricted to specific objects or events that experimental psychologists usually refer to as stimuli, but may include social or cultural contexts. Similarly, when it is referred to the reaction to an emotion, the concept covers not only observable physical activities (e.g., behaviors, facial expressions, and physiological responses) but also mental activities (e.g., thoughts, preferences, attitudes, and decision-making). This three-stage structuring of emotion as a process has been used to ensure the consistency and clarity of the rest of the review.

To integrate unconscious emotion into this structure, a line must be drawn at each stage to indicate the border of the conscious; the area beneath this border then represents the unconscious. In order to define the border between the conscious and the unconscious, it is necessary to a more complex question: what does "unconscious" mean? Cognitive psychologists usually equate "the unconscious" with subliminal information processing (Kihlstrom et al. 2000). By this definition, unconscious processing deals only with low-intensity, unsophisticated mental activities below the threshold of perception, such as subliminal perception. Social psychologists use a broader definition of the unconscious mind (Bargh and Morsella 2008). They consider the unconscious to be those mental processes that the individual cannot access through introspection (Wilson 2003). In order to accommodate most of the theories, the later definition is used as the criterion of conscious emotional process. The way in which this criterion defines the border between the conscious and unconscious at each stage is as follows: At the first stage, emotion elicitation, the border is defined in terms of the degree to which the individual is able introspectively to recognize stimulus information. At the second stage, the momentary emotional state, the differentiation between the conscious and the unconscious depends on whether the individual can recognize introspectively the emotion that is being experienced. At the third stage, emotional reaction, the border is defined by the degree to which the individual is aware introspectively of the changes in his or her own behavior, thought, or bodily sensations. It should be noted that these three stages can be either conscious or unconscious, respectively, to constitute the whole process, i.e., a conscious activity at one stage may be related to unconscious activity at another stage and vice versa. For example, a stimulus that is not consciously recognized might elicit a conscious emotion; similarly, an unconscious emotion may produce a conscious reaction. In the remaining sections of this review, the mental content at each stage will be specified in terms of this structure according to the assumptions of the theory under discussion.

Theories of Emotion

Many theories of emotion suggest that an emotion emerges when an individual makes sense of something as being in some way related to or caused by the situation (Gendron and Barrett 2009). Rather than thinking of emotion as a *state*, it has been suggested that emotion should be considered as a process of conceptualizing the world (Barrett 2012). While the human mind has a variety of cognitive functions such perception, learning, and memory, there are higher mental mechanism operating beyond cognition, such as emotion (Kihlstrom 2008). Most psychologists agree that the activation of emotion is effortless and automatic, not controlled (Kihlstrom 2008; Barrett 2012). However, the key question is: what is the source of emotion activation? Contemporary theories of emotion can be categorized into four different perspectives: evolution based, body based, situation based, and mind based.

Evolution-Based Theories

The functional-evolutionary perspective defined emotion in a teleological fashion by stating that emotions are evolved phenomena with important survival functions that have been selected for because they have solved certain problems that humans have faced as a species. As results of natural selection, to some degree all humans should share a common pattern of emotional expression and the innate capability to interpret each other's emotions, especially through facial expression. This branch of theories holds a biological view and asserts that emotions are *natural kinds* (Barrett 2006a). It assumes that emotions are universally manifested by humans, selected for over the course of evolution and biologically primitive; the fundamental elements of emotional life are thus conceived as an ancestral heritage. In the view of the modern evolutionary theorists, basic emotions are dedicated neural programs or circuits that are hardwired into the brain (Lang 2010); this implies that all emotional reactions are survival-related, preprogrammed physiological reactions and physical behaviors.

Putting this theoretical view into the unified structure, the sources of emotion activation appear to be utilitarian functions with survival value (see Fig. 2). The brain circuits are responsible for making sense of the external situation and generate

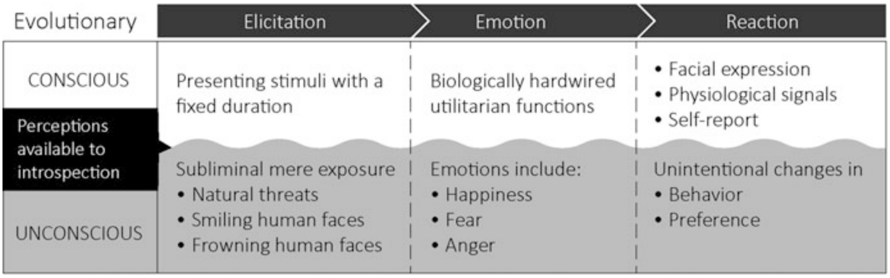


Fig. 2 Evolution-based perspective

emotions accordingly to drive survival mechanisms. For example, external threats trigger the relevant brain circuits and thus generate the emotions of fear or anger, which in turn produces an impulse to flee or fight. The direct causation of behavior is the primary function of emotion. Since emotions are defined as natural kinds and the brain circuits are hardwired, emotion activation itself is more like an instinct. Such instincts could be accurately defined as fixed action tendencies corresponding to utilitarian functions, meaning that most emotional reactions are observable as spontaneous, homologous facial expressions and bodily reactions. Although the contemporary evolutionary perspective considers emotions to be related to social context, facial expressions were considered as a natural adaptation to the social environment, and social events were conceived as survival challenges from an evolutionary perspective (Öhman 1986). In evolutionary emotion research, a common approach to emotion elicitation is the presentation of affective pictures, specifically images of facial expressions and natural threats. As for measuring emotional reactions, the evolutionary perspective holds that all emotions can be manifested through facial expression and behaviors as it assumes that emotions are preprogrammed and universally recognized.

From the evolution-based perspective, the activation of basic emotional systems is to some extent independent of conscious awareness of the current situation. With regard to the unified structure, since emotion activation is considered biologically coded and hardwired in the brain circuits, the border between conscious emotion and unconscious emotions can only be *the threshold of perception*: conscious emotions are emotions triggered by stimuli that are perceived consciously, while unconscious emotions are emotions induced by stimuli that are exposed below the threshold of perception. This has led to considerable research into subliminal processing of affective stimuli. Researchers who take an evolutionary perspective argue that some mental activities can occur without being available for introspection, such as implicit perception and implicit cognition (Kihlstrom 2009). This claim was substantially supported and implicit processing characterized by empirical studies on the *mere exposure effect* or *subliminally priming effect* (Zemack-Rugar et al. 2007; Bornemann et al. 2012). Emotional reactions related to subliminal priming are not measured directly from facial expression or physical behaviors as in studies on conscious emotions; rather they are assessed from indirect measures such as changes in preferences and changes in behavior when performing a specific task which is held to be significantly related to affect (Winkielman et al. 2005).

