Pietro Michelucci *Editor*

Handbook of Human Computation

Springer

Handbook of Human Computation

Pietro Michelucci
Editor

Handbook of Human Computation

Springer

Editor
Pietro Michelucci
ThinkSplash, LLC
Fairfax, VA, USA

ISBN 978-1-4614-8805-7 ISBN 978-1-4614-8806-4 (eBook)
DOI 10.1007/978-1-4614-8806-4
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013954962

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Foreword: Making a Difference

Mary Catherine Bateson

Mary Catherine Bateson is a cultural anthropologist and author.

A volume of papers on human computation (HC) has been needed to lay the foundation of a field and establish a framework in which researchers can effectively build on each other's work. It is also likely to set off alarm bells in many quarters. Yet there is a possibility that the thinking collected here will constitute an important step toward solving a fundamental ethical problem in human society, namely, the increasingly widespread conviction that "nothing I can do will make any difference." Kant's Categorical Imperative[1] was an attempt to solve the problem by eliminating the question of scale and proposing that an action be evaluated as if it were universal, but this has not proved particularly effective in ever larger populations. The problem of taking responsibility for individual and local actions is most severe at the global level. Thus, for instance, individuals have difficulty believing that leaving an extra electric light burning in their suburban backyard is connected to the likelihood of lethal storms thousands of miles away. Exactly the same kind of reasoning discourages voters from going to the polls for local elections. How will people learn that what they do "counts"? By counting.

We badly need models of interdependence and connectivity that will convey to those who work with them the conviction that individual voices and actions *count*, a message conveyed through many different modalities, both in science and in popular culture. At the same time, the very term "human computation," accompanied by fascinating analogies to insect communities, may suggest a dystopic loss of individual autonomy and value. Human computation for socially useful goals will depend on giving individuals a sense of agency – a sense that they indeed can make a difference.

Agency has been the central issue for patient communities, so that enrolling patients as active collaborators in research has been an important new model for citizen science. One of the earliest examples of citizen science was the St. Louis baby tooth collection organized by Barry Commoner, in which scientists "took over the tooth fairy"[2] to demonstrate the dangers of nuclear testing in the atmosphere. The demonstration that Strontium 90 was being transferred in mothers' milk was a significant element in the banning of atmospheric testing, but so no doubt was the engagement it evoked in the parents.

There is a long history, going back to the Greeks and Romans, of attempting to use voting (an early form of human computation), with various modifications, to create a sense of agency that supports responsibility, and some of the hazards are known. Experience suggests, for instance, that plebiscites are easily manipulated by autocrats (as in the rise of fascism), so that it makes more sense to vote for individuals who are then able to deliberate together about issues and act systemically as surrogate decision makers in a second round of voting than it does to decide policy by majority popular vote. Other variations such as proportional representation also attempt to avoid the dangers of simple majority rule. Voter initiatives may appear to increase democracy but when overused may lock in dysfunctional policies. And at the same time, voters are increasingly taking the libertarian position that all legislation and regulation is pernicious. A central promise of human computation, already partially realized, is the possibility of creating an awareness of the vast number of decisions we all make every day, including the decision involved in where attention is focused from minute to minute,[3] along with information about the aggregate effect of those decisions and how they are shifting.

At the same time, information about new ideas and emerging patterns needs to be accessible and individual voices need to be audible. Human computation may run the risk of simply reinforcing existing trends, which may be negative, by facilitating conformity. The popularity of SUVs and violent movies and games tends to be self-reinforcing, and the most popular restaurant in town may not be the most pleasant place to go on a Saturday night. Thus, simply waiting to see what "goes viral" on YouTube or Twitter is not sufficient. A noteworthy variation on regenerative feedback, however, is Kickstarter.com, which works like a chain letter to raise funds for nonprofit projects.

A significant effort related to human computation is the effort to create interactive contexts for the expression of greater diversity of knowledge and imagination. Interdisciplinary conferences (such as the Macy conferences on Cybernetics and on Group Process after World War II) can be seen as an example of taking a group of individuals and turning them into a thinking system, a kind of superorganism.[4] With the decline in support for exploratory interdisciplinary work, there has been a rise in designs for interactive processes, such as America Speaks, the 21st Century Town Hall Meeting format devised by Carolyn Lukensmeyer,[5] and Laura Chasin's Public Conversations Project,[6] as well as research on conflict resolution and mediation[7] simultaneously alas with the steady increase in what Deborah Tannen calls the *Argument Culture, in which issues are approached antagonistically.*[8] Such innovative techniques can be regarded as forms of computation.

Human beings change in response to their habitual interactions, and there is already concern about deleterious effects of electronic communication, which will play a major role in human computation as we move forward. Much of human computation depends on persuading large numbers of individuals, acting separately, to contribute personal information, which is then combined, both processes facilitated by electronic technology. But it is important to notice that the implicit message of such an operation is *membership in a larger whole*. Any living system processes quantities of material and information, in ways that affect the state of that system and other systems to which it is connected, and attending to such processes potentially creates a sense of unity and an awareness of the reality of interdependence.

We know today that our entire planet can be looked at as a living system[9] with some capacity for self-regulation, and that the circulation of water and atmospheric gases is such that disruption or pollution in one place on the planet has measurable effects elsewhere. Indeed, earth systems are far more closely integrated than the present human capacity to respond to them, even in the preparation for and response to major disasters. The emphasis on individual autonomy that underlies American culture is a product of the circumstances under which Europeans settled the North American continent, but it is descriptively inaccurate for the human condition and inhibits effective cooperation in problem solving and humanitarian relief as we experience and attempt to mitigate the global effects of climate disruption. Arguably, then, if increased reliance on human computation shifts attitudes away from the fetish of individual autonomy and teaches us, by implication, to recognize that we are connected parts of a larger whole, this is a goal to be pursued. Perhaps too, the awareness of inescapably "making a difference," for better or for worse, by our individual choices will come to be seen as an essential aspect of human dignity.

Mary Catherine Bateson

Notes

1. "There is, therefore, only one categorical imperative. It is: Act only according to that maxim by which you can at the same time will that it should become universal law." Kant I, Foundations of the metaphysics of morals (trans: Beck LW, ed: Wolff RP, section 2, p. 44.
2. Bateson MC (1972) Our own metaphor: a personal account of a conference on conscious purpose and human adaptation. New York: Alfred A. Knopf, pp 140–141.
3. Jackson M (2009) Distraction: the erosion of attention and the coming dark age. Amherst: Prometheus Books.
4. Heims SJ (1991) The cybernetics group. Cambridge, MA: MIT Press.
5. Lukensmeyer C (2007) The change handbook: the definitive resource on today's best methods for engaging whole systems. San Francisco: Burrett-Koehler.
6. Fostering dialogue across divides: a nuts and bolts approach. www.publicconversations.org/docs/resources/Jams_website.pdf
7. Fisher R (1991) Getting to yes: negotiating agreement without giving in. New York: Penguin Books.
8. Tannen D (1998) The argument culture: moving from debate to dialogue. New York: Random House.
9. Lovelock J (1995) A new look at life on earth. Oxford/New York: Oxford University Press.

Preface

In all of your deliberations in the Confederate Council, in your efforts at law making, in all your official acts, self-interest shall be cast into oblivion. Cast not over your shoulder behind you the warnings of the nephews and nieces should they chide you for any error or wrong you may do, but return to the way of the Great Law which is just and right. Look and listen for the welfare of the whole people and have always in view not only the present but also the coming generations, even those whose faces are yet beneath the surface of the ground – the unborn of the future Nation.

– Great Law of the Haudenosaunee[1]

Why a Book About Human Computation?

In the new techno-culture of buffered sociality, in which young people spend more time wearing earbuds and texting frenetically than having real live conversations in a café, we consider the mounting existential challenges that our children and subsequent generations will face. Though human computation may not be a panacea, it does represent an opportunity for us to draw together more effectively as a global people to address such challenges. However, there is a practical issue.

The problem that exists today is that human computation (HC) research is fragmented across isolated communities. That is, HC is developed and implemented in multifarious ways across diverse fields of inquiry and application; yet each of these efforts occurs as an offshoot of some other discipline or as a novel method in some

[1] The Haudenosaunee league of Native American nations is known in Western culture as the Iroquois. However, "Iroquois" is a French transliteration of a derogatory name used historically by a competing tribe. The correct and proper name, "Haudenosaunee," means "people of the longhouse," which implies that member nations should live together as families in the same longhouse (Wikipedia 2013).

applied domain. But there is very little cross-fertilization, due to typical aversions to crossing community boundaries, philosophical differences, and terminology confusion. One even gleans cultural differences, such as an emphasis in Eastern cultures on systems that support collective rather than individual stakeholders. Thus, this book responds to the need for a clear, comprehensive, current, and interdisciplinary treatment of HC.

Rather than just reporting on the state of practice, we have challenged the confines of our conceptual comfort zones and engaged in bold analysis and risky ideation – something humans still do much better than machines. Ultimately, we have sought to collectively assess the state of the art and anticipate future directions, presenting the combination as a foundation and inspiration for future work and unlikely collaborations.

The Collaboration Has Already Begun

It has been both a tremendous honor and an exercise in humility to collaborate with such a talented, globally distributed (see Fig. 1), and remarkably genuine community of over 115 authors and editors. Perhaps it is the promise of human computation that draws out the humanity in us, that somehow echoes the mantra "we want to own our destiny." Indeed, the goals of this book have already begun to be realized as a consequence of its very development. Authors have formed new, respectful cross-disciplinary relationships, spawning new ideas, many of which appear on the pages of this book. From this chrysalis, we hope to nurture the emergence of human computation as a formal discipline, a charter for which is conveyed in the final chapter of the book.

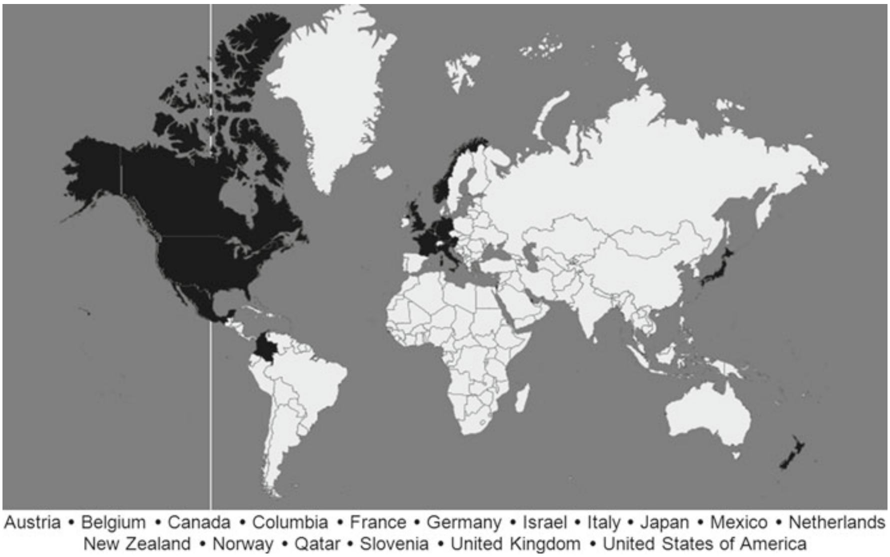


Fig. 1 Geographic representation of handbook contributors

About the Editor-in-Chief

Pietro Michelucci

Pietro Michelucci is an Indiana University–trained cognitive scientist and mathematical psychologist. He has been supporting the US Government as a science consultant since 2006. In that capacity, he developed a 2009 SBIR solicitation called "Massively Distributed Problem Solving" toward developing "a problem-solving system that combines an automated optimizing infrastructure with myriad distributed human processing agents." He also conducted a pilot study to assess the affordances of shared sensing, collective reasoning, and coordinated action on group efficacy. He speaks on various topics related to human computation and to advocate for its advancement as a formal discipline. More recently, in response to a recommendation from the Presidential Committee for the Advancement of Science and Technology (PCAST), he has been serving on the executive committee of an NITRD working group to develop a federal cross-agency initiative in social computing. Dr. Michelucci is currently founding the first-ever journal of human computation.

About the Section Editors

Matthew Blumberg

Editor, *Foundations*

Matthew Blumberg is an entrepreneur focusing on digital media and technology. He is founder and Executive Director of GridRepublic, a nonprofit organization that uses "volunteer computing" to provide supercomputing resources to public interest research (www.gridrepublic.org).

Haym Hirsh

Editor, *Applications*

Haym Hirsh is Dean of Computing and Information Science at Cornell University. His research has focused on foundations and applications of machine learning, data mining, information retrieval, and artificial intelligence, especially targeting questions that integrally involve both people and computing. Most recently, these interests have turned to crowdsourcing, human computation, and collective intelligence. From 2006 to 2010, he served as Director of the Division of Information and Intelligent Systems at the National Science Foundation, from 2010 to 2011, he was a Visiting Scholar at MIT's Center for Collective Intelligence at the Sloan School of Management and from 2011 to 2013 he was Chair of Computer Science at Rutgers University. Haym received his B.S. from the Mathematics and Computer Science Departments at UCLA and his M.S. and Ph.D. from the Computer Science Department at Stanford University.

Kshanti A. Greene

Editor, *Techniques and Modalities*

Kshanti received her Ph.D. in Computer Science from the University of New Mexico (UNM). She specializes in developing models that integrate and utilize intelligence derived from large groups of people. Her ambition is to provide tools to society that enable communities to self-organize to make decisions, form policy, and solve the problems that face them. Dr. Greene has been PI on DARPA, Army, and Navy research grants relevant to human computation and cofounded the Social Logic Institute to apply this research toward social and environmental welfare.

Michael Witbrock

Editor, *Infrastructure and Architecture*

Michael Witbrock is VP for Research for Cycorp, one of the world's leading AI research companies, and founder and CEO of Curious Cat Company, the maker of Curious Cat, a mobile social guide and companion based on strong AI. In his youth, he was a founder of NZ.COM, one of the very first web-based travel advice sites. He has also acted as CTO of Envigence d.o.o., a Slovenian Green Infrastructure firm, and as principal scientist at Lycos, one of the first web search engines. To encourage the positive use of AI technology, Michael acts as an advisor and is a frequent speaker at European agencies working in the area of eGovt, and as a scientific advisor to European research projects. Michael has a Ph.D. in Computer Science from Carnegie Mellon University and a B.Sc. (Hons) in Physiological Psychology from the University of Otago, in his home country of New Zealand.

Remco Chang

Editor, *Algorithms*

Remco Chang is an Assistant Professor in the Computer Science Department at Tufts University. He received his B.S. in Computer Science and Economics from Johns Hopkins University in 1997, M.Sc. from Brown University in 2000, and Ph.D. in Computer Science from UNC Charlotte in 2009. Prior to his Ph.D., he worked for Boeing, developing real-time flight tracking and visualization software, followed by a position at UNC Charlotte as a research scientist. His current research interests include visual analytics, information visualization, human–computer interactions, and computer graphics.

Caroline Ziemkiewicz

Coeditor, *Algorithms*

Dr. Caroline Ziemkiewicz is an Associate Research Engineer in the Analytics, Modeling, and Simulation Division at Aptima, Inc., where she contributes expertise in visualization and cognition-driven visual analytics. Her research interests cover topics in cognition-motivated visual information design and understanding users of complex analytical interfaces. Specific research interests include visual metaphors, information structure in design, individual differences in analytical task performance, and user-adaptive interfaces. Dr. Ziemkiewicz holds a Ph.D. in Computer Science and a Graduate Certificate in Cognitive Science from the University of North Carolina at Charlotte, and a B.A. in Computer Science from Ithaca College.

Winter Mason

Editor, *Participation*

Dr. Mason received his B.Sc. in Psychology from the University of Pittsburgh in 1999 and his Ph.D. in Social Psychology and Cognitive Science from Indiana University in 2007. He worked as a Visiting Scientist at Yahoo! Research from 2007 to 2011, when he joined the Stevens Institute of Technology as an Assistant Professor. Dr. Mason's research can be described as "Computational Social Science" and is focused on social networks, social media, crowdsourcing, and group dynamics. Methodologically, his research spans traditional psychological methods including laboratory experiments, new methods such as online data collection with crowdsourcing, and computer science methods such as data mining.

Kristina Lerman

Editor, *Analysis*

Kristina Lerman is a Project Leader at the Information Sciences Institute and holds a joint appointment as a Research Associate Professor in the Computer Science Department of the USC Viterbi School of Engineering. Her research focuses on applying network- and machine learning–based methods to problems in social computing.

Dan Thomsen

Editor, *Policy and Security*

Mr. Thomsen's career has focused on computer security research ranging from high-assurance operating systems and multilevel database security to security policy management. Mr. Thomsen has also contributed research into the effective communication of ideas from one person to another. At SIFT, Mr. Thomsen works to improve the effectiveness of computer security using artificial intelligence and human factors research to manage the complexity of modern computer security mechanisms. Mr. Thomsen also leads an effort in human computation using games to increase the chance of innovation for solving problems that have never been solved before.

Pietro Michelucci

Editor, *Impact*

See section "About the Editor-in-Chief."

Acknowledgments

Perhaps firstly, I am grateful to the reader. The future of human computation lies with you; whether as a researcher, developer, analyst, or even a participant, you will ultimately have the potential to be instrumental in the development of this technology and practice. Furthermore, by reading from these pages, you will be better equipped to approach and understand the implications of your involvement so that you can make conscientious and informed decisions.

This book was built on the generosity, brilliance, creativity, diligence, patience, and sacrifice of many kind people. Andrew Spencer at Springer-Verlag took the initiative to connect me to his States-based counterpart, Melissa Fearon, with whom I first began this journey. Aside from being a very nice person, Melissa has been outstanding to work with. She is a veritable font of publishing wisdom and has been quick to develop creative solutions to sundry issues. I am grateful to Melissa for having the courage to take a chance on this speculative project. Courtney Clark, Patrick Carr, and Jennifer Malat have kept me on the straight and narrow to ensure adherence to an aggressive publication timeline that would best serve the HC community.

I could not have hoped for a more industrious and visionary group of editorial collaborators. Each editor cast a distinctive, sometimes unexpected, but always beneficial framing to his or her section. It was a true privilege and education to work with each of them.

Perhaps the most personally gratifying aspect of this enterprise has been interacting with the other contributors. Despite stresses of work, academia, and myriad competing obligations, the chapter authors have been consistently thoughtful and responsive in their writing and correspondence, demonstrating genuine investment and ownership in this shared product. It is no small reflection of this level of dedication that the book is being published exactly one year from its November 2012 launch.

I also wish to acknowledge the population of "would-be" authors, whose circumstances precluded active participation, but who nonetheless recommended colleagues or engaged in fruitful discourse. And in particular, I want to thank Edith Law for elevating my sensitivity to "cultural nuances" that helped me more effectively draw collaborators across community boundaries.

Mary Catherine Bateson has honored this book and community with her foreword. I am grateful to her for her patient and enlightening conversation and for seeking not just human but *humanistic* computation.

Finally, my deepest and most heartfelt gratitude goes to Pamela K. Michelucci, whose selflessness, grace, and wisdom made this endeavor possible.

Yours collectively,

Pietro Michelucci
Editor-in-Chief

Contents

Part I Foundations

Foundations in Human Computation 3
Matthew Blumberg

Patterns of Connection 5
Matthew Blumberg

Human Computation and Divided Labor 13
David Alan Grier

Ant Colonies as a Model of Human Computation 25
Melanie Moses, Tatiana Flanagan, Kenneth Letendre, and Matthew Fricke

Parallels in Neural and Human Communication Networks 39
L.J. Larson-Prior

The Psychopathology of Information Processing Systems 51
Matthew Blumberg and Pietro Michelucci

Information and Computation 61
Carlos Gershenson

Epistemological Issues in Human Computation 71
Helmut Nechansky

Synthesis and Taxonomy of Human Computation 83
Pietro Michelucci

Part II Application Domains

Human Computation in the Wild 89
Haym Hirsh

Human Computation for Disaster Response 95
Patrick Meier

The Virtuous Circle of the Quantified Self: A Human Computational Approach to Improved Health Outcomes 105
Paul Wicks and Max Little

Knowledge Engineering via Human Computation 131
Elena Simperl, Maribel Acosta, and Fabian Flöck

Human Computation in Citizen Science 153
Chris Lintott and Jason Reed

Human Computation as an Educational Opportunity 163
Carole R. Beal, Clayton T. Morrison, and Juan C. Villegas

Search and Discovery Through Human Computation 171
Albert Yu-Min Lin, Andrew Huynh, Luke Barrington, and Gert Lanckriet

Human Computation in Electronic Literature 187
Scott Rettberg

Human Computation for Information Retrieval 205
Christopher G. Harris and Padmini Srinivasan

Human Computation-Enabled Network Analysis for a Systemic Credit Risk Rating 215
François Bry

Innovation via Human Computation 247
Lisa Purvis and Manas Hardas

Human Computation for Organizations: Socializing Business Process Management 255
Marco Brambilla and Piero Fraternali

Solving Wicked Problems 265
Dan Thomsen

Part III Techniques and Modalities

Introduction to Techniques and Modalities 279
Kshanti A. Greene

Social Knowledge Collection 285
Yolanda Gil

Location-Based Games for Citizen Computation 297
Irene Celino

Augmented Reality Interfaces in Human Computation Systems 317
Mark Billinghurst

Pervasive Human Computing 333
Joel Ross

Building Blocks for Collective Problem Solving 347
Kshanti A. Greene and Thomas A. Young

Adaptive Agents in Combinatorial Prediction Markets 367
Anamaria Berea

Risks and Rewards of Crowdsourcing Marketplaces 377
Jesse Chandler, Gabriele Paolacci, and Pam Mueller

Designing Systems with Homo Ludens in the Loop 393
Markus Krause

Human-Computer Interaction Issues in Human Computation 411
Stuart Reeves

Collective Action and Human Computation 421
Jasminko Novak

Cultural Evolution as Distributed Computation 447
Liane Gabora

Collective Search as Human Computation 463
Winter Mason

Organismic Computing 475
Pietro Michelucci

Part IV Infrastructure and Architecture

Infrastructure and Architecture for Human Computer Intelligent Collaboration 505
Michael Witbrock

Interactive Crowds: Real-Time Crowdsourcing and Crowd Agents 509
Walter S. Lasecki and Jeffrey P. Bigham

The Semantic Web and the Next Generation of Human Computation 523
Dominic DiFranzo and James Hendler

Conversational Computation 531
Michael Witbrock and Luka Bradeško

Modeling Humans as Computing Resources 545
Yu-An Sun and Christopher Dance

Service Oriented Protocols for Human Computation 551
Daniel Schall

CyLog/Crowd4U: A Case Study of a Computing Platform for Cybernetic Dataspaces 561
Atsuyuki Morishima

Multiagent Environment Design for Pervasive Human-ICT Systems: The SAPERE Approach 573
Gabriella Castelli, Marco Mamei, Alberto Rosi, and Franco Zambonelli

The "Human Sensor:" Bridging Between Human Data and Services 581
Neal Lathia

Part V Algorithms

Algorithms: Introduction 597
Remco Chang and Caroline Ziemkiewicz

The Wisdom of Crowds: Methods of Human Judgement Aggregation 599
Aidan Lyon and Eric Pacuit

Balancing Human and Machine Contributions in Human Computation Systems 615
R. Jordan Crouser, Alvitta Ottley, and Remco Chang

Constructing Crowdsourced Workflows 625
Peng Dai

Distributed Intelligent Agent Algorithms in Human Computation 633
Edmund H. Durfee

Human-Based Evolutionary Computing 641
Jeffrey V. Nickerson

Algorithms for Social Recommendation 649
Ido Guy

Part VI Participation

Participation 675
Winter Mason

Methods for Engaging and Evaluating Users of Human Computation Systems 679
Jon Chamberlain, Udo Kruschwitz, and Massimo Poesio

Participating in Online Citizen Science: Motivations as the Basis for User Types and Trajectories 695
Jason T. Reed, Ryan Cook, M. Jordan Raddick, Karen Carney, and Chris Lintott

Cultivating Collective Intelligence in Online Groups 703
Anita Williams Woolley and Nada Hashmi

Human Computation and Collaboration: Identifying Unique Social Processes in Virtual Contexts 715
Alecia M. Santuzzi, Christopher J. Budnick, and Derrick L. Cogburn

Game Theory and Incentives in Human Computation Systems 725
Arpita Ghosh

Part VII Analysis

Analysis: An Introduction 745
Kristina Lerman

Social Informatics: Using Big Data to Understand Social Behavior 751
Kristina Lerman

Computational Analysis of Collective Behaviors via Agent-Based Modeling 761
Lilian Weng and Filippo Menczer

Stochastic Modeling of Social Behavior on Digg 769
Tad Hogg

Activation Cascades in Structured Populations 779
Aram Galstyan

Synchrony in Social Groups and Its Benefits 791
Qi Xuan and Vladimir Filkov

Psychosocial and Cultural Modeling in Human Computation Systems: A *Gamification* Approach 803
Antonio Sanfilippo, Roderick Riensche, Jereme Haack, and Scott Butner

Part VIII Policy and Security

Introduction to Security and Policy Section 819
Dan Thomsen

Labor Standards 823
Alek Felstiner

Exploitation in Human Computation Systems 837
James Caverlee

Big Data, Dopamine and Privacy by Design 847
Thomas W. Deutsch

Privacy in Social Collaboration 857
Elena Ferrari and Marco Viviani

Applying Security Lessons Learned to Human Computation Solving Systems 879
Dan Thomsen

Part IX Impact

The Impact of Human Computation 893
Pietro Michelucci

From Human Computation to the Global Brain: The Self-Organization of Distributed Intelligence 897
Francis Heylighen

Superorganismic Behavior via Human Computation 911
Theodore P. Pavlic and Stephen C. Pratt

Gaming the Attention Economy 961
Daniel Estrada and Jonathan Lawhead

Human Cumulative Cultural Evolution as a Form of Distributed Computation 979
Paul E. Smaldino and Peter J. Richerson

Human Computation and Conflict 993
Juan Pablo Hourcade and Lisa P. Nathan

The Role of Human Computation in Sustainability, or, Social Progress Is Made of Fossil Fuels 1011
Bonnie Nardi

Human Computation: A Manifesto 1021
Pietro Michelucci

Index 1039

Contributors

Maribel Acosta Karlsruhe Institute of Technology, Germany

Luke Barrington Digital Globe Corporation, USA

Carol R. Beal University of Arizona, USA

Anamaria Berea George Mason University, USA

Jeffrey P. Bigham University of Rochester, Pittsburgh, USA

Mark Billinghurst HITLab – University of Canterbury, Christchurch, New Zealand

Matthew Blumberg GridRepublic, USA

Luka Bradeško Institut Jožef Stefan/Curious Cat Company, Slovenia

Marco Brambilla Politecnico di Milano, Dipartimento di Elettronica, Informazione e Bioingegneria, Milano, Italy

François Bry LMU Munich, Germany

Christopher J. Budnick Northern Illinois University, USA

Scott Butner Pacific Northwest National Laboratory (PNNL), USA

Karen Carney Adler, Planetarium, USA

Gabriella Castelli University of Modena and Reggio Emilia (UNIMORE), Italy

James Caverlee Texas A&M, USA

Irene Celino CEFRIEL – ICT Institute – Politecnico di Milano, Italy

Jon Chamberlain University of Essex, UK

Jesse Chandler Princeton University, USA

Remco Chang Tufts University, Medford, USA

Derrick L. Cogburn American University/Syracuse University, USA

Ryan Cook Adler Planetarium, USA

R. Jordan Crouser Tufts University, Medford, USA

Peng Dai Google Inc., 1600 Amphitheater Pkwy, Mountain View, CA, USA

Christopher Dance Xerox Research Centre Europe, France

Thomas W. Deutsch IBM Research, San Jose, USA

Dominic DiFranzo Rensselaer Polytechnic Institute, USA

Edmond H. Durfee University of Michigan, Ann Arbor, USA

Daniel Estrada University of Illinois, Urbana-Champaign, Champaign, USA

Alek Felstiner Law Clerk, United States District Court for the District of Columbia, Washington, DC, USA

Elena Ferrari Dipartimento di Scienze Teoriche e Applicate (Department of Theoretical and Applied Sciences), Università degli Studi dell'Insubria (University of Insubria), Varese, Italia

Vladimir Filkov UC Davis, USA

Tatiana Flanagan University of New Mexico, Albuquerque, New Mexico

Fabian Flöck Karlsruhe Institute of Technology, Germany

Piero Fraternali Politecnico di Milano, Dipartimento di Elettronica, Informazione e Bioingegneria, Milano, Italy

Matthew Fricke University of New Mexico, Albuquerque, New Mexico

Liane Gabora U British Columbia, Kelowna, Canada

Aram Galstyan USC – Information Sciences Institute, USA

Carlos Gershenson Universidad Nacional Autónoma de México, Mexico

Arpita Ghosh Cornell University, USA

Yolanda Gil USC – Information Sciences Institute, USA

Kshanti A. Greene Social Logic Institute, USA

David A. Grier The George Washington University, USA

Ido Guy IBM Research, Haifa, Israel

Jereme Haack Pacific Northwest National Laboratory (PNNL), USA

Manas Hardas Spigit Inc, Pleasanton, CA, USA

Christopher G. Harris SUNY Oswego, Oswego, New York, USA

Nada Hashmi Massachusetts Institute of Technology, USA

James Hendler Rensselaer Polytechnic Institute, USA

Francis Heylighen The Free University of Brussels/Principia Cybernetica, Belgium

Haym Hirsh Cornell University, USA

Tad Hogg Institute for Molecular Manufacturing, USA

Juan Pablo Hourcade University of Iowa, Des Moines, USA

Andrew Huynh University of California, San Diego, USA

Markus Krause Leibnitz University, Germany

Udo Kruschwitz University of Essex, UK

Gert Lanckriet University of California, San Diego, USA

L.J. Larson-Prior Washington University in St. Louis, St. Louis, USA

Walter S. Lasecki University of Rochester, Rochester, USA

Neal Lathia Cambridge University, UK

Jonathan Lawhead Columbia University, New York, USA

Kristina Lerman USC – Information Sciences Institute, USA

Kenneth Letendre University of New Mexico, Albuquerque, New Mexico

Albert Yu-Min Lin UCSD/National Geographic Society, USA

Chris Lintott University of Oxford, UK

Max Little Aston University, Aston, UK

Aidan Lyon University of Maryland, USA

Marco Mamei University of Modena and Reggio Emilia (UNIMORE), Italy

Winter Mason Stevens Institute of Technology, USA

Patrick Meier Harvard Humanitarian Initiative/Qatar Computing Research Institute, Qatar

Filippo Menczer Indiana University, USA

Pietro Michelucci ThinkSplash LLC, Fairfax, USA

Atsuyuki Morishima University of Tsukuba, Japan

Clayton T. Morrison University of Arizona, USA

Melanie Moses University of New Mexico, Albuquerque, New Mexico

Pam Mueller Princeton University, Princeton, USA

Bonnie Nardi UC Irvine, Irvine, USA

Lisa P. Nathan University of British Columbia, Vancouver, Canada

Helmut Nechansky Nechansky, Vienna, Austria

Jeffrey V. Nickerson Stevens Institute of Technology, USA

Jasminko Novak University of Applied Science Stralsund, Berlin, Germany

Alvitta Ottley Tufts University, Medford, USA

Eric Pacuit University of Maryland, USA

Gabriele Paolacci Erasmus University, Rotterdam, Netherlands

Theodore P. Pavlic Arizona State University, USA

Massimo Poesio University of Essex, UK

Stephen C. Pratt Arizona State University, USA

Lisa Purvis Spigit, USA

M. Jordan Raddick Johns Hopkins University, USA

Jason T. Reed Adler Planetarium, USA

Stuart Reeves University of Nottingham, UK

Scott Rettberg University of Bergen, Bergen, Norway

Peter J. Richerson UC Davis, Davis, USA

Roderick Riensche Pacific Northwest National Laboratory (PNNL), USA

Alberto Rosi University of Modena and Reggio Emilia (UNIMORE), Italy

Joel Ross University of Puget Sound, USA

Antonio Sanfilippo Pacific Northwest National Laboratory (PNNL), USA

Alecia M. Santuzzi Northern Illinois University, USA

Daniel Schall Siemens Corporate Technology, Austria

Elena Simperl University of Southampton, UK

Paul E. Smaldino Johns Hopkins University, Baltimore, USA

Padmini Srinivasan University of Iowa, USA

Yu-An Sun Xerox Corporation, USA

Dan Thomsen Smart Information Flow Technologies, USA

Juan C. Villegas Universidad de Antioquia, Medellín, Medellín, Colombia

Marco Viviani University of Insubria, Italy

Lilian Weng Indiana University, USA

Paul Wicks PatientsLikeMe, USA

Michael Witbrock Cycorp Inc./Curious Cat Company, USA

Anita Williams Woolley Carnegie Mellon University, USA

Qi Xuan UC Davis, USA

Thomas A. Young Social Logic Institute, USA

Franco Zambonelli University of Modena and Reggio Emilia (UNIMORE), Italy

Caroline Ziemkiewicz Aptima, USA

Introduction

A more descriptive title for this book would have been "The application, design, infrastructure, and analysis of heterogeneous multi-agent distributed information processing systems and their political, societal, and ethical implications," but as brevity is the soul of wit, I decided to go with simply *Handbook of Human Computation*.

Human computation means different things to different people. To some, it means using a computer to combine answers from many people into a single best answer. To others, it means taking a problem that is too big for any one person and splitting it into smaller, more manageable pieces that can be delegated to many people. Human computation can be the analysis of human behavior in a social network to better understand the spread of ideas or to predict outcomes on the world stage. And possibly it even represents an opportunity to recognize or engineer a new life-form with superhuman intelligence. Regardless of which of these things human computation might be, they all involve interconnected humans and machines that process information as a system, and they all serve a purpose.

What This Book Is Not

Though you will find much discussion of crowdsourcing herein, this is not a handbook of crowdsourcing. Crowdsourcing does not require computation; the term derives simply from "outsourcing to crowds." The individual contribution of each crowd member need not be computational nor give rise to computational analysis or output. Crowdsourcing is, however, a common method for engaging many participants in human computation; so they often coincide.

Nor is this a handbook of social computing. Social computing is defined as the intersection of social behavior and computational systems (Wikipedia 2013). However, social behavior is not a prerequisite for human computation. In fact, a workflow process may elicit human input, transform that input, and then pass the result to another human, in a pipeline that involves no social behavior or interaction whatsoever, yet is very much a manifestation of human computation. Thus, human computation subsumes social computing.

Then What Do We Mean by Human Computation?

To answer that question, we must first consider what we mean by "computation." Computation in this context refers not just to numerical calculations or the implementation of an algorithm. Computation refers more generally to *information processing*. This definition intentionally embraces the broader spectrum of "computational" contributions that can be made by humans, including creativity, intuition, symbolic and logical reasoning (though we humans suffer so poorly in that regard), abstraction, pattern recognition, and other forms of cognitive processing. As computers themselves have become more capable over the years due to advances in artificial intelligence and machine learning techniques, we have broadened the definition of computation to accommodate those capabilities. Now, as we extend the notion of computing systems to include human agents, we similarly extend the notion of computation to include a broader and more complex set of capabilities.

With this understanding of computation, we can further generalize our notion of human computation to encompass not only computation by an individual human but also machine-mediated computation by groups of individuals (e.g., pipelined problem solving systems), aggregate analytic results by groups that result from individual information processing (e.g., prediction markets), distributed networks of human sensors (e.g., mash-ups), and many other varieties of information processing that derive from the computational involvement of humans in simple or complex systems.

While this is what is meant by human computation for the purpose of establishing conceptual guideposts for this handbook, it is itself among the directives of the handbook to not only formally define this space of research and practice but to explore the past, present, and future scope of this frontier.

Why Is Human Computation Important?

Each of this book's many contributors may have a distinct answer to this question. My short answer is the following. As a species, we face multifarious challenges stemming directly and indirectly from our use of technology, and many of these challenges pose an existential threat to humanity. I believe that one promising avenue of recourse is to use technology to help us cooperate more effectively to solve the problems we have created. Thus, I believe our very survival depends upon the rapid advancement of human computation as a theoretical and applied science, to help us mitigate the effects of climate change, cure disease, end world hunger, protect human rights, and resolve conflicts.

Synopsis of Sections

Though the high-level structure of the book is ordinal by design, the following section synopsis will help point the reader who has specific areas of interest to the section of most immediate relevance. For the armchair reader, you may embark on a guided tour of human computation by beginning at page one. But if you happen to have a mercurial spirit, just open the book to a random chapter and see where that might lead you.

Foundations

The foundations section, edited by Matthew Blumberg, seeks to cast new light on the subject matter by asking basic questions, like "What is thinking?" "What is information?" and even "What is mental disease?" Answers come in novel forms that recast the interrelationship of foundational disciplines toward a deeper understanding of human computation.

Applications

The applications section, edited by Haym Hirsh, seeks to convey the value proposition of human computation by examining recent examples of how people have been brought together in new ways to achieve desired outcomes. This section surveys a broad range of human computation applications, in domains such as disaster relief, archaeology, medicine, science, education, literature, finance, innovation, business management, and others.

Techniques and Modalities

This section, edited by Kshanti A. Greene, catalogs an expansive and growing list of human computation techniques – that is, repeatable methods defined jointly by their applications, interaction paradigms, and/or computational methods. It is essentially a set of "design patterns" for human computation that facilitates modeling a new HC system on prior work.

Infrastructure and Architecture

The infrastructure and architecture section, edited by Michael Witbrock, seeks to balance the logistics of humans as computational resources with goals of actualization and empowerment. Thus, it covers the broad space of computational structures such as state space, communication protocols, human device drivers, reward structure programmability, as well as HC-specific interaction modeling techniques that are sensitive to the quality of human experience.

Algorithms

This section, coedited by Remco Chang and Caroline Ziemkiewicz, describes a variety of "systematic and general ways to treat humans as computational units" as well as new methods for formalizing the properties of human computation algorithms. Thus, this section may be useful for assessing, identifying, and constructing algorithms to fit specific use cases.

Participation

This section, edited by Winter Mason, explores a range of factors and associated techniques that influence the decision to participate in human computation activities. Importantly, it also considers dynamics that affect the quality of participation.

Analysis

This section, edited by Kristina Lerman, considers several analytic methods that can be used to predict emergent collective behavior and to inform the design of future human computation systems. These analytic methods are also considered in the context of quality control and performance assessment.

Policy and Security

This section, edited by Dan Thomsen, examines near-term ethical, regulatory, and economic considerations relevant to the emergence and growing prevalence of human computation and associated labor markets. It also delves into security and privacy issues germane to HC systems, along with relevant technical and policy-based solutions.

Impact

The impact section, which I had the privilege of editing, is a collection of forward-thinking essays on the near- and long-term implications of human computation on individuals, society, and the human condition. It asks hard questions and considers carefully the potential risks and rewards associated with the advancement of this new technology. It attempts to characterize a future with pervasive human computation and considers how we might prepare for it.

Bon Voyage!

Whatever your interest in human computation might be, by reading from this book you will hear from a coalescent community of communities and perhaps begin to understand our place in the world in a new way.

Fairfax, VA, USA Pietro Michelucci

Reference

Wikipedia (2013) Social computing. In: Wikipedia, the free encyclopedia. Retrieved from http://en.wikipedia.org/w/index.php?title=Social_computing&oldid=553413728

Part I
Foundations

Foundations in Human Computation

Matthew Blumberg

The current state-of-the art in Human Computing all too often involves large batches of some mundane but nevertheless computationally intractable problem (find the blue dot; read the words; fit the puzzle pieces); and is undertaken by a developer who realizes that large numbers of people might by various means be induced to each perform modest numbers of these tasks before getting bored and moving on to something else. And if enough people can do enough of these tasks useful things can be accomplished.

But this section—and this *Handbook* generally—seeks to encourage thinking beyond such a "Virtual Sweatshop" model; and to replace it with the aspiration to create massively large scale thinking systems, systems which might some day be used to address problems at an order of complexity beyond the competence of any individual person.

Moving in this direction—opening this avenue of investigation—involves giving thought to some basic ideas: what is computing? What is thinking? What is information? This direction benefits from ideas about the nature of communication; about complex systems and the emergent properties of such systems; about control of complex systems. Ideas about networks, about collaboration, about minds, about ecosystems, about culture—and a great many other topics.

In many of these instances, the best and deepest thought has been done in domains which might on their face seem distant from software development: Epistemology, Psychology, Cybernetics, Biology, Anthropology, Economics, and so on.

This chapter is not in any sense a comprehensive collection of "Foundational" concepts; it is more a diverse set of interesting tidbits, a taste. We aspire to continue an ongoing flow of such illuminating ideas as a regular feature in a forthcoming *Human Computation* journal. But the chapters that follow embody some introductory discussions:

M. Blumberg (✉)
GridRepublic, USA
e-mail: mblumberg@picador.net

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_1,

Patterns of Connection (Matthew Blumberg)—Drawing on ideas from Marvin Minsky, this chapter explores the nature of Mind, and the extent to which Mind emerges from particular patterns of connection. This is used to illustrate the concept of "Cognitive Architecture", which is proposed as a central concept in Human Computing.

The History of Human Computation (David Alan Grier)—The idea of organizing groups of people to perform cognitive work precedes computers and the Internet. This fascinating chapter traces the origins of these ideas back to Charles Babbage's early analysis of factories at the dawn of the industrial era.

Biological Networks as Models for Human Computation (Melanie Moses, Tatiana P. Flanagan, Kenneth Letendre, G. Matthew Fricke)—Notions of Mind have traditionally reflected to the technology of the day; advancing technology has lead, curiously, to ever more powerful metaphors. At various points, the mind was a garden, a factory, a computer. Recent trends return to biology: this chapter explores biological networks as an instructive model.

From Neural to Human Communication (Linda Larson-Prior)—If one wants to learn to organize a thinking system, a natural place to look to for guidance is the brain. This chapter considers both neural and human communication in order to better understand the potential for computation as an emergent behavior of a system.

Pathology in Information Systems (Pietro Michelucci, Matthew Blumberg)—Mental Illness in Humans can be viewed as a specific case of the more general phenomena of pathology in information systems. Thus Human Computing systems—and groups of people generally—may become pathological: large scale political failures like the Inquisition or Fascism being an example; as potentially are smaller scale systems like dysfunctional families. This chapter speculatively explores these issues, proposing this as a domain for future inquiry, so as to develop means to prevent, diagnose, and repair such systemic pathologies—i.e., to develop means to debug complex systems.

Information Theoretic Analysis and Human Computation (Carlos Gershenson)—This chapter introduces concepts of Information Theory in the context of Human Computing systems. What is Information? What is Computing? How does one talk about Networks?

Epistemological Issues in Human Computation (Helmut Nechansky)—The field of Epistemology brings to bear centuries of thought about the nature of Knowledge. This chapter takes as a start the view of Knowledge as "an individual model of an external world", and explores the use of such models for decision-making. Implications for Human Computing are considered.

Synthesis and Taxonomy of Human Computation (Pietro Michelucci)—As this *Foundations* section demonstrates, a wide range of fields contribute to the growing body of work in human computation. Each field, though, has its own set of concepts and associated words (e.g., social computing, distributed thinking, crowdsourcing, etc.) This chapter draws from these various disciplines—and from the diverse contributions found in this volume—in an effort to organize the concepts and provide a common conceptual framework.

Patterns of Connection

Matthew Blumberg

Background

My interest in Human Computation– described here as “Distributed Thinking”– dates back to 2008 and the FIFA World Cup final. The truth is (being American) I didn’t watch. I only read about it the next day. It was quite a match, apparently– eventually won 1-0, on a 73rd-minute goal by Wayne Rooney of Manchester United. But what was most notable in the coverage– to me– was the comment that the match had been watched, live, by 700 million people.

A soccer game being about 90 min long, this amounts to more than *a billion hours* of human attention– focused on a bouncing ball. That’s about 120,000 *person-years* of attention– compressed into 90 min.

Which raised the question: what could be done with all that cognition? Could it be harnessed for constructive purposes? What knowledge and tools and methods would be required?

Crowdsourcing

A number of web-based projects have emerged which draw on the aggregated intellectual skills of large numbers of people over the Internet. These projects represent the “state of the art” in Human Computation—exciting efforts to harness many minds in order to do intellectual work that would otherwise be impossible. A few key examples follow (there are of course many others):

M. Blumberg (✉)
GridRepublic, USA
e-mail: www.gridrepublic.org

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_2,

- *Clickworkers (2001)*—People were shown images of the surface of Mars, and asked to help map it by drawing circles around the craters. (Computers aren't good at this sort of pattern recognition, but people are.[1])
- *Stardust@home (2006)*—A NASA probe dragged a volume of gel through the tail of a comet; the comet particles were quite few and small, and searching for them in the large volume of gel was a challenge. The Stardust team posted nearly a million images of small sections of the volume online, and people were asked to search through these and to find characteristic tracks of particles. This collective effort considerably accelerated the search for the "needles" in the "haystack"
- *Galaxy Zoo (2007)*—People are shown images of galaxies, and asked to categorize them by visual features: spiral, disk, etc.; the goal is to build a celestial almanac. (As above, computers aren't good at this sort of image analysis.)
- *ESP Game (2003)*—Pairs of people are shown an image at the same time, and each starts typing descriptive words. When both have entered the same word, they "win" (and the system presumes to have learned a useful "tag" for use in categorizing the image).
- *Ushahidi (2008)*—People in and around crisis situations submit reports by web and mobile phones. These are aggregated (and organized temporally and geospatially), to give an accurate and unmediated view of the emerging situation.[2]
- *eBird (2002)*—Bird watchers throughout the world submit observations, creating a real-time database of bird distribution and abundance.
- *Iowa Electronic Market (1995)*—People buy and sell "contracts" in a (not-for-profit) Futures market, as a tool for predicting outcomes of elections, Hollywood box office returns, and other cultural phenomena.
- *FoldIt (2008)*—People solve 3D visual puzzles, as a means to solve problems in protein structure prediction.
- *Phylo (2010)*—People search for matching patterns in sequences of DNA, represented as strings of colored blocks.
- *EteRNA (2010)*—People solve visual puzzles related to the folding of RNA molecules

The above represents a fairly wide range of objectives and activities—thought it may be observed that all follow a certain pattern, one which is presently characteristic of what is commonly referred to as Crowdsourcing:

- In each project above, all users perform the same task repetitively (i.e., all users draw circles to mark craters, or place a pin to mark traces of comet, or find matching patterns in strings of colored blocks.)
- In most cases, the task is quite simple; it is the vast quantity that must be slogged through which requires the crowd input.

[1] Possibly of interest: see article in this volume by Jordan Crouser and Remco Change, discussing relative strengths of humans vs computers.

[2] Possibly of interest: See article in this volume on crowdsourcing disaster relief by Ushahidi founder Patrick Meier. Human Computation for Disaster Response.

- Tasks are single-user: interaction among participants while performing the work is not required.[3]
- There is no parceling of task-type based on user expertise (At most, users of measured skill—ie users who have returned validated results—might get harder versions of the task at hand.)

In sum, with these tasks, there is no "higher level" thinking being done by the "Crowdsourcing" system. All of the tasks completed by the public (individually and collectively) could plausibly have been done by the project organizers—in most cases better.[4] The projects are really a means of collecting and applying large quantities of unskilled labor. This of course is useful; but much more is possible.

The discussion below seeks to make the case that it is possible to create "Thinking" systems—systems created of many minds, and capable of sophisticated problem solving….

Distributed Thinking

In order to contemplate what a large scale thinking system might look like, it is useful to have a notion of what *Thinking* is.

As a point of reference, consider the model proposed by Marvin Minsky in *Society of Mind* (1988). In Minsky's model "minds are built from mindless stuff".

Minsky hypothesizes that a Mind—that thinking—is made up of many small processes (which he calls "agents"); that these are simple; that they are not especially intelligent in and of themselves—And that *it is the way that these things are connected* that creates intelligence, as a sort of emergent property of the "thinking" system.

Picking Up a Cup of Tea

For example, if one wanted to pick up a cup of tea there might be several processes involved (several "agents"):

- Your GRASPING agents want to keep hold of the cup
- Your BALANCING agents want to keep the tea from spilling
- Your THIRST agents want you to drink the tea
- Your MOVING agents want to get the cup to your lips

---

[3] ESP game is an exception here; sort of.

[4] A notable exception is FoldIt: In the case of FoldIt, it turned out that a public participant was unusually good at the task, better than subject area experts. This fact alone highlights the sophistication of that project. I.e., FoldIt serves to demonstrate the example that when projects are sufficiently advanced, they may draw in "savants", persons unusually good at the particular task—better in some cases than the project organizers themselves. And/or, projects may empower novel combinations of intellectual skills of persons otherwise unknown the project organizers.

… These would all be independent processes, performed in parallel, competing for resources in various ways—and collectively producing the behavior of picking up and drinking the cup of tea.

Stacking Blocks

Another illustration, a slightly more complicated cognitive problem– Imagine you had a pile of blocks, and you wanted to pile them up in a stack. You might hypothesize the existence of a "mental program" to do this, call it "Builder" (Fig. 1):

Fig. 1

BUILDER

In the Minsky view of the mind, this program would be composed of smaller applications, for instance (Fig. 2):

Fig. 2

```
              BUILDER
       ___________|___________
      /           |           \
  BEGIN          ADD          END
```

And each of these "programs" or "agents" would themselves be composed of smaller functions. And each of these, of possibly smaller… Until you got down to some list basic "primitive" functions from which all the others are built (Fig. 3):

Fig. 3

```
                 ADD
       ___________|___________
      /           |           \
    FIND         GET          PUT
    /           /   \        /   \
  SEE      GRASP     MOVE      RELEASE
  / | \    / | \     / | \      / | \

           -- PRIMITIVES --
```

What's interesting about this approach is that if you took from the previous chart describing "Builder" only the list of the Agents themselves, you wouldn't know anything about what the Builder does. It's only when you put the things into a structure that it becomes possible to contemplate that they might do something useful (Fig. 4):

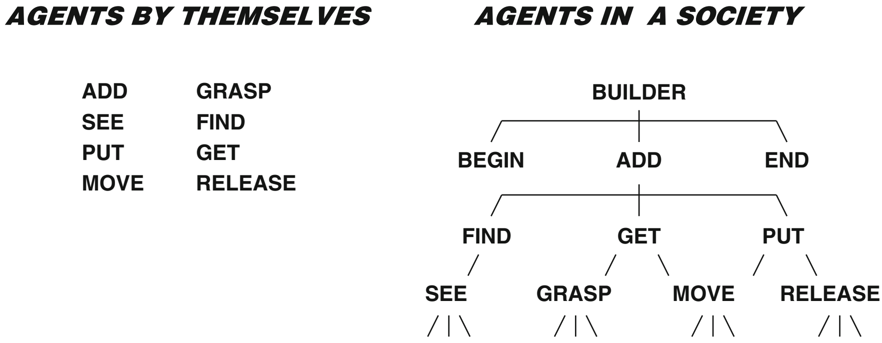


Fig. 4

This brings us to the first essential point of this essay: *Intelligence is created not from intellectual skill, but from the patterns within which intellectual skills are connected.*

The Minsky "Society of Mind" model is but one example; in general, patterns of organization which result in emergent "intelligent" behavior may be referred to as *"Cognitive Architectures"*.

From Crowdsourcing to Intelligent Systems

With an eye towards imagining a system which has a higher level of intelligence than its individual participants, and following Minsky's Cognitive Architecture– it's perhaps interesting to imagine what the set of "primitives" (the basic, unintelligent functions from which more complicated processes might be built) could be. Perhaps:

- **Pattern Matching/Difference Identification**
- **Categorizing/Tagging/Naming**
- **Sorting**
- **Remembering**
- **Observing**
- **Questioning**
- **Simulating/Predicting**
- **Optimizing**
- **Making Analogies**
- **Acquiring New Processes**

…This is not meant as a comprehensive list, just some illustrative examples. Note that none of these functions are especially complicated in and of themselves (though several are to varying degrees computationally intractable). Most are, in a wide range of contexts, quite parallelizable.

As food for thought, consider that many of the previously listed crowdsourcing projects provide quite nice templates for several of these very activities:

- **Pattern Matching/Difference Identification—**As noted, in *Clickworkers*, participants identified circles in a database of images; in *Stardust@home*, participants identified characteristic traces of comet dust in a database of images; in a range of other projects participants mark features on satellite images to generate or enrich maps, etc.
- **Categorizing**—In *Galaxy Zoo*, participants are shown images of galaxies, and asked to categorize them, by visual features: spiral, disk, etc.—and this is used to build up a structured database of astronomical objects.
- **Tagging/Naming—**In *ESP Game* participants create useful tags for image search (*In fact the system was licensed by Google to improve their image-search functionality).
- **Observing—**In *Ushahidi*, in *eBird*, and many other projects, distributed observations are entered into a shared central database
- **Simulating/Predicting—**In *Iowa Electronic Market*, and a wide range of subsequent "Prediction Markets", participants engage in a process which has been shown to effectively predict the outcome of a range of events.
- **Optimizing—**In *FoldIt* participants are asked to optimize the shape of an object according to certain parameters.
- **Etc…**

Following the earlier discussion, while it may be the case that any individual one of these systems is useful and interesting, it is the potential of *putting these things together into systems*– into intelligent patterns, into Cognitive Architectures– where really interesting things may become possible.

A Speculative Example

Imagine creating a drug discovery pipeline using Distributed Thinking –

By way of context, note that one method of drug discovery is {1} to identify a mutant or malformed protein which has been implicated in a specific pathology. And then {2} to find some other protein that binds to this deviant but nothing else– this is akin to sticking a monkey wrench into a running machine: the goal is to muck up the works, to cause that process to fail. And this can be quite effective.

Given a target identified by lab work, one could imagine subsequently breaking the process of discovering such "monkey-wrench" proteins into a sequence of steps– like, "docking" to see what candidate proteins stick to your target; "similarity analysis" to see which proteins are like which other proteins (to find alternative avenues of exploration); "optimizing" (to improve marginally useful candidates); "cross screening" (to see if a candidate has side effects, by checking whether it docks with anything it's not supposed to); and so on… (Fig. 5).

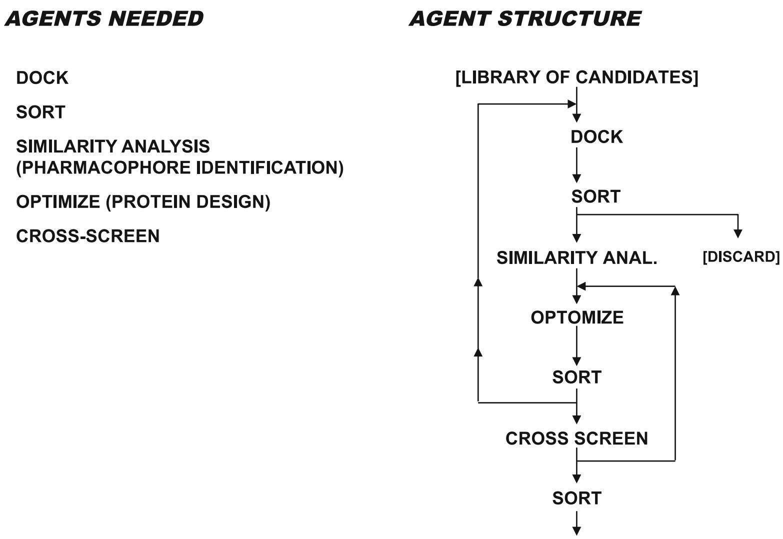


Fig. 5

All of these individual steps/processes could be imagined in terms of systems on par with existing crowdsourcing applications. And one could imagine linking these functions—these agents– into a fairly elaborate workflow,[5] the collective function of which would be to seek out and create promising drug candidates.[6]

Summary

A great deal has been done with Crowdsourcing. Current examples share a number of features however, most notably insofar as each supports only a single type of task. The discussion presents the idea of “Cognitive Architectures”—patterns into which individual systems, each performing specific (and potentially mundane) tasks, might be interconnected to collectively create a higher level of cognition.

[5] A small but important step in the evolution from current-generation Crowdsourcing to Distributed Thinking would be adoption of a standard means to integrate individual projects (individual functions) into more complex workflows. It is hoped that developers of such projects—and especially developers of middleware like BOSSA and PyBOSSA—will provide APIs that enable others to submit inputs and collect outputs, so the output of one project might be used as the input for another. For instance, a Phylo-style DNA project might input sequences into a FoldIt style structure prediction application.

[6] Of course there are numerous technical reasons why these steps are incomplete or may be impractical or infeasible at the moment. As noted it is a speculative example, a broad-stroke illustration.

The goal is to raise the prospect that "Apollo Project" challenges might be met by the application of sufficient attention, properly structured—It's all a matter of the patterns by which we connect ourselves and our information.

References

Minsky M (1988) The society of mind. Simon and Schuster, New York. ISBN 0-671-65713-5

Websites

Clickworkers: Originally http://clickworkers.arc.nasa.gov/top. Later revised to http://clickworkers.arc.nasa.gov/hirise. Current version: http://beamartian.jpl.nasa.gov/welcome
Stardust@home: http://stardustathome.ssl.berkeley.edu
Galaxy Zoo: www.galaxyzoo.org
ESP Game: originally: http://www.espgame.org; subsequently included in images.google.com
Ushahidi: www.ushahidi.com
eBird: ebird.org
Iowa Electronic Markets—tippie.uiowa.edu/iem
FoldIt—fold.it
Phylo—phylo.cs.mcgill.ca/eng
EteRNA—eterna.cmu.edu/

Human Computation and Divided Labor

The Precursors of Modern Crowdsourcing

David Alan Grier

Nature of Crowdsourcing

Though it is often to be a new phenomenon, one that is deeply tied to new technology. Jeff Howe, who identified and named the phenomena, claimed that it was a revolution "intertwined with the internet." (Howe 2008) However, it is actually a very old idea, one that has many historical antecendents in the twentieth, nineteenth and even eighteenth century. To understand crowdsourcing, we need to go back to Charles Babbage, the early nineteenth century mathematician.

Babbage was perhaps the first to understand that computation of any form was merely a form of divided labor. Babbage, of course, was not the first to discover divided labor. The concept of divided labor opens Adam Smith's 1776 book, The Wealth of Nations. "The greatest improvements in the productive powers of labour," Smith wrote in his first chapter, "and the greater part of the skill, dexterity, and judgment, with which it is anywhere directed, or applied, seem to have been the effects of the division of labour."(Smith 1776)

However, crowdsourcing is not merely any form of divided labor but the single form of divided labor that is untouched by modern information technology, the division of work by skill. Traditionally, economists have identified five ways of dividing labor. Any task can be divided by time, place, person, object and skill. You can create tasks by identifying the time when it must be done, the place where it must be done, the people with whom it must be done, the object on which the work is done and finally, the skill needed for the task.(Barnard 1936) Of those five methods, the first four can be mediated by modern information processing technology.

This technology can used to move work so that it need not be done at a specific time or place. It can move data, the object on which the work must be done, from one place to another. Finally, it can also be used to establish communications

D.A. Grier (✉)
Center for International Science and Technology Policy,
Elliott School of International Affairs, George Washington University, USA
e-mail: grier@gwu.edu

P. Michelucci (ed.), *Handbook of Human Computation*, 13
DOI 10.1007/978-1-4614-8806-4_3, © Springer Science+Business Media New York 2013

between any team of people on any part of the globe. The one thing that it cannot do is to change the skill of individual workers, though it can connect workers with different skills to work on the same project.

Crowdsourcing moves beyond the mere division of labor by skill and looks at the problem of how to combine best the skills of workers with the capabilities of information technology. It considers how to divide work and assign some tasks in order to get the right skills doing the right pieces of the job. As such, it is an example of what production managers call "refactoring work." The current forms of work that we identify as crowdsourcing are merely ways of refactoring work in a way that can use workers flexibly and that gets the right skills to the right part of a production. Charles Babbage was among the first scholars to look at this problem and certainly prepared the foundation for crowdsourcing.

Babbage and the First Scholar of Crowdsourcing

As a starting point for the study of crowdsourcing, Babbage has a much better perspective than Smith. Smith wrote at the start of the industrial era and focused on the four forms of divided labor that are easily handled by information technology: the division of labor by time, place, person, and object. Furthermore, he had a limited understanding of the potential of machines. He wrote of machines as tools. "A great part of the machines made use of in those manufactures in which labour is most subdivided, were originally the invention of common workmen, who, being each of them employed in some very simple operation, naturally turned their thoughts towards finding out easier and readier methods of performing it." (Smith 1776)

Writing more than 50 years after Smith, Babbage had a better understanding of the division of labor by skill and the role that machines might have in such a division. Babbage is generally remembered as a nineteenth century mathematician who designed computing machines.(Hyman 1982) In fact, Babbage is a much broader scholar, who was interested in chemistry, astronomy, and economics as well as mathematics. Perhaps the best way to understand Babbage is to recognize that he identified himself as an "analytical mathematician" during his years at Cambridge University and formed a club called the Analytical Society. (Grier 2010)

By labeling himself as an Analytical, Babbage was first identifying with a school of European mathematicians, such as Leonhard Euler or Joseph Louis Lagrange, who approached the study of calculus in a certain way. However, Babbage broadened his definition to analysis to include almost any activity that divided work into small pieces, create a symbol for those pieces and manipulated those symbols mechanically. (He named his second computing machine the "Analytical Engine" because it was capable of manipulating mathematical symbols in such a way.) (Grier 2011)

Because of his analytical background and his interest in machinery, Babbage studied the organization of factories and production. The result of this work he published, On the Economy of Machinery and Manufacturers, in 1831. The book combines broad principles of industrial organization with surprisingly detailed comments on industrial tasks. He gives principles of using machinery and mixes them with comments on cutting glass and splitting wood. In it, he builds upon Smith's work and moves beyond the division of labor by time, place, object and person to the division of labor by skill.

In considering the division of labor, Babbage realized that the division of labor by skill had more economic impact than the other four forms of divided labor. "That the master manufacturer, by dividing the work to be executed into different processes, each requiring different degrees of skill or of force," he wrote can purchase exactly that precise quantity of both which is necessary for each process"(Babbage 1831) He argued that if you did not divide work by skill, the manufacture would have to hire people who had all the skills necessary for the job. Such individuals, he observed, would be unlikely to perform all skills equally well and would be more expensive than workers who had only a single skill. This observation is generally known now as Babbage Rule. (Braverman 1975)

The Progenitor to Crowdsourcing: Dividing Mental Labor

In the Economy of Machinery and Manufacturers, Babbage applied the ides of divided labor to clerical tasks and calculation, categories of work that he identified as mental labor. It may, he wrote, "appear paradoxical to some of our readers that the division of labour can be applied with equal success to mental as to mechanical operations." He argued that not only was such work governed by the principles of Adam Smith but that showed that it showed that manufacturing was "founded on principles of deeper root than may have been supposed."(Babbage 1831)

At the time, both Great Britain and France did scientific calculation, one of Babbage's forms of mental labor, with methods that were quite similar to modern crowdsourcing. Beginning in 1767, British Nautical Almanac used freelance workers to prepare its annual volume of astronomical tables. These workers were generally teachers or clerics in the British Isles, though at least one worker was the widow of a cleric and the other was a teacher who lived in North America and communicated with the Almanac office through the slow and irregular North Atlantic mails. (Grier 2005)

The workers for the Almanac would get their assignments in much the way that crowd workers would get their assigns from the markets at oDesk or eLance. The director of the Almanac would determine which charts needed to be calculated and describe the nature of the calculations. He offered these calculations to anyone who was qualified and willing to do them. The workers would accept the tasks and do them at their homes. Most used this job to supplement their income. (Grier 2005)

In writing about calculation, Babbage argued that since it was governed by the same economic laws as physical labor, it would be pulled into the same forms of production as had word working or pottery. He noted that economic forces "causes large capitals to be embarked in extensive factories." In a crude way, this argue presages the argument, made 100 years later by Ronald Coase, for the existence of organized companies. "The main reason," argued Coase, "why it is profitable to establish a firm would seem to be there is a cost" to making all decisions in the market. (Coase 1937)

Indeed, in 1831, the Nautical Almanac was in the process of moving its production into a single office and eliminating freelance computation. Babbage had been on the committee that had reviewed the Almanac and had recommended the new computing factory model. He also watched as a second computing office, that at the Royal Observatory at Greenwich, also adopted factory models for its calculation. (Grier 2005)

Babbage got many of his ideas about mental labor, the organized processing of information by studying the computing office of the French civil surveying office, or Bureau Cadastre. By any measure, the Bureau Cadastre followed factory precepts. It operated a single computing office in Paris and employed no freelancers. However, it served as a model for later efforts that were much closer to crowdsourcing and it also taught Babbage about the division of labor by skill and how to utilize machinery to minimize costs. (Grier 2005)

The Bureau Cadastre operated a computing office from 1791 to about 1795 under the direction of the engineer, Gaspard de Prony. The Revolutionary French Government had assigned this office the task of creating trigonometric tables for surveying and navigation. In particular, they wanted these tables based not on the standard units that divided a circle into 360° but a new division that divided each quarter circle into 100 grads. (Daston 1994) (Rogel 2010)

De Prony divided the calculations by skill. He created three groups of workers. The first group was a small office of well-trained mathematicians. The Author de Roegel argues that this group may have had about six individuals, including the mathematician Andrien-Marie Legendre. It identified the equations that would be used in the calculation. The second group was less skilled than the first. It took the equations and used them to compute some of the basic values of the trigonometric functions. This group was called calculateurs. The third group was the least skilled. They took the basic values from the calculateurs and interpolated intermediate values between them. De Roegel notes that this group was the largest of the three. It had at least 15 workers and might have had as many as 60. De Prony once claimed that it had 150 workers. (Rogel 2010)

In writing about the Bureau Cadastre, Babbage was primarily interested in the problem of refactorization, of dividing labor and utilizing machines for some of the tasks. The "possibility of performing arithmetical calculations by machinery may appear to non-mathematical readers to be rather too large a postulate," he explained. However he would "remove a small portion of the veil which covers that apparent mystery."(Babbage 1831) He argued that his first computing machine, the Difference Engine could do exactly the kind of interpolation that was done at the Bureau Cadastre. "The ease and precision with which it works leave no room to doubt its success," he added. (Babbage 1831)

The Bureau Cadastre operated for only 3 years before it was disbanded. "The division of labour cannot be successfully practiced unless there exists a great demand for its" products, Babbage noted, "and it requires a large capital to be employed in those arts in which it is used."(Babbage 1831) Indeed, few organizations could afford to support any information processing office, much less a scientific computing office. During the rest of the nineteenth century, most of the scientific computing was done on a small scale. A single scientist would do the work, aided by a student, a child, or a spouse. The few large computing organizations, such as the American Nautical Almanac or the Harvard Observatory, tended to build computing factories because they were able to do more work with their resources. They also tended to look closely at how they could substitute machinery for labor. None tried to build a complex computing machine like Babbage's Difference Engine. However, many of them were able to expand their capacity by using small, mass produced adding machines. (Grier 2005)

Resurrection of the Bureau Cadastre as a Crowdsourced Organization

In 1938, the American Government created a computing organization that was based on the model of the Bureau Cadastre and used methods that were much closer to those of modern crowdsourcing. This organization, called the Mathematical Tables Project, followed the outlines of the Bureau Cadastre. It had three divisions. The first were senior mathematicians who identified the calculations. The second was a planning committee who created worksheets to guide the work. The third was group of clerks who completed the worksheets. In general, the members of this last group had limited mathematical skills. They were usually asked only to do addition or subtraction. (Grier 2005)

Unlike the Bureau Cadastre, the Mathematical Tables Project used market mechanisms to manage its workers. It hired senior mathematicians as freelancers to identify the calculations. It used a two-stage market to engage the clerks in the third group. Unlike modern crowdsourcing operations, it was restricted only to a crowd that lived close to its base of operations in New York City. Still, within the limits of the communications technology of the time, it operated much as a complex crowdsourcing company of the twenty first century. (Grier 2005)

At its founding, the Mathematical Tables Project represented a retreat from the practices of its day. Most organizations wanted long term relationships with clerical employees. They wanted the employees to learn more about the organization, gain skill in their job and become more efficient. The most prominent expert on organizing office work in that age was William Henry Leffingwell, whose book, A Textbook of Office Management, was widely read by office directors. In it he argued that office workers needed to be permanent members of the staff. If large numbers of workers were leaving their jobs after a short time, they represented "a serious loss." (Leffingwell 1932)

Leffingwell was a student of Frederick Winslow Taylor, the mechanical engineer who invented the concept of scientific management. Taylor's system involved dividing work into small tasks, analyzing these tasks and setting goals for the workers and using a task market to pay the workers. The workers would be rewarded for each task completed. However, Taylor did not want the workers to gain control of production through the task market. He was often critical of factories that used a task market and didn't attempt to set standards for production. In reviewing one factory, he argued that "The workmen together had carefully planned just how fast each job should be done, and they had set a pace for each machine throughout the shop." (Taylor 1911)

Yet 1938 was not the easiest year in which to apply the ideas of scientific management. The United States had been in a depression for 9 years and had recently seen a sharp rise in unemployment. The Administration of Franklin Roosevelt had set the goal of getting jobs for workers. "Our greatest primary task is to put people to work," Roosevelt had explained to the nation. He wanted to find jobs for people even if the work was not always profitable. "The joy, the moral stimulation of work," he added, "no longer must be forgotten in the mad chase of evanescent profits." (Roosevelt 1933)

The Mathematical Tables Project was therefore organized as a work relief effort. It had to be flexible. It had to make use of workers when they were available and be ready to train new workers when they arrived. To do this, it used a two-stage market. It used one market to get workers. That market was run by the Works Progress Administration, the financier of the project. Each day, the project would tell the main Works Progress Administration how many workers it could use and accepted the workers that came from that office. (Grier 2005)

The Project operated a second market within its office. This market was represented by a rack of worksheets. Each worker would take sheets from the rack, complete the calculations and return the sheets to the rack. They had to complete a minimal number of worksheets each day to be paid. (Grier 2005)

Though the project followed the crowdsourcing model, it pushed to refactor labor and move towards a factory model, much as Charles Babbage had observed 100 years before. The leaders pushed to acquire calculating machines and punched card equipment, arguing that these devices made the group more efficient. The manager of the organization, Arnold Lowan, argued that such machinery allowed handicapped workers to do more. It "has been found from actual work records over an extended period of time," explained one report, "that one armed operated using the new Frieden calculator was able to produce 40 % more work an unimpaired worker using a calculator which is not fully automatic." (Grier 2005)

Lowan also reduced the size of the organization and strived to retain workers. He was motivated partially by ambition and partially by rising labor costs. He desperately wanted the organization to be accepted by American scientific institutions. For the first 2 or 3 years of operation, he regularly wrote to university scientists and begged them to give him something to compute. Rarely did he receive a reply much less a problem. Furthermore, the nation's scientific leadership was skeptical of the group. They argued that the unemployed were not prepared to do scientific work and so the Mathematical Tables Project could not be expected to produce valid results. (Grier 2005)

However, by 1941, Lowen felt the pressure of rising labor costs more than desire to build a respectable organization. The preparation for the second world war required large numbers of workers and had raised the cost of labor. Lowan, who once could rely on a large pool of inexpensive workers, now had to try to keep every worker he could find. As Leffingwell argued, he tried to keep workers and give them an opportunity to build skill. By the start of the war, he was offering mathematics classes to his workers in order to keep them engaged in the process. (Grier 2005)

Shortly after Pearl Harbor, the group split in two. The Navy took one group and had them prepare navigation charts. The Office of Scientific research and development took the other and had them to general-purpose scientific calculation. This second group was the most active computing organization of the war. Still, the combined size of the two was a small fraction of the original organization. The Navy group had roughly 50 workers while the other group had 25. At its inception, the project had 450 workers. (Grier 2005)

During the war, both parts of the Mathematical Tables Project worked to systematize their operations and move away from a management model that resembled crowdsourcing. The Naval Section of the project moved quickly towards this goal. It produced only one kind of calculation, navigation tables for the new LORAN radio navigation system. The leader of the section, the mathematician Milton Abramowitz, devoted a great deal of time to studying the algorithm that produced the tables. He discovered a way of reusing information and several steps that could be simplified. Finally, as Babbage had done 100 and 10 years before, he explored ways of substituting machine work for human labor. He first introduced adding machines into the process and later, found a way to do a substantial fraction of the calculations with punched card machines. The punched card machines actually used a more complicated algorithm than the hand computation, but it produced results that required substantially less review for errors. (Grier 2005)

The other section of the Mathematical Tables Project also worked to simply operations, remove market management techniques and substitute machines for human labor. As it was a general purpose computing office, it was driven less by a single, repeated calculation than by the need to be able to address many different kinds of problems in a short period of time. The mathematical leader, Gertrude Blanch, found that the project received many requests that simply required too much effort to prepare for the large group of modestly skilled clerks. (Grier 2005)

Initially, Blanch tended to do many of these special jobs herself, spending an extra evening or weekend had her adding machine. However, by the fall of 1942 or the winter of 1943, she received too many requests to be able to handle them herself. As others had before he, she worked to improve the skills of her workers and extend their capacity through adding machines. In this work, she was added by the mathematician Cornelius Lanczos, who had once served as Albert Einstein's research assistant. Blanch and Lanczos ran a series of classes to train the workers. These classes began with the basic properties of arithmetic and ended with college level course on numerical analysis. (Grier 2005)

Even though Blanch moved her office away form crowdsourced management methods during the war, she still occasionally used the methods of crowdsourcing for sensitive or secret calculations. Both the Office of Scientific Research and Development and the U. S. Army regularly asked for computations that it wished to keep secret. These calculations included radar tables, bombing plans, shock wave propagations, and, most famously, a serious of calculations for the plutonium bomb being designed at Los Alamitos. (Grier 2005)

The Office of Scientific Research and Development considered the staff of the Mathematical to be a security risk. Many of them came from social groups that the Army considered to be unreliable or had had had dubious associations during the 1930s. Blanch, for example, lived with a sister who recruited for the Communist Party. For these calculations, Blanch would receive the request from mathematicians outside of the project. These requests would have no reference to the physical problem behind the calculation or any hint of the physical units of the various elements. Blanch would convert these requests to worksheets which further obscured the calculations. (Grier 2005)

Other War Time Crowdsourcing Efforts

Though the Mathematical Tables Project moved away from crowdsourcing during the war, other organizations embraced methods for raising funds or producing goods. In some ways, the second world war was a war of amateurs, a war that asked people to undertake roles that they had not done before. Women moved into factories, shop stewards became factory managers, factory managers became entrepreneurs. In this environment, organizations regularly turned to their employees for innovation in much the same way that companies turn to crowds for the same ideas. (Grier 2005)

The methods of crowdsourced innovation in the 1940s were called “suggestion systems” and these processes were symbolized by the suggestion box. Though such systems fell into disfavor during the 1950s, they were a common practice in the 1930s and 1940s. They had been developed during the first world war by the National Cash Register Company and had been promoted during the 1920s by the National Association for Suggestion Systems. Among the organizations that used suggestion systems were Swift Meats, United Airlines, People’s gas and Light, Firestone Rubber and Westinghouse. (National Association for Suggestion Systems 1944)

The National Association for Suggestion Systems published books that described how to design and operate such systems. The theory behind these books was quite similar to the theories of open innovation. It posited that the employees of a company had untapped knowledge about the company’s products and production methods. It presented ways of soliciting ideas from employees, curating and developing those ideas, testing the ideas in practice and rewarding the ideas. (National Association for Suggestion Systems 1944)

During the war, many, many organizations also used ideas that were similar to the modern idea of crowdfunding. These organizations ranged from the Federal government, which sold low value War Bonds, to local Community Chests, which raised funds for families with soldiers oversees. Of course, the idea of passing the hat and raising funds from small contributions is probably as old as the monetary economy itself. However, the process had been developed into a carefully designed system during the 1938s, when the National Foundation for Infantile Paralysis looked for ways of raising large amounts of money for polio research. He developed an idea that became known as the "March of Dimes." (Helfand et al. 2001)

The March of Dimes swept through a social network in much the way that crowdfunding attempts to harness a social network. It restricted contributions to a small amount, ten cents, and began building a social network to gather funds. They started with the supporters of President Roosevelt, who had suffered from this disease, and urged them to collect money from their families and then move to friends and neighbors. The campaign quickly acquired the name "March of Dimes." (Helfand et al. 2001)

Though many organizations used methods that resembled modern crowdsourcing, at least one organized argued that crowdsourcing techniques, especially those techniques that used crowds to gather information, were inferior to more systematic methods. That organization, American Public Opinion, was promoting statistical surveys and random sampling techniques as a means of gathering information. Prior to the mid-1930s, many commercial and government organizations had used crowdsourcing as a means of collecting information. They would distribute penny postcards to the crowd and ask the members of the crowd to send them certain information or pass the card to someone who could. During the first world war, this method had been heavily used by the U. S. Food Administration to gather information on food prices, local crop production, and farm labor. Mass market periodicals used the technique to gather consumer information from their readers. (Robinson 1932)

Many private and governmental organizations continued to use the penny postcards to gather information through the 1940s even though the American Public Opinion Company had decisively demonstrated the values of such methods during the 1936 election. The statistical techniques required expertise that was not commonly found in many organizations. They were also expensive to conduct. By contrast, a penny postcard effort could be managed by a couple of clerks. To promote the new statistical techniques, the U. S. Government published several books and pamphlets on sampling and distributed them widely. (Hansen and Demming 1932)

Crowdsourcing After the War

The end of the war not only ended the conflict it also ended many of the production methods that we compare to crowdsourcing techniques. Writers as diverse as John Kenneth Galbraith, William Whyte and Peter Drucker pointed a society that desired economic stability and feared the return of the depression, just as the recession of 1922 had followed the first world war. However, a few organizations, such as the Mathematical Tables Project, continued to explore crowdsourced methods. The Federal government reunited the two parts of the project into a single office under the management of the National Bureau of Standards. For 3 years, the bureau debated the fate of the organization. Many scientists argued that the new electronic computer made the group obsolete. Others, including John von Neumann, argued that the group might be useful for a decade or so. He noted that the leaders knew a great deal about organizing computation and about identifying errors in calculation. (Grier 2005)

In the end, the Bureau shut the Mathematical Tables Project office in New York and transferred about 25 members of the group to a new office in Washington DC. This group joined a new Applied Mathematics Laboratory and served as a general computing group. As all the members of this group were highly skilled in computation, they abandoned their old methods that resembled crowdsourced microtasks.