Body-Based Theories

Contrary to earlier psychologists who believed that emotions were mental events that caused physical changes in the body, some researchers asserted that emotion emerges immediately *after* one perceives an internal physical sensation that is triggered automatically and adaptively by the external environment (James 1884). This counterintuitive proposition led to a new approach to research on emotion but also generated considerable theoretical debate specifically about the temporal relationship between the physical sensation and the emotional experience. Yet, it is still valid to assume that emotion emerges through making the internal physical sensation meaningful.

The theories of psychological constructionists consider emotion to be an act of sensemaking directed by the internal sensory or affective state (Gendron and Barrett 2009). In contrast to the basic emotion theorists, psychological constructionists claim that emotions are not biologically given but are constructed through the process of making sense of physical changes arise in the body (Barrett 2006b). The basis of the sensemaking process is also not a hardwired instinct; rather it is socially shared conceptual knowledge (Barrett 2012). The psychological constructionist view considers that emotions are socially related but *cannot* be reduced to mere social situations, as it holds that the presence of an individual mind is necessary to the sensemaking process. Sensemaking is the core concept of psychological constructionist approach to emotion, as Barrett states: "Human brains categorize continuously, effortlessly, and relentlessly...Via the process of categorization, the brain transforms only *some sensory stimulation* into information... To categorize something is to render it meaningful" (Barrett 2009).

Damasio (1996) proposed a similar hypothesis, namely, that conscious emotions are actually feelings – emotions can be consciously *felt* through perception of the bodily sensation. Similar to the basic emotion theory, he defined emotions as biological life-regulation phenomena that are reflections of cognitive actions in human bodies toward a given situation. According to Damasio, what people perceive as emotional *feelings* are composite perceptions of these biological phenomena; in other words, feelings are images of emotions, rather than emotions themselves. Feelings and emotions are therefore not always identical, and not all emotions can be perceived as feelings. Some empirical evidence on significant changes in physiological data has supported this hypothesis. These data occur when an individual is performing certain decision-making tasks. Changes in an individual's physiological signals – *somatic markers* is Damasio's term – indicate when a risky choice is being made although the individual is unaware of them (Bechara et al. 2005). He has put forward a promising theoretical perspective that is

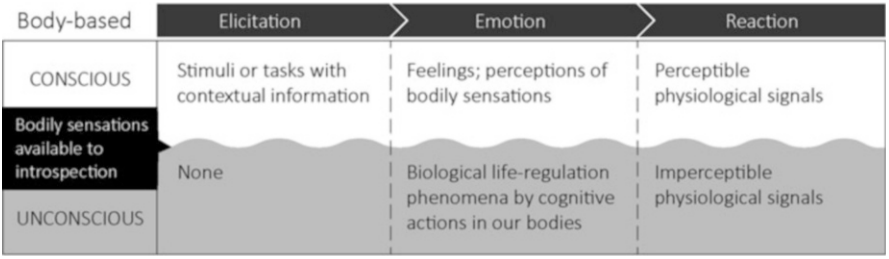


Fig. 3 Body-based perspective

complementary to the psychological constructionist view of emotion in that the bodily sensation is the target of the sensemaking process.

Based on the above theories, being able to make sense of physical sensations in introspection appears to be the primary difference between conscious and unconscious emotions. As the psychological constructionists suggest, the sensemaking process relies on a connection to the situation in which the emotion emerges. In other words, when a mismatch or loss of connection occurs during the sensemaking process, the emotion is unavailable for introspection and thus becomes unconscious. Studies of emotion elicitation from the body-based perspective usually use stimuli with rich contextual information or simulated scenarios, such as images, emotion words, and simulated games, but there have been no investigations using stimuli that are unavailable for introspection (see Fig. 3). To track the emotional reaction, measures of physiological signals and brain states are frequently used, because physiological responses are at the heart of the body-based perspective. Some of the physiological changes in the autonomic nervous system (ANS) are available for attentive introspection, such as heart rate and respiration rate, but other changes such as changes in brain states cannot be detected without biometric equipment.

Situation-Based Theories

The contemporary situation-based perspective is embodied in the *appraisal theory* proposed by Arnold (1960). Arnold used the term "appraisal" to describe an essential cognitive mechanism: a given situation is evaluated, and the results of this evaluation then cause emotion. Frijda (1988) described appraisals as assessments of situational meaning: "Input some event with its particular meaning; out comes an emotion of a particular kind." Although appraisals are necessarily intentional, their meaning analyses are automatic cognitive mechanisms specifically dedicated to emotion, and they need not be available to awareness. Social constructionists also take a situation-based perspective but one that differs from that of the cognitive theories of emotion. Rather than looking inwards, social constructionists treat emotions as cultural products whose meaning is based on learned social rules

(Averill 1980). Social constructionists emphasize the social functions of emotion and how social contexts cause one's emotional state. For them, cultures provide the content of the appraisals that form the basis of emotions, which are constructed within a culture to serve specific social purposes. Unlike the psychological constructionists who hold that emotions are socially related but not socially specific, social constructionists hold that culture plays a central role in the organization of emotions; they consider that emotions can be reduced to the social situation in which they occur (Gendron and Barrett 2009). Rather than considering emotion as an entity, social constructionists define emotion as a dynamic process that emerges while an individual is interacting with others within a given social context. They believe that emotions have a social function and that to understand emotions, one has to consider what a specific emotion accomplishes socially (Averill 1980). This is the fundamental distinction between these two constructionist views: for social constructionists, the target of the sensemaking process is the external social situation, whereas psychological constructionists regard bodily sensations as the target of the sensemaking process.

Taking a similar approach to appraisal theory, some social psychologists argued that emotion might also be inaccessible to the conscious. The concept of the *adaptive unconscious* (Wilson 2003) was put forward. These authors regarded the sensemaking process as an inferential or self-attribution process by which an individual makes sense of a given situation in terms of his or her bodily sensations. Making inferences about the world is not always rational and intellectual; it may be intuitive, effortless, and adaptive (Gigerenzer 2007). This inferential process can take place either consciously or unconsciously and cause conscious or unconscious emotion accordingly. Surprisingly, some empirical results have revealed that there can be differences between conscious and unconscious emotions related to a given situation (Schachter and Singer 1962). "The conscious system is quite sensitive to personal and cultural prescriptions about how one is supposed to feel… People might assume that their feelings conform to these prescriptions and fail to notice instances in which they do not" (Wilson 2004). Although physical sensations are an important reference for the sensemaking process, the target of it appears to be the situation that the individual finds him- or herself.

According to the situation-based theories discussed above, the distinction between conscious and unconscious emotions is based on whether an individual is able consciously to identify the connection between his or her emotion and the given situation (see Fig. 4). In terms of the unified structure in this review, the emotion elicitation process relies on the external situation, more specifically a social circumstance. As for the emotional reaction, like the evolutionary-based research on emotions, research in the situation-based tradition usually utilizes self-reports of preferences or decision-making tasks to capture conscious emotional responses, while physical behaviors are used as an index of unconscious emotion. This is probably because researchers with a situation-based perspective tend to use connections with the external world as the criterion for distinguishing conscious and unconscious emotion.

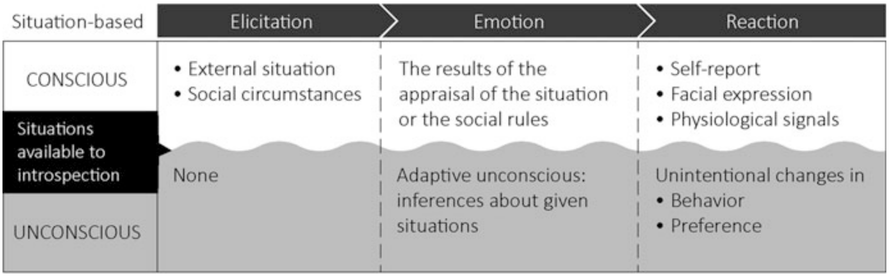


Fig. 4 Situation-based perspective

Mind-Based Theories

Social psychologists place focus on groups and tend to look externally, to the situation for explanation, whereas the psychoanalytic tradition focuses on individuals and uses internal data to explain inward subjective experience (Chancer 2013). The psychoanalytic tradition stands in opposition to mainstream psychological research; rather than trying to deconstruct emotions, it tends to approach human mind as a whole. In psychoanalytic theories, the unconscious mind is the primary source of psychological content, which is deeper than the conscious mind. Rather than looking for connections to the body or the environment, psychoanalytic theories seek for psychological meanings within the deeper levels of the human mind. The importance of the unconscious was first advocated by Freud (1922); in Freud's original model, the metaphor of the dark cave was applied to the human mind. He argued that the unconscious was a complex dynamic that suppressed primal desires, and these desires can only be represented as anxiety. Since mental content (memories, thoughts, and emotions) in the unconscious was assumed to be inaccessible to conscious introspection, Freud likened psychoanalysis to an archeological dig in which the psychological content of the unconscious would be uncovered (Wilson 2004). Freud assumed that emotions were the outlets for instinctual drives and that emotions emerged when the expression of these instinctual drives was blocked (Freud 1922). Emotions can be felt at the surface of the conscious; however, the cause of these emotions remains buried deep in the unconscious and unreachable for introspection.

The Freudian concept of the unconscious, although it is well known outside psychological science, has not been supported by empirical studies, but it continues to provide inspiration for research in the psychoanalytic tradition. In addition to Freud's concept of the individual unconscious, Jung (1959) proposed an unconscious specifically focused on the psychological content that was common to all human minds. Jung's interpretation of observational data from his patients was that there is a universal pattern to the psychological content of the mind that can only be expressed through symbolic content. Jung argued further that the universal pattern in the unconscious mind contained some contents and modes of behavior that are

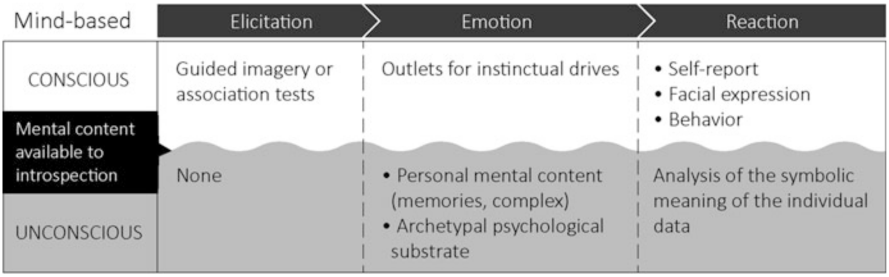


Fig. 5 Mind-based perspective

identical in all human beings and thus constituted the common psychological substrate of a universal nature that is present in every individual. All human beings, sharing essentially the same biological equipment (e.g., the brain and the central nervous system), would show a tendency to perceive common meanings embodied in symbolic content at an unconscious level. Jung's theory has been used for analyzing media content, especially movies (Chang et al. 2013). Although psychoanalysis has often been criticized for the unfalsifiability of its hypotheses because it relies largely on the interpretation of the symbolic meaning of individual data, there has recently been new interest in psychoanalytic theories in neuroscience; surprisingly, new evidence has been found in neuroscience to support some psychoanalytic hypotheses (Shevrin et al. 2013).

Since the psychological content of the unconscious mind cannot be directly accessed or described, the psychoanalytic tradition usually utilizes meaning analysis of its symbolic content expressed in different ways by individuals (e.g., association tests and guided imagery). Nevertheless, it is clear that the source of emotion activation is psychological content within the mind (see Fig. 5). The key distinction between conscious and unconscious emotion is whether the individual can make sense of his or her psychological content (e.g., thoughts and memories) at a conscious level. Conscious emotions are directly available for introspection because individuals can recognize their psychological content, whereas unconscious emotion can only be expressed indirectly through symbolic content that seems irrelevant or irrational but is implicitly associated with previous life events.

Instead of eliciting specific emotions by providing specific situations or preselected stimuli, the psychoanalytic tradition applies a different approach, using techniques that allow subjects to relax and respond freely to the guiding cues or symbolic content in a task. This is probably because many psychoanalytic theories are developed in clinical and therapeutic practice. The psychoanalytic tradition uses the analysis of the symbolic meanings of the data provided by the subject to infer emotional reactions. These data usually are not directly informative about emotions but are a rich source of information that is interpreted by the facilitator on the basis of his or her experience. Measuring emotion in psychoanalytic research is challenging by its nature because the cause of certain emotions is related to personal life events and difficult to justify.

Models of Emotion

The above four perspectives represent four different metaphysical hypotheses about the nature of emotion. This section provides a review of models of emotion that are widely accepted and often used in other fields of study. Models of emotion are generalized from emotion theories to allow all kinds of emotional qualities to be classified and conceptualized in a more concrete way. Therefore, models of emotion are considered as *direct* representations, in that they have been developed as a kind of operational definition that can be applied directly for practical uses, for example, emotion recognition for video game play.

Discrete and Dimensional Models of Emotion

Psychological models of emotion are mainly based on two views of emotion. The first view holds that emotions are discrete and fundamentally distinct constructs, while the alternative view is that emotions can be represented dimensionally. Some researchers have combined these two perspectives to provide an integrated view (Hamann 2012). It appears that these two models both have advantages and disadvantages in terms of context of use in research.