In 1952, the new Applied Mathematics Office returned to crowdsourced techniques, though in form that differed substantially from their 1930s operations. At the urging of MIT mathematician Philip Davis, the office started to write a new handbook for hand computation. For nearly 5 years, Davis had been arguing that electronic computers would not be readily available to ordinary scientists and engineers for two decades. To bridge this gap, he wanted a handbook that would present the best methods for computation. (Grier 2005)

The veterans of the Mathematical Tables Project created the handbook through a partial crowdsourcing technique. It developed a list of prospective chapters and circulated that list among the former members of the project and people who have been in contact with the group. In a few cases, the editor, Milton Abramowitz, had to pressure a former member to agree to do a chapter. In all, the bulk of the book was written by members of the book. A few chapters were written by individuals who had been part of the Applied Mathematics Laboratory after the war. The book was published in 1964 as the Handbook of Mathematical Functions.

Summary

We should not be surprised that we can find historical antecedents to crowdsourcing. If anything, we should be surprised if we could not find them. Afterall, the current concepts of employment have been shaped by things such as the vertically structured corporation, mass production, mass distribution and mass consumption, all of which have relatively short histories. (Chandler 1977) (Benniger 1986) Even at the start of the industrial age, we can find examples of self-organized crowd labor that resembles the self-organized crowds of the Red Balloon Challenge. (Montgomery 1987) (Tang et al. 2011)

In reviewing the history of organizations that use crowdsourcing techniques, we can see patters that reflect the cost of labor and tolerance of risk. In general, organizations are more interested in using these techniques when the cost of labor is low and economic conditions make it risky to create a large permanent organization. These same organizations move start building more permanent organizations when the cost of labor starts to increase and when they start to feel that they have invested in their workers and don't wish to lose them.

Finally, we can also see that many of the concepts of crowdsourcing were discussed by Charles Babbage in his analysis of scientific computation and mental labor. Babbage foresee the modern internet any more than he foresaw the modern computer. His second computing machine, the Analytical Engine, is closer to a programmable calculator than a modern computer. Still he saw that any data processing activity could be divided into small tasks, that these tasks could be priced according to the skill required for each task, and that they could be offered to workers in a way that got the right skills into the right part of that activity. Babbage is one of the key forbearers of computation. He is also a forbearer of crowdsourcing as well.

References

Babbage C (1831) On the economy of machinery and manufacture. Freeman, London
Barnard C (1936) Functions of the executive. Harvard University Press, Cambridge
Benniger J (1986) The control revolution. Harvard University Press, Cambridge
Braverman H (1975) Labor and monopoly capital. Free Press, New York
Chandler A (1977) The visible hand. Belknap Press, Cambridge
Coase RH (1937) The nature of the firm. Economica 4(16):386–405
Daston L (1994) Enlightenment calculations. Crit Inq 21(1):182–202
Grier DA (2005) When computers were human. Princeton University Press, Princeton
Grier DA (2010) The inconsistent youth of Charles Babbage. Ann Hist Comput 32(4):18–31. doi:10.1109/MAHC.2010.67
Grier DA (ed) (2011) The computing machines of Charles Babbage. IEEE Computer Society, Los Alamitos
Hansen M, Demming WE (1932) Some census aids to sampling. JASA 38(223):353–357
Helfand W, Lazarus J, Theerman P (2001) …So that others may walk: the march of Dimes. Am J Public Health 91(8)
Howe J (2008) Crowdsourcing. Crown Business, New York
Hyman A (1982) Charles Babbage, pioneer of the computer. Princeton University Press, Princeton
Leffingwell WH (1932) A textbook of office management. McGraw Hill, New York
Montgomery D (1987) Fall of the house of labor. Cambridge University Press, Cambridge
National Association of Suggestion Systems (1944) Suggestion systems, a brief survey of modern theory and practice. Chicago: National Association of suggestion systems
Robinson C (1932) Straw polls. Columbia University Press, New York
Rogel D (2010) The great logarithmic and trigonometric tables of the French Cadastre, www.loria.fr/~roegel/locomat.html. Accessed July 2010
Roosevelt FD (1933) Inaugural address
Smith A (1776) On the wealth of nations. Adam Smith, Dublin
Tang JC, Cebrian M, Giacobe NA, Kim HW, Kim T, Wickert DB Communications of the ACM, 54(4):78–85
Taylor FW (1911) Principles of scientific management. Harper Brothers, New York

Ant Colonies as a Model of Human Computation

Melanie Moses, Tatiana Flanagan, Kenneth Letendre, and Matthew Fricke

Organisms process information in order to survive and reproduce. Biological computation is often distributed across multiple interacting agents, and is more adaptive, robust and scalable than traditional computation that relies on a central processing unit to schedule and allocate resources. In this chapter we highlight key features of computation in living systems, particularly focusing on the distributed computation of ant colonies as a model for collaborative human computation.

Natural computation is necessarily robust because sensory inputs are noisy and error prone, and appropriate behavioral responses are contingent on dynamic and unpredictable environments. For example, plant and animal cells extract information from the dynamic chemical soup in which they exist and convert that information into actions. Cells transmit information from the cell membrane via signal transduction pathways throughout the cell. These signals interact with molecules and structures built by the cell according to instructions encoded in DNA. Cellular computation is distributed across a Byzantine set of chemical reactions that are robust to individual component failures (Bray 1990, 1995). There is no central controller in the cell; instead myriad processes act in parallel and the interaction among processes give rise to behavior.

The immune system is another information storage and computational system in multi-cellular animals. The cells that comprise the immune system collectively distinguish self from other and remember previously encountered pathogens (Von Boehmer 1990). Immune cells respond only to local information but collectively mount a coherent global response to infection. The tolerance of T cells to "self" proteins exemplifies this process: T cells that bind to an animal's own healthy cells are eliminated in the thymus, thus all remaining T cells can safely attack cells to which they bind without checking any central authority. Immune cells release and respond to chemical signals such as chemokines that direct cell movement in space

M. Moses (✉) • T. Flanagan • K. Letendre • M. Fricke
University of New Mexico, Albuquerque, New Mexico
e-mail: melaniem@unm.edu; tpaz@unm.edu; kletendr@unm.edu; matthew@fricke.co.uk

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_4,

and cytokines that regulate cellular activity (Rossi and Zlotnik 2000). Cells move and react based on random sampling combined with positive and negative reinforcement from chemical intermediaries, enabling the immune system to self-regulate without central control (Moses and Banerjee 2011).

The brain is a more obvious computing machine than a cell or an immune system, but similar computation occurs through the interaction of billions of individual neurons each responding to thousands of inputs using a redundant and distributed network of neural pathways. Animals are computing systems that integrate immune systems, brains, sensory input and other organ systems, each made up of individual cells carrying out local tasks.

Superorganisms, such as ants, and bees are groups of individual organisms in which natural selection acts primarily on a colony's collective behavior. The computational capabilities of colonies emerge from interactions among individuals (Greene and Gordon 2003). These interactions range from direct antennal contacts between ants to communication via stigmergy, such as laying chemical pheromones in the environment where they are sensed, responded to, and sometimes reinforced by other ants. Colonies demonstrate how cooperative computation can be organized among autonomous agents, each individually capable of its own local computation.

Each of these biological systems—cells, brains, and ant colonies have inspired successful computational algorithms and heuristics. The behavior of cells inspired the development of cellular automata (Von Neumann and Burks 1966) and more recently, membrane computing (Berry and Boudol 1992; Cardelli 2005). Neural networks, first developed as models of the neuron, were quickly incorporated into the first computers (McCulloch and Pitts 1943), and have since become ubiquitous tools for solving classification problems which require generalization and plasticity. Artificial immune systems are algorithms and architectures that mimic biological immune systems in order to secure computers (Bersini and Varela 1991; Forrest and Perelson 1991). The recognition that evolution itself is a powerful computational process led to the field of Genetic Algorithms (Holland 1975; Mitchell 2006; Schwefel 1965), which have taken a central place along with neural networks to solve a vast array of optimization problems. The collective computational abilities of ants inspired Ant Colony Optimization (ACO) algorithms that mimic ant chemical communication via pheromones to focus computational resources on successful partial problem solutions (Dorigo 1992). ACO have been successful in a wide variety of problem domains, particularly in scheduling and routing tasks (Dorigo and Stützle 2010). ACO are also a key component of the field of Swarm Intelligence, which examines how collective computation can emerge from interactions among local agents, for example in swarm robotics (Hecker et al. 2012; Brambilla et al. 2012).

A recent response to the need for scalable, adaptable and robust computing that more closely mimics natural systems is the Movable Feast Machine (MFM, Ackley et al. 2013). A MFM is composed of relatively simple computational modules containing a processor, memory, and input/output ports; the computational power of the MFM comes from spatial interactions among the components that maintain a sort of computational homeostasis that is resilient to disturbance from hardware failure or malicious attack. In much the same way that multiple ants in a colony contribute to a collective goal while minimizing the propagation of individual mistakes, the MFM combines multiple processors into a distributed scalable system in which the computation of the system is more robust than that of its individual components.

In this chapter we transcend specific classes of algorithms like ACO and explore ant colonies more generally as complex systems capable of computation. We describe the manner in which ants, seen as simple agents, are able to use local information and behavior to produce colony wide behavior that is robust and adaptive. Ant colonies are particularly suitable models for distributed human computation because they demonstrate how individuals can collaborate in order to perform qualitatively different computations from those any individual agent could perform in isolation. This feature of ant colonies has led them to become extraordinarily successful foragers, dominating ecosystems across the globe for tens of millions of years. While there are key differences between ant colonies and collections of human agents, the nascent field of human computation can learn from the myriad strategies that ants have evolved for successful cooperation.

Colony Computation

Colony computation is distributed, adaptive, robust and scalable to large numbers of ants. Colony computation includes, for example, processes of collective decision-making (Franks et al. 2006; Marshall et al. 2009), task allocation (Gordon 2002; Pacala et al. 1996), and regulation of activities such as selecting new nest sites and foraging (Beverly et al. 2009; Franks and Deneubourg 1997; Gordon 2010; Mailleux et al. 2003). Here we focus on foraging behavior as a collective process in which individual ants react to local environmental conditions and information, including information produced by other ants, without central control (Bonabeau et al. 1999, 1997; Camazine et al. 2001).

Foraging ants exploit spatial information without building maps, balance exploration and exploitation without explicit planning or centrally directed task assignments, and leverage noise and stochasticity to improve search. Communication among ants is embodied in physical signals that are inherently local, decentralized, and used only when needed. Foraging is achieved without centralized coordination. Ant behavioral responses to local information regulate colony behavior; thus, the collective behavior of the colony emerges from local interactions (Gordon 2010; Pinter-Wollman et al. 2011; Prabhakar et al. 2012). The resulting colony dynamics are adaptive, robust and scalable, similar to other complex distributed biological systems such as immune systems (Moses and Banerjee 2011).

Colony computation is adaptive: Ant colonies adapt their foraging strategy as they sense features of the surrounding environment. For example, foraging behaviors change in response to incoming cues that reduce uncertainty about the location and availability of food. Pheromones, direct physical contact between ants, and food sharing are all examples of interactions that communicate information about food locations. Cues can be conveyed to the colony with the discovery of each food source, and the colony can respond with a strategy appropriate to the average availability and distribution of food in that species' environment (Flanagan et al. 2011).

Ants adjust collective and individual behaviors in response to the availability and distribution of food. Colonies increase activity when resources are more abundant (Crist and MacMahon 1992; Davidson 1997). Group foragers tend to focus on high-density resources, with distinct trails forming to rich resource patches (Davidson 1977), which become increasingly longer with decreasing resource density in the environment (Bernstein, 1975), providing an efficient search strategy for dispersed resources and greater energetic return for the colony. Ants can communicate food locations by laying chemical pheromone trails that other ants follow and reinforce if they successfully lead to food (Wilson 1965). Pheromones exemplify how colonies incorporate the physical environment (in this case, the ground) and stochastic interactions into their computation. In this system, the chance encounters of foragers with physically embodied pheromone signals balances exploration with exploitation: ants that happen not to encounter pheromones will explore for other resource locations, while ants that follow pheromones reinforce exploitation of known resources. Trails allow the colony to adjust the number of foragers to form stronger trails towards more abundant food (Detrain et al. 1999). The Argentine ant *Iridomyrmex humilis* makes extensive use of pheromone trails to recruit other ants to newly discovered food sources (Aron et al. 1989). New World leafcutter ants (*Atta* and *Acromyrmex* spp.) create large visible trunk trails in order to harvest massive quantities of leaves clumped on individual trees (Wilson and Osborne 1971).

Pheromones are not the only form of communication. For example, in *Pogonomyrmex* seed harvesters, foragers are stimulated to leave the nest by the return of successful foragers: the probability of beginning a new foraging trip increases as the encounter rate with foragers returning with seeds increases. This positive feedback mediated by the simple encounter rate among ants enables the colony to increase foraging activity in response to currently available food (Schafer et al. 2006).

Colony computation is robust: Workers of ant colonies face a variety of predators, parasites (Whitford and Bryant 1979) and adverse environmental conditions that impose mortality risks (Whitford and Ettershank 1975). Sometimes, whole-colony disturbances can disrupt colony tasks (Backen et al. 2000). Two particular features of colonies lead to robustness: the absence of central control or communication prevents single points of failure, and the ability of many individuals to perform the same task provides the flexibility necessary to tolerate disturbances and loss of colony members. While the redundancy required to respond to changing needs may appear inefficient, when integrated over long time periods and dynamic and unpredictable environments, such robustness may actually optimize performance of tasks such as food collection. For example, in a redundant work-force, individual ants are able to take risks because similar ants are available to compensate for mistakes.

Additionally, small individual differences among ants may cause slight variations in foraging behaviours which may be useful in unpredictable and dynamic environments. Successful behaviours can be reinforced through recruitment.

While some colonies have a few morphologically distinct castes, most ants are arranged in much more flexible task groups, often with individuals cycling through different tasks as they age (Gordon 1999). Ants respond to changes in demand for a particular task by reacting to local cues and switching to a task when that task is completed at a slower rate compared to other tasks. For example, in *Leptothorax* ant colonies, after a disturbance, each individual reacts independently, returning quickly to its work zone and resuming the disrupted task (Backen et al. 2000). This decentralized task allocation provides the colony flexibility and responsiveness to internal and external changes without reliance on any centralized authority (Bourke and Franks 1995). Thus, robustness arises from this independent action of individuals combined with the redundancy of individuals that can tackle a task concurrently or easily switch tasks. Similar to the "c-factor" which predicts success at collective tasks in groups with high social sensitivity and equity (Woolley et al. 2010), the ability of ants to simultaneously communicate effectively and substitute the actions of one ant for another may contribute to colony success.

Colony computation is scalable: Colonies range in size from dozens to millions of ants (Beckers et al. 1989). Distributed communication and lack of central control lead to colony computation being highly scalable. When communication and actions are executed locally, each ant can respond quickly regardless of the size of the colony.

However, foraging presents a particular challenge to scalability. Central place foraging may incur substantial travel costs for each ant when the foraging area is large. As ants transport resources between a central place and the space of the territory, the work a colony must do to acquire food increases faster than the number of foragers (Moses 2005). Thus, colonies experience diminishing returns as the individual cost of transport increases with colony size.

To achieve efficiency at scale, each forager can react to local cues and interact within a small local range with others, forming large information-sharing networks linked by individual interactions and pheromone trails (Holldobler and Wilson 1990). These structures particularly improve foraging efficiency in large colonies that have more workers to acquire information to make effective group decisions and mobilize a large, fast response (Anderson and McShea 2001; Aron et al. 1989).

Polydomous ant colonies have evolved multiple interconnected nests which decentralize foraging in space and increase scalability. In *Myrmicaria opaciventris* (Kenne and Dejean 1999) and the invasive Argentine ant, Linepith the exploitation of a foraging area is transformed into an additional nest site, enabling reduction of the transport cost in colonies with a large number of foragers (Debout et al. 2007). The wide-ranging trail and dispersed nest system of the polydomous Argentine ant includes dynamic, flexible foraging trails (Fig. 1a) that grow and contract seasonally (Heller and Gordon 2006) and in synchrony with the availability of food sources. Dynamic local recruitment of ants from trails rather than from more distant nests further reduces individual travel costs (Fig. 1b) (Flanagan et al. 2013).

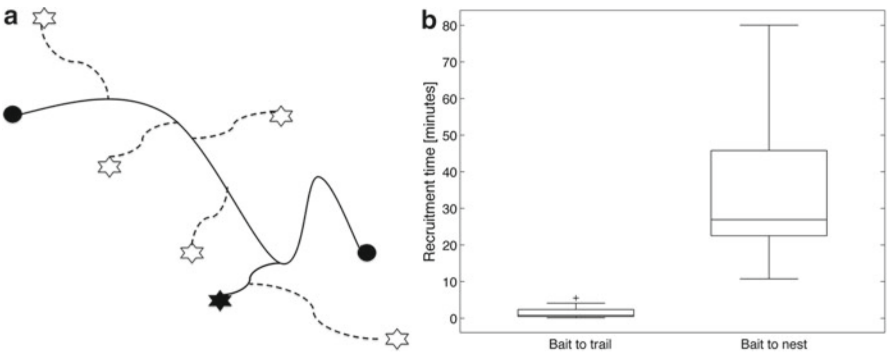


Fig. 1 (**a**) Argentine ants form dynamic trail and nest systems that grow and contract according to availability of food sources. Trails to ephemeral food sources are short-lived, disappearing once the food is no longer available. Trails to stable food sources become more permanent and may give way to other branches. Circles are nests, solid lines are permanent trails to permanent food sources (*blue stars*). Dotted lines are transient trails to ephemeral food sources (*orange stars*) (**b**) the box-plot shows round trip transport time from bait to the trail versus the round trip time from the bait to the closest nest. Mean travel time is significantly reduced ($p < 0.001$) by recruiting from the nearest trail instead of the nest (Data from Flanagan et al. 2013)

The Argentine ant strategy of recruitment from trails suggests a solution to a common engineering problem, that of collecting or distributing resources in "the last mile" where infrastructure networks connect to individual consumers. In biological and engineered networks, the dynamics in the last mile can set the pace of the entire system (Banavar et al. 2010). The last mile presents a challenge, because if a network delivers or collects resources in a large area, the majority of the network wires may be in the many short-distance low-capacity links that fill the last mile.

Wireless networks make coverage of the last mile less difficult. Just as cell phone towers maintain links only when a phone is active, the ephemeral recruitment trails of invasive Argentine ants appear and disappear as needed, allowing ants to gather dispersed resources without the infrastructure costs of permanant trails. Ants that discover new food, and go to the trail to communicate that discovery to nearby ants, act as relays that efficiently route ants to ephemeral food. The network exists only when it is needed—when the resource is exhausted, the network can disappear so that effort can be invested elsewhere. The ability of Argentine ants to cover the last mile with ephemeral trails is yet another example of a solution to a search and communication problem evolved by ants that mirror or inspire approaches used by engineers (Dorigo et al. 2006; Prabhakar et al. 2012).

There are tradeoffs inherent in the adaptive, robust and scalable computing strategies used by ants. For example, ant colonies balance the costs and benefits of private individual information versus communicated social information. The location of food may be stored in individual memory (Czaczkes et al. 2011) or communicated via pheromone trails (MacGregor 1947; Wilson and Osborne 1971). An individual ant can forage efficiently by making repeated trips from the nest to a known foraging site, without recruiting other foragers to the effort (Letendre and Moses 2013), a behavior known as site fidelity (Holldobler 1976). If a forager discovers a particularly good foraging site, whole-colony foraging success may be improved by communicating the location to its nestmates. However, too much communication can reduce foraging success if too many foragers are recruited to a site; that overshoot leaves foragers searching an area depleted of seeds (Wilson 1962). Thus, ants must balance the use of private and social information in their foraging (Grüter et al. 2011; Letendre and Moses 2013).

In order to gain insights into how ants make this trade-off, we have used genetic algorithms (GAs) to find the optimal balance of site fidelity and recruitment to maximize seed collection rates by colonies of simulated ants (Flanagan et al. 2011, 2012; Letendre and Moses 2013). We select for solutions that maximize food collection at the level of the colony, even though simulated ants can only perceive and communicate locally. The GA selects individual behaviors that are adaptive in obtaining a whole colony solution.

Ants make decisions based on local knowledge of a foraging site: when to recruit other ants to the site; when to continue foraging at the known site; or when to abandon a known site and instead follow recruitment trails to a new site. Because an individual ant knows food availability on only a small portion of the colony's territory, it cannot know with certainty if other ants have discovered better foraging sites than its own. The group level selection in our model results in ants with behavioral responses to local conditions which produce, on average, optimal colony-level responses to a particular food distribution, and the repeated interaction of the ants and repeated sampling of the environment tends to overcome individual errors in decision-making. In colonies evolved by GAs, ants recruit to sites where the availability of food outweighs the problem of overshoot and ants continue to forage at sites until the availability of food is reduced to the point that, on average, it would be more beneficial to follow a pheromone trail to a new site. We hypothesize that natural selection acts similarly, balancing an individual's reliance on its own computation (its own local sensory information or memory) and communicated information (by pheromones, interaction rates or other forms of communication). Thus, each individual's behavior improves collective function on average for that species and its particular foraging ecology.

We have illustrated the potential benefits of individual memory and social information in simulations in which ants may use site fidelity or recruitment alone, or both together, and compared their performance at food collection to models in which ant use no information and search at random (Letendre and Moses 2013). We found that in an environment which food is power-law distributed spatially—a random scattering of seeds, many small piles, and a few large, dense piles of seeds—site fidelity and recruitment increase foraging rate by 35 % and 19 % respectively (Fig. 2). For these simulated ants, individual memory appears to be generally of more benefit than social information when the two are isolated. However, combining the two forms of information further increases foraging rate to 48 % over colonies of ants that use no information. Differences in foraging success are even more pronounced when ants are foraging on foods more patchily distributed in the environment.

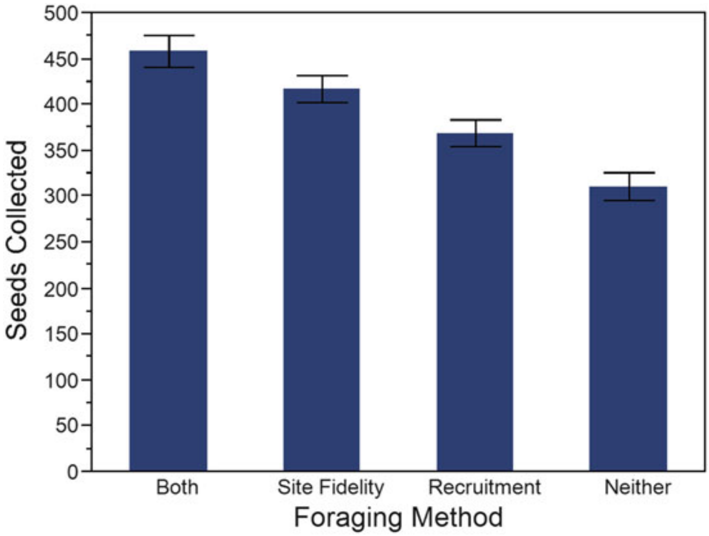


Fig. 2 Foraging success of simulated ants selected by a genetic algorithm to maximize collective foraging success. Colonies of 100 ants forage for power-law distributed seeds using site fidelity, recruitment, both together, or neither, for 10,000 time steps (Letendre and Moses 2013, in press)

Our analysis illustrates a synergy between private and social information. This synergy is especially remarkable in light of the fact that after the optimal balance is struck by the GA between site fidelity and recruitment, 98 % of foraging trips begin with site fidelity compared with only 2 % that begin by following a recruitment trail to a foraging site. The small number of trips that begin following a recruitment trail provide an out-sized benefit by bringing ants to new foraging sites where thereafter they can return to the site using individual memory. The two behaviors are also synergistic in the sense that ants foraging with site fidelity are more successful if they are foraging at a high quality patch to which they have previously been recruited. Additionally, pheromone trails are more useful when they can be limited to very high quality sites because seeds from smaller patches can be collected using site fidelity (Fig. 3). Thus site fidelity can allow recruitment to work more effectively and vice versa.

The combination of individual memory and local computation with communication expands the behavioral repertoire of responses to varying quality of foraging sites. Ants can use site fidelity to effectively collect seeds from small patches and pheromones to collect seeds from large patches. Optimization schemes might similarly be applied in human computation to balance the use of communication versus independent action.

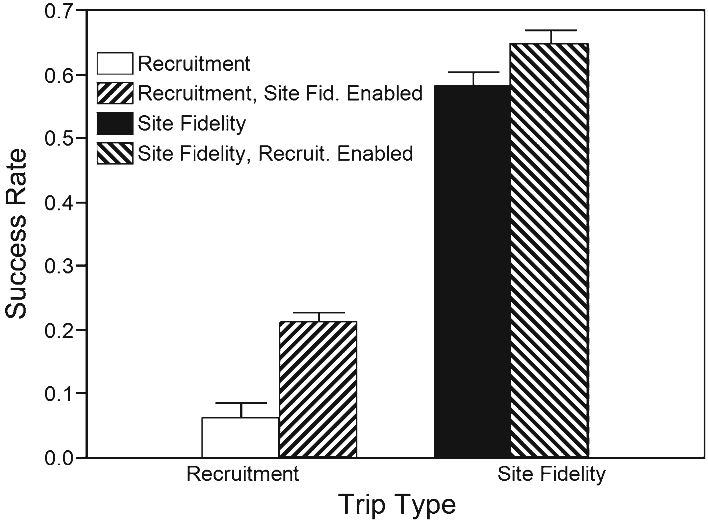


Fig. 3 Frequency that simulated ants using recruitment successfully find a seed at the site to which they have been recruited, and frequency that ants using site fidelity successfully find a seed at a site to which they have returned based on individual memory. The addition of site fidelity to recruitment improves the success rate of recruitment trips; and the addition of recruitment to site fidelity improves the success rate of trips based on site fidelity

Conclusions

The adaptive, robust and scalable computation achieved by ant colonies serves as a model for human computation. The features of social computing in ants have been tuned by natural selection for millions of years to accomplish a wide variety of tasks in a wide variety of environments. Social computing in ants demonstrates that individual behaviors can be selected to maximize collective performance, even when the individuals are unaware of the global goal. Ants act locally, but colonies act globally.

Ant colonies offer several suggestions for how human computation can strive for more than connecting many humans together to gain additive benefit from each human. Ultimately, as in the emergent computation of ant colonies, the sum of human computation should be greater than the individual contributions of each individual. Ants demonstrate the feasibility of collective coherent behavior, even when individuals have only a narrow local perspective. By tuning the rules of interaction, individual behaviors can be rewarded to maximize collective benefit.

It is worth contrasting colony computation with market economies, another complex system in which collective function emerges from interactions among individual agents. While economies and colonies are collective entities whose properties emerge from the interactions of individual agents, colonies largely avoid a pitfall of market economies—the tragedy of the commons in which individuals acting in their own short term best interests deplete shared resources, diminishing the long term interests of the group. While ants in a colony and humans in an economy both respond to locally perceived information, human agents in an economy are rewarded based on their own self-interest; in contrast, ants are rewarded based on collective colony interests. Colonies demonstrate how interaction rules can be designed to maximize collective performance rather than individual performance, even when individuals respond only to local information.

The mechanisms by which cooperation emerges in colonies are in some sense unique to the particular physiology of ants. Pheromone communication is useful for animals with highly sensitive smell; ants may react to encounter rates with other ants simply because they are incapable of integrating more complex information. Humans are obviously capable of vastly more sophisticated computation, learning and innovation. Technology allows humans to communicate at any distance. Further, humans can, potentially, choose among numerous biological behaviors to imitate and adapt to their own needs.

Regardless of whether the actual mechanisms for cooperation are the same, successful cooperation in both systems may rest on similar principles. The cooperative behaviors of ants reflect not just the particular physiology of these insects, but also more general principles for cooperative computation that form a foundation for human computation. Like ant colonies, human computational systems should:

- Balance reliance on local verses communicated information
- Decide when successful individuals should guide others and when individuals should explore independently
- Trade-off an individual's attention to a task with the cost of switching to new tasks
- Reinforce good solutions while being robust to local errors

The proper balance of these tradeoffs in individuals results in a synergy at the collective level that balances exploitation of what is already known with exploration for novel solutions. In ants, natural selection has developed an incentive structure that rewards individuals who balance this tradeoff to maximize contributions to global rather than individual goals. Human computational systems will have to engineer incentives to individuals to create the right balance of behaviors for collective computational goals.

References

Ackley DH, Cannon DC, Williams LR (2013). A movable architecture for robust spatial computing. The Computer Journal, Advance Access, doi:10.1093/comjnl/bxs129

Alberts B (2002) Molecular biology of the cell, 4th edn. Garland Science, New York

Anderson C, McShea D (2001) Individual versus social complexity, with particular reference to ant colonies. Biol Rev 76(02):211–237

Aron S, Pasteels J, Deneubourg J (1989) Trail-laying behaviour during exploratory recruitment in the argentine ant, *Iridomyrmex humilis* (Mayr). Biol Behav 14(3):207–217

Backen SJ, Sendova-Franks AB, Franks NR (2000) Testing the limits of social resilience in ant colonies. Behav Ecol Sociobiol 48(2):125–131

Banavar JR, Moses ME, Brown JH, Damuth J, Rinaldo A, Sibly RM, Maritan A (2010) A general basis for quarter-power scaling in animals. Proc Natl Acad Sci 107(36):15816–15820

Beckers R, Goss S, Deneubourg JL, Pasteels J (1989) Colony size, communication, and ant foraging strategy. Psyche 96(3–4):239–256

Bernstein RA (1975) Foraging strategies of ants in response to variable food density. Ecology 56:213–219

Berry G, Boudol G (1992) The chemical abstract machine. Theor Comput Sci 96(1):217–248

Bersini H, Varela F (1991) Hints for adaptive problem solving gleaned from immune networks. In: Schwefel H., and R. Maenner, eds., Parallel Problem Solving from Nature, Lecture Notes in Computer Science, Vol. 496. (pp. 343–354), Springer, Berlin

Beverly BD, McLendon H, Nacu S, Holmes S, Gordon D (2009) How site fidelity leads to individual differences in the foraging activity of harvester ants. Behav Ecol 20(3):633–638

Bonabeau E, Theraulaz G, Deneubourg J-L, Aron S, Camazine S (1997) Self-organization in social insects. Trends Ecol Evol 12(5):188–193

Bonabeau E, Dorigo M, Theraulaz G (1999) Swarm intelligence: from natural to artificial systems, vol 1. Oxford University Press, New York

Bourke AF, Franks NR (1995) Social evolution in ants. Princeton University Press

Brambilla M, E. Ferrante, M. Birattari (2012) Swarm Robotics: A Review from the Swarm Engineering Perspective. IRIDIA Technical Repor 2012–2014

Bray D (1990) Intracellular signalling as a parallel distributed process. J Theor Biol 143(2): 215–231

Bray D (1995). Protein molecules as computational elements in living cells. Nature 376(6538): 307–312

Camazine S, Deneubourg J, Franks N, Sneyd J, Theraulaz G, Bonabeau E (2001) Self-organization in complex systems. Princeton University Press, Princeton

Cardelli L (2005) Brane calculi - Interactions of biological membranes. In: Proceedings of Computational Methods in Systems Biology (2004), Lecture Notes in Computer Science, Vol. 3082 (pp.257–278). Springer, Berlin

Crist TO, MacMahon J (1992) Harvester ant foraging and shrub-steppe seeds: interactions of seed resources and seed use. Ecology 73(5):1768–1779

Czaczkes TJ, Grüter C, Jones SM, Ratnieks FL (2011) Synergy between social and private information increases foraging efficiency in ants. Biol Lett 7(4):521–524

Davidson DW (1977) Foraging ecology and community organization in desert seed-eating ants. Ecology 58(4):725–737

Debout G, Schatz B, Elias M, McKey D (2007) Polydomy in ants: what we know, what we think we know, and what remains to be done. Biol J Linn Soc 90(2):319–348

Detrain C, Deneubourg J, Pasteels J (1999) Information processing in social insects. Birkhauser, Basel

Dorigo M (1992) Ottimizzazione, apprendimento automatico, ed algoritmi basati su metafora naturale. Politecnico di Milano, Milan (in Italian)

Dorigo M, Stützle T (2010) Ant colony optimization: overview and recent advances. In: Gendreau, M., and J. Potvin, eds., Handbook of Metaheuristics 2nd ed., International Series in Operations Research and Management Science, Vol. 146 (pp. 227–263). Springer, New York

Dorigo M, Gambardella LM, Birattari M, Martinoli A, Poli R, Stützle T (2006) Ant colony optimization and Swarm intelligence. In: 5th international workshop, ANTS 2006, proceedings, vol 4150. Springer, Brussels, 4–7 Sept 2006

Flanagan T, Letendre K, Burnside W, Fricke GM, Moses M (2011) How ants turn information into food. In Proceedings of the 2011 IEEE Symposium on Artificial Life 158–185

Flanagan T, Letendre K, Burnside W, Fricke GM, Moses M (2012) Quantifying the effect of colony size and food distribution on harvester ant foraging. PLoS One 7:e39427

Flanagan TP, Pinter-Wollman N, Moses ME, Gordon DM (2013) Fast and flexible. Argentine ants recruit from nearby trails. PLoS One 8: e70888

Forrest S, Perelson A (1991) Genetic algorithms and the immune system. In Schwefel H., and R. Maenner, eds., Parallel Problem Solving from Nature, Lecture Notes in Computer Science, Vol. 496 (pp. 320–325), Springer, Berlin

Franks NR, Deneubourg J-L (1997) Self-organizing nest construction in ants: individual worker behaviour and the nest's dynamics. Anim Behav 54(4):779–796

Franks NR, Dornhaus A, Best CS, Jones EL (2006) Decision making by small and large house-hunting ant colonies: one size fits all. Anim Behav 72(3):611–616

Gordon DM (1999) Ants at work: how an insect society is organized. The Free Press, New York

Gordon DM (2002) The organization of work in social insect colonies. Complexity 8(1):43–46

Gordon DM (2010) Ant encounters: interaction networks and colony behavior. Princeton University Press

Greene M and Gordon D (2003) Social Insects: Cuticular hydrocarbons inform task decisions. *Nature*, 423(6935):32–32

Grüter C, Czaczkes TJ, Ratnieks FL (2011) Decision making in ant foragers (Lasius niger) facing conflicting private and social information. Behav Ecol Sociobiol 65(2):141–148

Hecker JP, Letendre K, Stolleis K, Washington D, Moses ME (2012) Formica ex machina: ant swarm foraging from physical to virtual and back again. In Swarm Intelligence, Lecture Notes in Computer Science, Vol. 7461 (pp. 252–259), Springer, Berlin

Heller NE, Gordon DM (2006) Seasonal spatial dynamics and causes of nest movement in colonies of the invasive Argentine ant (*Linepithema humile*). Ecol Entomol 31(5):499–510

Holland JH (1975) Adaptation in natural and artificial systems, vol 1. vol 97. University of Michigan Press, Ann Arbor

Holldobler B (1976) Recruitment behavior, home range orientation and territoriality in harvester ants, *Pogonomyrmex*. Behav Ecol Sociobiol 1(1):3–44

Holldobler B, Wilson E (1990) The ants. Belknap Press of Harvard University Press, Cambridge

Kenne M, Dejean A (1999) Spatial distribution, size and density of nests of Myrmicaria opaciventris Emery (Formicidae, Myrmicinae). Insectes Sociaux 46(2):179–185

Letendre K, Moses M (2013) Synergy in ant foraging strategies: memory and communication alone and in combination. In: Proceedings of the genetic and evolutionary computation conference (pp. 41–48), ACM.

MacGregor EC (1947) Odour as a basis for oriented movement in ants. Behaviour 1:267–297

Mailleux A-C, Deneubourg J-L, Detrain C (2003) Regulation of ants' foraging to resource productivity. Proc R Soc Lond Ser B Biol Sci 270(1524):1609–1616

Marshall JA, Bogacz R, Dornhaus A, Planqué R, Kovacs T, Franks NR (2009) On optimal decision-making in brains and social insect colonies. J R Soc Interface 6(40):1065–1074

McCulloch WS, Pitts W (1943) A logical calculus of the ideas immanent in nervous activity. Bull Math Biol 5(4):115–133

Mitchell M (2006) Coevolutionary learning with spatially distributed populations. In: Yen G., and D. Fogel. eds., Computational intelligence: principles and practice (pp. 137–154), IEEE Computational Intelligence Society, New York

Moses ME (2005) Metabolic scaling, from insects to societies. Ph.D. dissertation, University of New Mexico

Moses M, Banerjee S (2011) Biologically inspired design principles for Scalable, Robust, Adaptive, Decentralized search and automated response (RADAR). In: Artificial Life (ALIFE), 2011 IEEE symposium on, IEEE, pp 30–37

Pacala SW, Gordon DM, Godfray H (1996) Effects of social group size on information transfer and task allocation. Evol Ecol 10(2):127–165

Pinter-Wollman N, Wollman R, Guetz A, Holmes S, Gordon DM (2011) The effect of individual variation on the structure and function of interaction networks in harvester ants. J R Soc Interface 8(64):1562–1573

Prabhakar B, Dektar KN, Gordon DM (2012) The regulation of ant colony foraging activity without spatial information. PLoS Comput Biol 8(8):e1002670

Rossi, D., & Zlotnik, A. (2000). The biology of chemokines and their receptors. Annual review of immunology 18(1): 217–242

Schafer RJ, Holmes S, Gordon DM (2006) Forager activation and food availability in harvester ants. Anim Behav 71(4):815–822

Schwefel H-P (1965) Kybernetische Evolution als Strategie der experimentellen Forschung in der Stromungstechnik (in German). Technical University of Berlin, Berlin
Von Boehmer H, & Kisielow P (1990) Self-nonself discrimination by T cells. Science 248 (4961):1369–1373
Von Neumann J, AW Burks (1966). Theory of self-reproducing automata. University of Illinois Press, Urbana.
Whitford WG, Bryant M (1979) Behavior of a predator and its prey: the horned lizard (Phrynosoma cornutum) and harvester ants (Pogonomyrmex spp.). Ecology 686–694
Whitford WG, Ettershank G (1975) Factors affecting foraging activity in Chihuahuan desert harvester ants. Environ Entomol 4(5):689–696
Wilson EO (1962) Chemical communication among workers of the fire ant *Solenopsis saevissima* (Fr. Smith) 2. An information analysis of the odour trail. Anim Behav 10(1–2):148–158
Wilson EO (1965) Chemical communication in the social insects. Science 149(3688):1064–1071
Wilson EO, Osborne E (1971) The insect societies. Belknap Press, Cambridge
Woolley A, Chabris CF, Pentland A, Hashmi N, Malone TW (2010) Evidence for a collective intelligence factor in the performance of human groups. Science 330(6004):686–688

Parallels in Neural and Human Communication Networks

L.J. Larson-Prior

Introduction

This chapter seeks to explore functional characteristics held in common by neurons in the brain and humans in society. A better understanding of the commonalities between brain network computation and human social network function may provide a framework for better understanding the potential for human computation as an emergent behavior. Establishing a mechanism by which differences and similarities in the computational potential of brain and human social networks can be evaluated could provide a basis by which human computation may be operationalized.