The discrete model of emotion proposes that there are a limited number of distinct types of emotion, each with characteristic properties (Barrett et al. 2007). This model was originally inspired by one of the evolution-based perspective theories – the basic emotion theory. In accordance with the assumption of this theory, the discrete representation of emotion encompasses a limited number of basic emotions (i.e., happiness, sadness, anger, fear, and disgust) that are universal, biologically inherited, and unique so that each class of discrete emotion has a distinctive physiological and neural profile. Other emotions are either conceptualized as blends of the basic emotions (e.g., contempt is a blend of anger and disgust) or given a different status (e.g., shame is a complex social emotion). The discrete emotion model is mainly characterized by its reliance on linguistic forms. Emotional words such as “anger” are universally understandable references to the corresponding emotional qualities. This aspect of the model has drawn criticism from other researchers for the way in which it confuses the nature of emotion and global terms for labeling emotions (Barrett 2006b). It is also claimed that this model does not make clear predictions on the eliciting conditions for discrete emotions (e.g., the behaviors associated with fear can range from freezing to vigilance to flight, instead of a one-on-one relationship), which has implications for its predictive validity. Moreover, the defining criteria for basic and nonbasic emotions are also unclear.

Unlike the taxonomic approach of the discrete emotion model, the dimensional model of emotion is based on the idea that emotions can be described by several independent and meaningful dimensions (Russell 1979). The current version of dimensional model, known as the circumplex model, makes two primary dimensions: "unpleasant–pleasant" (valence) and "activation–deactivation" (arousal) as the *core affect*. The exact nature of the dimensions and the existence of a circumplex distribution of central emotional states have been debated. Core affect is defined as "the most elementary consciously accessible affective feelings (and their neurophysiological counterparts) that need not be directed at anything" (Russell and Barrett 1999); different prototypical emotional episodes or distinct emotions induced by specific objects can therefore be represented by a combination of the two dimensions. Some researchers have proposed the inclusion of a third, less prominent dimension "in control–dominated" (dominance) (Lang et al. 2008) to distinguish emotions that are in almost the same position in two-dimensional affective space, e.g., anger and fear. There are a number of drawbacks to dimensional theories, such as the absence of attempts to predict theoretically the determinants of differences between emotions and the lack of an explanatory mechanism that predicts response patterning. Moreover, not all researchers agree on the number and nature of dimensions needed in a framework for classifying emotion.

Building Models with Physiological Signals

Physiological signals are important indirect representations of emotion because they are tightly bound to emotional reactions. Observations of physiological activations in the autonomous nervous system enable researchers to identify and label corresponding psychological states. Psychophysiological instruments have the advantage of being language independent and can be used with individuals from different cultural backgrounds, and individuals do not need to be interrupted with questions during measurement. It has also been suggested that physiological observations could be used to identify unconscious emotional states that are inaccessible through introspective reports (Fairclough 2009). Recent years have seen growing trends toward the application of machine learning techniques for training predictive models of classification based on data obtained from physiological signals. These predictive models can also be considered as a kind of representational model of the emotional responses induced by given classes of affective stimuli, i.e., emotion profiles. However, unlike the discrete model and the dimensional model, which both have meaningful classification rules, a predictive model is a "black box" with parameters that it is difficult to relate to psychological concepts (Fairclough 2009). The classification rate of the predictive model only reflects the homogeneity of the stimuli in each category and the physiological heterogeneity of the groups of responses. Although a considerable body of research has attempted to relate physiological signals to meaningful classification rules such as those used by the discrete and dimensional models, the results obtained of empirical studies still lack consistency, and the variation in results is difficult to explain (Kreibig 2010). What researchers may find more psychologically meaningful is the classification of the affective stimuli that serve to elicit emotion in the unified structure. Many researchers training predictive models strive to achieve a higher recognition rate; developing a well-classified collection of affective stimuli is equally important but has often been overlooked.

The discrete and the dimensional models have probably gained widespread acceptance because they come close to the folk psychological understanding of emotion, the way in which individuals consciously recognize their own emotional states. However, these two models do not take into account the possibility that emotion may be unconscious. Even if unconscious emotions were to exist, it is unclear that one could simply add one more dimension termed "conscious–unconscious" to the existing dimensional models as it has not yet been confirmed that unconscious emotions and conscious emotions are homologous (e.g., do unconscious joy and conscious joy have the same quality of joy?). Researchers can use physiological signals as the basis for an alternative model of emotion because the modeling algorithms do not have psychological meanings. The model derived from physiological signals can only be interpreted on the basis of the psychological meaning of the classification of the affective stimuli used for emotion elicitation.

Measures of Emotion

The most common method of measuring emotion in empirical studies is self-report techniques: subjects report their emotional states using introspection. In the case of conscious emotion, most researchers have considered introspective judgments about emotional feelings to be fundamentally valid or correct because it is assumed that human beings are capable of conceptualizing their own emotions in a universally recognized way. Although it is recognized that there are some cultural differences, the expression and interpretation of emotions is considered to be universal at a certain level. Empirical studies have also indicated that self-reports of emotion are likely to be more reliable than other measures although they are limited in some conditions. For example, some people may be unaware of or incapable of reporting their emotional feelings. The Self-Assessment Manikin (SAM) was based on the dimensional model; it was developed by mapping three meaningful dimensions – valence, arousal, and dominance – to Likert scales (Bradley and Lang 1994). Although some researchers have raised concerns about the validity of self-reports of emotions, the SAM scale has been proved effective and is widely used in several research domains. However, self-report techniques are less effective for capturing unconscious emotions because unconscious emotions are unavailable for introspection. To address this, cognitive science has seen an increase in the use of indirect indicators to assess a larger range of human emotions.

As discussed in the section of models of emotion, physiological measures are the most commonly used indirect measures of emotion due to the fact that certain physiological changes coincide with the appearance of specific emotions. Biosensors have proved a powerful tool and have allowed researchers to monitor almost exhaustively the physiological activities of human body in a convenient way. Physiological measurement offers several advantages: emotional experiences can be monitored continuously, physiological signs are not language – or memory – dependent, and the measurement process does not interfere with the emotional experience. Using physiological signals also allows researchers to utilize computational algorithms to build models of psychological events (Fairclough 2009). A computational model obtained using this approach can be used directly for emotion recognition in practical applications (i.e., make predictions about emotion recognition). Advances in computational algorithms have enabled researchers to keep improving the quality of models based on the same data. Although psychophysiological measurements seem to offer several advantages, some limitations should be noted. Firstly, physiological measurements relate to a broad range of cognitive and affective activities as well as emotional reactions. The extraction of useful information from physiological data is not as intuitive as direct representational methods. Most psychophysiological approaches use supervised classification methods based on introspective judgments about emotional states. In dealing with unconscious emotions, i.e., emotions that are unavailable for introspection, researchers have to consider the risk of falling into epistemological traps when interpreting physiological data (i.e., does the measure index the emotion it is assumed to be related to?). Secondly, previous studies have demonstrated recognition accuracy for various sets of emotions ranging from 30 % to 90 % across different experimental settings (Novak et al. 2012). It appears that physiological responses are influenced not solely by psychological and affective stimuli but by diverse factors such as physical activity, ambient temperature, and individual physiological differences.

Rather than physiological measurements, social psychologists usually look into the changes in behavior, attitude, preference, and decision-making. The social psychological approach has a long history and is widely used in several research domains; it is particularly important in behavioral science. The advantage of this approach is that changes in social psychological measure are intrinsically meaningful, unlike physiological changes. Results are usually self-explanatory and can be interpreted intuitively. The key to this approach is experimental design and inference. Because the social psychological approach focuses specifically on changes, the experimental design must consider the baseline from which changes are to be measured. It is difficult to define a baseline because this leads to a debate on the nature of emotion (e.g., discrete vs. dimensional). Furthermore, because there is not a simple one-to-one causal mapping between behavior and emotion, it is hard to infer the exact emotional cause of a specific behavioral change (e.g., fear and anger both may cause the behavior of fighting). The context and the motivation for the experimental task are also critical factors as they may affect how subjects understand and perform the tasks. This means that data obtained using this approach is usually only valid in the given context and interpretations cannot be generalized.

To obtain data on emotional experiences, the psychoanalytic tradition focuses on the symbolic meaning of the content generated by clients in imagery tasks. Jung (1959) claimed that there is inborn psychological substrate embedded in every human's unconscious mind which cannot be directly accessed or described by the conscious. However, this substrate – which he called *archetypes* – can be manifested through symbolic content in various forms. The concept of archetypes, although not all psychoanalysts accept it, was the basis of the symbolic tradition in psychoanalysis. Within the symbolic tradition, many therapeutic techniques were developed to help clients express their emotional feelings indirectly by guided construction of imagery or association tasks using specific stimulus words, graphics, or objects. For example, the *guided affective imagery* (GAI) technique uses a narrative script to guide the client and a set of imagery targets such as meadows, trees, and houses as cues to enable clients to indirectly express their emotions (Leuner 1969). The psychoanalytic approach differs from mainstream psychological approaches in that it takes a bottom-up approach. Rather than inducing and recognizing predefined emotions such as joy and sadness, psychoanalysts tend to pose a neutral scenario and observe how people generate symbolic content related to it to express their emotions implicitly. This approach requires that the subject be deeply relaxed (e.g., hypnotized) to facilitate the emergence of unconscious emotions. While the subjective nature of this approach is often criticized by scientific positivists, there is some empirical evidence for the therapeutic efficacy of the psychoanalytical approach (Shedler 2010).

An Integrated Overview of Unconscious Emotion

According to the above four theoretical perspectives and various representations of emotion, a theoretical overview of emotion research is provided (see Fig. 6). In general, emotion research encompasses three complementary building blocks: theories, models, and measures of emotion. Although these three building blocks share some common ground, they reflect differences in scientific aims and diverse research questions. These three building blocks form a "triangulation" of emotion research (for more discussions about triangulation, see (Moran-Ellis 2006)). They are mutually constraining yet also support each other. When a theoretical perspective is proposed to explain the underlying mechanisms of emotion, researchers can conduct empirical studies to assess the validity of the theory using various measures. On the other hand, data collected from different measures may also be useful for developing new models and theories. While theoretical perspectives concentrate on the ontology and functionality of emotion, models of emotion are intended exhaustively to describe and

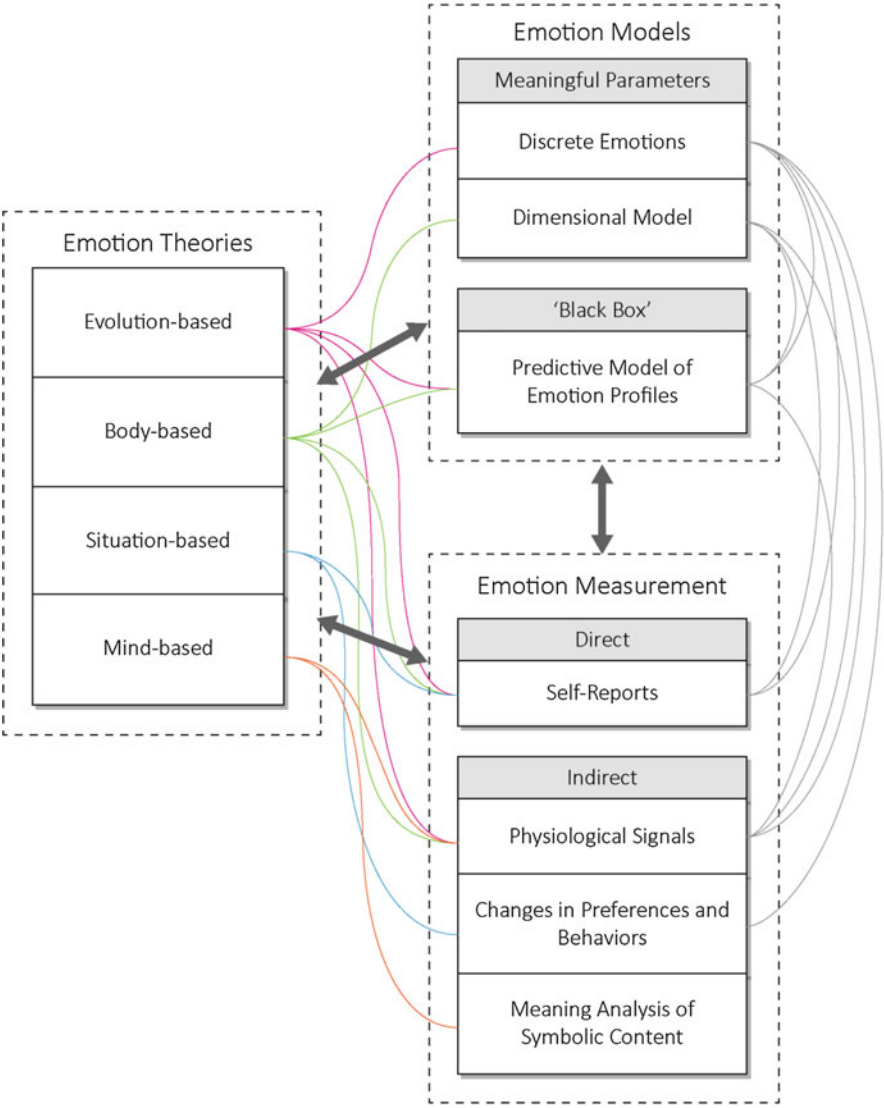


Fig. 6 An integrated overview of emotion research. *Thick arrows* represent relationships among the three building blocks; *thin lines* represent approaches that are often used in mainstream research on emotions

classify all possible emotions. When it comes to unconscious emotion, it is important to reconsider which theoretical perspective, model, and measures can be used to confirm or refute the existence of this phenomenon and interpret experimental results without falling into epistemological traps.