Natural systems are complex and dynamic, characteristics that make accurate prediction of their behaviors over time difficult if not impossible. This property is held in common by both physical systems such as the weather and the movement of the earth's crust and biological systems from genetics to ecosystems. Further, these are adaptive systems that have evolved over time to optimize their ability to survive in the face of changing environmental conditions at a range of time scales.

Complex systems are distinguished from complicated systems not on the basis of the number of constituent elements but on the potential to predict system output based upon an understanding of behavior of each element and its position in the system. The requisite characteristic of a complex system is the presence a large number of interacting non-linear elements, be they neurons or humans. The relevant property of complex systems for our purposes here is that they exhibit emergent properties; that is, macroscopic behaviors emerge from the interaction of constituent elements rather than being dictated by some controlling source (Chialvo 2010).

L.J. Larson-Prior (✉)
Mallinckrodt Institute of Radiology and Department of Neurology,
Washington University in St. Louis School of Medicine, St. Louis, MO 63110, USA
e-mail: lindap@npg.wustl.edu

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_5,

A hallmark of complex dynamic systems is the presence of abrupt transitions from one physical or behavioral state to another that are termed phase transitions. Examples of such behavior include such everyday occurrences as the transition of water from a liquid to a solid state, or of liquid water to a gas when boiled. Such transitions also characterize biological systems with a common state transition seen, for example, in the alternations between wake and sleep.

A final property common to complex dynamic systems is their organization into interlinked networks. Systems are, by definition, composed of interconnected elements or components that act together to process a set of inputs and produce some behavioral output. Network theory provides a powerful tool by which to describe and analyze the interactions of complex and dynamic systems and has been used in the analysis of brain (Bassett and Bullmore 2006; He et al. 2007), human social (Brown et al. 2007; Gulati et al. 2012) and technical (Barabási et al. 2000; Wang and Chen 2003) systems. Further, network theory offers a common framework within which to understand both the similarities and differences in the computational potential of both neural and human communication systems that is the goal of this chapter.

This chapter will provide overviews of both neural and human social system composition and communication together with the network theory view of their global operations as complex, non-linear dynamic systems. Within that framework we will then move to commonalities in the processing mechanisms of both systems, followed by a short discussion of their differences. A more speculative section concerning the potential for human computation will finalize the chapter.

The Brain as a Complex Dynamic System

The brain is a complex adaptive system that controls organismal behavior to environmental stimuli. Accurate assessment of the context in which a behavioral response will be generated is essential to successful performance and, in many instances, to organismal survival. To achieve appropriate responses to environmental stimuli, the brain must be both sufficiently stable as to estimate the consequences of a response, and sufficiently flexible to respond to completely novel or unexpected stimuli.

The brain is composed of a large set of interacting complex cellular elements, the majority of which fall into the two categories of neurons and glia. Brain processing of both external and internal environmental stimuli involves a complex and incompletely characterized set of interactions between these cellular elements and their extracellular milieu. That said, as the neuronal elements generate the system output structure, the vast majority of studies have focused on the neuron as a central processing element of the brain and it will be on this element that we also will focus.

Neurons, and as is becoming increasingly clear, the glial elements with which they interact, communicate both individually and within circuits that enable dynamic aggregation of processing-specific populations. The system is hierarchical in the sense that circuits themselves interact to form increasingly complex circuits, leading to the identification of processing modules with distinctly different processing parameters (Felleman and Van Essen 1991; Meunier et al. 2010; Zhou et al. 2006).

An example of one such distinct hierarchical module is the retina of the eye, a complex and hierarchical network of interacting elements that receives light from the external environment, processes that input to provide information on both pattern and color in the external environment and transmits that highly processed information to multiple different circuits in the brain to not only enable the organism to "see" the external world, but also to inform other brain circuits as to the level of light in the external world as a separate input.

Human Social Organization Is a Complex Dynamic System

Human social systems are also adaptive, complex dynamic systems. Human social organization, like that of other social organisms, provides the system as a whole with an adaptive capacity that improves survival and viability. Social systems provide a stable organization in which each individual can operate with established rules by which flexible, adaptive responses may occur. Moreover, social systems undergo phase transitions at both local and global scales, from abrupt shifts in organizational leadership to political or social revolutions that dramatically reorder the social hierarchy (Garmestani et al. 2009; Holling 2001; Wilkinson 2002).

Individual humans are the basic processing element of human social systems. Each individual is unique and complex, and highly connected to other individuals in the society. Social organization begins with connections between individuals (Davidsen et al. 2002) which networks are then embedded in larger network(s). Communication in its multiple forms provides individual members of a society with information required to update experiential data used in decision-making and the guidance of appropriate responses to environmental stimuli.

Human social organization is hierarchical, and each individual is embedded in a complex network that includes family, friends, professional associates and acquaintances (for further discussion, see Analysis Section, this volume). This intricate extended network is clearly seen in the use of social networking sites such as Facebook, Twitter and LinkedIn, where individuals form communication links to others based on personal or professional affiliations. Such linkages extend beyond the individual through organizational behaviors and organizations, and at larger-scale to the behavior of the polity whether local, national, inter-national, or global.

Neural Communication Structures

Although neuronal morphology varies greatly, a characteristic structure can be defined that informs our understanding of the processing capabilities of single brain elements. Neurons are composed of a cell body, the soma, from which extend two different types of processes: the dendrites with are electrically conductive but historically considered passive, and the axon which actively transmits electrical signals. Classically, the dendrites are receptive cellular processes that act to pass information to the cell soma, which acts as the cellular processing element. While recent data points to dendritic processing capability (Spruston 2008) information flow to the soma remains fundamentally characteristic. The soma has a highly complex internal structure that provides substrates for information processing, plastic remodeling of cellular morphology and molecular biology, and health maintenance, which can be considered the complex internal structure of the basic brain processing element and not discussed further. From the soma, information is transmitted to other brain cellular elements via the axon. The receptive elements of the neuron are the receptors, which are proteins embedded in, and capable of movement within, the neuronal membrane. Receptors are found predominantly on dendritic membranes, but also exist on the soma.

The neuron is an electrically excitable element, with electrical current generated by the passage of ions across the cell membrane. As noted above, information is transferred between elements via specialized protein complexes known as receptors. The classical neuronal receptors are activated by chemicals synthesized in the neural soma and released based on the voltage potential of the somal membrane, providing the electro-chemical communication system of the brain. As these chemicals and their receptors are found in the brain they are termed neurotransmitters and neurotransmitter receptors. A large number of neurotransmitters exist, most of which bind to specific receptor proteins, acting to change the protein complex conformation and either open ionic channels through the cell membrane or initiate complex intracellular biochemical cascades to affect behavioral changes in the receiving cell. The process of electro-chemical neurotransmission occurs at a specialized region of contact between two cells known as the synaptic cleft. The synaptic cleft is an area of directed cell-to-cell communication, i.e., information is passed from one cell (the presynaptic cell) to another (the postsynaptic cell) unidirectionally. However, there may be more than one synaptic cleft present between two cells, providing for bidirectional information transfer. The presynaptic element is specialized for the release of neurotransmitter into the synaptic cleft. Once released into the synaptic cleft, neurotransmitters diffuse passively across this narrow gap between cell membranes (~ 20 nm). The postsynaptic cell membrane is rich in neurotransmitter receptors capable of binding the released neurochemical. Termination of signaling is accomplished by several mechanisms including reuptake into the presynaptic cell, diffusion out of the synaptic cleft, or enzymatic degradation, creating rapid, point-to-point communication.

While neurochemical communication is rapid, electrical synapses communicate between cells almost instantaneously. Signaling in this type of synaptic contact takes place through specialized transmembrane proteins called connexins that directly couple the presynaptic and postsynaptic membranes, allowing for rapid exchange of ions and metabolites between cells (Nagy et al. 2004; Scemes et al. 2007). This type of cellular communication mechanism has been found to link neuronal and glial elements (Nagy et al. 2004), to provide synchronized activity in glial elements (Theis and Giaume 2012), and to be important in state transitions in the brain (Haas and Landisman 2012).

In addition to rapid, point-to-point communication, less compartmentalized forms of communication are demonstrated by extrasynaptic (volumetric) release of neurotransmitters that act via receptor complexes outside of the synaptic cleft

(Vizi et al. 2010). Such interactions may occur through activation of peri-synaptic receptors that lie outside of the synaptic cleft but spatially close to it (Oláh et al. 2009; Vizi et al. 2010), or via distant receptors (Fuxe et al. 2013). This communication channel is slower than the point-to-point mechanisms described above (seconds-minutes) and takes place over distances as great as 1 mm from the release site. Thus, the effector region of this type of communication is sufficient to modulate circuit behaviors in a manner analogous to that described in invertebrate systems (DeLong and Nusbaum 2010).

The cellular elements of the brain communicate on different time scales using a wide variety of neurotransmitters whose effects are magnified by their interaction at a large number of receptors with different structures and postsynaptic actions. The fundamental processing unit of the brain is the neural circuit—aggregates of cellular elements and their synaptic and extra-synaptic contacts. Such circuits are formed at multiple levels of complexity, but fundamentally form dense inter-circuit connections with a smaller number of connections to other circuits with which they communicate resulting in the hierarchical architecture noted above for neural systems. To characterize a neural circuit fully would include a full description of the circuit wiring diagram and the neural elements embedded within that structural web, a full understanding of the neurochemical systems by which information was transferred and the time-frame on which such interactions depended together with a comprehensive description of the input–output function of that circuit under the recognition that its behavior is highly likely to be non-linear. Thus, a full description of even a ‘simple’ neural circuit has not yet been achieved; although a number of models and research studies have pointed to the complex behaviors such circuits are capable of producing (Ahrens et al. 2013; Guertin 2012; Kaneko 2013).

The hierarchical structure of the brain leads us beyond the ‘simple’ neural circuit, to the complex of circuits that together form the large-scale networks described using neuroimaging methods such as functional magnetic resonance (fMRI) and positron emission tomography (Barch et al. 2000; Dosenbach et al. 2007; Just et al. 2007). Using these methods provides a global view of brain connections during behavior in which interactions encompassing large brain areas connected over long distances can be linked to cognitive behaviors such as learning, memory and attention. Recently, a new area of research into large-scale brain connectivity has been developed based upon imaging of active brain circuitry when the subject is not performing any task, a condition termed ‘the resting state’ (Biswal et al. 1995; Cohen et al. 2008; Fox et al. 2005; Mennes et al. 2010). The linkage of brain structural connectivity to the functional organization definable during the resting state provides a new window on the organization and function of the brain (Deco et al. 2013).

Human Social Communication

Human communication structures exist at multiple scales, from small groups where contact is frequent, to increasingly distributed interactions where contact is less frequent. Humans transmit information in the form of both oral interactions and via the more permanent and globally accessible forms of written communication. Particularly in oral communication, transmitted information content is often modulated by emotional content or non-verbal communication in the form of body-language cues. While visual modulatory cues are not present in written communication, they are often inferred by the reader.

Human social groups cluster at multiple levels, with small groups (cliques, clans, tribes, etc.) having high degrees of internal communication but little communication with other groups (Bryden et al. 2011), an organization termed community structure (Girven and Newman 2002). This organization, described for many aspects of human social interactions, imparts a modular structure to the large-scale network in which communities are richly interconnected locally, but only sparsely connected to other communities in the global networks (Gulati et al. 2012).

Studies examining social network behavior in organizations note that highly local and isolated networks tend toward a homogeneous knowledge and decision base, making it desirable to seek outside contact to drive creativity and innovation (Gulati et al. 2012). The current emphasis on knowledge as a commodity in modern society has led to an increased interest in better understanding the means by which knowledge is disseminated in human social networks (Dupouët and Yıldızoğlu 2006; Morone and Taylor 2004). Human actors can accumulate knowledge by individual learning or through processes of interactive learning, processes that can be carried out both under formal learning conditions such as educational institutions or under informal conditions. An interesting result of simulation studies suggests that widely divergent levels of knowledge within a network tends to lead to a gap in knowledge dissemination, leading to community divisions into a highly knowledgeable, a group that is attaining greater knowledge at a slower rate, and a marginalized group that could be considered ignorant (Morone and Taylor 2004). Moreover, this division does not arise from community structure per se, as communities in which knowledge levels are not highly variable tend to disseminate knowledge efficiently and more equitably (Morone and Taylor 2004).

A sea change in human communication mechanisms was driven by the global introduction of computer-enhanced methods such as email, communication platforms such as Facebook and Twitter, and the interactive informational 'bloggersphere'. An important feature of social communication networks is the interrelationships between them—such that the network of friends, colleagues, and trade-partners influence responses of any individual agent to all networks to which that agent belongs (Szell et al. 2010). While social media can be seen to provide an unprecedented mechanism for the global exchange of knowledge, information, and opinion, to fully comprehend its reach requires a much fuller understanding of these complex inter-relationships.

As is true of the brain, the hierarchical and dynamic properties of human social—and, by extension, economical, technological and political—interactions lead to unpredictable emergent behaviors at multiple levels. Network theory provides a method by which such complexities may be evaluated in both space and time.

Network Theory Links Neural and Social Communication Systems

We have seen that the brain is a complex dynamic system (Amaral et al. 2004) consisting of on the order of 10^{11} neurons and 10^{15} synaptic connections (Sporns et al. 2005). In common with other complex dynamic systems, the brain exhibits critical dynamics (Chialvo 2010; Poil et al. 2008) and scale-free behavior (as explained below). Human social systems are also complex dynamic systems, with a global population of approximately 7×10^9 human beings according to the US Census Bureau (www.census.gov).

Complex systems exhibit non-random linkages over multiple temporal and spatial scales, a relationship captured by the popular 'six degrees of freedom' concept (Watts 2004). Although not without controversy, many such systems are described as scale-free or scale-invariant and follow power law distributions (Kello et al. 2010). Scale-free systems are characterized by the property of criticality; that is, they sit on the cusp between completely predictable (rigid) and completely unpredictable (chaotic) behavior. This is precisely the state we noted above as useful for a system that needs to be both highly adaptive and yet stable; these properties have been described in brain networks at multiple scales, from local and large-scale circuits (Fiete et al. 2010; Kitzbichler et al. 2009; Rubinov et al. 2011) to cognitive behaviors as complex as language (Kello et al. 2010; Steyvers and Tenenbaum 2005), online collaborative interactions (Woolley and Hashmi 2013—this volume), and the phase shifts from wake to sleep (Bedard et al. 2006; Zempel et al. 2012).

Scale-free systems share a common architecture described in the seminal paper of Watts and Strogatz (1998) as a small world network. In this architecture, network elements (termed nodes) are linked by connections (termed edges) such that the majority of connections are local while there are only sparse linkages between distant elements (Butts 2009; Watts and Strogatz 1998). This architecture confers several important properties to the system, and points to interesting system behaviors. As it is this architecture that links human social organization and behavior to that of the brain network, a brief description of some of these properties will be provided along with references for those interested in learning more.

A characteristic of small world networks is the presence of hub elements—elements that are richly connected to other network elements—while the majority of elements are more sparsely connected (Eguiluz et al. 2005). This organizational feature has been shown to be present in the brain for both structural and functional linkages (Collin et al. 2013; van den Heuvel et al. 2012), and has formed the basis for designation of a set of linking hubs labeled as 'rich club' elements. The same feature has been shown to be critical to human social interactions, from dissemination of information via communication (Opsahl et al. 2008; van den Heuvel et al. 2012; Vaquero and Cebrian 2013) to the diffusion of disease epidemics (Christakis and Fowler 2008; Pastor-Satorras and Vespignani 2001; Zhang et al. 2011).

These hub elements are critical to communication in small world networks as they provide the links between modules or communities in the global network. While many studies have relied upon analysis of network interactions in stable periods, the interactions described are dynamic, with both the structure of local communities and the links that bind them in flux on multiple time scales. No single node, whether human or neural, is embedded in only a single community, so that its behaviors are the result of both its structural embedding and the multirelational networks in which it operates.

The Computational Power of Human Social Communication

The concept of harnessing human elements for computation is not new (Grier 2005), and the practice of using humans as computational elements can be found as early as the eighteenth century. Modern computing has been argued to have developed from the intersection of scientific problem solving, technological innovation, and the social practice of computing teams (Rall 2006). Human computers calculated solutions to problems, often using pen and pencil but in later periods augmented with simple adding machines. In some instances, the human computers were well trained, but this was not always the case (Grier 1998, 2005; Rall 2006). While the period of human computers focused on calculating solutions to problems, as has been noted by others, the modern view of human computation rests on a partnership between electronic—or perhaps quantum—computers and humans in which each provides a unique skill set (Heylighen 2013).

One similarity remains as essential to the new view of human computation as it was to earlier views and that is the need to clearly and carefully define the problem at hand and the solution space within which it resides. While crowd-sourcing and citizen science are clear paths toward social modes of computation, they do not erase the need for expert knowledge and successful implementation of human computation will require a solid understanding of the social interrelationships needed to interleave expert and unskilled team members. This is not to suggest that, for example, all such teams are comprised of non-expert members—teams may also be composed of teams of interlinked experts in different arenas. However, regardless of the team composition, from the sheer number of individuals and computers involved to the skill sets of individual agents, social interaction and cultural biases must be understood to optimize any solution. Network analysis is one tool that may aid in this endeavor.

References

Ahrens MB, Huang KH, Narayan S, Mensh BD, Engert F (2013) Two-photon calcium imaging during fictive navigation in virtual environments. Front Neural Circuit 7:104

Amaral LA, Diaz-Guilera A, Moreira AA, Goldberger AL, Lipsitz LA (2004) Emergence of complex dynamics in a simple model of signaling networks. Proc Natl Acad Sci USA 101(44): 15551–15555

Barabási A-L, Albert R, Jeong H (2000) Scale-free characteristics of random networks: the topology of the world-wide web. Phys A Stat Mech Appl 281(1):69–77
Barch DM, Braver TS, Sabb FW, Noll DC (2000) Anterior cingulate and the monitoring of response conflict: evidence from an fMRI study of overt verb generation. J Cogn Neurosci 12(2):298–309
Bassett DS, Bullmore E (2006) Small-world brain networks. The Neuroscientist 12(6):512–523
Bedard C, Kroeger H, Destexhe A (2006) Does the 1/f frequency scaling of brain signals reflect self-organized critical states? Phys Rev Lett 97(11):118102
Biswal B, Zerrin Yetkin F, Haughton VM, Hyde JS (1995) Functional connectivity in the motor cortex of resting human brain using echo planar MRI. Magn Reson Med 34(4):537–541
Brown J, Broderick AJ, Lee N (2007) Word of mouth communication within online communities: conceptualizing the online social network. J Interact Mark 21(3):2–20
Butts CT (2009) Revisiting the foundations of network analysis. Science 325(5939):414–416
Chialvo DR (2010) Emergent complex neural dynamics. Nat Phys 6(10):744–750
Christakis NA, Fowler JH (2008) The collective dynamics of smoking in a large social network. N Engl J Med 358(21):2249–2258
Cohen AL, Fair DA, Dosenbach NU, Miezin FM, Dierker D, Van Essen DC, Schlaggar BL, Petersen SE (2008) Defining functional areas in individual human brains using resting functional connectivity MRI. Neuroimage 41(1):45
Collin G, Sporns O, Mandl RC, van den Heuvel MP (2013) Structural and functional aspects relating to cost and benefit of rich club organization in the human cerebral cortex. Cereb Cortex, epub ahead of print, PMID 23551922, DOI: 10.1093/cercor/bht064
Davidsen J, Ebel H, Bornholdt S (2002) Emergence of a small world from local interactions: modeling acquaintance networks. Phys Rev Lett 88(12):128701
Deco G, Jirsa VK, McIntosh AR (2013) Resting brains never rest: computational insights into potential cognitive architectures. Trends Neurosci 36(5):268–274
DeLong ND, Nusbaum MP (2010) Hormonal modulation of sensorimotor integration. The J Neurosci 30(7):2418–2427
Dosenbach NU, Fair DA, Miezin FM, Cohen AL, Wenger KK, Dosenbach RA, Fox MD, Snyder AZ, Vincent JL, Raichle ME (2007) Distinct brain networks for adaptive and stable task control in humans. Proc Natl Acad Sci 104(26):11073–11078
Dupouët O, Yıldızoğlu M (2006) Organizational performance in hierarchies and communities of practice. J Econ Behav Organ 61(4):668–690
Eguiluz VM, Chialvo DR, Cecchi GA, Baliki M, Apkarian AV (2005) Scale-free brain functional networks. Phys Rev Lett 94(1):018102
Felleman DJ, Van Essen DC (1991) Distributed hierarchical processing in the primate cerebral cortex. Cereb Cortex 1(1):1–47
Fiete IR, Senn W, Wang CZ, Hahnloser RH (2010) Spike-time-dependent plasticity and heterosynaptic competition organize networks to produce long scale-free sequences of neural activity. Neuron 65(4):563–576
Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME (2005) The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proc Natl Acad Sci USA 102(27):9673–9678
Fuxe K, Borroto-Escuela DO, Romero-Fernandez W, Agnati LF (2013) Volume transmission and its different forms in the central nervous system. Chin J Integr Med 19(5):323–329
Garmestani AS, Allen CR, Gunderson L (2009) Panarchy: discontinuities reveal similarities in the dynamic system structure of ecological and social systems. Pap Nat Res 166. Ecology and Society 14(1):15. [online] URL: http://www.ecologyandsociety.org/vol14/iss1/art15/
Grier DA (1998) The math tables project of the work projects administration: the reluctant start of the computing era. Ann Hist Comput, IEEE 20(3):33–50
Grier DA (2005) When computers were human. Princeton University Press, Princeton
Guertin PA (2012) Central pattern generator for locomotion: anatomical, physiological, and pathophysiological considerations. Front Neurology 3:183

Gulati R, Sytch M, Tatarynowicz A (2012) The rise and fall of small worlds: exploring the dynamics of social structure. Organ Sci 23(2):449–471

Haas JS, Landisman CE (2012) Bursts modify electrical synaptic strength. Brain Res

He Y, Chen ZJ, Evans AC (2007) Small-world anatomical networks in the human brain revealed by cortical thickness from MRI. Cereb Cortex 17(10):2407–2419

Heylighen F (2013) From human computation to the global brain: the self-organization of distributed intelligence. In: Michelucci P (ed) The handbook of human computation. Springer, New York

Holling CS (2001) Understanding the complexity of economic, ecological, and social systems. Ecosystems 4(5):390–405

Just MA, Cherkassky VL, Keller TA, Kana RK, Minshew NJ (2007) Functional and anatomical cortical underconnectivity in autism: evidence from an FMRI study of an executive function task and corpus callosum morphometry. Cereb Cortex 17(4):951–961

Kaneko T (2013) Local connections of excitatory neurons in motor-associated cortical areas of the rat. Front Neural Circuits 7:75

Kello CT, Brown GD, Ferrer-i-Cancho R, Holden JG, Linkenkaer-Hansen K, Rhodes T, Van Orden GC (2010) Scaling laws in cognitive sciences. Trends Cogn Sci 14(5):223–232

Kitzbichler MG, Smith ML, Christensen SR, Bullmore E (2009) Broadband criticality of human brain network synchronization. PLoS Comput Biol 5(3):e1000314

Mennes M, Kelly C, Zuo X-N, Di Martino A, Biswal B, Castellanos FX, Milham MP (2010) Interindividual differences in resting state functional connectivity predict task-induced BOLD activity. Neuroimage 50(4):1690

Meunier D, Lambiotte R, Bullmore ET (2010) Modular and hierarchically modular organization of brain networks. Front Neurosci 4

Morone P, Taylor R (2004) Knowledge diffusion dynamics and network properties of face-to-face interactions. J Evol Econ 14(3):327–351

Nagy JI, Dudek FE, Rash JE (2004) Update on connexins and gap junctions in neurons and glia in the mammalian nervous system. Brain Res Rev 47(1):191–215

Oláh S, Füle M, Komlósi G, Varga C, Báldi R, Barzó P, Tamás G (2009) Regulation of cortical microcircuits by unitary GABA-mediated volume transmission. Nature 461(7268):1278–1281

Opsahl T, Colizza V, Panzarasa P, Ramasco JJ (2008) Prominence and control: the weighted rich-club effect. Phys Rev Lett 101(16):168702

Pastor-Satorras R, Vespignani A (2001) Epidemic spreading in scale-free networks. Phys Rev Lett 86(14):3200–3203

Poil SS, van Ooyen A, Linkenkaer-Hansen K (2008) Avalanche dynamics of human brain oscillations: relation to critical branching processes and temporal correlations. Hum Brain Mapp 29(7):770–777

Rall DN (2006) The 'house that Dick built': constructing the team that built the bomb. Soc Stud Sci 36(6):943–957

Rubinov M, Sporns O, Thivierge J-P, Breakspear M (2011) Neurobiologically realistic determinants of self-organized criticality in networks of spiking neurons. PLoS Comput Biol 7(6):e1002038

Scemes E, Suadicani SO, Dahl G, Spray DC (2007) Connexin and pannexin mediated cell—cell communication. Neuron Glia Biol 3(3):199

Sporns O, Tononi G, Kötter R (2005) The human connectome: a structural description of the human brain. PLoS Comput Biol 1(4):e42

Spruston N (2008) Pyramidal neurons: dendritic structure and synaptic integration. Nat Rev Neurosci 9(3):206–221

Steyvers M, Tenenbaum JB (2005) The large-scale structure of semantic networks: statistical analyses and a model of semantic growth. Cogn Sci 29(1):41–78

Szell M, Lambiotte R, Thurner S (2010) Multirelational organization of large-scale social networks in an online world. Proc Natl Acad Sci 107(31):13636–13641

Theis M, Giaume C (2012) Connexin-based intercellular communication and astrocyte heterogeneity. Brain Res

van den Heuvel MP, Kahn RS, Goñi J, Sporns O (2012) High-cost, high-capacity backbone for global brain communication. Proc Natl Acad Sci 109(28):11372–11377

Vaquero LM, Cebrian M (2013) The rich club phenomenon in the classroom. Sci Rep 3:1174

Vizi E, Fekete A, Karoly R, Mike A (2010) Non-synaptic receptors and transporters involved in brain functions and targets of drug treatment. Br J Pharmacol 160(4):785–809

Wang XF, Chen G (2003) Complex networks: small-world, scale-free and beyond. Circuits Syst Mag IEEE 3(1):6–20

Watts DJ (2004) Six degrees: the science of a connected age. WW Norton, New York

Watts DJ, Strogatz SH (1998) Collective dynamics of 'small-world' networks. Nature 393(6684): 440–442

Wilkinson D (2002) Civilizations as networks: trade, war, diplomacy, and command-control. Complexity 8(1):82–86

Woolley AW, Hashmi N (2013) Cultivating collective intelligence in online groups. In: Michelucci P (ed) The handbook of human computation. Springer, New York

Zempel JM, Politte DG, Kelsey M, Verner R, Nolan TS, Babajani-Feremi A, Prior F, Larson-Prior LJ (2012) Characterization of scale-free properties of human electrocorticography in awake and slow wave sleep states. Front Neurol 3:76

Zhang J, Xu X-K, Li P, Zhang K, Small M (2011) Node importance for dynamical process on networks: a multiscale characterization. Chaos An Interdiscip J Nonlinear Sci 21(1): 016106–016107

Zhou C, Zemanová L, Zamora G, Hilgetag CC, Kurths J (2006) Hierarchical organization unveiled by functional connectivity in complex brain networks. Phys Rev Lett 97(23):238103

The Psychopathology of Information Processing Systems

Matthew Blumberg and Pietro Michelucci

Introduction

If we organize human participants into systems modeled loosely on cognitive architectures (see Blumberg (2013), Heylighen (2013), Pavlic and Pratt (2013), all this volume), it is conceivable that such systems will exhibit dysfunction, just as humans do. And it therefore will be necessary to develop methods of thinking about, diagnosing, and treating (e.g. debugging) such issues.

Two approaches will be explored: the first views mental illness from the standpoint of communications theory and logical structure—essentially viewing mental illness as failure in information processing. The second views mental dysfunction from the standpoint of brain chemistry. Each maps somewhat differently to information processing systems—suggesting different modes of analysis and means of intervention.

The goal of this chapter is to discuss the nature of such systemic pathologies speculatively—we aim not to provide a rigorous analysis, but rather to begin a conversation.

M. Blumberg (✉)
GridRepublic, USA
e-mail: mblumberg@picador.net

P. Michelucci
ThinkSplash LLC, Fairfax, VA, USA

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_6,

Part I: Interaction Dysfunction

Schizophrenia in Persons

What Is Schizophrenia?

Gregory Bateson describes a Schizophrenic as "a person who does not know what kind of message a message is." (Bateson et al. 1956). In broad terms, this means understanding the condition as being rooted in a failure to discern context; more rigorously it means understanding the condition as a specific patterned failure to keep straight the "logical type" of messages. From this perspective, what manifests as mental illness is at root a pathology of information processing.

The "Theory of Logical Types" (Whitehead and Russell 1927) is a formal way to describe what one might intuitively describe as "levels of abstraction" within a set of information. In this formalization, each item of a set is a member of a "class". A critical distinction is that a class cannot be a member of itself; nor can one of the members of the class be the class itself. That is, a "class" represents a higher level of abstraction—a higher logical type—than its "members". In other words: information is hierarchical.

A few examples illustrate this idea of "logical types" (Bateson 1979):

- The name is not the thing named but is of different logical type, higher than the thing named.
- The injunctions issued by, or control emanating from, the bias of a house thermostat is of higher logical type than the control issued by the thermometer. (The *bias* is the device on the wall that can be set to determine the temperature around which the temperature of the house will vary.)
- The word *tumbleweed* is of the same logical type as *bush* or *tree*: It is not the name of a species or genus of plants; rather, it is the name of a class of plants whose members share a particular style of growth and dissemination.
- *Acceleration* is of a higher logical type than *velocity*.

The use of Logical Typing can be seen to be fundamental to human communication. "Play", "non-play", "fantasy", "sacrament", "humor", "irony", and "learning" are all examples of *classes* of communication. In all these cases, proper interpretation of a specific message[1] depends upon proper identification of the Logical Type to which it belongs—i.e., proper identification of the *context* of the message.

By way of example, children may "Play" at fighting. While such activity may have many of the outward markers of "real" fighting, the participants can nevertheless engage without anger or malice. But if a participant for some reason comes to see the context of such interaction as "real" fighting, the meaning of events changes, and the actions may degrade to harmful violence.

[1] Note that "message" here is used in a general way: including verbal utterance, non-verbal action, absence of utterance or action, etc. Anything that can or should be taken as being of consequence in a social interaction.

Returning to our thesis, Bateson defines a schizophrenic as one who (a) has difficulty in assigning the correct communicational mode to the messages he receives from other persons; (b) has difficulty in assigning the correct communicational mode to those messages which he himself utters or emits nonverbally; and/or (c) has difficulty in assigning the correct communicational mode to his own thoughts, sensations, and percepts. Most generally—"He has special difficulty in handling signals of that class whose members assign Logical Types to other signals."

In "Toward A Theory Of Schizophrenia", Bateson looks specifically at the role "Double Binds", situations in which messages implicit at different levels of communication conflict with one another. These experiences can be very difficult for affected persons to define, because conflict exists between *different levels of the same interaction*. For instance,

> A young man who had fairly well recovered from an acute schizophrenic episode was visited in the hospital by his mother. He was glad to see her and impulsively put his arm around her shoulders, whereupon she stiffened. He withdrew his arm and she asked, 'Don't you love me anymore?' He then blushed, and she said, 'Dear you must not be so easily embarrassed and afraid of your feelings.' (Bateson et al. 1956)

I.e., in this example the young man can become confused by the conflict between meanings implicit at different levels of the same interaction: the mother on the one hand stiffening when hugged, while at same time asking "don't you love me anymore?"; and this is made worse when the mother provides the young man with an incorrect attribution for his inner confusion ("Dear, you must not be so easily embarrassed by your feelings.")

The means by which such apparently small communication issues can lead to large scale pathology is in some sense analogous to how the flow if a river creates a canyon: not by brute force, but by the slow and persistent effect of the water's friction over time. Thus in the above example, imagine that young man, growing up, has been subject to millions of similarly muddled communications over his life—always in the vital emotional context of a parent-child relationship. The young man, seeking to make sense of the world, may thus learn to muddle the logical type of messages—as a means to adapt, to make sense of things. (It is, in a sense, a perfectly rational response to an irrational environment). And the child, having so learned, may take these interpretative habits into secondary relationships as well (including his relationship with himself).

Schizophrenia in Social Groups

In the above view, certain patterned failures in communication and interpretation lead to behavioral dysfunction. We wish to introduce, speculatively, the notion that similar patterns of communication within social groups will lead to similar dysfunction—at the group level. I.e., that there exist pathologies of information systems generally, which can be exhibited at multiple scales, individual or group.

From the point of view described earlier, {1} Schizophrenia in an individual is understood to be rooted in the individual's inability "to know what kind of message a message is"; and {2} that this confusion is often related to repeated experience of double-binds. With this in mind, consider:

A characteristic feature of contemporary media has been the blending of news and entertainment; of advertising and news; of opinion and fact, of expert and amateur. That is: one who turns to media for information about the world literally does not know what kind of messages he is getting.

Similarly, American political discourse is frequently characterized by double-binds, often with each of the two primary political parties articulating conflicting messages about the nature of reality. For instance, a terrorist attack creates a bind, putting into conflict essential principles civil liberties with the desire for security.

The increasing unreliability of information type (is it news? entertainment? advertising?), frequently experienced with particular acuteness around issues presented by institutional parties as double-binds (ex the requirements of civil liberties vs security), can thus be recognized as a communications pattern very much like that experienced by the individual schizophrenic. And so it is perhaps not surprising that the adventurous cultural analyst might perceive comparable *symptoms* to be exhibited at the cultural level. I.e., one might consider contemporary culture to be exhibiting certain "schizophrenic" patterns—for instance: pervasive belief in conspiracy theories[2] (i.e. paranoia); a government with reduced capacity to take consistent action; economic dysfunction.

As problems from this point of view are understood to be rooted in deficiencies in communications, interventions would be similarly focused. Examples might include actions to strengthen the journalistic establishment; to alter forms of political discourse, especially to either minimize or explicitly recognize the double-binds; and perhaps to introduce "therapeutic" double-binds, the resolution of which would require denial of one or another element of a larger bind.

Part II: Organismal Dysfunction

Next we apply an organismic view to complex, distributed information-processing systems, endowing them with agency, such as goal-directed behavior, and a tendency toward homeostasis—an equilibrium state. This view permits us to adapt extant pharmacological treatment models of human behavioral dysfunction to neurosis in these distributed systems.

[2] "About half the American public endorses at least one kind of conspiratorial narrative"—"Conspiracy Theories, Magical Thinking, and the Paranoid Style(s) of Mass Opinion". J. Eric Oliver and Thomas J. Wood, *Working Paper Series*, University of Chicago, 2012.

An Associative View

In general, thought-processes in humans are influenced heavily by fundamental drives. According to drive theory (Seward 1956), any disturbance to the equilibrium state in a person *drives* the person to engage in thought processes leading to behaviors that restore homeostasis, to ensure survival. For example, dehydration gives rise to thirst, which drives a person to seek water (the goal state). This leads to a series of thoughts about how to obtain water that might involve planning and decision-making (see Busemeyer and Townsend 1993). Though drive theory is oft criticized for not addressing secondary reinforcers (such as money), it is still a useful general framework for this discussion because it provides a context for understanding the role of stress in reinforcing thought processes.

When an organism's equilibrium is disrupted, the distance between the current state and goal state increases, which causes stress. Stress places an organism into a heightened state of arousal, which can increase associative learning by causing connections between neurons to form more easily. This is generally considered adaptive because it enables more rapid experiential learning for lessons most relevant to survival.

Post-Traumatic Stress Disorder (PTSD)

However, extreme stress due to trauma and the consequent sudden and heightened learning can have deleterious effects, such as post-traumatic stress disorder (PTSD). In such cases, the brain states that coincided with the trauma are formed indelibly. The resultant associations are so strong that if those brain states or similar ones are reproduced by other means (external or internal), it can trigger an association to the traumatic event that stimulates a stress response comparable to the original experience. This stress response then further reinforces the association of the trigger event to the traumatic event and may even cause new associations to form that are unrelated to the original event. This self-perpetuating cycle creates an associative "gravity well" that can eventually link so many aspects of daily experience to stress that a person becomes effectively paralyzed by anxiety.

Consider, for example, a person who is mugged at gunpoint by someone wearing a ski mask. Subsequently, when the victim is approached in a new context by an actual skier wearing a ski mask that resembles the one worn by the mugger, he experiences tremendous anxiety. And since the post-traumatic anxiety is experienced in the novel context of ski slopes, the victim creates new stress associations to that context and, consequently, avoids skiing.

Fear Circuits

The networks of association between brain states and stress response are sometimes referred to as "fear circuits". These may be phylogenetic in origin, such as the innate fear of seeing one's own blood, or ontogenetic, such as the learned fear of hospitals (Bracha 2006). Importantly, these associations are generated by and apply to perceptual states, regardless of whether they correspond to an external world state, a dream state, or wakeful thought processes. Indeed, the Ancient Greek philosopher Epictetus (2004) made the prescient observation that "what bothers people is not what happens, but what they think of it."[3]

Obsessive-Compulsive Disorder (OCD)

While PTSD has a multi-faceted clinical presentation, including such symptoms as blackouts (memory loss), it is the manifestation of persistent, intrusive thoughts, or "obsessions" that is most germane to this discussion. Obsessive-compulsive disorder[4] (OCD) is often the diagnosis given to the presentation of such chronic rumination. The same notion of fear circuits attributed to PTSD applies also to OCD, but does not require a traumatic precursor event, and may involve more abstract concepts.

For example, the perception that a country is moving toward civil war could generate anxiety in a person that leads to obsession. The increased level of stress induces hyper-associative learning, such that any new thoughts would be more likely to be connected to the concept of civil war. For example, a typical shortage of food at the grocery store could be misconstrued as a sign of stockpiling in anticipation of a food shortage due to war.[5] This might link any food-related concepts to civil war such that any future meal preparation would activate the civil war fear circuit. Furthermore, food preparation itself would be incorporated into an expanding and self-reinforcing civil war fear circuit.

Treating OCD in Persons

Today, there are two accepted treatments for OCD: a behavioral treatment called cognitive-behavioral therapy [6](CBT) and pharmacological treatment. Herein, we focus on the latter. The most effective drugs for treating OCD are serotonin reuptake inhibitors (Simpson 2010), suggesting that a serotonin deficiency may be responsible for obsessive behavior. Serotonin is a neurotransmitter, a chemical messenger that supports communication among neurons in the brain. What is most relevant here is that these drugs are used in the treatment of all anxiety disorders, not just OCD.

---

[3] Special thanks to Ernesto Michelucci for re-popularizing this simple quote, which has deep implications for the human condition.

[4] Not to be confused with clinical perfectionism, which is sometimes referred to as obsessive-compulsive *personality* disorder (OCPD).

[5] Indeed, according to the Batesonian model described above, this could be described as simply another example of context misinterpretation.

[6] CBT involves overt associative remapping via exposure and response prevention.

Though the specific mechanism by which serotonin alleviates OCD symptoms is not well-understood, it is the conjecture of this author that reducing anxiety attenuates the hyper-associative growth and reinforcement of fear circuits, thereby disrupting the rumination cycle. Without constant reinforcement, the fear circuits diminish over time at a normal rate of memorial decay. This interpretation suggests a computational proxy, discussed below, that might be effective for treating obsession in distributed cognitive systems.

A Problem-Solving Superorganism

A superorganism in its most general definition is simply an organism consisting of many organisms. The present discussion, however, is interested in superorganisms consisting of a technology-mediated collective of human (and possibly machine) agents, functioning collectively as a distributed information processing system. We also assume for this discussion that this system, like all organisms, seeks homeostasis. Thus, it has drives related to maintaining an equilibrium state and engages in goal-directed behaviors resulting from those drives. One example of such a system would be a massively distributed problem solving (Michelucci 2009) system in which very large numbers of people contribute to solving complex problems that exist in the real world (see Greene and Young (2013), this volume).

Let's further consider that communication among humans in this superorganism is mediated by a software-based workflow or cognitive architecture. Thus, the quality and quantity of information that flows among information processing agents is both monitored and influenced by the automated control system. Presuming the control system more heavily weights factors that lead most directly to a solution state, associations relevant to those factors would be reinforced most heavily.

Obsession in Superorganisms

So how would OCD manifest in such a superorganism. Consider that in such a problem-solving system "stress" would be characterized by a systemic assessment of distance between the current state and solution state for whatever problems are being addressed. For example, if the solution state is a stable earth climate, then the level of stress in the superorganism might correspond to the perceived distance between the current state and solution state. Thus, if agents within the system perpetuate the belief that industrial carbon dioxide emissions are causing climate change, the system would strengthen the association of carbon dioxide emissions to stress, leading to increased activity around solving the sub-problem of carbon dioxide emissions.

It is easy to imagine how such an association could then lead to further associations to carbon-dioxide emission such as human respiration, which itself could subsequently lead to the more general observation that all animal respiration adds the "stress" of climate change. While this association may be valid in terms of first-order effects, it would ignore systemic effects. And if the associations were made too strongly due to the perceived influence of carbon dioxide emissions on climate change-related stress in the system, the problem-solving resources of the system could become pathologically overcommitted to resolving that carbon dioxide sub-problem. In other words, the system could be obsessed with reducing carbon dioxide emissions to the negligence of a more holistic solution that takes a more balanced view of the multifaceted nature of the problem.

Treating OCD in Superorganisms

Indeed, such group-based obsession occurs also in natural social systems, as described in Part I, though in such cases, the only recourse may be behavioral—that is, policy-based. Engineered systems, however, afford a recourse that might not otherwise exist. Access to the controlling software would make it possible to both observe and adjust the rules that govern the strengths of associations among individual agents in the distributed problem solving system. Decreasing the extent to which systemic stress influences collaborative activities among agents could help restore balance to distributed thought processes. Indeed, it is conceivable that just as with humans, whereby minor changes in neurotransmitter levels can give rise to significant changes in behavior, small calibrations to parameters that govern association strength in distributed problem solving algorithms could resolve obsessive behavior in superorganisms. Whether agents within the system acting would make such calibrations as implementers of an executive function, by a completely automated homeostatic algorithm, or by some external "superorganism psychiatrist" depends upon the evolution of these systems and the co-development of treatment models.

When Superorganisms Are Not Organisms

We should not ignore the possibility that other sorts of pathology may exist in superorganisms that don't in humans because superorganisms are fundamentally different than humans—they are themselves composed of interconnected highly complex organisms. Indeed, superorganisms are a different *logical type* than humans. A superorganism is a system of complex systems, which could give rise to entirely new and unprecedented classes of behavior dysfunction. Since we do not know what to look for, we may not at first be cognizant of the emergence of such dysfunction. And once we do become aware, we may need to develop new treatment models and methods specific to those needs. Given the potential impact of such dysfunction, it would behoove us to minimize the potential for disruptive surprise by developing our understanding of superorganismic behavioral pathology in close parallel with the development of superorganisms themselves.

Conclusion

The reader may or may not accept various portions of the above speculation, or may consider the discussion far too incomplete in its presentation for serious consideration. The point we hope will nevertheless be of interest, however, is that at an individual level, mental pathology can be seen to result from patterned defects in communication and learning; and that similar defects within a culture or future engineered social system may result in similar behavioral patterns at the larger group level. Indeed, the latter may contribute significantly to the former.

If true, one may aspire to develop means to identify and diagnose such information processing defects; and to develop interventions to prevent, minimize, or eliminate the defects or their symptoms.

Acknowledgments The authors would like to express their sincere gratitude to Mary Catherine Bateson for her penetrating observations and practicable feedback, all of which materially improved the conceptual exposition of this chapter.

References

Bateson G (1979) Mind and nature: a necessary unity. Hampton Press, Cresskill

Bateson G, Jackson D, Haley J, Weakland J (1956) Toward a theory of schizophrenia. Behav Sci 1(4):251–254

Bracha HS (2006) Human brain evolution and the "neuroevolutionary time-depth principle:" implications for the reclassification of fear-circuitry-related traits in DSM-V and for studying resilience to warzone-related posttraumatic stress disorder. Prog Neuro-Psychopharmacol Biol Psychiatry 30(5):827–853. doi:10.1016/j.pnpbp.2006.01.008

Busemeyer JR, Townsend JT (1993) Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. Psychol Rev 100(3):432–459

Epictetus (2004) Discourses. Courier Dover Publications, N. Chemsford, MA

Greene K, Young T (2013) Building blocks for collective problem solving. In: Michelucci PE (ed) Handbook of human computation. Springer, New York

Heylighen F (2013) From human computation to the global brain: the self-organization of distributed intelligence. In: Michelucci PE (ed) Handbook of human computation. Springer, New York

Michelucci P (2009) Massively distributed problem solving. URL: http://www.dodsbir.net/sitis/archives_display_topic.asp?Bookmark=36687, WebCite Cache: http://www.webcitation.org/6AeZMC8Nv. Accessed 5 July 2013

Pavlic T, Pratt S (2013) Superorganismic behavior via human computation. In: Michelucci P (ed) The handbook of human computation. Springer, New York

Seward J (1956) Drive, incentive, and reinforcement. Psychol Rev 63:19–203

Simpson HB (2010) Pharmacological treatment of obsessive-compulsive disorder. Curr Top Behav Neurosci 2:527–543

Whitehead AN, Russell B (1927) Principia mathematica. Cambridge University Press, Cambridge, UK

Information and Computation

Carlos Gershenson

Introduction

Before delving into the role of information theory as a descriptive tool for human computation (von Ahn 2009), we have to agree on at least two things: what is human, and what is computation, as human computation is at its most general level computation performed by humans. It might be difficult to define what makes us human, but for practical purposes we can take an "I-know-it-when-I-see-it" stance. For computation, on the other hand, there are formal definitions, tools and methods that have been useful in the development of digital computers and can also be useful in the study of human computation.

Information

Information has had a long and interesting history (Gleick 2011). It was Claude Shannon (1948) who developed mathematically the basis of what we now know as *information theory* (Ash 1990). Shannon was interested in particular on how a message could be transmitted reliably across a noisy channel. This is very relevant for telecommunications. Still, information theory has proven to be useful beyond engineering (von Baeyer 2005), as anything can be described in terms of information (Gershenson 2012).

C. Gershenson (✉)
Departamento de Ciencias de la Computación, Centro de Ciencias de la Complejidad, Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Universidad Nacional Autónoma de México, A.P. 20-726, 01000 México, D.F. México
e-mail: cgg@unam.mx

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_7, © Springer Science+Business Media New York 2013

A brief technical introduction to Shannon information H is given in Appendix A. The main idea behind this measure is that messages will carry more information if they reduce uncertainty. Thus, if some data is very regular, i.e. already certain, more data will bring few or no new information, so H will be low, i.e. few or no new information. If data is irregular or close to random, then more data will be informative and H will be high, since this new data could not have been expected from previous data.

Shannon information assumes that the meaning or decoding is fixed, and this is generally so for information theory. The study of meaning has been made by semiotics (Peirce 1991; Eco 1979). The study of the evolution of language (Christiansen and Kirby 2003) has also dealt with how meaning is acquired by natural or artificial systems (Steels 1997).

Information theory can be useful for different aspects of human computation. It can be used to measure, among other properties: the information transmitted between people, novelty, dependence, and complexity (Prokopenko et al. 2009; Gershenson and Fernández 2012). For a deeper treatment of information theory, the reader is referred to the textbook by Cover and Thomas (2006).

Computation

Having a most general view, computation can be seen simply as the transformation of information (Gershenson 2012). If anything can be described in terms of information, then anything humans do could be said to be human computation. However, this notion is too broad to be useful.

A formal definition of computation was proposed by Alan Turing (1936). He defined an abstract "machine" (a Turing machine) and defined "computable functions" as those which the machine could calculate in finite time. This notion is perhaps too narrow to be useful, as Turing machines are cumbersome to program and it is actually debated whether Turing machines can model all human behavior (Edmonds and Gershenson 2012).

An intermediate and more practical notion of computation is *the transformation of information by means of an algorithm or program.* This notion on the one hand tractable, and on the other hand is not limited to abstract machines.

In this view of computation, the algorithm or program (which can be run by a machine or animal) defines rules by which information will change. By studying at a general level what happens when the information introduced to a program (input) is changed, or how the computation (output) changes when the program is modified (for the same input), different types of dynamics of information can be identified:

- Static. Information is not transformed. For example, a crystal has a pattern which does not change in observable time.
- Periodic. Information is transformed following a regular pattern. For example, planets have regular cycles which in which information measured is repeated every period.

- Chaotic. Information is very sensitive to changes to itself or the program, it is difficult to find patterns. For example, small changes in temperature or pressure can lead to very different meteorological futures, a fact which limits the precision of weather prediction.
- Complex. Also called critical, it is regular enough to preserve information but allows enough flexibility to make changes. It balances robustness and adaptability (Langton 1990). Living systems would fall in this category.

Wolfram (2002) conjectured that there are only two types of computation: universal or regular. In other words, programs are either able to perform any possible computation (universal), or they are simple and limited (regular). This is still an open question and the theory of computation is an active research area.

Computing Networks

Computing networks (CNs) are a formalism proposed to compare different types of computing structures (Gershenson 2010). CNs will be used to compare neural computation (information transformed by neurons), machine distributed computation (information transformed by networked computers), and human computation.

In computing networks, nodes can process information (compute) and exchange information through their edges, each of which connects the output of node with the input of another node. A computing network is defined as **a set of nodes** N **linked by a set of edges** K **used by an algorithm** a **to compute a function** f (Gershenson 2010). Nodes and edges can have internal variables that determine their state, and functions that determine how their state changes. CNs can be stochastic or deterministic, synchronous or asynchronous, discrete or continuous.

In a CN description of a **neural network** (NN) model, *nodes* represent neurons. Each neuron i has a continuous state (output) determined by a function y_i which is composed by two other functions: the weighted sum S_i of its inputs $\bar{x}_i$ and an activation function A_i, usually a sigmoid. Directed *edges* $i\,j$ represent synapses, relating outputs y_i of neurons i to inputs x_j of neurons j, as well as external inputs and outputs with the network. Edges have a continuous state w_{ij} (weight) that relates the states of neurons. The *function* f may be given by the states of a subset of N (outputs $\bar{y}$), or by the complete set N. NNs usually have two dynamical scales: a "fast" scale where the network function f is calculated by the functional composition of the function y_i of each neuron i, and a "slow" scale where a learning *algorithm* a adjusts the weights w_{ij} (states) of edges. There is a broad diversity of algorithms a used to update weights in different types of NN. Figure 1 illustrates NNs as CNs.

Digital machines carrying out **distributed computation** (DC) can also be represented as CNs. *Nodes* represent computers while *edges* represent network connections between them. Each computer i has information H_i which is modified by a program $P_i(H_i)$. Physically, both H_i and P_i are stored in the computer memory, while the information transformation is carried out by a processor. Computers can share

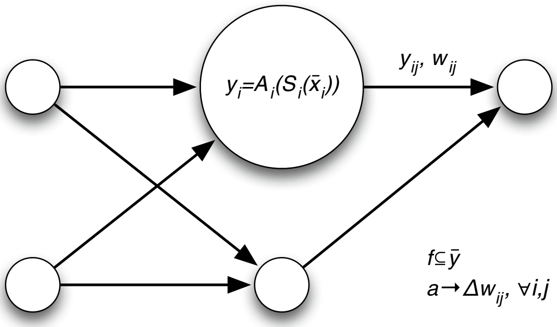


Fig. 1 A NN represented as a CN

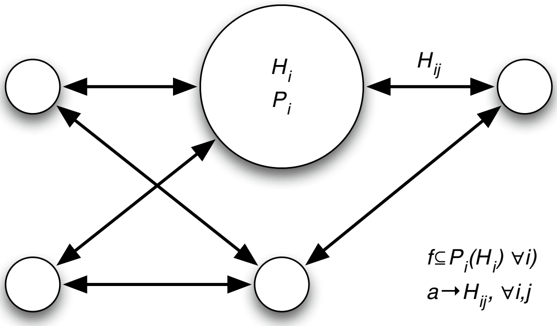


Fig. 2 A DC system or a HC system represented as a CN

information H_{ij} across edges using a communication protocol. The *function f* of the DC will be determined by the output of $P_i(H_i)$ of some or all of the nodes, which can be seen as a "fast" scale. Usually there is an algorithm *a* working at a "slower" scale, determining and modifying the interactions between computers, i.e. the network topology. Figure 2 shows a diagram of DC as a CN.

Human computation (HC) can be described as a CN in a very similar way than DC. People are represented as *nodes* and their interactions as *edges*. People within a HC system transform information H_i following a program $P_i(H_i)$. In many cases, the information shared between people H_{ij} is transmitted using digital computers, e.g. in social networks, wikis, forums, etc. In other cases, e.g. crowd dynamics, information H_{ij} is shared through the environment: acustically, visually (Moussaïd et al. 2011), stigmergically (Doyle and Marsh 2013), etc. The function *f* of a HC system can be difficult to define, since in many cases the outcome is observed and described only a posteriori. Still, we can say that *f* is a combination of the computation carried out by people. An algorithm *a* would determine how the social links change in time. Depending on the system, *a* can be slower than *f* or vice versa.

In DC, the algorithm *a* is centrally determined by a designer, while in most HC systems, the *a* is determined and executed by people (nodes) themselves.

Using information theory, we can measure how much information H_{ij} is transmitted between people, how much each person receives and produces, and how much the entire system receives and produces. In many cases, machines enable this transmission and thus also facilitate its measurement. Comparing the history of information transfers and current information flows can be used to measure the novelty in current information.

Examples

Social Networks

A straightforward example of human computation can be given with online social networks. There are key differences, e.g. links are bidirectional in Facebook (my friends also have me as their friend) and unidirectional in Twitter (the people I follow do not necessarily follow me, I do not necessarily follow my followers). People and organizations are represented with their accounts in the system as nodes, and they receive information through their incoming links, They can share this information with their outgoing links and also produce novel information that their links may receive. People can decide how to create or eliminate social links, i.e. a is decided by individuals.

These simple rules of the information dynamics on social networks are able to produce very interesting features of human computation (Lerman and Ghosh 2010), which can be described as functions f. For example, non-official news can spread very quickly through social networks, challenging mass media dominated by some governments. On the other hand, false rumors can also spread very quickly, potentially leading to collective misbelief. Nevertheless, it has been found that the dynamics of false rumors spreading is different from that of verifiable information (Castillo et al. 2011).

Describing social networks as CNs is useful because interactions are stated explicitly. Moreover, one can relate different scales with the same model: local scale (nodes), global scale (networks), and meso scales (modules); and also temporal scales: fast (f) and slow (a). Information theory can be used to detect novelty in social interactions (high H values in edges), imitation (low H values in edges), unusual patterns ("fake" information), correlations (with mutual information), and communities (modules (Newman 2010)).

Wikipedia

Wikipedia gives a clear example of the power of human computation. Millions of people (nodes) from all over the world have collaboratively built the most extensive encyclopedia ever. The sharing of information is made through editable webpages on a specific topic. Since these pages can potentially link more than two people (editing the webpage), the links can be represented as those of a hypernetwork (Johnson 2009), where edges can link more than two nodes (as in usual networks). The information in pages (hyperedges) can be measured, as it changes over time with the editing made by people linked to them. The information content delivered by different authors can be measured with H. When this is increased, it implies novelty. The complexity of the webpages, edits, and user interactions can also be measured, seen as a balance between maximum information (noise) and minimum information (stasis) (Fernández et al. 2013).

The function f of Wikipedia is its own creation, growth, and refinement: the pages themselves are the output of the system. Again, people decide which pages to edit, so the algorithm a is also decided by individuals.

Traditionally, Wikipedia—like any set of webpages—is described as a network of pages with directional edges from pages that link to other pages. This is a useful description to study the structure of Wikipedia itself, but it might not be the most appropriate in the context of human computation, as no humans are represented. Describing Wikipedia as a CN, the relationships between humans and the information they produce collaboratively is explicit, providing a better understanding of this collective phenomenon.

Conclusions

Concepts related to information and computation can be applied to any system, as anything can be described in terms of information (Gershenson 2012). Thus, HC can also benefit from the formalisms and descriptions related to information and computation.

CNs are general, so they can be used to describe and compare any HC system. For example, it is straightforward to represent online social networks such as Facebook, Twitter, LinkedIn, Google+, Instagram, etc. as CNs. As such, their structure, functions, and algorithms can be contrasted, and their local and global information dynamics can be measured. The properties of each of these online social networks could be compared with other HC systems, such as Wikipedia.

Moreover, CNs and Information Theory can be used to design and self-monitor HC systems (Gershenson 2007). For example, information overload should be avoided in HC systems. The formalisms presented in this chapter and in the cited material can be used to measure information inputs, transfers, and outputs to avoid not only information overload, but also information poverty (Bateson 1972).

In our age where data is overflowing, we require appropriate measures and tools to be able to make sense out of "big data". Information and computation provide some of these measures and tools. There are still several challenges and opportunities ahead, but what has been achieved so far is very promising and invites us to continue exploring appropriate descriptions of HC systems.

Appendix

Shannon Information

Given a string X, composed by a sequence of values x which follow a probability distribution $P(x)$, information (according to Shannon) is defined as:

$$H = -\sum P(x) \log P(x). \tag{1}$$

For binary strings, the most commonly used in ICT systems, the logarithm is usually taken with base two. For example, if the probability of receiving ones is maximal ($P(1) = 1$) and the probability of receiving zeros is minimal ($P(0) = 0$), the information is minimal, i.e. $H = 0$, since we know beforehand that the future value of x will be 1. Information is zero because future values of x do not add anything new, i.e. the values are known beforehand. If we have no knowledge about the future value of x, as with a fair coin toss, then $P(0) = P(1) = 0.5$. In this case, information will be maximal, i.e. $H = 1$, because a future observation will give us all the relevant information, which is also independent of previous values. Equation 1 is plotted in Fig. 3. Shannon information can be seen also as a measure of uncertainty. If there is absolute certainty about the future of x, be it zero ($P(0) = 1$) or one ($P(1) = 1$), then the information received will be zero. If there is no certainty due to the probability distribution ($P(0) = P(1) = 0.5$), then the information received will be maximal. Shannon used the letter H because equation 1 is equivalent to Boltzmann's entropy in thermodynamics, which is also defined as H. The unit of information is the bit. One bit represents the information gained when a binary random variable becomes known.

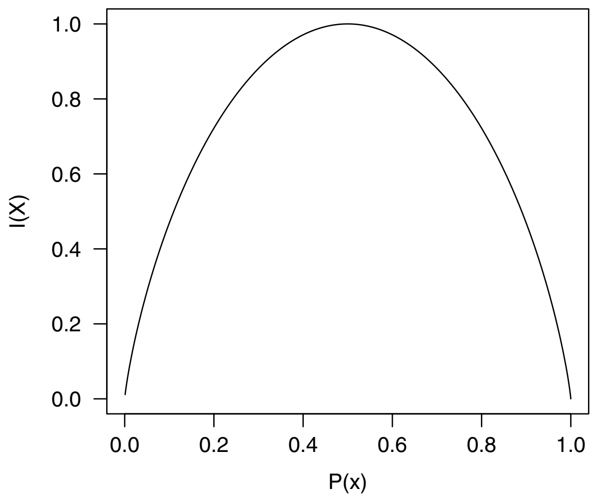


Fig. 3 Shannon's information $H(X)$ of a binary string X for different probabilities $P(x)$. Note that $P(0) = 1 - P(1)$

A more detailed explanation of information theory, as well as measures of complexity, emergence, self-organization, homeostasis, and autopoiesis based on information theory can be found in Fernández et al. (2013).

Acknowledgements I should like to thank Matthew Blumberg and Pietro Michelucci for useful advice. This work was partially supported by SNI membership 47907 of CONACyT, Mexico.

References

Ash RB (1990) Information theory. Dover, New York

Bateson G (1972) Steps to an ecology of mind. Ballantine, New York

Castillo C, Mendoza M, Poblete B (2011) Information credibility on twitter. In: Proceedings of the 20th international conference on world wide web. WWW '11. ACM, New York, pp 675–684. http://doi.acm.org/10.1145/1963405.1963500

Christiansen MH, Kirby S (2003) Language evolution, vol 3. Oxford University Press, Oxford/New York

Cover TM, Thomas JA (2006) Elements of information theory. Wiley-Interscience. http://www.elementsofinformationtheory.com/

Doyle MJ, Marsh L (2013) Stigmergy 3.0: from ants to economies. Cogn Syst Res 21:1–6. http://dx.doi.org/10.1016/j.cogsys.2012.06.001

Eco U (1979) A theory of semiotics. Indiana University Press, Bloomington

Edmonds B,Gershenson C (2012) Learning, social intelligence and the Turing test – why an "out-of-the-box" Turing machine will not pass the Turing test. In: How the world computes: Turing centenary conference and 8th conference on computability in Europe, CiE 2012, Cambridge, 18–23 June 2012. In: Cooper SB, Dawar A, Löwe B (Eds) Proceedings. Lecture notes in computer science, vol 7318/2012. Springer, Berlin/Heidelberg, pp 182–192. http://arxiv.org/abs/1203.3376

Fernández N, Maldonado C, Gershenson C (2013) Information measures of complexity, emergence, self-organization, homeostasis, and autopoiesis. In: Prokopenko M (Ed) Guided self-organization: inception. Springer (in press). http://arxiv.org/abs/1304.1842

Gershenson C (2007) Design and control of self-organizing systems. CopIt Arxives, Mexico. http://tinyurl.com/DCSOS2007, http://tinyurl.com/DCSOS2007

Gershenson C (2010) Computing networks: a general framework to contrast neural and swarm cognitions. Paladyn. J Behav Robot 1(2):147–153. http://dx.doi.org/10.2478/s13230-010-0015-z

Gershenson C (2012) The world as evolving information. In: Minai A, Braha D, Bar-Yam Y (Eds) Unifying themes in complex systems, vol VII. Springer, Berlin/Heidelberg, pp 100–115. http://arxiv.org/abs/0704.0304

Gershenson C, Fernández N (2012) Complexity and information: measuring emergence, self-organization, and homeostasis at multiple scales. Complexity 18(2):29–44. http://dx.doi.org/10.1002/cplx.21424

Gleick J (2011) The information: a history, a theory, a flood. Pantheon, New York

Johnson J (2009) Hypernetworks in the science of complex systems. Imperial College Press, London

Langton C (1990) Computation at the edge of chaos: phase transitions and emergent computation. Physica D 42:12–37

Lerman K, Ghosh R (2010) Information contagion: an empirical study of the spread of news on digg and Twitter social networks. In: Proceedings of 4th international conference on weblogs and social media (ICWSM), Washington, DC

Moussaïd M, Helbing D, Theraulaz G (2011) How simple rules determine pedestrian behavior and crowd disasters. PNAS 108(17):6884–6888. http://dx.doi.org/10.1073/pnas.1016507108
Newman M (2010) Networks: an introduction. Oxford University Press, Oxford
Peirce CS (1991) Peirce on signs: writings on semiotic by Charles Sanders Peirce. University of North Carolina Press, Chapel Hill
Prokopenko M, Boschetti F, Ryan AJ (2009) An information-theoretic primer on complexity, self-organisation and emergence. Complexity 15(1):11–28. http://dx.doi.org/10.1002/cplx.20249
Shannon CE (1948) A mathematical theory of communication. Bell Syst Tech J 27:379–423, 623–656. http://tinyurl.com/6qrcc
Steels L (1997) The synthetic modeling of language origins. Evol Commun 1(1):1–34
Turing AM (1936) On computable numbers, with an application to the entscheidungsproblem. Proc Lond Math Soc Ser 2 42:230–265. http://www.abelard.org/turpap2/tp2-ie.asp
von Ahn L (2009) Human computation. In: 46th ACM/IEEE design automation conference, 2009, DAC '09, San Francisco, pp 418–419
von Baeyer HC (2005) Information: the new language of science. Harvard University Press, Cambridge. http://www.hup.harvard.edu/catalog.php?isbn=9780674018570
Wolfram S (2002) A new kind of sciene. Wolfram Media. http://www.wolframscience.com/

Epistemological Issues in Human Computation

Helmut Nechansky

Defining Epistemology

Traditional Epistemology is the branch of philosophy concerned with individual human knowledge, its base, its content and its validity. The focus on the individual results from the fact that there is no knowledge without an individual carrier.

There is no unequivocal definition of knowledge yet, but there is some shared understanding that knowledge is determined by the following four aspects: (1) The human senses and the human mind form its structural base; This base determines (2) what can become its possible content; (3) This content may or may not amount to a representational model corresponding to the world external to the individual carrier; (4) If the content does amount to a valid representational model can be confirmed by repeated observation of a correspondence with the external world, by observation of predicted states, and by goal-orientated actions leading towards predicted goal-states.

Social epistemology adds that the knowledge of any single individual depends on and is interrelated with the knowledge of other individuals, since any human is born, brought up and mostly lives in a social world. Therefore individual knowledge cannot be studied alone.

There are many other branches of epistemology we cannot mention here due to limitations of scope and space. The Internet Encyclopedia of Philosophy (2013) and the Stanford Encyclopedia of Philosophy (2013) offer the easiest access to this wide field, while Goldman (1999) offers an interesting discussion. However, this chapter provides a context sufficient for understanding the role of epistemology in human computation.

H. Nechansky (✉)
Nechansky—Engineering Efficiency, Vienna, Austria
e-mail: hn@nechansky.co.at

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_8,

The Role of Epistemology

Epistemology as a branch of philosophy may seem outdated in a time of 'knowledge society', of cloud computing, and human computation. But this is not the case, since we do not yet have an unequivocal, agreed on, scientific definition what actually constitutes 'knowledge'. So any dealing with knowledge is ultimately still a philosophic endeavor.

And knowledge is the base of all our actions. Questions about this base arise often: How do we know? What can we know? Is this knowledge valid? Is it complete, i.e. sufficient to reach a goal? These are *epistemological* questions at the core of all human endeavors. Usually they do not get the attention they would deserve. And the more complex the systems, on which we rely, become, the more important become *answers* to these questions.

A Cybernetic Approach to Main Aspects of Epistemology

Cybernetics is the general theory of control in technical, biological and sociological systems. Control is pursuing and maintaining a goal-value, i.e. a certain physical state, against a changing environment, i.e. against physical influences disturbing that state. The process of control consists of (a) observing the environment with sensors, (b) comparing the sensor data with a goal-value and (c) deciding for an action to achieve that goal. Standard example for that process is a temperature controller, which aims at a desired room temperature as goal-value; to achieve that it (a) observes the current temperature, (b) compares it with the desired room temperature and (c) decides between the actions "heating" or "cooling" to achieve that.

In the following we will consider humans as complex controller structures. Here the brain has *in principle* controller functions similar to a temperature controller, but in much larger numbers and much more complex forms. Primarily the brain has to enable survival by maintaining some existential goal-values (necessary air, water and food supply; the body temperature). To achieve that it has (a) to observe the state of the environment, (b) to compare if that state serves the existential goal-vales, and (c) to decide for actions to enable that. Secondarily the brain has additional controller functions, which enable making a model of the environment, and, based on that, making predictions, concepts and setting long-term and short-term goal-values. To realize these future goal-values the brain again carries out the controller functions of (a) observing the environment, (b) checking if it corresponds to the goals and (c) deciding for actions to make it so.

Of course, the preceding description of brain functions is a crude simplification (for some important underlying complexities see Nechansky 2012a, b, 2013a, b), but we do maintain that the brain has primarily controller functions. For reasons of brevity we consider here just a few of these controller functions, each illustrating an epistemological problem. We will first describe these controller functions for individuals (illustrating the core problems of traditional epistemology) and then analyze how they work when two individuals interact (illustrating the core problems of social epistemology). Then we equip these two individuals with connected computers and discuss the resulting options. And this will be the base to finally place human computation within epistemology.

Epistemological Aspects of the Individual

To illustrate basic epistemological aspects of human reasoning we present here a complex controller structure (see Fig. 1): Here sensors provide input data used to model aspects of the environment, which are relevant to the given goal-values. Then these models are used for two purposes: internally, occasionally, to modify the goal-values and externally, continuously, to make decisions for goal-orientated actions. In more detail this structure carries out the following functions:

Sensor Inputs: Sensors allow the observation of certain physical aspects of the environment, and turn these into internal sensor data, which somehow represent and map them.

Modeling Decisions: Under this heading we summarize all decisions that have to do with sensor data. This includes what is usually called ‘learning’, but goes beyond that.

Primarily modeling decisions are about what to *ignore*. Humans are permanently confronted with more stimuli than they can observe. So they have to decide to pay

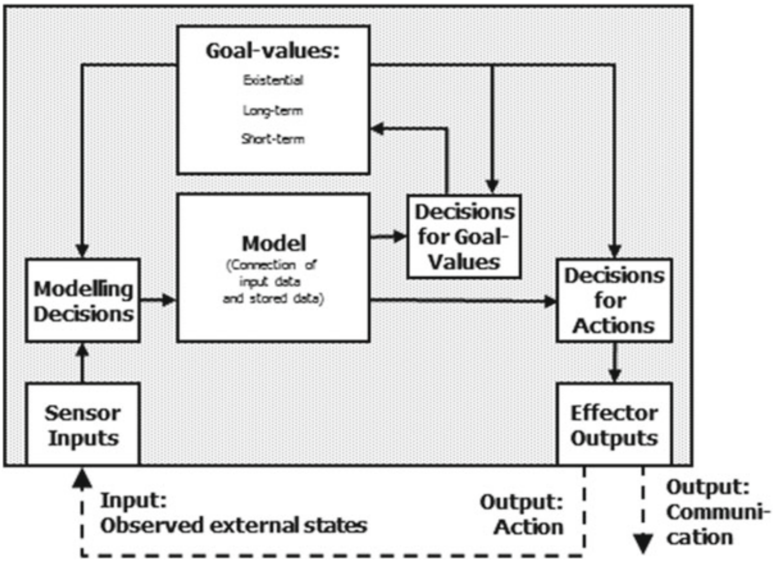


Fig. 1 Epistemological aspects of the individual: the external loop with modeling decisions and decisions for actions and the internal loop with decisions for goal-values

attention to some, and to ignore others considered irrelevant in relation to their *goal-values* (see below). And from the stimuli humans pay attention to they produce permanently more sensor data than they can use for active data processing. So they have to ignore sensor data considered irrelevant.

Secondarily modeling decisions are about what do to with the data considered to be relevant. These decisions are about (a) storing actual sensor data; (b) retrieving stored data for comparison with actual sensor data for pattern recognition; (c) connecting actual and stored sensor data to patterns, sequences and more complex *models* (see below) representing aspects of the environment considered relevant in relation to goal-values; (d) replacing partial or whole models that did not serve the realization of goal-values (the last two points form the core of 'learning'); (e) using models to make predictions of possible future states and events; (f) starting to make new models in relation to newly set long term or short term goal-values.

The decisive point here is that sensor inputs alone are of no value for the individual. Active decisions are required to use them. In these decisions the sensor inputs are evaluated in relation to *goal-values* (see below), i.e. how valuable the data are to pursue certain goals.

Goal-values: All decision processes we discuss here aim at what we call summarizing 'goal-values'. These are all the objectives, which humans partly have to maintain and partly want to achieve.

We distinguish the following three individual goal-values: We are born with just a few fixed (a) existential goal-values (necessary air, water and food supply; the body temperature) plus basic emotions about what is good or bad in relation to these goal-values. While we grow up we learn external states that serve these existential goal-values, for better or worse. Based on that we make *models* (see below) and predictions, which lead to *decisions for goal-values* (see below). These decisions set (b) long-term goal-values (e.g. learning a profession, participating in a human computation project) and (c) short-term goal-values (e.g. how to be successful now in that profession, or project).

Models: Models are the result of previous *modeling decisions* (see above) about the use of actual sensor data and stored data in relation to certain *goal-values* (existential, long-term or short-term).

Models consist primarily of stored relevant sensor data, which represent previous observations in the form of (a) patterns mapping single external states; (b) sequences of patterns representing external events; (c) interrelated sequences of patterns interconnecting events to whole stories experienced in the past.

Models are secondarily organized as plans and concepts consisting of chains of causes and effect leading towards certain goal-values. So there are large numbers of models standing side by side, representing chains of cause and effect to serve existential goal-values (like eating), long-term (like professional conduct), and short-term goal-values (like the necessary steps of a project). There is generally a hierarchy with hierarchically higher goal-values requiring higher priority of models (e.g. eating has to occasionally interrupt professional conduct, which in turn determines the necessary steps of a project).

Models allow (a) identifying current sensor inputs as corresponding to certain patterns or as being part of a previously observed sequence of patterns; (b) deriving predictions of possible future states and events from known sequences or interrelated sequences of patterns.

Models can be confirmed by repeated observations, by observing predicted states, and by goal-orientated actions leading towards predicted goal-states.

Model based predictions are used for two different types of further decisions, leading to two different kinds of feedback loops—one internal and the other external:

Decisions for goal-values: Predictions may be occasionally used to set long-term or short term goal-values (see above). This is an *internal* feedback loop. Here primarily existential goal-values are applied to make decisions for long-term goal-values (e.g. trying to make a living within a certain profession). Then secondarily these long-term goals are used to make decisions for short-term goals. So decisions for goal-values are primarily related to the existential goal-values and create secondarily a hierarchy of subordinated goal-values, by adding, changing or deleting long-term and short term goals.

This is the most important and least understood process of individual epistemology. It determines the entire further behavior of the individual: The previously set goal-values determine directly what is considered important in *modeling decisions* (see above), i.e. which models are made, and indirectly which predictions become possible and which *decisions for actions* (see below) are made.

Decisions for actions: Normally predictions derived from models are just used to trigger one of the effectors (muscles generally, but mainly arms, hands, legs, feet or mouth) to take an appropriate physical action or to start a communication. This is the usual *external* feedback loop, trying to change the external world in some way towards a goal-value.

So the goal-value (whether existential, long-term, or short-term) currently applied determines which action to choose (e.g. to eat, work or communicate, etc.).

Effector Outputs: Decisions for actions trigger the effectors to cause external effects, either physical actions or communication, i.e. primarily words, addressing other individuals.

In summary, these two feedback loops work as follows: Humans make observations of their environment. Based on that, they make primarily models that serve their existential goals. From learning what serves these needs best they secondarily derive models to serve long-term and short-term wants. The sum of these goal-values for needs and wants determines their *modeling decisions* and their *decisions for actions*, i.e. their entire further individual behavior.

Aspects of Social Epistemology

Now let us apply this controller model of a human to the interaction of two individuals (see Fig. 2). This illustrates the problems of social epistemology:

An interaction starts when individual A acts towards B. B observes these actions and evaluates the corresponding sensor inputs in relation to currently important

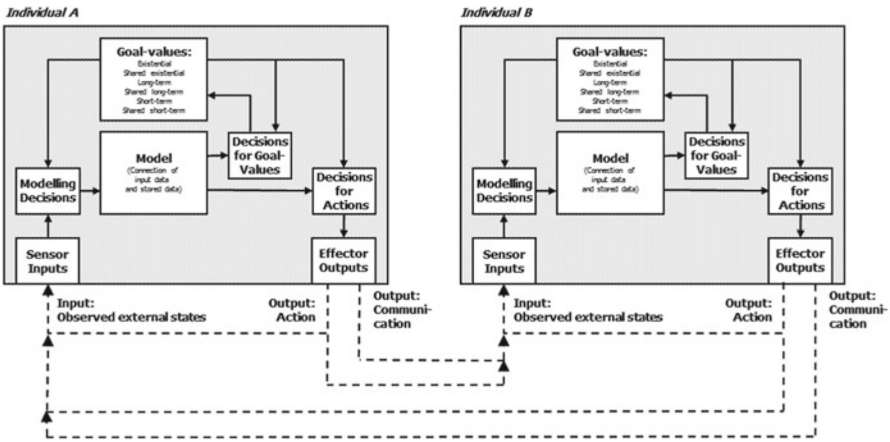


Fig. 2 Aspects of social epistemology: interactions can lead to individual decisions for shared goal-values

goal-values (existential, long-term, short-term). Making a *modeling decision* (see above) B develops a model of A's behavior, predicts its usefulness or danger, and decides for an appropriate action. Then A runs the same process in relation to B's response.

Repetition of this basic exchange may cause at some point in time a *decision for goal-values* (see above) in A and/or in B: Repeated usefulness of A's behavior will cause B to consider A as predictably 'good' or 'interesting'. Then B may decide to add goal-values referring to A to the list of B's already given individual goal-values. Now A may, but need not do the same.

Ideally, of course, this process leads to the development of *shared goal-values* (existential, long-term or short-term), which all interacting parties agree on and add to their individual goal-values. The basic form of this process is realized, of course, in the upbringing of a child. Here the parents serve the needs of the child. So the child will develop shared goal-values with the parents.

Let us mention that the development of shared goal-values may happen spontaneously (e.g. when people face the same problem or threat).

Or this process may be skipped, because interacting people already came independently to shared goal-values (e.g. the same interests or profession).

But mostly, shared goal-values result just from stipulating reciprocally advantageous exchanges of goods, or services, or money and labor, etc.

On the other hand shared goal-values may be propagated by manipulation (A may control the data available to B, using e.g. advertising, censored news, political propaganda, etc.; thus A can limit what may enter into B's models); or they may be enforced to a certain degree (A may have power to control B's access to important resources, like income, etc., or may even be able to apply force; thus A can make B subordinate to and serve his or her goal-values).

The general constraint on developing shared goal-values is the scarcity of goods or societal positions (A and B cannot eat the same bread, or fill the same position in a hierarchy, etc.). Therefore individual goal-values do remain important.

Once parties do share goal-values this will lead to similarities in *modeling decisions* (see above). So they will consider similar data as relevant, will remember and store similar data, and will be interested in making similar models containing certain sequences of cause and effect and enabling particular predictions. Shared goal-values will lead, too, to similar *decisions for actions* (see above).