Research into Unconscious Emotions in Media Content

Human experience has been promoted to a new research agenda in recent years (Hassenzahl and Tractinsky 2006). Considering media content as information that flows continuously through the audience's mind, the emotional experiences that arise along with the story are of great interests to researchers in related fields.

Mapping Media Experience with the Storyline

Media content stimulates the process of meaning making and the emergence of emotional experiences. A traditional view toward the relationship between the audience and the media is similar to an input–output mechanical process. From this perspective, emotional experiences are nothing more than the outcome of the viewing experience. In recent years, researchers have made considerable progress moving from passive viewing experiences to a multimodal interaction that takes into account affective states of the audience. Various applications of physiological technology are employed to measure and recognize human emotion, so that digital media systems are empowered to react not only to the audience's conscious behaviors but also to their affective states (Chang et al. 2014a). A nice example of new paradigms for applying affective computing in developing digital media is *enactive media* (Tikka 2010). The theoretical foundation of the idea originates from enactivism in cognitive science, which takes a "groundless" stance that resists the mind–body dualism, and argues that the mind, body, and world are interrelated and interdependent (Varela et al. 1992). Extending enactivism to the cinema experience, cinema is considered as a metaphorical externalization of embodied mind; the audience's mind and the image of the cinema emerge and interplay with and within each other. The audience and the cinema have to be put together into a systematic context that mutually and continuously creates its own meaning.

In the view of enactivism, meaning making is an automated process, and the emotional experiences can be seen as its representation of the interaction between the mind and the story (Chang et al. 2014a). In order to study the relationship between emotion and the media content, the storyline of the media content plays the key role in structuring the whole emotional experience. In the field of mythology, considerable efforts have been devoted to the study of universal storyline across various cultures and religions (Campbell 1973; Propp 1984). This knowledge has been used in screenwriting in today's screenwriting (Field 1984; McKee 1997) and has inspired the making of numerous classic movies, such as the Star War series (Tiffin 1999).

Since the structure of stories is suggested to be universal to a certain level, research into the emotional experiences built upon this storyline is thus a promising direction to proceed. This typical kind of experience was called "archetypal experience," and evidence has suggested that it introduces unique physiological patterns that can be recognized with significance (Ivonin et al. 2015). While the preliminary findings are promising, various genres of movies are yet to be studied. While Hollywood movies are globally well known, Bollywood movies created its own genre and are getting more popular in recent years, especially in Asia and other regions. They provide a great resource for this research direction and might lead to interesting findings.

Potential Challenges and Promising Approaches

Regarding the unconscious emotions in media content, some preliminary studies have revealed that specific narratives in media content would induce unique physiological patterns while the audience is unaware of it (Ivonin et al. 2013, 2014, 2015; Chang et al. 2014b). It is promising to explore further with this direction. For example, how could media content influence the audience's attitude, preference, decision-making, and goal setting? Are these effects conscious or unconscious? Numbers of interesting research topics are worthy researching. In order to take the initiative, it is necessary to point out below the potential challenges and outline some promising approaches that could be used in future studies of unconscious emotion in media content.

The first challenge that researchers encounter is the choice of guiding theory or model for future empirical research. As illustrated earlier, the existing theories represent four complementary perspectives about emotion (i.e., the brain, body, situation, and mind). There are considerable overlaps between them, and none can be excluded from a discussion of emotion. Rather than seeing these four theoretical perspectives as conflicting conceptualizations, it is better to consider them as different epistemological approaches to the same phenomena and look for evidence that might lead to new theories and models. As some researchers have pointed out, there are limits to behaviorist explanations of psychological phenomena, and it is time to look for new paradigms in psychology (Kohler 2010). Given that the existing theories of emotion have been well researched, it might be useful to explore human emotions using an integrated approach. Several new ideas have been proposed. For example, an integrated approach based on the "extended mind" considered the environment as an extension of the mind (Clark and Chalmers 1998); thus as individuals change the environment, they are also changing how they experience the world. However, these new theoretical approaches are more like philosophical propositions rather than predictive, explanatory theories. There is a long way to go before they can form the basis of empirical, scientific studies. When it comes to the concept of unconscious emotion, the real challenge is not proving it exists, but providing a good explanatory account of the experimental data to improve current theory in this area.

The other critical challenge is a reevaluation of the basis of measurements of emotion. Because self-reports have been considered the most valid measure of conscious emotion, many studies have regarded the results of introspection as a fundamentally true account of experienced emotions. The accuracy of self-reports relies on their direct relationship with individual emotional experience; all the other measures are indirect and inferential. Unfortunately, these effective measures are unsuitable for assessing unconscious emotion because by definition unconscious emotions only take place below the threshold of introspection. It is epistemologically impossible to assess unconscious emotions through direct measures, which means that there is no easy foundation for models and theories. Since unconscious emotion cannot be accessed directly, it is necessary to use indirect measures. It should also be noted that conscious and unconscious emotions are interrelated and may coexist in most cases (Reingold and Merikle 1990). To extract information about unconscious emotion from the overall emotional experience appears to be the most difficult task of all. Recent years have growth in the use of mixed-methods approaches to complex phenomena (Moran-Ellis 2006). It is suggested that researchers should integrate direct and indirect measures in order to investigate phenomena that cannot be measured directly. This would help distinguish conscious and unconscious emotions and perhaps produce less biased interpretations of empirical data.

What Has Been Missing?

In the overview, it can be seen that many potential theoretical approaches have not yet been used. In particular, mind-based theories have long been marginalized by mainstream scientific studies because the unconscious was, and sometimes still is, treated as a taboo in science. Researchers tend to prefer externally rather than internally oriented explanatory accounts of a given phenomenon, partly because contemporary science is deeply influenced by Newtonian notion of "objectivity" (Chancer 2013). The legacy of the behaviorist tradition has been the emphasis on using observations of human behavior to understand human psychology and a negligence of the concept of mind. Nevertheless, it has recently been asserted that researchers should consider humans outside the machine paradigm (Kohler 2010). The fast, subjective, and irrational aspect of the human mind has been noted by some leading scientists (Kahneman 2003; Kihlstrom 2008). While the best method of investigating unconscious phenomena in empirical studies remains unclear, psychoanalysis might provide a useful starting point for those seeking new paradigms.

The procedures associated with the experimental psychological approach usually encompass both qualitative and quantitative research methods. While emotion recognition has largely relied on quantitative measures such as physiological signals and questionnaires like SAM, inferences and attributions relating to emotion activation can only be understood using qualitative methods. Social psychology looks for explanations in the external world (i.e., situation-based accounts), whereas psychoanalysis searches for answers from the subjective world (i.e., mind-based accounts). Social psychologists often combine quantitative data and qualitative analysis to investigate how social circumstances influence emotional experience. However, little attention has been paid to psychoanalytic studies combining qualitative and quantitative methods. This approach has been missing from contemporary research on emotion, and it seems a promising direction, in particular for researchers who are interested in subjectivity and the unconscious aspects of the human mind.

Conclusion

As demonstrated in the overview, many new theoretical approaches have yet to be applied into practice. Research into emotions in media content might even more complex because it is difficult to put media content directly into the context of psychological experiments. Nevertheless, it is a whole new research agenda when taking into account the continuity of media content and the unconscious part of emotional experience. This chapter started with the motivation to research into the relationship between myths and the unconscious. There are numbers of resemblance across different cultures and religions particularly in their myths and stories. Nowadays, these narratives still profoundly influence the human society, shaping the world as it is today. While most scientific research fully concentrates on how media content influences audience at a conscious level, it is also important to look into the irrational part of the media experience that might emerge along with the narratives. This chapter serves as a theoretical basis for researchers who intend to delve into this research direction and new inspirations for inventing new genres of digital media.

Recommended Reading

M.B. Arnold, *Emotion and personality* (Columbia University Press, New York, 1960)

J.R. Averill, A constructivist view of emotion, in *Emotion: Theory, Research, and Experience*, ed. by R. Plutchik, H. Kellerman (Academic, New York, 1980), pp. 305–339

J.A. Bargh, E. Morsella, The unconscious mind. Perspect. Psychol. Sci. **3**, 73–79 (2008)

L.F. Barrett, Are emotions natural kinds? Perspect. Psychol. Sci. **1**, 28–58 (2006a). doi:10.1111/j.1745-6916.2006.00003.x

L.F. Barrett, Solving the emotion paradox: categorization and the experience of emotion. Personal Soc. Psychol. Rev. **10**, 20–46 (2006b). doi:10.1207/s15327957pspr1001_2

L.F. Barrett, The future of psychology: connecting mind to brain. Perspect Psychol. Sci. **4**, 326–339 (2009). doi:10.1111/j.1745-6924.2009.01134.x

L.F. Barrett, Emotions are real. Emotion **12**, 413–429 (2012). doi:10.1037/a0027555

L.F. Barrett, B. Mesquita, K.N. Ochsner, J.J. Gross, The experience of emotion. Annu. Rev. Psychol. **58**, 373–403 (2007). doi:10.1146/annurev.psych.58.110405.085709

A. Bechara, H. Damasio, D. Tranel, A.R. Damasio, The Iowa Gambling Task and the somatic marker hypothesis: some questions and answers. Trends Cogn. Sci. **9**, 159–162 (2005). doi:10.1016/j.tics.2005.02.002

B. Bornemann, P. Winkielman, E. van der Meer, Can you feel what you do not see? Using internal feedback to detect briefly presented emotional stimuli. Int. J. Psychophysiol. **85**, 116–124 (2012). doi:10.1016/j.ijpsycho.2011.04.007

M.M. Bradley, P.J. Lang, Measuring emotion: the self-assessment manikin and the semantic differential. J. Behav. Ther. Exp. Psychiatry **25**, 49–59 (1994)

J. Campbell, *The Hero with a Thousand Faces* (Princeton University Press, Princeton, 1973)

L.S. Chancer, Sociology, psychoanalysis, and marginalization: unconscious defenses and disciplinary interests. Soc. Forum **28**, 452–468 (2013). doi:10.1111/socf.12033

H.-M. Chang, L. Ivonin, M. Díaz et al., From mythology to psychology: identifying archetypal symbols in movies. Technoetic Arts **11**, 99–113 (2013). doi:10.1386/tear.11.2.99_1

H.-M. Chang, L. Ivonin, M. Diaz et al., Enacting archetypes in movies: grounding the unconscious mind in emotion-driven media. Digit. Creat. **26**, 154–173 (2014a). doi:10.1080/14626268.2014.939985

H.-M. Chang, L. Ivonin, M. Diaz et al., Unspoken emotions in movies. Informatik-Spektrum **37**, 539–546 (2014b). doi:10.1007/s00287-014-0823-3

A. Clark, D.J. Chalmers, The extended mind. Analysis **58**, 10–23 (1998)

A.R. Damasio, The somatic marker hypothesis and the possible functions of the prefrontal cortex. Philos. Trans. R. Soc. Lond. B Biol. Sci. **351**, 1413–1420 (1996). doi:10.1098/rstb.1996.0125

A. Dijksterhuis, Think different: the merits of unconscious thought in preference development and decision making. J. Pers. Soc. Psychol. **87**, 586–598 (2004). doi:10.1037/0022-3514.87.5.586

A. Dundes, The flood as male myth of creation, in *The Flood Myth*, ed. by A. Dundes (University of California Press, London, 1988), pp. 167–182

S.H. Fairclough, Fundamentals of physiological computing. Interact. Comput. **21**, 133–145 (2009). doi:10.1016/j.intcom.2008.10.011

S. Field, *Screenplay: The Foundations of Screenwriting* (Dell Publishing, New York, 1984)

S. Freud, The unconscious. J. Nerv. Ment. Dis. **56**, 291–294 (1922)

N.H. Frijda, The laws of emotion. Am. Psychol. **43**, 349–358 (1988). doi:10.1037/0003-066X.43.5.349

M. Gendron, L.F. Barrett, Reconstructing the past: a century of ideas about emotion in psychology. Emot. Rev. **1**, 316–339 (2009). doi:10.1177/1754073909338877

G. Gigerenzer, *Gut Feelings: The Intelligence of the Unconscious* (Penguin Group, New York, 2007)

S. Hamann, Mapping discrete and dimensional emotions onto the brain: controversies and consensus. Trends Cogn. Sci. **16**, 458–466 (2012). doi:10.1016/j.tics.2012.07.006

M. Hassenzahl, N. Tractinsky, User experience – a research agenda. Behav. Inf. Technol. **25**, 91–97 (2006). doi:10.1080/01449290500330331

L. Ivonin, H.-M. Chang, W. Chen, M. Rauterberg, Unconscious emotions: quantifying and logging something we are not aware of. Pers. Ubiquitous Comput. **17**, 663–673 (2013). doi:10.1007/s00779-012-0514-5

L. Ivonin, H.-H. Chang, M. Díaz, et al., Beyond cognition and affect: sensing the unconscious. Behav. Infect. Technol. 1–19 (2014). doi:10.1080/0144929X.2014.912353

L. Ivonin, H.-M. Chang, M. Diaz et al., Traces of archetypal experiences in introspective reports and physiological responses. PLoS One **10**, e0124519 (2015). doi:10.1371/journal.pone.0124519

C.E. Izard, Emotion theory and research: highlights, unanswered questions, and emerging issues. Annu. Rev. Psychol. **60**, 1–25 (2009). doi:10.1146/annurev.psych.60.110707.163539