Shared goal-values will only lead to similar, but not equal, models, as long as the parties rely just on their individual modeling decisions and model making. Only if they cooperate to make externalized mutual verbal concepts, plans, computer programs or mathematical models, they can get to increasingly equal or even unequivocal models.

In summary social epistemology is about human interactions, which make individuals activate their *internal feedback loop for decisions for goal-values*, the process least understood in individual epistemology. The best result is that A and B end up with *individual as well as some shared some goal-values*. And whenever they apply shared goal-values in their current decisions for actions, they will cooperate.

Individuals and Computers: Structures and Interactions

Now let us introduce computers into the relationship of the individuals A and B (see Fig. 3). We characterize computers as controller structures, too, which, of course, differ from humans:

The main differences are: (a) Computers work usually with *fixed* goal-values set by the programmer (we show that in Fig. 3 with the bold arrows directly setting goal-values). So (b) computers lack the *internal feedback loop for making decisions for goal-values*. (In machine learning we occasionally allow computers to make decisions for short-term goals. But we definitely do not want a computer to change its long-term goal-values by itself, so that e.g. a computer programmed to analyze climate data decides on its own to analyze some other data.)

The *external feedback loop* of humans and computers is widely similar: Computers also have sensor inputs (via a keyboard, sensors or data lines). They apply *modeling programs* (matching human *modeling decisions*, but with fixed goal-values) and derive *models* from them (containing here mainly data and mathematical functions, which represent external patterns and sequences), which are used to make predictions. Based on these predictions *programs deciding for actions* are applied (again matching human *decisions for actions* with fixed goal-values).

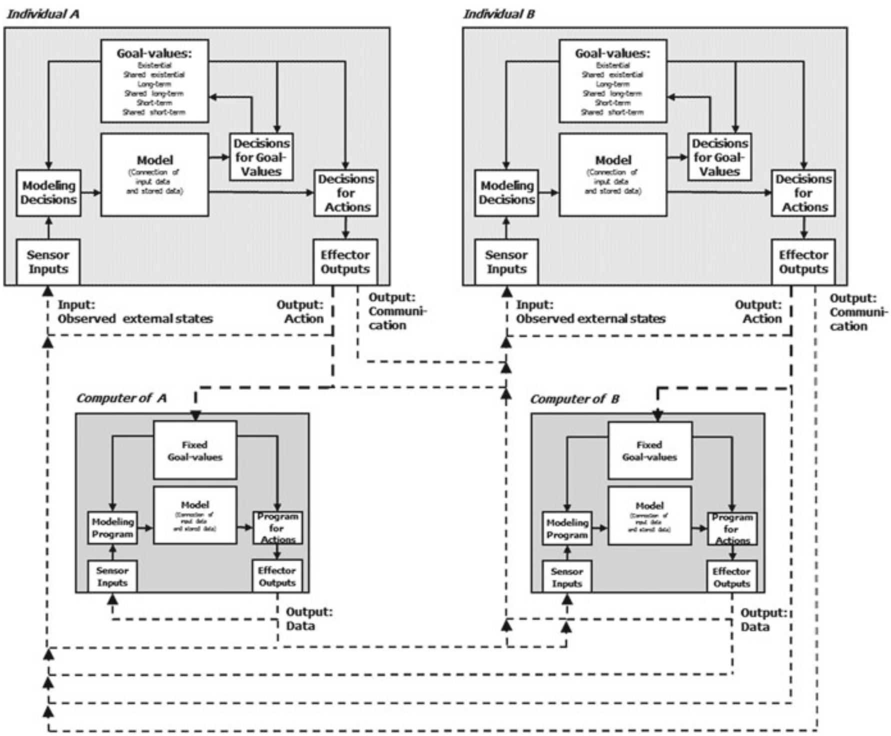


Fig. 3 Individuals and computers: interaction channels added to the context of social epistemology

The *effector outputs* are actions like sending data to other computers, controlling some technical device, or making printouts for human users, etc.

If we take this basic scenario with two individuals and two computers n—times, to match a network, we will get hierarchies (Nechansky 2008) of individuals and computers. We cannot detail that here. We can only assert that this does not change the involved basic epistemological processes.

The Epistemic Processes of Human Computation

Now let us apply all we developed above to human computation:

A human computation project starts with an *initiation phase*, when an initiator defines the long-term goals and short-term tasks. Since human computation is generally applied to problems that require some human contribution, reaching these goals includes tasks that computers cannot yet perform. So the usual advantage of computers we emphasized in section "Individuals and Computers: Structures and Interactions", that they can be directly programmed to work towards a goal, is not available here.

Therefore collaborators have to be sought. The long-term goals and short-term tasks have to be communicated to them. They have to agree on them. And then they have to make *individual decisions for values*, accepting them as *shared goal-values*. We cannot overemphasize that:

1. The success of a human computation projects depends widely on the precise descriptions of long-term goals and short-term tasks, so that the collaborators can understand them and can later make the appropriate individual *modeling decisions* (see above).
2. So the decisive step of a human computation project is a successful finalization of the basic process of social epistemology, as discussed in section "Aspects of Social Epistemology", leading to the acceptance of shared goal-values. Persons focused too much on computation may easily overlook that.

   Dividing a project into subprojects may, but need not weaken that requirement: Now shared goal-values are just needed for the subprojects. But some people might deny contributing, because they do not share the goal-values of the whole project (e.g. a pacifist might deny to contribute to a subproject of a military project).

Once collaborators are found the *computation phase* of the project can start. It may take various forms (see e.g. Quinn and Bederson 2011), which may use any of the possible interconnections between individuals and their computers shown in Fig. 3.

After data acquisition and data distribution, the decisive step is, of course, the human evaluation of the data. Here the short term task of the human contributors is to make *modeling decisions*, judging if data meet the goal-values of the project (e.g. if pictures contain certain patterns, or data sets belong to a certain category, etc.). As emphasized above, the quality of this step depends primarily on clarity of goal-values, i.e. the preceding initiation phase. But it is important, too, that the collaborators do not have any conflicts of interest, i.e. that no other competing long-term and short-term goal-values influence them in their *modeling decisions*. So the success of the project depends to a large degree on the precise consideration and crafting of goal-values the collaborators can fully agree on. More on the importance of goal setting, and its interrelation with motivation and task performance, can be found in Locke's and Lathan's (1990, 2002) classic works on organizational psychology.

After the collection of the results from the collaborators questions of quality control arise. Some evaluation of the results by the initiator must be performed to check if the contributors acted as expected. Since computation is not directly available for obvious reasons, this can only be done indirectly, with approximate use of computers, applying statistics, employing experts, or another round of human computation. Anyway the understanding of the decisive *individual modeling decisions* of the contributors remains vague. So the validity of results obtained using this method remains in question.

However the results may be aggregated in computer models. These models can be confirmed with the usual options of repeated observations, observation of predicted states, and goal-orientated actions leading towards predicted goal-states.

Conclusion

Human computation projects are aiming at goal-values formulated by an initiator. Their critical phase is the initiation, when collaborators have to be found to subscribe to these goals. At that point clearness, specificity, and completeness are the base for the future success. Unfortunately these are difficult to ensure and hardly to measure.

So human computation projects are firmly intertwined in the loops which form the core of individual and social epistemology, i.e. how to interact and communicate with other individuals to make them decide for shared goal-values and cooperate towards them. The results achieved in human computation depend on the success of these processes.

If the results produced by human computation projects contribute to creating further shared goal-values is still another question. That may happen, if these results impress individuals because they show an important relation to their previous goal-values (existential, long-term or short-term), so that they make new decisions for goal-values, set new shared goal-values and start cooperating towards them. Of course, that requires again running through the core processes of individual and social epistemology. We can never escape these loops.

Given the experiences with the precise computer models of climate change, we should not be overly optimistic that human computation projects will lead towards new shared goal-values. All these climatological models predict a threat to the *existential goal-values of all humans*. But not even these threatening results have led to widespread decisions for new goal-values among the endangered people. These decisions always remain individual ones. We can try to influence them, as discussed above, but we cannot directly activate the loops of individual and social epistemology to achieve shared goal-values.

References

Goldman A (1999) Knowledge in a social world. Oxford University Press, Oxford

Internet Encyclopedia of Philosophy (2013) http://www.iep.utm.edu. Accessed 12 Apr 2013

Locke EA, Latham GP (1990) A theory of goal setting and task performance. Prentice Hall, Englewood Cliffs

Locke E, Lathan G (2002) Building a practically useful theory of goal setting and task motivation. Am Psychol 57(9):705–717. doi:10.1037//0003-066X.57.9.705

Nechansky H (2008) Elements of a cybernetic epistemology: decisions, control and principles of societal organization. Kybernetes 37(1):83–93. doi:10.1108/03684920810851005
Nechansky H (2012a) Elements of a cybernetic epistemology: sequence learning systems. Kybernetes 41(1/2):157–176. doi:10.1108/03684921211213007
Nechansky H (2012b) Elements of a cybernetic epistemology: pattern recognition, learning and the base of individual psychology. Kybernetes 41(3/4):444–464. doi:10.1108/03684921211229514
Nechansky H (2013a) Elements of a cybernetic epistemology: elementary anticipatory systems. Kybernetes 42(2):185–206. doi:10.1108/03684921311310567
Nechansky H (2013b) Elements of a cybernetic epistemology: complex anticipatory systems. Kybernetes 42(2):207–225. doi:10.1108/03684921311310576
Quinn AJ, Bederson BB (2011) Human computation: charting the growth of a burgeoning field. In: CHI'11, proceedings of the SIGCHI conference on human factors in computing systems, pp 1403–1412. doi:10.1145/1978942.1979148
Stanford Encyclopedia of Philosophy (2013) http://plato.stanford.edu. Accessed 12 Apr 2013

Synthesis and Taxonomy of Human Computation

Pietro Michelucci

Introduction

Human Computation is an emerging, multidisciplinary field spanning communities. Broadly, it refers to human participation in computational systems and the information and capabilities that arise from that. Beyond this general definition, however, there is a tendency for multiple and sometimes conflicting perspectives, as well as confusion. Therefore, this chapter seeks to characterize the conceptual space of human computation by defining key terminology within an evolving taxonomy.

Previous efforts have sought to flesh out the conceptual space of human computation (Law and Von Ahn 2011) and related terminology (Quinn and Bederson 2011). The present effort seeks to update this body of work in the context of new research and broader multidisciplinary context.

Key Concepts

Two key concepts are described here that provide a context for interpreting and understanding the definitions that follow.

Goals and Intentionality

Human computation (HC) systems are purposeful. They are driven by outcomes that derive from individual behavior, such as enjoyment from playing a game (see Celino; Ghosh; Sanfilippo et al., all this volume) or payment for completing a task (see Chandler

P. Michelucci (✉)
ThinkSplash LLC, Fairfax, Virginia, USA
e-mail: pem@thinksplash.com

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_9,

et al., this volume). They are also driven by outcomes that derive from collective behavior or interactions, such as the advancement of science that results from citizen science projects (see Lintott, this volume). Furthermore, the locus of intentionality in human computation systems may be individual or collective. For example, an individual may launch a crowdsourcing campaign to satisfy a personal objective. Or a system's behavior may be driven by goals that are defined collaboratively by system participants.

Two related ideas emerge from this conceptual framing: emergent HC and engineered HC. Emergent HC refers to systems in which collective behavior is a natural consequence of individual behaviors; and may help inform a deeper understanding of individual behaviors in the context of system dynamics. Engineered HC refers to the notion of overtly creating a context in which the interaction of individuals within will give rise to desired systemic behavior. Though the emergent/engineered dichotomy is being introduced in this volume, the underlying concept is relevant both to understanding the scope of human computation and the relatedness of the terms that follow. Estrada and Lawhead (this volume) introduce the related concepts of natural, stable, and disruptive human computation, which also seem to be useful concepts for further partitioning the space of HC systems.

Computation = Information Processing

The relationship between computation and information processing has been a subject of some controversy. These terms have been differentiated on the basis of historical usage in theoretical contexts (see Piccinini and Scarantino 2010). However, the construal of computation as being equivalent to information processing seems to best fit the practical context of human computation.

In HC, "computation" refers not just to numerical calculations or the implementation of an algorithm. It refers more generally to *information processing*. This definition intentionally embraces the broader spectrum of "computational" contributions that can be made by humans, including creativity, intuition, symbolic and logical reasoning, abstraction, pattern recognition, and other forms of cognitive processing. As computers themselves have become more capable over the years due to advances in AI and machine learning techniques, we have broadened the definition of computation to accommodate those capabilities. Now, as we extend the notion of computing systems to include human agents, we similarly extend the notion of computation to include a broader and more complex set of capabilities.

It is this sense of computation that is intended in the definitions that follow.

Key Terminology

This chapter seeks to define key terms, which have been selected on the basis of prevalence in the book and broad usage across sub-disciplines. These definitions derive from prior work, lively collegial discourse, and the application of basic inference to a growing set of related concepts. It goes without saying that the meaning of terms evolves through usage. For maximal relevance herein, current popular usage *as applied to the study and practice of human computation* exerts considerable bias on these definitions. For this reason, you may discover that in some cases canonical meanings have been deprecated. Given the diversity of the community, context-based usages, and dynamic nature of the conceptual space in a rapidly growing field, it is unlikely that this set of definitions will meet with unilateral agreement. However, this chapter seeks to represent the most common views and, in certain cases, multiple views when there are divergent semantic tracks. For brevity of exposition, we do not belabor etymology, but instead seek to provide the reader with an accessible point of reference.

Glossary

Term	Definition
Collective Action	Human computation in which individual behaviors contribute to a collective product that benefits all members of the collective (see Novak, this volume).
Collective Intelligence	A group's ability to solve problems and the process by which this occurs.
Crowdsourcing	The distribution of tasks to a large group of individuals via a flexible open call, in which individuals work at their own pace until the task is completed (see Chandler, this volume).
Distributed Cognition/Collective Cognition	"The use of information technologies to make distributed information processing by humans much more powerful, focused and efficient" (see Heylighen, this volume).
Distributed Intelligence	The problem-solving capacity of distributed cognitive systems (see Heylighen, this volume).
Distributed Problem Solving	The application of massively distributed cognitive systems to solving problems (see Greene and Thomas, this volume).
Distributed Thinking	The effective distribution and coordination of information processing tasks among human computational agents informed by cognitive architecture (see Blumberg, this volume).
Human Computation/ Distributed Human Computation	1. The design and analysis of multi-agent information processing systems in which humans participate as computational elements. 2. The subset of systems theory in which the systems are composed of machines and humans connected by communications networks.
Organismic Computing	Augmented human collaboration characterized by shared sensing, collective reasoning, and coordinated action (see Michelucci, this volume).

Participatory Sensing	The human-use of sensor-enhanced devices for spatially distributed data collection, enabled by pervasive computing (see Lathia, this volume).
Social Computing	Information processing that occurs as a consequence of human social interaction, usually assumed to occur in an online medium. Note: there is some debate in the field about how to classify systems in which behavior relies upon social knowledge or judgment but does not involve social interaction among participants.
Social Informatics/ Social Network Analysis	The use of big data to understand social behavior (see Lerman, this volume); in Social Network Analysis the "big data" is presumed to originate from behavioral data derived from technology-mediated social systems.
Superorganism	1. Individual organisms functioning together to support the objectives of the collective (see Pavlic and Pratt, this volume). 2. "A collection of agents which can act in concert to produce phenomena governed by the collective" (Kelly 1994).

Conclusion

This synthesis of key concepts in human computation is a snapshot. It is expected that the usage of these terms and related concepts will evolve with the discipline. Thus, this glossary should be revisited and refined by the community as necessary to best support fluid communication and broad comprehension across sub-disciplines.

Acknowledgments The author wishes to gratefully acknowledge useful discussion and feedback from Matthew Blumberg, Irene Celino, Daniel Estrada, Kshanti Greene, Neal Lathia, Jonathan Lawhead, Stuart Reeves.

References

Kelly K (1994) Out of control: the new biology of machines, social systems and the economic world. Addison-Wesley, Boston, p 98

Law ELM, Von Ahn L (2011) Human computation. Morgan & Claypool, San Rafael

Piccinini G, Scarantino A (2010) Computation vs. information processing: why their difference matters to cognitive science. Stud Hist Philos Sci Part A 41(3):237–246. doi:10.1016/j.shpsa.2010.07.012

Quinn AJ, Bederson BB (2011) Human computation: a survey and taxonomy of a growing field. In: Proceedings of the SIGCHI conference on human factors in computing systems. ACM, New York, pp 1403–1412. doi:10.1145/1978942.1979148

Part II
Application Domains

Human Computation in the Wild

Haym Hirsh

One of the backbones of human society has been finding ways to organize human labor to achieve desired outcomes. The advent of computing has allowed us to bring to bear the ideas and tools of computing to this task, giving rise to what we are now calling "human computation." Unlike mechanical computers, which are sufficiently developed and formalized that we can write down on paper an abstract representation of an algorithm and have reasonable expectations about its behavior, human computation bottoms out at fallible, unpredictable people, and, at least at present, no amount of talking or theorizing replaces the need to see what happens when you pull people together in some new way in service of some human-computation-based effort. We're still in the early years of human computation, and our growing understanding of the field is occurring by people building real systems with real people achieving real outcomes.

Furthermore, computing has also provided us a lens that reveals in hindsight that the earliest examples of human computation predate computing, and gives us the language for seeing these efforts in a new, more uniform "human computation" light.

- Britain's 1714 Longitude Act established the Longitude Prize, giving a cash prize to those advancing the technology of measuring a ship's longitude while at sea (Sobel 1995). This early example of the "competition" design pattern of human computation, followed by such examples as (Masters and Delbecq 2008) Sweden's 1734 prize for a method for stopping the progress of fires, France's 1775 Alkali Prize to produce alkali from sea salt, Napolean's 1795 prize for preserving food and 1810 prize for a flax spinning machine, 1833's prize from the Société d'encouragement pour l'industrie nationale for the invention of large-scale commercial hydraulic turbines, 1852's Guano Prize from the Royal Agricultural Society of Britain for a fertilizer as effective as Peruvian guano, and

H. Hirsh (✉)
Cornell University, Ithaca, NY, USA
e-mail: haym.hirsh@cornell.edu

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_10,

the 1863 prize from the Phelan and Collender billiard ball company for a non-ivory billiard ball anticipated today's GoldCorp Challenge, Netflix Challenge, Innocentive, TopCoder, and other examples that bring people together to achieve outcomes through competition.

- The idea of partitioning a job into appropriately sequenced small pieces and doling the pieces out to multiple "micro-work" laborers can be found in how Alexis-Claude Clairaut, Joseph-Jérôme de Lalande, and Nicole-Reine Lepaute went about computing the next arrival of Halley's Comet in 1757 (Grier 2005). Similar ideas can be found in Lewis Fry Richardson's (1922) proposal for predicting weather in "a large hall like a theatre" with tens of thousands of "computers" (people) "at work upon the weather of the part of the map where each sits, but each computer attends only to one equation or part of an equation" (Richardson 1922), and whose ideas can be found in the implementation of such a scheme beginning in 1938 in the Mathematical Tables Project (Grier 2005). These predate and still have lessons for the human computation systems such as Amazon Mechanical Turk, which now allow us to write programs that call human labor as subroutines in their work (Grier 2011).
- The "collection" design pattern of human computation harnesses a distributed workforce to create the elements out of which some larger desired outcome is assembled and is now found in myriad examples of human computation, from Amazon reviews to citizen science. Its early origins can be seen in Friedrich W.A. Argelander's 1844 "Appeal to the Friends of Astronomy" for the organized observation of variable stars by the world's amateur astronomers; the Smithsonian's establishment in 1849 of the Meteorological Project that set up a network of over 100 volunteer weather observers across the United States ultimately giving rise to the US National Weather Service; the initiation in 1858 by the Philological Society of what would become the *Oxford English Dictionary* whose contents were based on the voluntary contributions of thousands of English speakers (Winchester 2004); and Wells W. Cooke's 1881 initiation of a project to document the arrival and departure dates of migratory birds across the U.S., ultimately encompassing thousands of volunteers and including a partnership with the U.S. Lighthouse Board and the establishment of a reporting network of lighthouse keepers across the country. Indeed, one could view the development of scientific journals—initiated in 1665 with the creation of the French *Journal des sçavans* and the English *Philosophical Transactions of the Royal Society*—as also reflecting the collection design pattern for human computation.

These examples each reflect a new way of thinking about how people can be brought together to achieve desired outcomes, and, importantly, are largely intertwined with the technological innovations of their day, but yet they stayed fragmented through history until we had the framing of computation to let us see the patterns and let them suggest new opportunities for the future. Indeed the following section in this volume, called "Techniques and Modalities" (see Greene 2013), takes this very approach in identifying and describing human computation design patterns to enable their reuse.

The vibrant development of human computation over the last few decades has continued through the development of fielded systems that integrally involve people, increasing the leverage we can gain by studying and learning from this growing body of human computation applications. The goal of the chapters in this "applications" section of the handbook is to assemble a record of recent human computation applications to help further drive our understanding of the field.

For example, the widespread access to computing and communications technologies has created a wave of human computation innovations in service of humanitarian aid and disaster response. Patrick Meier starts this section by presenting six examples of human computation applications in this area over the period of 2010–2013. In addition to documenting these wonderful examples of human computation in the service of societal good, Meier also considers what we more general lessons we can learn from them and identifies directions for the future that they suggest.

A second example arises in the medical sector, where human computation innovations are changing the face of healthcare, often driven by the patients themselves bypassing the traditional medical enterprise. Caring for one's health, particularly in the face of life-changing illness, continues to motivate those impacted by illness to push the envelope of what technology and network-based social interaction can achieve in health and medicine. Wicks and Little present a tour of some of the most important human computation innovations taking place in the medical sector. Again, importantly, they learn from this history of success to suggest what implications these examples may have for the future.

A third example occurs as we attempt to get the diverse knowledge of our world into computer-based form. Whereas the World Wide Web contains semi-structured information largely crafted for human consumption, the goal of the Semantic Web is to create a parallel infrastructure that stores information in ways that include some sense of the meaning of the information in computer-manipulable form. Getting vast amounts of semantically represented information in accurate, online form requires massive effort. Simperl, Acosta, and Flöck's chapter provides a comprehensive survey of how people have built a range of human computation systems to facilitate various facets of this work. Their chapter shows how different human computation design patterns, particularly those of games-with-a-purpose and paid micro-labor, have had particular traction in this domain. They also suggest directions for the future, especially in terms of going beyond the generation of new systems and instead reusing and coupling the different ideas developed thus far.

Three chapters in this section concern "citizen science," the process of scientific inquiry that in whole or part includes participants who are not professional scientists and often have far more limited training than professional scientists. Lintott and Reed present an overview of human computation in citizen science, especially from the perspective of their work on Galaxy Zoo, which has hundreds of thousands of participants and contributed new knowledge via dozens of scientific publications. They furthermore document their insights arising from their generalizing beyond Galaxy Zoo in the creation of the Zooniverse platform, which now includes dozens of projects in domains ranging from astronomy to zoology, especially so as to be able to scale to increasing numbers of people and use worker effort wisely, support open-ended investigation by participants, leverage complementary functionalities of machine learning, and ultimately stay in tune with motivations and knowledge of the people who participate in such projects.

Beal, Morrison, and Villegas complement such consideration of human computation in citizen science by also considering the learning opportunities that participation in such projects can provide. They focus on a case study, the Biosphere 2 Evapotranspiration Experiment, which brings middle and high school students to a project studying the loss of water from soil and the leaves of vegetation while also providing them with educational experiences in this domain.

In a series of related case studies, Lin et al. consider the application of distributed human computation to the problem of search and discovery, and in particular, toward the use of collective perception to find loosely-defined things. In this context they discuss first their experiences in the “Expedition: Mongolia” project, in which tens of thousands of participants contributed more than two million pieces of information to detect archaeological anomalies within massive quantities of high-resolution multi-spectral imaging data. They then describe subsequent related efforts in disaster assessment and search and rescue. The chapter concludes with tantalizing ideas about how to enhance the existing approach by tightly integrating human inputs with machine learning methods.

The transformative opportunities for computing and communications technologies have not been lost on those in the creative arts, where numerous innovative human computation ideas have been and are being explored. Rettberg’s chapter provides an overview of key examples of human computation in electronic literature and digital art. Moreover, Rettberg focuses on lessons that appear when human computation is viewed from a digital literary perspective, especially in terms of the statements they make about the relative roles of and relationship between computing and people. Rettberg also shows us that the end goal of some of these examples of human computation are not be the direct outcome of their organized labor but rather they are particularly designed to make a point, to serve as a meta-critique of the values that may underlie human computation.

As we develop a greater number of human computation systems, gaining a better understanding of their relative strengths and weaknesses—across different methods, and in comparison to possible automated methods—grows in importance. Harris and Srinivasan use the task of query refinement in information retrieval as a platform to study the relative benefits of two forms of human computation: micro-labor markets and games-with-a-purpose. They show that for this task human computation beats automation, and that games yield better results than micro-labor markets.

This section next presents three papers suggesting directions for future applications of human computation. First, François Bry presents an approach to credit risk rating that turns not only to lenders but also debtors in assessing the risk faced in the market. Purvis and Hardas next propose a human computation perspective on innovation, formulating the network of people involved in innovation in a way that captures many of the social elements that people bring to bear within organized human labor. Finally, Brambilla and Fraternali present a human computation perspective to integrating social interaction and business process management, where social interactions are treated as extensions to business process models.

The section concludes with Thomsen's provocative chapter that considers the application of human computation to "wicked problems"—tasks that are so difficult humans can't determine if a proposed solution will solve the task. His goal is nothing less than to seek the creation of human computation systems that solve problems that could not otherwise be previously solved. His chapter discusses the various characteristics that will be necessary to build applications capable of tackling wicked problems with human computation.

References

Greene K (2013) Introduction to techniques and modalities. In: Michelucci P (ed) The handbook of human computation. Springer, New York
Grier DA (2005) When computers were human. Princeton University Press
Grier DA (2011) Error identification and correction in human computation: lessons from the WPA. In: Proceedings of the 2011 human computation workshop, AAAI Press
Masters WA, Delbecq B (2008) Accelerating innovation with prize rewards: history and typology of technology prizes and a new contest design for innovation in African agriculture. Intl Food Policy Res Inst 835
Mayall RN (1961) "The Story of the AAVSO," Review of Popular Astronomy 55(513):4–9
Richardson LF (1922) Weather prediction by numerical process. Cambridge University Press
Sobel D (1995) Longitude: the true story of a lone genius who solved the greatest scientific problem of his time. Walker and Company, New York
Winchester S (2004) The meaning of everything: the story of the Oxford English dictionary. Oxford University Press

Human Computation for Disaster Response

Patrick Meier

Introduction

Disaster-affected communities are increasingly using social media to communicate during major disasters. One consequence of this is the rise of Big (Crisis) Data. Recent empirical studies reveal that a small but critical-and-growing fraction of tweets posted during a disaster contain important information for disaster response.[1] Finding the proverbial needle in this growing "haystack" of crisis information has rapidly become a major challenge for the international humanitarian community. Social media use during Hurricane Sandy in 2012 produced a "haystack" of half-a-million Instagram photos and 20 million tweets over just a few days. The year before, over 300,000 tweets were posted every minute following Japan's devastating earthquake and Tsunami. There are at least two ways to manage this volume and velocity of data: (1) Artificial Intelligence and (2) Artificial Artificial Intelligence, or Human Computation.[2] The purpose of this chapter is to analyze the use of human computation for disaster response.

The chapter is structured as follows: the first section describes the use of human computation in response to six major humanitarian crises: Haiti Earthquake (2010), Libya Revolution (2011), Somali Crisis (2011), Hurricane Sandy (2012), Typhoon Pablo (2012) and Mali Crisis (2013). The human computation technologies used to support these disaster response efforts include CrowdCrafting, CrowdFlower, Humanitarian OpenStreetMap's Tasking Server, MapMill, Tomnod and Ushahidi. The groups engaged in deploying and using these technologies include the Standby

[1] See: "Debating the Value of Tweets for Disaster Response (Intelligently)," available online at: http://iRevolution.net/2012/12/17/debating-tweets-disaster.

[2] See TEDx Talk on "Crowdsourcing and Advanced Computing," available online at: http://iRevolution.net/2012/10/21/crowdsourcing-and-advanced-computing.

P. Meier (✉)
Qatar Computing Research Institute (QCRI), Qatar Foundation

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_11,

Volunteer Task Force (SBTF), the Humanitarian OpenStreetMap Team (HOT), the UN Office for the Coordination of Humanitarian Affairs (UN OCHA), the UN High Commissioner for Refugees (UNHCR) and the US Federal Emergency Management Agency (FEMA). The second section builds on these case studies to outline what the future of human computation for disaster response will look like. This section also highlights the use of mobile solutions, gamification and massively multiplayer online games to process humanitarian microtasks. The chapter concludes with a call to action—namely the launch of Big (Crisis) Data Philanthropy for Humanitarian Response in order to grant humanitarian organizations full access to social media data during major disasters.

Haiti Earthquake

Human computation was first used for disaster response following the devastating earthquake that struck Port-au-Prince on January 12, 2010. Graduate students at The Fletcher School (Tufts University) launched a live crisis map within hours of the earthquake to document both the extent of the damage and the disaster-affected population's urgent needs.[3] This information was initially sourced from social media such as Twitter and quickly complemented with reports from the mainstream media. In order to cope with the extensive live coverage of the disaster, Fletcher School students decided to crowdsource the real-time monitoring and processing of several hundred online sources. Within days, several hundred volunteers from Boston, Montreal, New York, London and Geneva answered the call. Together, they manually triaged and geo-referenced over 1,500 reports that were mapped using the Ushahidi platform. Ushahidi is a free and open source mapping software.

Several days after the earthquake, an SMS short code was set up and integrated with the Ushahidi platform. This short code (4636) enabled anyone in Haiti to text in his or her location and urgent needs.[4] Information about the short code was disseminated via community radio stations in Haiti and via Haitian Diaspora news channels. The team behind the Ushahidi software quickly developed a platform to crowdsource the translation of incoming text messages since the vast majority of these were written in Haitian Creole. Volunteers from the Haitian Diaspora were recruited via social media. Together, they translated some 10,000 text messages during the entire search and rescue phase. Two weeks later, the translation efforts were relocated to Haiti thanks to a partnership with the microtasking company CrowdFlower. This enabled Haitians to earn money for their translation work.

[3] See: "How Crisis Mapping Saved Lives in Haiti," available online at: http://newswatch.nationalgeographic.com/2012/07/02/crisis-mapping-haiti.

[4] See: "Ushahidi and the Unprecedented Role of SMS in Disaster Response," available online at: http://iRevolution.net/2010/02/20/sms-disaster-response.

These volunteer-based efforts in response to the Haiti Earthquake marked a watershed moment for the international humanitarian community and the new field of Humanitarian Technology. One first responder, the US Marine Corps, publicly stated that the live crisis map enabled them to save hundreds of lives.[5] Craig Fugate, the Administrator of the US Federal Emergency Management Agency (FEMA), referred to the crisis map as the most comprehensive and up-to-date information available to the humanitarian community.[6] As a result of these efforts, the Fletcher student who spearheaded the Haiti response proposed the launch of a global volunteer community for digital humanitarian response.[7] Together with several colleagues, he co-founded the Standby Volunteer Task Force (SBTF) in October 2010. Today, the SBTF includes over 1,000 digital volunteers based in over 80 countries around the world. Together, this award-winning network of pro-active volunteers have managed some of the most important live crisis mapping operations that have supported both humanitarian and human rights organizations over the past 3 years.[8]

Libya Revolution

One of the most important SBTF deployments remains the response to the Libya Crisis. The United Nations Office for the Coordination of Humanitarian Affairs (UN OCHA) officially activated the SBTF to request a live, crowdsourced social-media crisis map of the escalating situation in the country.[9] The SBTF launched the crisis map within an hour of the request. The volunteer network was able to do this because they had designed specific criteria and workflows beforehand to manage live crisis mapping requests. For example, the SBTF has specific activation criteria that must be met by the activating organization. In addition, the SBTF is composed of multiple teams each of which is responsible for the human computation of the information processing cycle. For example, the Media Monitoring Team is responsible for monitoring both social and mainstream media for the type of information requested by the activating organization. The Geo-Location Team is tasked with identifying the GPS coordinates for relevant reports identified by the Media Monitoring Team. The Mapping Team adds the tagged reports to the crisis map while the Analysis Team produces regular trends analyses.

---

[5] See: "How Crisis Mapping Saved Lives in Haiti," available online at: http://newswatch.nationalgeographic.com/2012/07/02/crisis-mapping-haiti.

[6] See: "How Crisis Mapping Saved Lives in Haiti," available online at: http://newswatch.nationalgeographic.com/2012/07/02/crisis-mapping-haiti.

[7] See: "Standby Crisis Mappers Task Force: Apply Now!" available online at: http://iRevolution.net/2010/09/26/crisis-mappers-task-force.

[8] Standby Volunteer Task Force: http://blog.standbytaskforce.com.

[9] Libya Crisis Map Deployment 2011 Report, available online at: http://blog.standbytaskforce.com/libya-crisis-map-report.

Thanks to these pre-designed human computation workflows and the use of Skype, SBTF volunteers were able to monitor well over 300 online sources and map thousands of relevant reports for an entire month, maintaining live coverage of the situation throughout. The fact that volunteers are also based in multiple time zones also meant that the map was updated around the clock. Because OCHA did not initially have any information officers on the ground in Libya and could obviously not rely on Libyan state media for accurate information, the live social media crisis map provided them with critical situational awareness during the early weeks of the crisis. Moreover, "OCHA did not have the idle capacity to gather, verify and process the enormous amount of available online information."[10] In an email to SBTV volunteers, OCHA wrote "The dedication and professionalism of the Task Force is commendable. Your efforts at tackling a difficult problem have definitely reduced the information overload; sorting through the multitude of signals is no easy task. The Task Force has given us an output that is manageable and digestible, which in turn contributes to better situational awareness and decision making."[11]

Somali Crisis

"Having a real-time map complete with satellite photos, of where everyone is at any one moment is almost as good as having your own helicopter."[12] The United Nations High Commissioner for Refugees (UNHCR) was in desperate need of such a map when the crisis in Somalia began to escalate in October 2011. A massive number of people had been displaced to the "Afgooye Corridor" just West of Mogadishu due to the worsening famine and Al Shabab's terrorist activities. While UNHCR had a couple estimates for the number of displaced individuals, they needed another way to validate these estimates. Getting an accurate figure for the number of Internally Displaced People (IDPs) is critical for disaster response. However, due to the volatile security situation brought about by Al Shabab, humanitarian organizations could not directly access IDPs in order to carry out on-the-ground surveys.

Live crisis maps, like helicopters, can provide a "bird's eye view" of an unfolding situation in real-time. So the SBTF recommended that UNHCR "take to the skies" and use satellite imagery to estimate the number of IDPs in the "Afgooye Corridor." The SBTF partnered with the satellite-imagery provider DigitalGlobe and Tomnod (c.f. Chapter by Luke Barrington) to microtask the analysis of satellite

[10] See: "The [unexpected] Impact of the Libya Crisis Map and the Standby Volunteer Task Force," available online at: http://blog.standbytaskforce.com/sbtf-libya-impact.

[11] Libya Crisis Map Deployment 2011 Report, available online at: http://blog.standbytaskforce.com/libya-crisis-map-report.

[12] See: "Maps, Activism and Technology: Check-In's with a Purpose," available online at: http://iRevolution.net/2011/02/05/check-ins-with-a-purpose.

imagery of Somalia. Tomnod is a microtasking platform specifically designed for the tagging satellite imagery. The imagery is sliced up into smaller pictures each of which is then displayed to volunteers on the Tomnod platform. Users were asked to tag any informal and permanent shelters they could see in each satellite image. Within 120 h, volunteers created over a quarter million tags after analyzing close to 4,000 images.[13] One of the advantages of microtasking platforms like Tomnod is the built-in quality control mechanisms that ensure a relatively high quality of output data. In the case of the Somalia project, each unique image was viewed by at least three different volunteers. Only when there was consensus between three volunteers vis-à-vis the type and location of a given shelter was that data point pushed to UNHCR. This triangulation mechanism yielded a count of 47,000 shelters in the Afgooye Corridor—a figure that the UN was able to use to estimate the approximate number of IDPs in the area.

After the completion if this human computation project for disaster response, the Deputy High Commissioner of UNHCR Alex Aleinikoff thanks SBTF volunteers via video.[14] The transcript: "[…] I've just learned about the wonderful work done by the Standby Task Force which has permitted us to count shelters in the Afgooye Corridor in Somalia through the volunteer work of folks like you around the world. This is such a wonderful project for us it provides enormously important information to UNHCR and helps to create a worldwide virtual community involved in helping refugees and internally displaced people. So I salute you for your work and for the time you have devoted to this project, it's important to us, it's important to people who have been forced from their homes and who are trying to create a new home and a new beginning, thank you."

Hurricane Sandy

Hurricane Sandy caused extensive damage along the Northeastern United States in October 2012. Within hours of the damage, the US Civil Air Patrol (CAP) flew a number of aircraft along the coastline to capture very high-resolution aerial imagery of the disaster-affected areas. According to a FEMA official working with Air Patrol at the time, "CAP imagery is critical to our decision making as they are able to work around some of the limitations with satellite imagery so that we can get an area of where the worst damage is. Due to the size of this event there is an overwhelming amount of imagery coming in, your assistance will be greatly appreciated and truly aid in response efforts. Thank you all for your willingness to help."

[13] See: "Crowdsourcing Satellite Imagery Analysis for UNHCR-Somalia: Latest Results," available online at: http://iRevolution.net/2011/11/09/crowdsourcing-unhcr-somalia-latest-results.

[14] See: "Thank You Video from UNHCR's Deputy High Commissioner," available online at: http://blog.standbytaskforce.com/thank-you-video-from-unhcrs-deputy-high-commissioner.

To rapidly analyze the tens thousands of pictures produced by CAP for damage assessment purposes, the Humanitarian Open Street Map Team (HOT) team customized the MapMill platform to microtask the analysis of the imagery.[15] Volunteers using MapMill would tag each picture as "OK" (no infrastructure damage), "Not OK" (some damage) or "Bad" (significant damage). The result? Nearly 6,000 volunteers analyzed over 30,000 images within the first week and provided almost 150,000 damage assessments in that time. About half of these volunteers produced around 80,000 assessments in the first 48 h alone. On average, every image was tagged or voted on 91 times. The resulting assessments were automatically shared with FEMA via their public GeoPlatform.[16] FEMA subsequently launched a service for people to type in their address and get the CAP image of their house or building.

The HOT network was launched shortly after the remarkable response carried out by OpenStreetMap (OSM) volunteers following the devastating Haiti Earthquake of 2010. Using aerial and satellite imagery provided by the World Bank, volunteers traced the most detailed street map of Port-au-Prince ever created—and they did this within a week. Some 700 volunteers made over 1.4 million edits to the map during the first 30 days following the earthquake.[17]

Typhoon Pablo

Typhoon Pablo devastated large regions of the Philippines in December 2012. Twenty-four hours after the typhoon made landfall, the UN Office for the Coordination of Humanitarian Affairs (OCHA) activated the Standby Volunteer Task Force (SBTF) to assess the damage. OCHA requested that the multimedia assessment be based on Twitter and the resulting analysis provided to the UN within 12 h. The SBTF partnered with the Qatar Computing Research Institute's (QCRI) Crisis Computing Team to collect over 20,000 tweets related to the Typhoon.[18] Next, the SBTF used the CrowdFlower microtasking platform previously employed in response to the Haiti Earthquake. This time, CrowdFlower workers were paid to rapidly identify all tweets that had links to either pictures or video footage. These relevant tweets were then uploaded to the free and open source CrowdCrafting microtasking platform where SBTF volunteers tagged each image and video if they depicted evidence of damage. Volunteers also used CrowdCrafting to microtask the geo-tagging of all relevant pictures and video footage. Twelve hours after OCHA's activation, the SBTF provided them with a detailed dataset of some 100

[15] See: "Crowdsourcing the Evaluation of Post-Sandy Building Damage Using Aerial Imagery," available online at: http://iRevolution.net/2012/11/01/crowdsourcing-sandy-building-damage.

[16] http://fema.maps.arcgis.com.

[17] See: "OpenStreetMap in the First Month After the Haiti Quake," available online at: http://www.maploser.com/2010/09/06/openstreetmap-in-the-first-month-after-the-haiti-quake.

[18] QCRI is a member of the Qatar Foundation: http://www.qcri.com.

georeferenced images and videos depicting the devastation resulting from Typhoon Pablo.[19] Note that like Tomnod, both CrowdFlower and CrowdCrafting also have built-in quality control mechanisms.

The OCHA team in Geneva used this data to create an official UN crisis map of the situation, which they immediately shared with their personnel in the Philippines. The map was also used by the Government of the Philippines and several other UN agencies. This crisis map of the typhoon was the first ever official UN information product based entirely on social media content. Following this deployment, QCRI's Crisis Computing Team developed a way to automatically identify tweets that link to pictures or videos. The SBTF plans to use this in future deployments to accelerate the processing of tweets. This doesn't mean that paid microtasking work has no role to play in digital humanitarian response. Microtasking platforms like Amazon Mechanical Turk and CrowdFlower have large, multinational and multi-lingual global workforces that will continue to be relevant for disaster-response human computation.

Mali Crisis

In January 2013, the Humanitarian OpenStreetMap Team (HOT) of volunteers began to map the transportation infrastructure, buildings and populated areas of Northern Mali to produce a basemap for humanitarian organizations monitoring the humanitarian crisis in the country. The volunteer network carries out these mapping assignments by tracing high (and low) resolution satellite imagery. Having access to the resulting map is particularly important for humanitarian logistics—that is, the delivery of goods and services to the disaster-affected population. This explains why open access to satellite imagery (and indeed other relevant data) is so important for disaster response. At the end of January, UN OCHA formally activated the HOT network to encourage volunteers to continue their mapping efforts and also expand them to include airports, health facilities, schools, water points, land use areas, etc.[20]

To carry out this work, OpenStreetMap volunteers used their own customized microtasking platform.[21] This tool places a grid of cells on top of the area that needs to be mapped. The platform can prioritize the microtasking work to focus on certain cells if specific areas are of particular importance to humanitarian organizations. For the Mali deployment, the HOT network traced roads, rivers, buildings, contour

[19] See: "How the UN Used Social Media in Response to Typhoon Pablo (Updated)," available onlineat: http://blog.standbytaskforce.com/how-the-un-used-social-media-in-response-to-typhoon-pablo-updated.

[20] See: "Mali Activation," available online at: http://hot.openstreetmap.org/updates/2013-02-01_Mali_Activation.

[21] See: "Open Street Map's New Micro-Tasking Platform for Satellite Imagery Tracing," available online at: http://iRevolution.net/2011/09/07/osm-micro-tasking.

of residential areas, water wells, health services and other points of interest.[22] At the time of writing, over 700,000 points had been added to the OSM database over a 6-week period. Each mapped object—such as a well or house—is represented by one or many points that trace the outline of said object.

The Future

As William Gibson famously noted, "The future is already here—it's just not evenly distributed." To get a glimpse of what the future holds for the use of human computation in disaster response, one should look back 2 years at the launch of SyriaTracker.[23] The project combines crowdsourced human intelligence with automated data mining in order to collect relevant information on the crimes and atrocities committed in Syria. The team behind SyriaTracker (all volunteers) use crowdsourcing to collect on the ground eyewitness accounts via email and Twitter. In addition, they repurposed Harvard University's HealthMap, which used data mining for rapid digital disease detection. SyriaTracker customized the platform to automatically monitor human rights violations in Syria by mining over 20,000 English-based sources of news that regularly cover the crisis. The team cross-references and triangulates the crowdsourced reports with the data mining results in an attempt to further verify the accuracy of the collected information. The US Agency for International Aid (USAID), the Office of US Foreign Disaster Assistance (OFDA) and several other agencies are making direct use of the SyriaTracker data in their own official crisis maps of Syria.[24]

SyriaTracker is the longest running crisis map ever. Why? Because the project is powered by human computation *and* data mining. Keeping this map up to date using volunteer-based human computation alone would be a Herculean task. Recall the "haystack" of half-a-million Instagram photos and 20 million tweets posted during Hurricane Sandy. Microtasking is no match for this volume and velocity of Big Crisis Data. Advanced computing techniques such as Artificial Intelligence and Machine Learning are needed to build hybrid approaches that combine the power of the crowd with the speed and scalability of automated algorithms.[25] QCRI is developing just such a system, a Twitter Dashboard for Disaster Response.[26]

[22] See: http://tasks.hotosm.org/#all/Mali.

[23] See: "Crisis Mapping Syria: Automated Data Mining and Crowdsourced Human Intelligence," available online at: http://iRevolution.net/2012/03/25/crisis-mapping-syria.

[24] See: "Why USAID's Crisis Map of Syria is So Unique," available online at: http://irevolution.net/2012/11/27/usaid-crisis-map-syria.

[25] See TEDx Talk on "Crowdsourcing and Advanced Computing," available online at: http://iRevolution.net/2012/10/21/crowdsourcing-and-advanced-computing.

[26] See: "Update: Twitter Dashboard for Disaster Response," available online at: http://iRevolution.net/2013/02/11/update-twitter-dashboard.

The platform enables users such as professional humanitarians to create their own automated classifier on the fly. A classifier is an algorithm that automatically classifies information. For example, if an earthquake were to strike Indonesia, OCHA could create a classifier to automatically detect tweets referring to infrastructure damage. Of course, the algorithm will not accurately tag all tweets, but the use of machine learning will ensure that the classifier improves over time, i.e., learns from its mistakes thanks to human supervision. To create these classifiers on the fly requires the use of microtasking—hence the importance a hybrid approach for disaster response.

The human computation component for disaster response still requires considerable improvement, however. Microtasking needs to become "Smart Microtasking," which means a system that adapts to the skill set of its users. For example, a user that is particularly adept at geo-tagging should be assigned such tasks whereas a user that is more efficient at the categorization of messages as health, shelter, food, etc., should be given those tasks. These "Smart Microtasking" solutions also need to have mobile solutions—that is, they must be easily accessible via smart phone app. In terms of interface, whether web-based or mobile-based, the microtasking platforms used for disaster response have thus far been devoid of any gamification features. This stands in stark contrast to other microtasking projects in the area of citizen science. Zooniverse, for example, has mastered the development of gamified microtasking platforms, which explains why they have hundreds of thousands of users (See Chapter by Chris Lintott). But Zooniverse's expertise and *savoir faire* has yet to crossover into the humanitarian space.

Lastly, there is huge untapped potential in leveraging the "cognitive surplus" available in massively multiplayer online games to process humanitarian microtasks during disasters.[27] The online game "League of Legends," for example, has 32 million players every month and three million on any given day.[28] Over 1 *billion* hours are spent playing League of Legends every month. Riot Games, the company behind League of Legends is even paying salaries to select League of Legend players. Now imagine if users of the game were given the option of completing microtasks in order to acquire additional virtual currency, which can buy better weapons, armor, etc. Imagine further if users were required to complete a microtask in order to pass to the next level of the game. Hundreds of *millions* of humanitarian microtasks could be embedded in massively multiplayer online games and instantaneously completed. Maybe the day will come when kids whose parents tell them to get off their computer game and do their homework will turn around and say: "Not now, Dad! I'm microtasking crisis information to help save lives in Haiti!"

[27] See: "Using Massive Multiplayer Games to Turksource Crisis Information," available online at: http://iRevolution.net/2010/03/24/games-to-turksource.

[28] See: "League of Legends Bigger Than Wow, More Daily Players Than Call of Duty," available onlineat:http://www.forbes.com/sites/jasonevangelho/2012/10/12/league-of-legends-bigger-than-wow-more-daily-players-than-call-of-duty.

Conclusion

Human computation has already played an invaluable role in disaster response. The future, however, belongs to hybrid methodologies that combine human computation with advanced computing. The success of these next-generation humanitarian technologies depends on a number of critical factors. The first is the availability of the data. Twitter's Terms of Service (ToS) restricts the number of downloadable tweets per day to a few thousand. Compare this with the 20 million tweets posted during Hurricane Sandy. Accessing the full Twitter Firehose of ~450 million daily tweets is prohibitively expensive. A possible solution? Big (Crisis) Data Philanthropy for Disaster Response.[29] Data philanthropy involves companies sharing proprietary datasets for social good. Call it Corporate Social Responsibility (CRS) for digital humanitarian response. Companies in this Data Philanthropy club would benefit from the publicity of supporting these positive and highly visible efforts. More importantly, their support would help to save lives. All that is needed is an agreed set of protocols that would provide humanitarian organizations with temporary emergency access to Big Crisis Data. The time to act is now. Both UN Secretary General Ban Ki Moon and UN Under-Secretary General for Humanitarian Affairs Valerie Amos have demonstrated the political will to have the humanitarian industry join the digital age. What we need now is the corporate will from Twitter and companies others to help save lives during the next major humanitarian disaster.

[29] See: "Big Data Philanthropy for Humanitarian Response," available online at: http://iRevolution.net/2012/06/04/big-data-philanthropy-for-humanitarian-response.

The Virtuous Circle of the Quantified Self: A Human Computational Approach to Improved Health Outcomes

Paul Wicks and Max Little

> *"What I've found to be most amazing about these forums thus far is the ability of patients to identify common side effects, formulate solutions, test them, and confirm their general efficacy all in a matter of days, when it would take researchers weeks or even months to generate the same knowledge."—Patient with ALS discussing potential treatments on the forum of the ALS Therapy Development Institute (ALSTDI, www.als.net)*

Introduction

Until recently, medical data was hand-written, inconsistently recorded, difficult to exchange between medical systems, and inaccessible to the patients it was written about. With the advent of electronic health records, disease registries, and patient portals, this state of affairs is changing rapidly. The *nature* of medical data collected is changing too, from a trained professional's observations of signs and symptoms to more objective measurement such as blood tests, genomic scans, imaging data, or even sensor data from medical devices. Patient self-report is also taking an increasingly prominent role as regulators and payers grant increasing authority to the experience of the patient (Basch et al. 2012).

The fact that data is held *about* a person is hardly new; governments, banks, insurers, and retailers have been collecting civic, financial, and behavioural data about us for a long time. But medical data has some unique attributes: of extreme local importance, it's considered highly private (often stigmatizing), can have high financial value, and when inaccurate has severe consequences.

P. Wicks, Ph.D. (✉)
PatientsLikeMe, Lichfield, UK
e-mail: pwicks@patientslikeme.com

M. Little, Ph.D.
Aston University, Aston, UK

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_12,

In the past decade, what was once a collection of dry, static observations silo'd away in a filing cabinet are now dynamic, interactive and fluid data that are perceptible, correctable, and influential on the behavior of the data's subject: the patient. That's because the real revolution of digital health data is that patients increasingly have the potential to see, generate, share, interpret, and alter their own data—"*Nothing about me without me*". Through technology and crowd sourcing, patients will increasingly gain the power to analyse data about themselves too, with the aim of creating value not only for themselves but also other patients like them. The tantalizing promise is not just that the cure to their disease may lie in their data but that they themselves might be the ones that discover it. In a world of crowd sourced medical computation, who cures cancer? We all do.

Patients Go Online

People with serious illnesses have been using the Internet to connect for a long time. Howard Rheingold documents an experience from 1986 when his young daughter was bitten by a tick that they weren't sure how to remove. It was late at night, and while his wife left a message at the pediatrician's he was able to log in to virtual community the "The WELL" and get the medical advice he needed before the pediatrician's office had even returned his wife's phone call (Rheingold 1993). One of the first online communities, the WELL was created by "Whole Earth Catalog" (WEC) founder Stewart Brand a year earlier and brought a technological platform to the 1960s counter-cultural tendencies originally nurtured by that group, such as distrust of authority, emphasis on do-it-yourself "tools", and the sharing of information.

As access to the Internet widened in the 1990s, increasing numbers of patients diagnosed with serious conditions (and their caregivers) took to the Internet to learn about their disease, connect with other patients, and share their experiences (Lester et al. 2004). Discussion groups with similar ground rules to The WELL flourished on pre-Web systems such as USENET, Compuserve and even email list-servs that allowed patients to organize under the banners of their diagnoses. Such patient groups typically preceded the adoption of the Internet by the "official" disease non-profits or health professionals by many years. In 1993, one group of researchers at Massachusetts General Hospital (MGH) surveyed the fragmented nature of the online field and attempted to address this divide by building a safe, moderated environment for people with neurological disorders to meet and communicate. The website's name was "BrainTalk" and it became an online home to tens of thousands of patients, a model for smaller disease-specific communities, and one of the first communities about which papers were written in the peer-reviewed scientific literature (Lester et al. 2004).

The technology of the day permitted systems like BrainTalk to operate as "bulletin boards" or "forums", less technically sophisticated than the social networks of today, but with rich narrative content and a strong sense of community. A member could register with an email address, pick a username to anonymise themselves, and enter key demographics such as age, sex, location, and diagnoses. Forum tools allowed patients to post new conversation "threads" and reply to these asynchronously at any time, but the fora were generally open to non-registered readers too, known as "lurkers".

In parallel to these neurologically focused message boards, caregiver activist Gilles Frydman founded the Association of Online Cancer Resources (ACOR) in 1995 for patients diagnosed with cancer. By creating over 200 support groups for patients with each of the specific subtypes of cancer and using the ubiquitous medium of email, ACOR has gone on to serve over 600,000 patients and caregivers.

Throughout the 1990s ACOR and other online health boards rapidly gained an international following, with topics on BrainTalk ranging from getting a diagnosis, how to communicate with healthcare professionals, tips to cope better with disease, and even alternative medicines (Lester et al. 2004). Anonymity was prevalent, which served to protect patients from identification but also made it difficult to verify who you were actually talking to. Healthcare professionals often lurked silently on communities like ACOR or BrainTalk, but for reasons of professional liability rarely chose to participate in discussions. Patients however, held no such reservations and shared crucial treatment tips with one another. For instance members of the epilepsy community on BrainTalk shared tips on clever ways to "hack" their daily doses of medication to be used to interrupt an ongoing seizure by grinding them up and administering the solution as a liquid to halt the ongoing damage of a severe seizure. Belatedly, professional bodies such as the American Medical Association (AMA) have recently produced "social media policies" that lay out the ground rules for how medical professionals could (if they desire) become a real part of such communities, (Policy 2011) but unfortunately the 20 year latency has not helped to foster online links between clinicians and patients. Left to their own devices, patients have taken up greater responsibility for their own care and that of their fellows.

> When I talk to my doctor, I hear myself asking questions that my online 'family' needs to know. It's as if all these other people—the members of my group—are asking questions through me. And whatever answers I hear from my doctor, I know I'll share with them on line.—Anonymous BrainTalk patient (Lester et al. 2004)

Early research literature focusing on the Internet was particularly concerned with the potential for poor and misleading information gathered online. However, thorough quantitative assessments from the BrainTalk group showed the actual level of misinformation was low: less than 6 % of forum posts on an open forum (Hoch et al. 1999). Others proposed theoretical harms that could result too, such as misunderstanding caused by the limited nonverbal cues available to participants, excessive dependence on a support group, emotional distress caused by reading "triggering" materials, breach of confidentiality, premature intimacy, excessive emotional intensity, and potentially unsafe relationships (Waldron et al. 2000). By contrast, Eysenbach suggested that researchers' focus on negative aspects of online communities and discussion of *potential* rather than *recorded* harms risked obscuring the potential benefits of such tools (Eysenbach 2003), and it is worth noting that all the potential harms noted above are just as feasible in an offline support group. From the perspective of BrainTalk patients for instance, few patients felt that inaccurate information affected them and the forums met an unmet need caused by the inability of health-care providers to answer questions or provide relevant information

(Hoch and Ferguson 2005). Relative to the commonplace harms visited upon patients in a hospital setting, for instance, the number of recorded cases of serious harm arising from patients using the Internet have been low (Crocco et al. 2002), though some subgroups such as those with mood disorders (Bessière et al. 2010) or eating disorders might be particularly vulnerable (Rouleau and von Ranson 2011).

While much of the progress in online communities appeared to have passed unnoticed by much of the medical profession during this period, a small cadre of clinicians, researchers, and activists calling themselves the "e-patient scholars" sought to redress the balance. In what became a manifesto, BrainTalk's director (and former medical editor of the Whole Earth Catalog) Dr. Tom Ferguson described an "e-patient" as one who is not just "electronic" but also equipped, enabled, empowered, and engaged in their own health care (Ferguson 2007). In a white paper completed posthumously after Dr. Ferguson lost his battle with multiple myeloma, the e-patient scholars laid out their anthropology of "citizens with health concerns who use the Internet as a health resource, studying up on their own disease… finding better treatment centers and insisting on better care, providing other patients with invaluable medical assistance and support, and increasingly serving as important collaborators and advisors for their clinicians."(Ferguson 2007)

In their white paper, Ferguson and his team lay out a number of startling anecdotes where patients interacting over the web were able to diagnose rare disease, avoid iatrogenic harms from the medical establishment, and support one another to plug gaps in the medical system (Ferguson 2007). While on an individual basis these stories were important, a constant refrain echoed from the traditional medical establishment: "The plural of anecdote is not data".

Patient Communities for Conducting Research: Early Opportunities and Limitations

From a human computation perspective this represented the greatest limitation of such systems at the time; forum posts were just stories—incomputable, subject to bias, dramatic license, or even outright confabulation. For the newly diagnosed patient (or "newbie"), entering such communities could be an overwhelming experience, with each forum having its own myriad social ties and histories, and each individual member having a rich offline history, only some of which was reflected online and could be hard to wade through. For instance an experienced forum member on BrainTalk might have tens of thousands of forum posts, and coming to understand where they were coming from on a given issue might require hours of reading. Therefore as they grew in scale, understanding narrative text risked becoming an inherently un-scalable proposition.

From the early online researcher's perspective, in the absence of modern techniques such as natural language processing, much of the existing textual information archived was unusable by researchers due to its sheer volume. Furthermore the unique nature of online interactions with its slang, emoticons, and hyperlinks didn't lend itself to existing forms of discourse analysis, never mind the ethical issues of conducting research as a "lurker". However two types of researchers that embraced online methods were able to quickly collect data in a scientifically rigorous framework; qualitative health services researchers and survey researchers.

For instance, in 2006 qualitative content analysis of over 5,200 email messages in ten ACOR lists was used to identified key themes and outcomes related to participation in the system (Michael Bowling et al. 2006). Like Ferguson's analysis of Braintalk and other sites, users of ACOR offered one another information about treatments, provided emotional support, advised one another on interacting with medical professionals, and offered many strategies for active coping (Meier et al. 2007). In 2005, oncology researchers created an online structured survey of fatigue and quality of life for patients with cancer of the bone marrow and were able to rapidly recruit a sample of over a thousand individuals through the ACOR mailing lists to validate their instrument (Mesa et al. 2007). This became a highly cited paper in the field including references in clinical trial designs and the development of new patient reported outcomes. Challenges from that era remain relevant today, however, such as the difficulty of calculating an accurate response rate and thereby accounting for response bias (Michael Bowling et al. 2006).

Although early days, credible scientific researchers were now successfully applying formal methods to extract useful data from content that had been previously construed as "purely anecdotal" or the purview of "internet users with too much time on their hands". To really take off as a research tool, however, the early online patient communities would have to find a way to maintain the benefit of textual narrative, strong relationships, emoticons, and hyperlinks, but also to support these with the objective data with which researcher were more familiar. Websites that patients found useful lacked credibility to researchers because they relied on "anecdote" or unsystematic clinical observations, which sit at the bottom of the pyramid of medical evidence for treatment decision making (Guyatt et al. 2000). In the layers above this are physiologic studies, observational studies (and systematic reviews thereof), randomized controlled trial (and systematic reviews thereof), and at the top of the pyramid the "N of 1 randomized trial" (Gabler et al. 2011). In order to climb the pyramid, online communities would take advantage of two converging technological trends: increased patient access to electronic medical records (EMRs), and the burgeoning availability of collaborative "Web 2.0" technologies that upgraded the level of measurement accessible to patients.

From Sharing Anecdotes to Controlling Their Data

Gimme my damn data; it's all about me so it's mine—E-Patient Dave

The traditional doctor's office visit involves the creation of structured data (the medical notes) from unstructured anecdote (the medical history). Historically, medical notes have served as an *aide memoire* for clinicians and a means of record keeping and communication with colleagues, but were never intended to be read by patients. The advent of electronic medical records (EMRs) means that barriers for patients to access them are lowering rapidly. Systems such as "My Health*e*Vet" within the Veteran's Administration (VA) have shown that most patients (84 %) find accessing their records useful, and about half felt it improved their communication with their healthcare provider (Nazi et al. 2013). While patients have been enthusiastic, physicians have shown less support and focused more on the potential for problems such as increasing their workload or changing how they would document things in the record (Ross et al. 2005). Within the United States, resistance is likely to be overcome to some extent by the Health Information Technology for Economic and Clinical Health (HITECH) act of 2009, which offers financial incentives to physicians that offer "meaningful" use of EMRs to their patients (Jha 2010). Such incentives may be needed to conquer institutional inertia; within the VA pilot, only 6 % of doctors had told their patients about the system (Nazi et al. 2013), and so widespread adoption will require continuous encouragement.

Some early patient adopters found individual benefits from their EMRs (with data managed by health providers) or personal health records (PHRs, with data controlled by patients, sometimes using imported health provider data). For example the now famous case of "E-patient Dave" started when cancer patient Dave deBronkart downloaded all of his medical records into the now defunct "Google Health". What he found was disturbing: incorrect dates, missing diagnoses, misdiagnoses, and most disturbingly of all, no mention of his allergy to steroids (deBronkart 2009). When it comes to research, scientists might do well to heed E-patient Dave's words of warning, but also his call to arms at TEDx Maastrict: *"Let patients help"*.

Patient, Know Thyself

In medical measurement, the ability of *objective* tools and measures to circumvent biases of human perception makes them preferred data sources wherever possible. However they require trained professionals with sophisticated equipment, and despite medical advances many conditions lack objective measures. In such cases, *subjective* measures may be applicable, though they are inevitably less reliable, repeatable, or sensitive.

A typical subjective clinician-lead tool is the clinical symptom assessment, which manifests as an interview between doctor and patient. For a wide range of illnesses, standardized measurement scales have been devised, often with accompanying training to ensure a level of consistency across clinical staff. Such measurements are the mainstay of many clinical approaches to studying and managing serious, chronic or progressive illnesses. The biggest limitations of clinical symptom reporting are that they are resource intensive (relying on expensive staff) and cannot be done frequently enough: typically, that means once or only a few times each year.

Another source of subjective data derives from the patient's perspective, unguided by a clinician such as a symptom diary or a patient-reported outcome questionnaire. Symptom diaries might be prescribed by clinicians managing a chronic asthma patient, for example, as a tool to tease out particularly complex interactions between environment, behaviour and disorder, which occur primarily outside of the care environment. Individually, symptom diaries may allow individuals to pinpoint behaviours or circumstances that precipitate worsening symptoms and at a group level became increasingly recognized as potentially valuable in clinical trials (Santanello et al. 1997). One significant limitation of these tools (particularly when completed on paper) is the "parking lot effect" which finds less diligent patients scrambling to complete their assigned homework in the minutes just before their next clinic visit (Stone et al. 2003).

Keeping with the topic of patient-reported measures, self-report questionnaires have historically been common in psychiatry, where a patient's own thoughts are the most reliable predictor of outcomes. Measures such as the Beck Depression Inventory (Beck et al. 1961), developed in the 1960s differed from earlier psychiatry models in that they took the patient's direct experience (and even the terminology they used for symptoms) and quantified them through simple scoring systems that mapped to theoretical models of disease (such as anhedonia, negative self cognitions, and somatic symptoms in the case of depression). Outside of psychiatry, self-report gained increasing prominence in the late 1980s as measures of "health-related quality of life" was increasingly recognized as an important adjunct to objective measures (Tarlov and Trust 1989) in conditions like human immunodeficiency virus (HIV) or cancer.

More recently a broader range of generic and disease-specific questionnaires have been developed, called "patient reported outcomes" (PROs), which have raised to a standard of reliability where appropriately developed (Food US. Drug Administration 2009) self-report questionnaires are increasingly used as endpoints in trials (Basch 2012), and indeed these tools have come to form a core feature of the next generation of online tools for medical human computation. Crucially, they provide patients themselves with access to the same standard of measurement as has traditionally been available only to medical professionals. This wider distribution of self-made and shareable tools would have been welcomed by the founders of the Whole Earth Catalog and has recently formed the basis for a more disruptive approach to computing outcomes in medicine: finally, we can let patients help.

Medicine 2.0

The Internet loves a buzzword, and in 2004 the term "Web 2.0" was coined to describe the plethora of Internet sites that allowed users (rather than central authorities) to collaborate and contribute dynamic (rather than static) user-generated content in entertainment (e.g. YouTube), photography (e.g. Flickr), knowledge (e.g. Wikipedia), and even friendship (E.g. Facebook) (Van De Belt et al. 2010). "Medicine 2.0" (or "Health 2.0") refers to the use of these Web 2.0 technologies (and philosophies) to increase patient participation and empowerment through the use of new information and communications technologies (with or without professional involvement), using social networking to develop a new type of health care collaboratively through more effective use of medical data (Van De Belt et al. 2010).

One community that exemplifies this movement is the website PatientsLikeMe. The company was founded in 2004 by brothers Ben and Jamie Heywood to help

Fig. 1 Patient profile of the inspiration for PatientsLikeMe, Stephen Heywood

find creative solutions for their brother Stephen Heywood, who was diagnosed with amyotrophic lateral sclerosis (ALS) aged just 29. A family of MIT graduates, they partnered with their friend Jeff Cole to create a site that took the scientific rigor of a clinical trial and matched it with the personal connectivity of an online dating site. Based near their *alma mater* in Cambridge Massachusetts and opened in 2006, the online ALS community had features of the older online communities like Braintalk such as a forum, but focused on structured, rather than unstructured data. ALS patients could enter their own PRO, the ALS Functional Rating Scale (Revised) (Cedarbaum et al. 1999), which was widely used in clinical trial research but not normally available to patients. Not only did they make it available but they helped patients to graph their displays visually over time, with the declining slope of their ALSFRS-R score profiled against the relative rates of decline of every other patient "like them" in the system (see Fig. 1). In addition, every member who completed this PRO was given a virtual avatar to represent them, known as the "stickman", which boiled down the technical questions of the ALSFRS-R into an easily understood set of iconography colour coded from green (an unaffected body region) to red (severe disability). Therefore a patient with severe problems speaking and swallowing (red head on their stickman) but who was still able to walk, breathe, and self-care (green legs, chest, and arms) would be able to quickly scan through the list of other patients and so quickly find a "patient like me".

Virtuous Circle

By using these newly acquired PRO tools to upgrade their level of data collection from anecdotal to observational, patients set a new benchmark in elevating their discourse to become closer to that of traditional health researchers. Learning more about themselves through PROs and visualization tools yielded benefits too, illustrated as a "virtuous circle" in Fig. 2. This diagram outlines the ways in which patients on PatientsLikeMe can not only track their progress with medical data, but use this data to connect with other patients who are most like them; they don't just have to listen to whoever is chattiest in the forum or logged on most recently, they could search for another ALS patients who was young at their age of onset, who lives in Massachusetts, or who had tried baclofen for stiffness. Tools which were unavailable even in the most advanced ALS clinic in the world were now in the hands of patients to collect their own data, form their own hypotheses, and eventually, develop their own research.

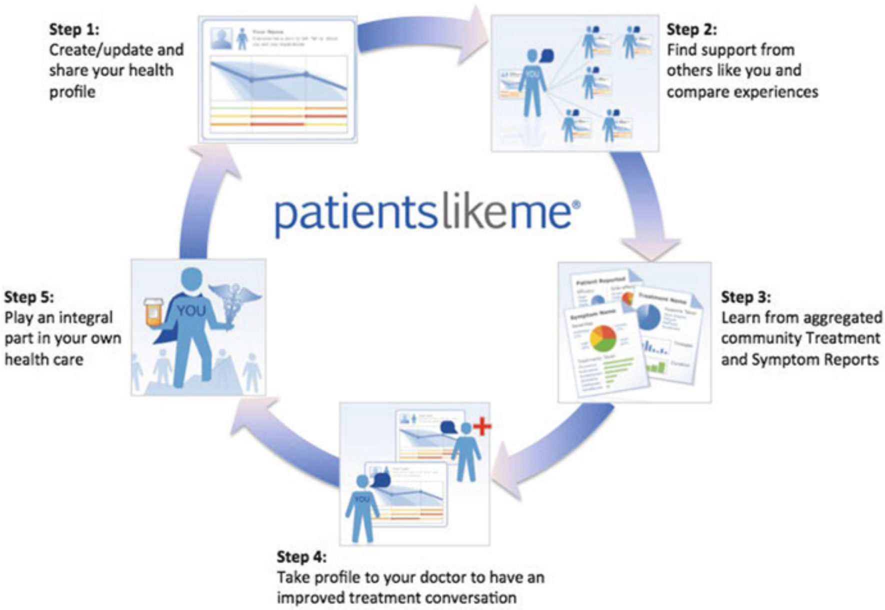


Fig. 2 The "virtuous cycle" of shared human computation underlying PatientsLikeMe

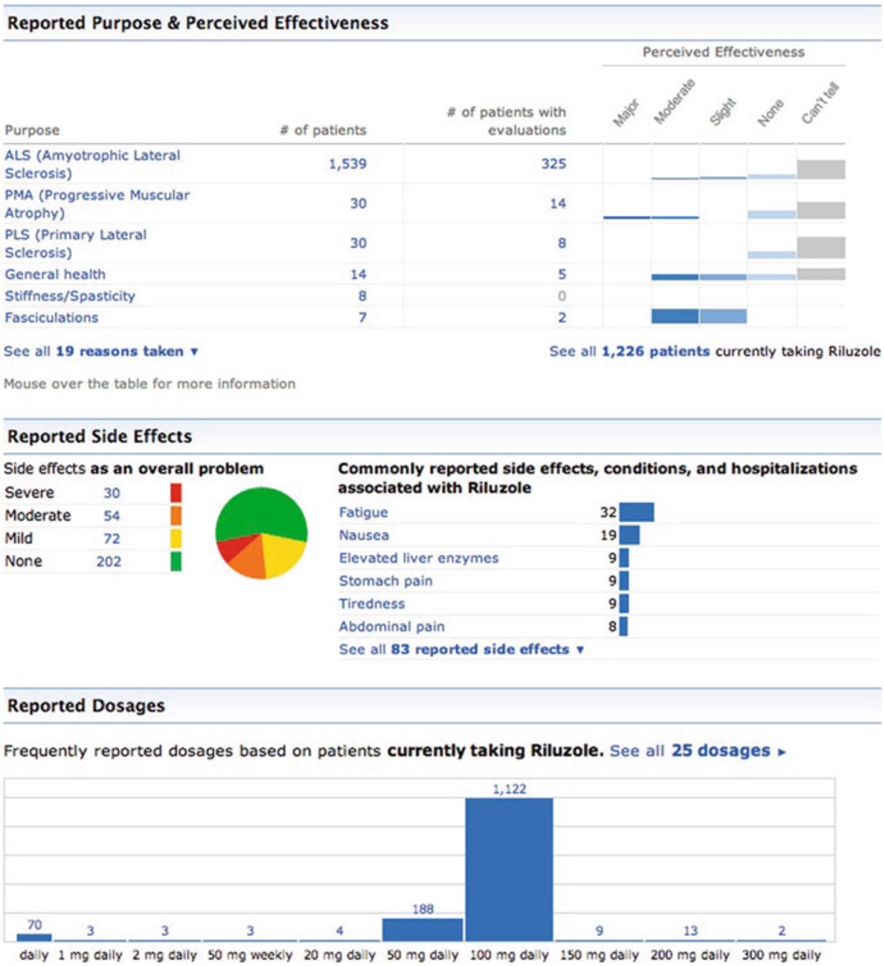


Fig. 3 Treatment report for the drug Riluzole® consisting of aggregated self-report data from individual ALS patients. Note that recommended dosage of Riluzole is 50 mg twice daily; this "real world" data shows outliers (300 mg) but also a low rate of erroneous entries (e.g. 1 mg daily)

Even without the desire to personally conduct their own human computation work, the site encouraged the interplay between the provision of social support in creating machine-readable data, and encouraged members to donate this data towards aggregated reports which allow members to see themselves in the context of, say, everyone else taking the same drug as them along with the side effects and dosage range (Fig. 3) or experiencing the same symptom including the severity and treatment options (Fig. 4).

Preliminary evidence for the virtuous cycle comes from two self-reported surveys in the peer-reviewed literature. The first was conducted in six communities

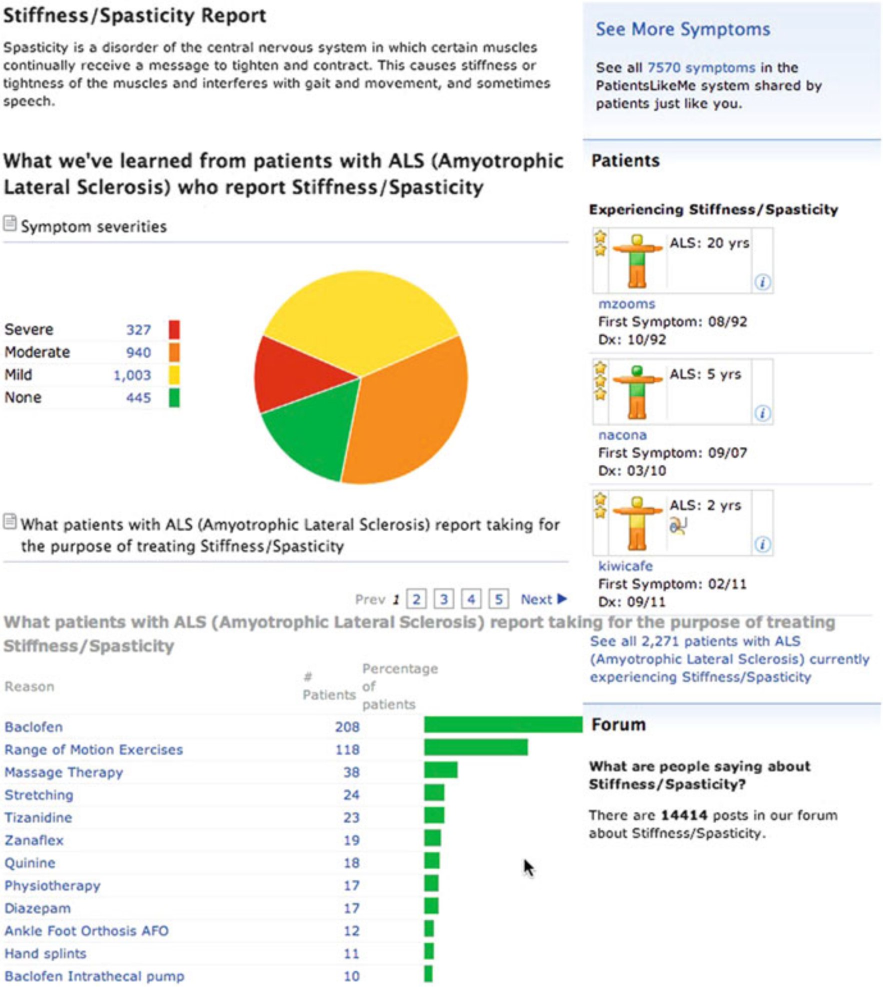


Fig. 4 Symptom report for stiffness and spasticity among ALS patients including perceived severity, recommended treatments, individual reports, and relevant forum-based discussions

(ALS, MS, Parkinson's disease, HIV, fibromyalgia, and mood disorders), and identified a number of perceived benefits to those engaged in the circle (Wicks et al. 2010). More than half of patients responding (57 %) found PatientsLikeMe to be helpful for understanding the side effects of treatments—in part because rather than the flat list of alphabetically listed side effects identified in trials that are reported in the prescribing information, the data available to patients comes from other patients like them, filtered through their unique experience but aggregated through visualization (Fig. 3). Most patients (72 %) reported value in using the system to learn about symptoms they experienced (Fig. 4)—by allowing patients not only to

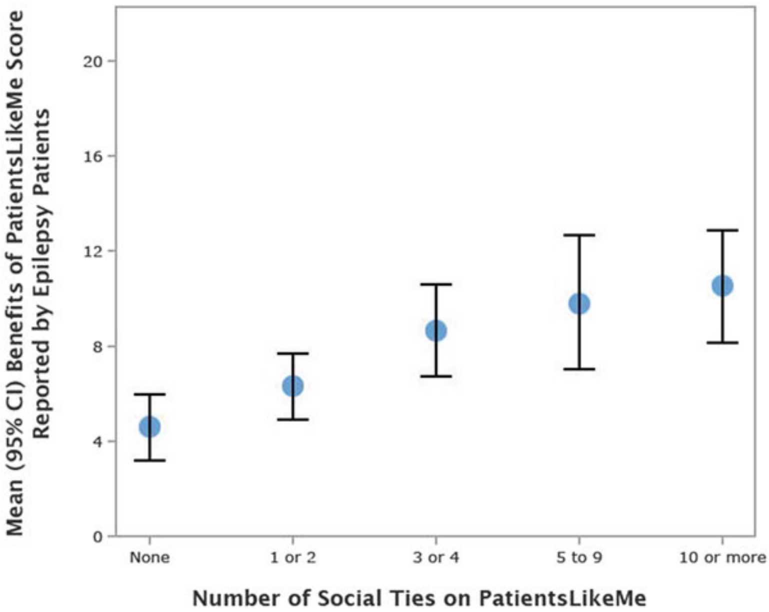


Fig. 5 "Dose effect curve for friendship"—Benefits experienced from using PatientsLikeMe (y-axis) against number of "connections" with other members in the community (x-axis) (Reproduced with permission from Wicks et al. 2012)

longitudinally track their own symptoms but also to use powerful search tools to help find and connect to other patients with similar experiences in order to learn from them. Perhaps most encouraging of all, a substantial minority (42 %) reported being more involved in their treatment decisions as a result of their use of the system and most patients (66 %) reported their healthcare professional team were supportive of their use of PatientsLikeMe.