W. James, What is an emotion? Mind **9**, 188–205 (1884)

C.G. Jung, *The Archetypes and the Collective Unconscious* (Princeton University Press, Princeton, 1959)

D. Kahneman, Maps of bounded rationality: psychology for behavioral economics. Am. Econ. Rev. **93**, 1449–1475 (2003). doi:10.1257/000282803322655392

J.F. Kihlstrom, The psychological unconscious, in *Handbook of Personality: Theory and Research*, ed. by O. John, R. Robins, L. Pervin, 3rd edn. (Guilford Press, New York, 2008), pp. 583–602

J.F. Kihlstrom, Unconscious cognition, in *Encyclopedia of Consciousness*, ed. by W.P. Banks (Academic, Oxford, 2009), pp. 411–421

J.F. Kihlstrom, S. Mulvaney, B.A. Tobias, I.P. Tobis, The emotional unconscious, in *Counterpoints: Cognition and Emotion*, ed. by E. Eich, J.F. Kihlstrom, G.H. Bower et al. (Oxford University Press, New York, 2000), pp. 30–86

G.S. Kirk, *The Nature of Greek Myths* (Penguin, New York, 1975)
A. Kohler, To think human out of the machine paradigm: Homo Ex Machina. Integr. Psychol. Behav. Sci. **44**, 39–57 (2010). doi:10.1007/s12124-010-9113-z
S.D. Kreibig, Autonomic nervous system activity in emotion: a review. Biol. Psychol. **84**, 394–421 (2010). doi:10.1016/j.biopsycho.2010.03.010
P.J. Lang, Emotion and motivation: toward consensus definitions and a common research purpose. Emot. Rev. **2**, 229–233 (2010). doi:10.1177/1754073910361984
P.L. Lang, M.M. Bradley, B.N. Cuthbert, International affective picture system (IAPS): affective ratings of pictures and instruction manual. In: Technical Report A-8. University of Florida, Gainesville, pp. 1–12 (2008)
H. Leuner, Guided affective imagery (GAI): a method of intensive psychotherapy. Am. J. Psychother. **23**, 4–22 (1969)
B. Malinowski, Myth in primitive psychology, in *Magic, Science and Religion and Other Essays* (Doubleday, New York, 1954), pp. 93–148
J.G. Manis, B.N. Meltzer, *Symbolic Interaction: A Reader in Social Psychology*, 3rd edn. (Allyn and Bacon, Boston, 1978)
R.A. Mar, K. Oatley, The function of fiction is the abstraction and simulation of social experience. Perspect. Psychol. Sci. **3**, 173–192 (2008). doi:10.1111/j.1745-6924.2008.00073.x
R. May, *The Cry for Myth* (W. W. Norton, New York, 1991)
R. McKee, *Story: Substance, Structure, Style, and the Principles of Screenwriting* (Harper-Collins Publishers, New York, 1997)
J. Moran-Ellis, Triangulation and integration: processes, claims and implications. Qual. Res. **6**, 45–59 (2006). doi:10.1177/1468794106058870
D. Novak, M. Mihelj, M. Munih, A survey of methods for data fusion and system adaptation using autonomic nervous system responses in physiological computing. Interact. Comput. **24**, 154–172 (2012). doi:10.1016/j.intcom.2012.04.003
K. Oatley, Why fiction may be twice as true as fact: Fiction as cognitive and emotional simulation. Rev. Gen. Psychol. **3**, 101–117 (1999). doi:10.1037/1089-2680.3.2.101
A. Öhman, Face the beast and fear the face: Animal and social fears as prototypes for evolutionary analyses of emotion. Psychophysiology **23**, 123–145 (1986). doi:10.1111/j.1469-8986.1986.tb00608.x
V. Propp, *Theory and History of Folklore* (University of Minnesota Press, Minneapolis, 1984)
M. Rauterberg, The three phases of life: an inter-cultural perspective. In: Proceedings of the 2nd International Conference on Culture and Computing. IEEE Computer Society, Los Alamitos, pp. 80–85 (2011)
E.M. Reingold, P.M. Merikle, On the inter-relatedness of theory and measurement in the study of unconscious processes. Mind Lang. **5**, 9–28 (1990)
J. Rottenberg, R.D. Ray, J.J. Gross, Emotion elicitation using films, in *Handbook of Emotion Elicitation and Assessment*, ed. by J.A. Coan, J.J.B. Allen (Oxford University Press, Oxford, 2007), pp. 9–28
J.A. Russell, Affective space is bipolar. J. Pers. Soc. Psychol. **37**, 345–356 (1979). doi:10.1037/0022-3514.37.3.345
J.A. Russell, L.F. Barrett, Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. J. Pers. Soc. Psychol. **76**, 805–819 (1999)
S. Schachter, J. Singer, Cognitive, social, and physiological determinants of emotional state. Psychol. Rev. **69**, 379–399 (1962). doi:10.1037/h0046234
K.R. Scherer, What are emotions? And how can they be measured? Soc. Sci. Inf. **44**, 695–729 (2005). doi:10.1177/0539018405058216
J. Shedler, The efficacy of psychodynamic psychotherapy. Am. Psychol. **65**, 98–109 (2010). doi:10.1037/a0018378
H. Shevrin, M. Snodgrass, L.A.W. Brakel et al., Subliminal unconscious conflict alpha power inhibits supraliminal conscious symptom experience. Front. Hum. Neurosci. **7**, 544 (2013). doi:10.3389/fnhum.2013.00544

J. Tiffin, Digitally remythicised: Star Wars, modern popular mythology, and Madam and Eve. J. Lit. Stud. **15**, 66–80 (1999)

P. Tikka, Enactive media – generalising from enactive cinema. Digit. Creat. **21**, 205–214 (2010). doi:10.1080/14626268.2011.550028

F.J. Varela, E.T. Thompson, E. Rosch, *The Embodied Mind: Cognitive Science and Human Experience* (MIT Press, Cambridge, MA, 1992)

T.D. Wilson, Knowing when to ask: introspection and the adaptive unconscious. J. Conscious Stud. **10**, 131–140 (2003)

T.D. Wilson, *Strangers to Ourselves: Discovering the Adaptive Unconscious* (Belknap, Cambridge, MA, 2004)

P. Winkielman, K.C. Berridge, J.L. Wilbarger, Unconscious affective reactions to masked happy versus angry faces influence consumption behavior and judgments of value. Pers. Soc. Psychol. Bull. **31**, 121–135 (2005). doi:10.1177/0146167204271309

Y. Zemack-Rugar, J.R. Bettman, G.J. Fitzsimons, The effects of nonconsciously priming emotion concepts on behavior. J. Pers. Soc. Psychol. **93**, 927–939 (2007). doi:10.1037/0022-3514.93.6.927

Index

R. Nakatsu et al. (eds.), *Handbook of Digital Games and Entertainment Technologies*,
DOI 10.1007/978-981-4560-50-4