One question arising from this study was the degree to which these benefits were only really accruing to those who engaged more deeply in the system, and therefore the cycle. A second study was created to replicate the original study in a newer community, epilepsy, and to build in an additional hypothesis to test whether the degree of social involvement was relevant. Within the epilepsy community a number of similar benefits were reported in terms of observations about treatments, symptoms, and management of their condition, as well as some intriguing condition-specific benefits which have triggered further study; 30 % of users felt they got better care as a result of using PatientsLikeMe, 27 % improved their medication compliance, 27 % reported reduced treatment side effects, 18 % felt they needed fewer ER visits, and 17 % reported that specifically from interacting with the site they had sought out an epilepsy specialist (Wicks et al. 2012). The epilepsy study also shed further light on the role of peer interaction in use of the site. In constructing a score of potential benefits experienced by epilepsy users, ranging from 0 to 20, the most predictive variable (even accounting for number of logins) was the number of social ties that a given patient had with other patients on the website (Fig. 5). Importantly then, it is not just the presence of data-tracking tools or even aggregated reports that was key in providing energy to the virtuous circle—it was interaction and engagement with other human actors who could interpret, contextualize, and help to synthesize diverse sets of data to address specific challenges. The authors referred to this finding as "a dose-effect curve for friendship". This finding is current being explored further in a more formal setting in collaboration with the Epilepsy Centers of Excellence (ECOE) of the VA.

Accelerating Research Through Human Medical Data Sharing

Since the site's early days, the PatientsLikeMe team included a number of scientists who worked alone and in harmony with external collaborators to begin climbing the pyramid of scientific credibility that could be achieved on the platform. An early study drew upon the experience of forum members experiencing a highly unusual symptom; uncontrolled outbursts of yawning—dozens, even hundreds of times per day, which in patients with a weakened jaw muscle due to the atrophy of ALS could become painfully dislocated. In response, the PatientsLikeMe team added a symptom "excessive yawning" to their standard battery of items and within a matter of weeks gathered data from 539 ALS patients and published the results, their first scientific output in a peer-reviewed article (Wicks 2007). By contrast, in prior studies using paper-and-pencil based methods it had taken a year's solid recruitment efforts just to recruit 104 patients from the largest ALS center in Europe (Wicks et al. 2007).

While "building a better mousetrap" for observational research was somewhat gratifying, the unique nature of online communities to enable human computation would help the team not only climb the credibility pyramid, but bring new entrants to participate. Cathy Wolf is a quadriplegic psychologist, writer, and poet who has lived with ALS for 17 years, and is only able to communicate via advanced technologies such as muscle sensors, eye gaze trackers, and even brain-computer interfaces. One day, as she used PatientsLikeMe to measure her decline in function on the ALSFRS-R scale, she scored a zero and realized that as far as researchers were concerned, she'd "bottomed out" of the scale. In response she wrote *"I have NOT bottomed out! If (researchers) can't think of objective measurements for PALS on the ventilator, let me educate him/her."* For instance, on the "communication" part of the scale, once a patient lost their ability to speak or write, they scored a zero. But as Cathy herself said *"there is a range of communication… Some talk, some use a physical keyboard, some use an onscreen pointing keyboard, some use multiple switch scanning, some single switch scanning. These are related to motor ability."* It became clear quickly that digital technology was allowing patients to have new experiences of disease that had never been measured before. And so, with Cathy as a co-author, PatientsLikeMe conducted the first study to survey patients who'd "bottomed out" of the traditional research scale to find out what they could still do. In all, they gathered data from 326 patients, many of whom were too sick to make the journey to hospital for traditional research visits, and together the team published a study that developed three new “extension” items called the ALSFRS-EX (Extension) which covered the remaining ability of patients to communicate emotion in their facial expressions; to manipulate switches with their fingers, and to move around inside their own homes even when they couldn’t walk outside (Wicks et al. 2009). In this way, the participation of citizen scientists enabled by an online platform allowed patients to be “participants” in research in the truest sense of the word.

In 2013, PatientsLikeMe was awarded a grant by the Robert Wood Johnson Foundation which will permit the development of an “Open Research Exchange” to allow developers of new PROs to prototype their questionnaires on PatientsLikeMe to more rapidly validate them with patient input. It is hoped that by accelerating the developing of PROs, patients with more conditions will be able to realize the same benefits as the ALS community has found in having a PRO they can control that is taken seriously by the wider medical community.

From Phenotype to Genotype

Around the time PatientsLikeMe was making strides in the phenotypic world of human computation, on the other coast of the United States in Mountain View California, 23andMe was doing the same for the genomic world. Founded in 2006 the company sold genetic tests normally only available to clinicians and researchers direct to the consumer (“DTC Genetics”) in order to provide entertaining insights (“how closely related are you to Cleopatra?”), support genealogy research (“what’s your maternal haplotype?”), and increasingly, support clinical research (“what sort of mutations do we find in individuals with Parkinson’s disease?”). The company caused ethical controversy at the time of its launch because in later versions of the product, consumers could reveal their risks of highly predictive single nucleotide polymorphisms for disease-causing genes such as BRCA-1 (breast cancer) and APOE-4 (Alzheimer’s disease).

Leaving such controversies aside for our purposes, the primary interest to medical human computation lies in the company’s commitment to combine genotypic and phenotypic data to find new discoveries. 23andMe first started establishing their scientific credibility by replicating benign known findings such as genetic variation underlying skin freckling or hair curl using online distributed methods (Eriksson et al. 2010). This replication would set the stage for later discoveries such as new reported associations between genes and human health traits like myopia (Kiefer et al. 2013). In support of further opportunities for human computation, participants in 23andMe are able to download their data and upload it to other “citizen science” communities. In this way many people can be “data donors” and leave the more complex analysis to those with the skills and expertise to do so (Swan et al. 2010). Although the advantage clearly lies with the organization itself to most rapidly make new discoveries, it is certainly possible that the next generation of health discoveries could originate from among their 200,000 members.

Supporting the expanded need for self-educating among their members, both 23andMe and PatientsLikeMe embrace “open access publishing” which allows a wider swathe of readers to access their scientific output than might otherwise be possible—in this way their members can more readily contribute data, ideas, and their own analyses to the human computational field. By contrast, the traditional medical establishment does research *to* patients—it extracts data *from* them, *blinds patients* in clinical trials as to their own treatment arm to maintain the integrity of the experiment, and then withholds the findings from the very people who participated by publishing their findings in closed-access journals. No wonder then, that as patients become more educated and engaged, they also become more dissatisfied with the status quo and less willing to be an obedient subject of centralized computation.

Whose Trial Is It Anyway?

Observational studies and correlational analyses are all well and good, but they never cured a patient of anything. The only way that Medicine 2.0 could effect major change in medicine was to climb the next layer of the pyramid to human research trials. The double-blind randomized placebo controlled trial (RCT) has been a gold standard of medicine since the 1950s. Randomizing one group of patients to receive active treatment and another to receive a sugar pill, (with neither patients nor healthcare professionals knowing who was in what group) was the only reliable way to factor out many biases which could cloud the quality of medical decision-making. For all the plaudits it has earned in medicine, however, patients themselves have not always been so enthusiastic.

In the early 1980s, people with HIV had no effective treatment and a bleak prognosis. In 1988 more than a thousand patients vocally expressed their anger and frustration to the US Food and Drug Administration (FDA) at their headquarters in Maryland about the maddeningly slow pace of RCTs to find effective treatments for their conditions. In the context of a rapidly lethal and infectious disease, waiting for early stage testing to be completed in healthy volunteers, rather than patients, felt like an unnecessary delay. After all, the patients reasoned, what safety issue that a drug has could be worse than HIV? Furthermore the idea that a doctor might intentionally provide a placebo that he knew would do nothing seemed particularly objectionable. Within a week the FDA updated their regulations to speed approvals for HIV research, but the seeds of patient revolution had already been sown.

Some HIV patients taking part in trials would swap pills or redistribute them amongst their fellow patients, even giving their medication to sympathetic pharmacists to try and decipher which were placebos (Murphy 2004). The groundswell of dissatisfaction among HIV patients was an early signal that patients could “hijack” a trial and even force regulators to speed their bureaucratic processes under enough pressure, but what happened next was truly revolutionary.

Gastrointestinal stromal tumor (GIST) is one of the most severe of the 200 or so cancers with which one could be diagnosed. Affecting the soft tissue of the gastrointestinal tract, GIST frequently metastasizes rapidly to the peritoneum and liver, is resistant to chemotherapy, and, left untreated, confers a median survival time of less than 2 years after metastasis. As a relatively rare disease with an incidence rate of only 6–15 cases per million people per year, recruiting sufficient patients to power a clinical trial has always been challenging, and so the role of non-profits in GIST has included not just the provision of information or support, but also assistance with clinical trial recruitment. In 2000 a large clinical study was initiated by the drug company Novartis® for their new drug Gleevec® with an aim to recruit some 800 patients with the disease to test for the drug's effect on survival and metastasis.

In addition to the trial data collected by Novartis, an Internet based patient non-profit, "The LifeRaft Group", set about collecting patient-reported questionnaires over the Internet from those taking Gleevec, their dosage, side effects, response to treatment, and via their caregivers, even their death. No participant was excluded from the study; it included all comers whether they were already in an authorized clinical trial or were receiving the drug from their doctor as part of routine care. Using retrospective self report data of all comers, the LifeRaft Group correctly anticipated the result; patients most recently reporting the lowest dose of Gleevec died after a median of 5 years, while the median patient most recently taking the higher dose were still alive at the time of survey (Call et al. 2010). Subsequently verified by traditional RCTs, the authors themselves were keen to point out that their data provided a "real-world complementary perspective to that seen in investigator-initiated randomized trials". It wasn't perfect, but it was a pivotal point showing that patient self-reported data had utility.

Nevertheless, in the case of GIST the data was submitted by a distributed group of patients but analysis remained in the hands of a centralized organization. Later in the 2000s, as tools for collaboration and analysis became more widely available, human medical computation seized upon a small finding to demonstrate its full potential.

A Patient-Lead Clinical Study Online

"Now, we monitor, watch and wait."—Leo Greene—ALS patient and journalist (http://www.dailybulletin.com/leosstory/ci_8089973)

In early 2008, an Italian group of clinicians published a study entitled "*Lithium delays progression of ALS*" in the prestigious *Proceedings of the National Academy of Sciences* (PNAS) (Fornai et al. 2008). In their study they compared 28 ALS patients on Riluzole®, the only approved drug for ALS (which provides 2–4 months additional lifespan (Miller et al. 2012)) to just 16 ALS patients on Riluzole and lithium carbonate. During the 15-month observation window a third of the Riluzole-only patients died, compared with none of the group supplementing their Riluzole with lithium. Even before the PNAS paper was officially published word spread through

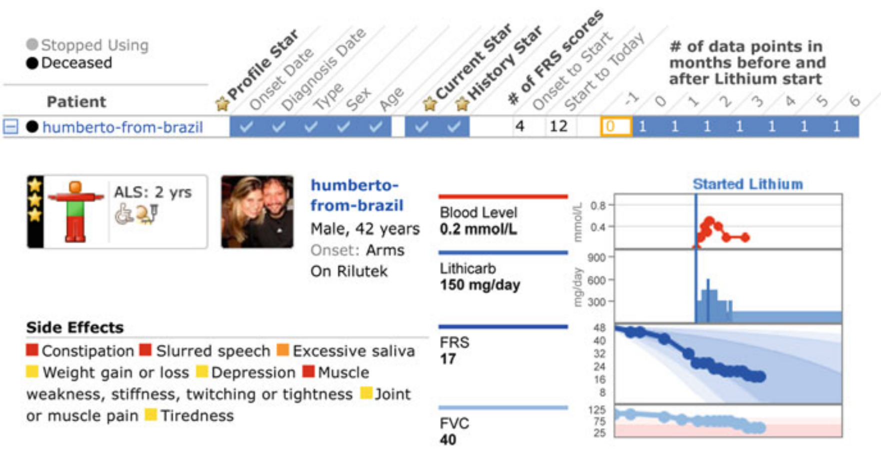


Fig. 6 PatientsLikeMe's ALS lithium study tool profile of advocate Humberto Macedo (Reproduced with permission from Wicks et al. 2011)

the community as enterprising ALS patients used "Google Translate" to interpret Italian-language conference abstracts describing the findings. As a widely available drug for the treatment of bipolar disorder, many patients with ALS begun sourcing the drug off-label from sympathetic doctors, in the hope that they might see the type of near-miraculous slowing of disease that Fornai et al. reported (Frost et al. 2008).

This time it was patients who lead the charge. ALS patient Humberto Macedo (living in Brazil) and ALS caregiver Karen Felzer (whose father suffered from ALS) collaborated to build a website where ALS patients could find out more about lithium, and links to a "Google Spreadsheet" that would allow patients who had obtained lithium off-label to track their progress using self-reported side effects, dosages, and even ALSFRS-R scores. Around this time the research team at PatientsLikeMe believed they could offer a more robust method of data capture and so modified their platform to collect more orderly structured data, such as ensuring that the ALSFRS-R was presented in a consistent fashion, and that side effects could be entered in a structured manner to allow later analysis (see Fig. 6).

In the space of a few months, there were over 160 ALS patients reporting their use of lithium with the tool; ten times the sample of the original PNAS paper. Furthermore, the open nature of the tools available such as Google Spreadsheets and PatientsLikeMe meant that patients themselves were extracting the data, visualizing it, and running their own statistical tests on the data to try and discern treatment effects. Although they lacked the statistical or methodological sophistication of a formal clinical trial, it was hoped that if Fornai et al.'s results were true, then even such crude measurement would discern a treatment effect quickly. For the first time in a decade the mood of the ALS community was ebullient and energized—an effective treatment was finally here.

Unfortunately however, the halting of progression failed to materialize. Patients worsened, some of the early advocates (sadly, including both Humberto

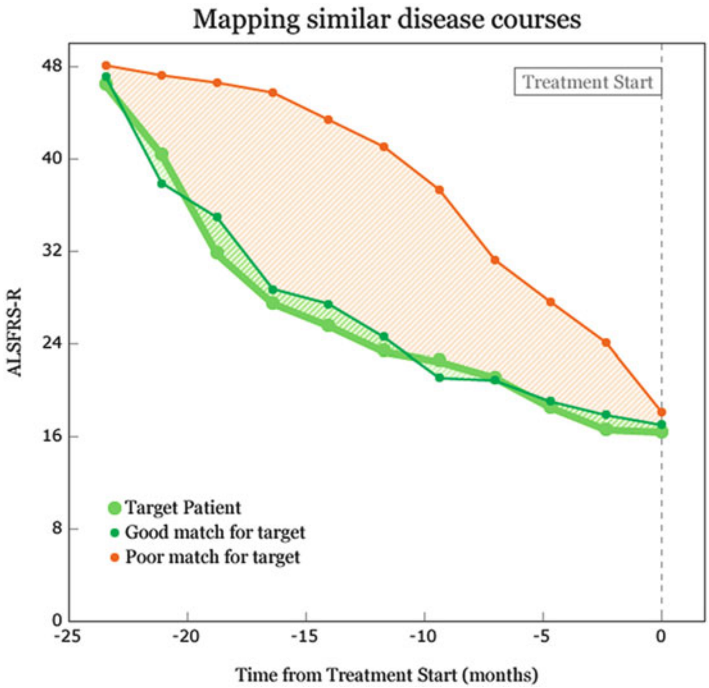


Fig. 7 Dots represent ALSFRS-R scores of two hypothetical patients progressing along different paths, who by traditional matching criteria would be considered comparable. The ellipse describes distance in progression curves; the PatientsLikeMe matching algorithm minimizes this area for each patient (Reproduced with permission from Wicks et al. 2011)

and Karen's father) passed away from complications of their ALS. The research team at PatientsLikeMe worked through a number of analytical approaches that would resolve the question as best as it could be worked through with the data to hand, finally culminating in a publication in *Nature Biotechnology* that described a novel matching algorithm, the disappointing results, and a de-identified copy of the entire ALS dataset so that others could try it for themselves (Wicks et al. 2011).

In order to account for the lack of a placebo arm in their open, self-reported clinical trial, PatientsLikeMe harnessed the collective power of the broader ALS community who were *not* self-experimenting with lithium by matching between three and five members of the ALS community with each lithium-taking patient. Unlike a traditional RCT that can only collect data at the study's baseline, the online community had already been passively submitting their ALSFRS-R outcome data for years before lithium was even identified. Therefore the researchers were able to match each lithium-taking patient with those non-lithium taking patients who were most similar to them along their entire disease course up until the point of deciding to take the drug (Fig. 7). This was the first truly patient-initiated study where an entire community donated their collective experiences to identify a potential cure for their disease. Unlike Gleevec in GIST, however, sadly lithium didn't work. The five traditional RCTs commissioned by government funders and non-profits around the world were all halted for futility—nobody ever replicated Fornai et al.'s findings again (Armon 2010).

How Continuous Automated Measurement Supports Human Computation

In each of the examples provided thus far, online medical discoveries have relied upon patient-reported data; whether it's survival in GIST through responding to a survey, the identification of traits through questions on 23andMe, or the completion of validated patient reported outcome measure on PatientsLikeMe. Because this data is easily provided by patients and, pending validation, can potentially rise to the same level of clinical relevance as a clinical measurement, they were an obvious place to start. But relative to these subjective measures, a truly automated, sensor-based objective measurement has the potential to add an additional level of sophistication: the measurement might be taken without any human intervention or initiation, for example, by a passive sensor like an accelerometer or GPS at the same point in time every day. This would allow surreptitious objective medical recording of a patient's health state such as their mobility, mood, or other physiological characteristics absent the "Hawthorne Effect" which means people tend to alter their behavior when they're being measured. Thanks in large part to tremendous advances in technology over the last century, such automated, objective, continuous symptom measurement is emerging outside of the intensive care unit or astronaut training center.

With the advent of "always on", objective, wearable monitoring devices such as the FitBit One, Nike+, and Jawbone UP, coupled with smartphone apps, it is now easier than ever to continuously record health-related measures such as pulse rate, activity, sleep duration and calorific intake. Enabled by this technology, a loose-knit global movement called "The Quantified Self" has emerged. These individuals use technology to record continuous objective data about their health, often sharing it freely with like-minded individuals to amass large-scale records of changing health data over time. Such data can be mined using statistical tools to detect changes in health status or even perform "n of 1" experiments (Swan 2012) which in a medical context are at the peak of the evidence pyramid. Correlating these sensor-derived data feeds with environment, social or genetic data may lead to insights that, if acted upon appropriately, could significantly alter the course of an individual's health or disease. It is not an unreasonable prediction that, as the monitoring technology becomes more ubiquitous and tangible, and individual and public health benefits become clear, we will likely see that continuous objective symptom recording will eventually become the norm for a significant fraction of the population.

Much as Moore's law predicts that the number of transistors on an integrated circuit doubles every 18 months or so, an extension of this law predicts that the price-to-performance ratios for many kinds of digital consumer products, such as smartphones, drops at a corresponding rate. This means that objective symptom monitoring technology will inevitably become so inexpensive and non-intrusive, that data from an entire population of millions could be recorded at almost negligible cost. A company like Apple or Google's access to smartphone data, for instance, would be unparalleled in human history and tantamount to a large-scale real-time sensor network. Wearable computing technologies such as Google Glass add a new dimension of head-mounted image or video capture—imagine tracking an outbreak of flu through facial recognition software in a population of citizens wearing such devices.

Glimpses of this new era of population-scale symptom recording are emerging through studies of the mass-scale details of telephone conversations held by mobile telephone companies, or Internet search firms. These studies have unearthed characteristic patterns of social interactions that appear to correlate with mental health and psychosocial disorders, such as the mapping of Google searches for mental health problems mapping to seasonal trends (Ayers et al. 2013). Based on billions of internet keyword searches from across the globe, the Google Flu Trends project has even been able to predict localized influenza outbreaks in real-time, with better accuracy and more rapidly than traditional influenza monitoring methods used by the US Centers for Disease Control (Dugas et al. 2013).

Because such observational data on symptoms has previously been unavailable, epidemiology has, historically, been unable to model the time variation of symptoms across a population. The arrival of ubiquitous personal digital sensing technology will likely change this situation, so that very large scale, classical epidemiological models will, for the first time, have the empirical data to make real-time predictions on the outcome of critical public health decisions.

The Ultra-low Cost, Global Reach of the Parkinson's Voice Initiative (PVI)

There is no simple blood test or other biomarkers for another neurological disease that requires careful monitoring: Parkinson's disease. Parkinson's is generally assessed in the clinic behaviorally, by asking patients to tap their fingers together in front of them or by observing the rate at which their limbs shake or how they walk. Research by one of the authors (Max Little) and collaborators, has demonstrated that it is possible to quantify the symptoms of Parkinson's disease on an objective, clinical scale, by a sophisticated combination of algorithms that analyze voice recordings and statistical machine learning (Little et al. 2009). Using lab-quality recordings, it was shown that this approach can achieve up to 99 % accuracy in replicating an expert clinical diagnosis of Parkinson's (Tsanas et al. 2012), and an error of less than the disagreement of two qualified experts, about the severity of symptoms (Tsanas 2010). The simplicity of recording the voice using a wide array of digital microphones available to most of the global population, raises the question of whether the standard, global telephone network could be used. In this way the potential for patients themselves to take charge of their own sensors and integrate them into their own daily management is far greater than when collection is tethered to sophisticated lab equipment.

To address this, one needs to ask: will this technology work outside the lab? For a technology to be ubiquitous, it should be possible to reproduce the results without using specialized hardware or controlled settings. The PVI is an attempt to test the accuracy of the voice-based Parkinson's algorithms on telephone-quality recordings collected in a largely uncontrolled way. Participants contributed to the project by calling a number in one of nine countries, and going through a short set of vocal exercises lasting about 3–5 min in total. At the end of the 6-month data collection phase of the project, a remarkable 17,000 participants had donated voice recordings in English, Spanish, French and Portuguese, achieved at a total collection cost of just $2,000. At the time of writing, the analysis phase is ongoing.

Other efforts by the same group have provided people with Parkinson's disease and healthy with Android smartphones. In order to crowd-source better algorithms to help distinguish patients from controls, the authors collaborated with the Michael J Fox Foundation (MJFF) to release the passive behavioral data collected alongside clinical and other demographic data to a "Kaggle" analysis competition for a grand prize of $10,000. The competition received over 20 novel submissions, of which 2–3 were deemed by the co-applicant to be 'high quality'. These submissions included diverse feature extraction and machine learning approaches for making predictions, with, in some cases, around 90 % accuracy in separating Parkinson's patients from healthy controls.

Once diagnosed, Parkinson's disease is particularly interesting because the drugs used to treat its symptoms are very effective; a moderately disabled patient who cannot move, speak, or think clearly when they are "off" as a result of their disease can be restored to an active and fluid "on" state through the use of dopamine-stimulating drugs such as levodopa or dopamine agonists. These drugs, however, have side effects such as uncontrollable movements and can wear off in effectiveness over time—therefore it's important to carefully manage the drug regimen and fine-tuning of the time of day and dosage of anti-Parkinsonian medication can optimize the proportion of "on" time during the day for several hours.

Sara Riggare is a woman who has lived with Parkinson's disease for several decades, and as part of her PhD studies at the Karolinska Institute is building smartphone applications that allow her to monitor her degree of disability objectively through a finger-tapping test which is prompted by a medication reminder. In this way Sara serves as a "patient-researcher" who intends to "co-produce" research and crowd source data and potentially management algorithms through the use of distributed data tools. Although an early pioneer, we propose that as subsequent generations are diagnosed with life-changing illnesses they will view it as their responsibility not just to participate in studies, but to design them, to run them, to publish them, to critique them, and to harness their learnings to manage their own condition day by day with the support of their healthcare providers.

Implications for the Future

In this chapter we have seen how the potential for medical human computation evolved from the unstructured qualitative discussions of the pre-web Internet to modern forms of scientific co-production and human computation that empower patients to truly participate in research. As the potential of these systems matures we believe that there are major gains to be made in simple-to-use analytics platforms that can de-mystify some of the more technically complex aspects of medical research such as statistics or hypothesis testing, and make clinical discovery for patients that live with the disease as common an activity as online shopping. These co-producers will no more need to analyse the statistical complexities underlying hypothesis testing than an online shopper needs to understand logistics chains.

There remain a number of major challenges to be addressed to attain this goal however. First is the issue of bias—to date it seems likely that the most active users of online systems are those patients who are younger and more educated (Bove et al. 2013). Although these can be addressed to some degree by over-sampling those who are under-represented, today's tools simply can't reach those who don't use the Internet or digital technology. This should get easier over time but there will probably be an inevitable "digital divide" that will remain unbridged for many people living with disease today.

Second is the issue of verification—until patient's electronic medical records can be securely authenticated at low cost it is impossible to confirm that someone self-reporting themselves as having ALS or Parkinson's disease truly does so. Although today there are few incentives for fraudulently pretending to have a serious disease like this, as the healthcare establishment begins to take more notice of such data, this is likely to change. It will be important not to lose some of the benefits that anonymity provides, however, and many patients remain afraid they will lose insurance cover or be discriminated against if they can be explicitly identified alongside their medical information. Until these policy failings are resolved there will be an inherent tension between identification and anonymity.

Third is the issue of privacy—the examples given in this chapter have concerned some of the most severe and disabling diseases a patient can experience—and so perhaps these individuals are less likely to mind the risks to their privacy against the severity of their diseases. But the worry for many developed health economies is not the rare lethal disease; it is the widespread chronic disease like diabetes, obesity, mood disorders, or back pain. It remains less clear whether patients with these disorders are as engaged with their health to submit regular data for research purposes nor whether they are willing to risk their privacy for the sake of conditions which may only be of mild or moderate intensity. Within this are very real concerns about discrimination, stigma, and loss of opportunity such as insurance coverage or hiring due to disclosures around health, which can only be addressed by legislation. Ferrari and Viviani explore these issues in more detail in the chapter "Privacy in Social Collaboration".

Finally there is the delicate issue "*cui bono*" (who benefits?). Patients donating their data to for-profit companies free of charge, analysts donating their cognitive surplus to improve the lives of people they'll never meet, and new organizations having access to big datasets that reveal more than we can possibly predict about ourselves—we approach this issue with hope and optimism based on the mission-lead nature of the organizations involved so far. But there is nothing to say that the tools described here couldn't also be used *against* patients—in raising insurance premiums on those who don't take their GPS-confirmed exercise, in refusing medical treatment to those who don't submit themselves to passive monitoring, in manipulating the prices of interventions to those who are shown to benefit the most through a quirk of genetics, perhaps even governments restricting the rights of people thought to be exposed to communicable diseases.

We agree with the patients who have themselves pioneered in this field; for now, the benefits outweigh the risks, but we must remain diligent and vigilant. The potential for empowering patients to join researchers in the quest to fight disease is incredible—we don't accelerate progress just by "standing on the shoulders of giants"—we accelerate progress by creating more giants.

References

Armon C (2010) Is the lithium-for-ALS genie back in the bottle? Neurology 75:586–587

Ayers JW, Althouse BM, Allem J-P, Rosenquist JN, Ford DE (2013) Seasonality in seeking mental health information on Google. Am J Prev Med 44(5):520–525

Basch E (2012) Beyond the FDA PRO guidance: steps toward integrating meaningful patient-reported outcomes into regulatory trials and US drug labels. Value Health 15(3):401–403

Basch E, Abernethy AP, Mullins CD, Reeve BB, Smith ML, Coons SJ et al (2012) Recommendations for incorporating patient-reported outcomes into clinical comparative effectiveness research in adult oncology. J Clin Oncol 30(34):4249–4255

Beck AT, Ward C, Mendelson M (1961) Beck depression inventory (BDI). Arch Gen Psychiatry 4(6):561–571

Bessière K, Pressman S, Kiesler S, Kraut R (2010) Effects of internet use on health and depression: a longitudinal study. J Med Internet Res 12(1):e6

Bove R, Secor E, Healy BC, Musallam A, Vaughan T (2013) Evaluation of an online platform for multiple sclerosis research: patient description, validation of severity scale, and exploration of BMI effects on disease course. PLoS One 8(3): e59707

Call J, Scherzer NJ, Josephy PD, Walentas C (2010) Evaluation of self-reported progression and correlation of Imatinib dose to survival in patients with metastatic gastrointestinal stromal tumors: an open cohort study. J Gastrointest Cancer 41(1):60–70

Cedarbaum JM, Stambler N, Malta E, Fuller C (1999) The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function. BDNF ALS Study Group (Phase III). (http://www.ncbi.nlm.nih.gov/pubmed/10540002) *Journal of the Neurological Sciences* 1999 Oct 31;169(1-2):13–21

Crocco AG, Villasis-Keever M, Jadad AR (2002) Analysis of cases of harm associated with use of health information on the internet. J Am Med Assoc 287(21):2869–2871

deBronkart D (2009) Imagine someone had been managing your data, and then you looked. e-patients.net

Dugas AF, Jalalpour M, Gel Y, Levin S, Torcaso F, Igusa T et al (2013) Influenza forecasting with Google Flu Trends. PLoS One 8(2):e56176

Eriksson N, Macpherson JM, Tung JY, Hon LS, Naughton B, Saxonov S et al (2010) Web-based, participant-driven studies yield novel genetic associations for common traits. PLoS Genet 6(6):e1000993

Eysenbach G (2003) The impact of the Internet on cancer outcomes. CA Cancer J Clin 53(6):356–371

Ferguson T (2007) e-patients: how they can help us heal healthcare. *Patient Advocacy for Health Care Quality: Strategies for Achieving Patient-Centered Care* 93–150

Food US. Drug Administration (2009) Guidance for industry: patient-reported outcome measures—Use in medical product development to support labeling claims. Fed Regist 74(235):65132–65133

Fornai F et al. (2008) Lithium delays progression of amyotrophic lateral sclerosis. *Proceedings of the National Academy of Sciences* 105.6: 2052–2057

Frost JH, Massagli MP, Wicks P, Heywood J (2008). How the social web supports patient experimentation with a new therapy: The demand for patient-controlled and patient-centered informatics. In AMIA Annual Symposium Proceedings (Vol. 2008, p. 217). American Medical Informatics Association

Gabler NB, Duan N, Vohra S, Kravitz RL (2011) N-of-1 trials in the medical literature: a systematic review. Med Care 49(8):761–768

Guyatt GH, Haynes RB, Jaeschke RZ, Cook DJ (2000) Users "Guides to the medical Literature XXV. Evidence-based medicine: principles for applying the users" Guides to Patient Care. J Am Med Assoc 284(10):1290–1296

Hoch D, Ferguson T (2005) What I've learned from E-patients. PLoS Med 2(8):e206

Hoch DB, Norris D, Lester JE, Marcus AD (1999) Information exchange in an epilepsy forum on the World Wide Web. Seizure 8(1):30–34

Jha AK (2010) Meaningful use of electronic health records. The J Am Med Assoc 304(15): 1709–1710

Kiefer AK, Tung JY, Do CB, Hinds DA, Mountain JL, Francke U et al (2013) Genome-wide analysis points to roles for extracellular matrix remodeling, the visual cycle, and neuronal development in myopia. PLoS Genet 9(2):e1003299

Lester J, Prady S, Finegan Y, Hoch D (2004) Learning from e-patients at Massachusetts general hospital. Br Med J 328:1188–1190

Little MA, McSharry PE, Hunter EJ, Spielman J, Ramig LO (2009) Suitability of dysphonia measurements for telemonitoring of Parkinson's disease. IEEE Trans Biomed Eng 56(4):1015

Meier A, Lyons EJ, Frydman G, Forlenza M, Rimer BK (2007) How cancer survivors provide support on cancer-related Internet mailing lists. J Med Internet Res 9(2):12

Mesa RA, Niblack J, Wadleigh M, Verstovsek S, Camoriano J, Barnes S et al (2007) The burden of fatigue and quality of life in myeloproliferative disorders (MPDs): an international Internet-based survey of 1179 MPD patients. Cancer 109(1):68–76

Michael Bowling J, Rimer BK, Lyons EJ, Golin CE, Frydman G, Ribisl KM (2006) Methodologic challenges of e-health research. Eval Program Plann 29(4):390–396

Miller RGR, Mitchell JDJ, Moore DHD (2012) Riluzole for amyotrophic lateral sclerosis (ALS)/ motor neuron disease (MND). Cochrane Database Syst Rev 3:CD001447–7

Murphy TF (2004) Case studies biomedical research ethics. MIT Press, Cambridge

Nazi KM, Hogan TP, McInnes DK, Woods SS, Graham G (2013) Evaluating patient access to electronic health records: results from a survey of veterans. Med Care 51(3 Suppl 1):S52–S56

Policy A (2011) Professionalism in the use of social media. J Am Med Assoc

Rheingold H (1993) The virtual community: homesteading on the electronic frontier. Addison-Wesley, Reading

Ross SE, Todd J, Moore LA, Beaty BL, Wittevrongel L, Lin C-T (2005) Expectations of patients and physicians regarding patient-accessible medical records. J Med Internet Res 7(2):e13

Rouleau CR, von Ranson KM (2011) Potential risks of pro-eating disorder websites. Clin Psychol Rev 31(4):525–531

Santanello NC, Barber BL, Reiss TF, Friedman BS, Juniper EF, Zhang J (1997) Measurement characteristics of two asthma symptom diary scales for use in clinical trials. Eur Respir J 10(3):646–651

Stone AA, Shiffman S, Schwartz JE, Broderick JE (2003) Patient compliance with paper and electronic diaries. Control Clin Trials 24(2):182–199

Swan M (2012) Crowdsourced health research studies: an important emerging complement to clinical trials in the public health research ecosystem. J Med Internet Res 14(2):e46

Swan M, Hathaway K, Hogg C, McCauley R (2010) Citizen science genomics as a model for crowdsourced preventive medicine research. *Journal of Participatory Medicine 2* (2010): e20

Tarlov AR, Trust PM (1989) The medical outcomes study: an application of methods for monitoring the results of medical care

Tsanas A (2010) New nonlinear markers and insights into speech signal degradation for effective tracking of Parkinson's disease symptom severity. Age (years)

Tsanas A, Little MA, McSharry PE, Spielman J, Ramig LO (2012) Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease. IEEE Trans Biomed Eng IEEE 59(5):1264–1271

Van De Belt TH, Engelen LJLPG, Berben SAA, Schoonhoven L (2010) Definition of health 2.0 and medicine 2.0: a systematic review. J Med Internet Res 12(2):e18

Waldron VR, Lavitt M, Kelley D (2000) The nature and prevention of harm in technology-mediated self-help settings: three exemplars. J Technol Hum Serv 17(2–3): 267–293

Wicks P (2007) Excessive yawning is common in the bulbar-onset form of ALS. Acta Psychiatr Scand 116(1):76; authorreply76–7

Wicks P, Abrahams S, Masi D (2007) Prevalence of depression in a 12-month consecutive sample of patients with ALS. Eur J Neurol 14(9):993–1001

Wicks P, Massagli MP, Wolf C (2009) Measuring function in advanced ALS: validation of ALSFRS-EX extension items. Eur J Neurol 16(3):353–359

Wicks P, Massagli M, Frost J, Brownstein C, Okun S, Vaughan T et al (2010) Sharing health data for better outcomes on PatientsLikeMe. J Med Internet Res 12(2):e19

Wicks P, Vaughan TE, Massagli MP, Heywood J (2011) Accelerated clinical discovery using self-reported patient data collected online and a patient-matching algorithm. Nat Biotechnol 29(5):411–414

Wicks P, Keininger DL, Massagli MP, la Loge de C, Brownstein C, Isojärvi J et al (2012) Perceived benefits of sharing health data between people with epilepsy on an online platform. Epilepsy Behav 23(1):16–23

Knowledge Engineering via Human Computation

Elena Simperl, Maribel Acosta, and Fabian Flöck

What Is Knowledge Engineering

Knowledge engineering refers to processes, methods, and tools by which knowledge in a given domain is elicited, captured, organized, and used in a system or application scenario (Studer et al. 1998). The resulting 'knowledge base' defines and formalizes the kinds of things that can be talked about in that particular context. It is commonly divided into a 'schema', also called 'ontology', and the actual data the application system manipulates. The data is described, stored, and managed as instantiations of the concepts and relationships defined in the ontology. With applications in fields such as knowledge management, information retrieval, natural language processing, eCommerce, information integration or the emerging Semantic Web, ontologies were introduced to computer science as part of a new approach to building intelligent information systems (Fensel 2001): they were intended to provide knowledge engineers with reusable pieces of declarative knowledge, which can be together with problem-solving methods and reasoning services easily assembled to high-quality and cost-effective systems (Neches et al. 1991; Schreiber et al. 1999). According to this idea, ontologies are understood as shared, formal domain conceptualizations; from a system engineering point of view, this component is strictly separated from the software implementation and can be thus efficiently reused across multiple applications (Guarino 1998).

The emergence of the Semantic Web has marked an important stage in the evolution of knowledge-driven technologies. Primarily introduced by Tim Berners-Lee

E. Simperl (✉)
Web and Internet Science Group, University of Southampton, Southampton, UK
e-mail: e.simperl@soton.ac.uk

M. Acosta • F. Flöck
Karlsruhe Institute of Technology, Institute AIFB, Karlsruhe, Germany
e-mail: maribel.acosta@kit.edu; fabian.floeck@kit.edu

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_13,

(2001), the idea of providing the current Web with a computer-processable knowledge infrastructure in addition to its original, semi-formal and human-understandable content foresees the usage of knowledge components which can be easily integrated into and exchanged among arbitrary software environments. In this context, the underlying knowledge bases are formalized using Web-based, but at the same time semantically unambiguous representation languages that are pervasively accessible and can (at least theoretically) be shared and reused across the World Wide Web. Although the combination of human-based computation and Semantic Web technologies yields promising results,[1] the implementation of such hybrid systems raises a whole set of new challenges which are discussed in detail in the chapter 'The Semantic Web and the Next Generation of Human Computation' of this book.

As a field, knowledge engineering is mainly concerned with the principles, processes, and methods that produce knowledge models that match this vision. It includes aspects related to knowledge acquisition, as a key pre-requisite for the here identification and organization of expert knowledge in a structured, machine-processable way, but also to software engineering, in particular when it comes to the actual process models and their operationalization. Last, but not least, knowledge engineering has strong ties to artificial intelligence and knowledge representation, in order to translate the results of the knowledge elicitation phase into structures that can be reasoned upon and used in an application system. In addition, it shares commonalities with several other areas concerned with the creation of models to enable information management, including entity relationships diagrams in relational data base systems engineering, and object-oriented programming, UML and model-driven architectures in software engineering. Each of these areas defines their specific way to capture domain knowledge, represent and exploit its meaning in the creation of innovative systems and applications.

In this chapter, we will look into several important activities in knowledge engineering, which have been the subject of human computation research. For each activity we will explain how human computation services can be used to complement existing automatic techniques, and give a short overview of the state of the art in terms of successful examples of systems and platforms which showcase the benefits of the general idea.

Why and Where Is Human Computation Needed

Many aspects of the knowledge engineering life cycle remain heavily human-driven (Siorpaes and Simperl 2010). Prominent examples include, at the technical level, the development of conceptualizations and their use in semantic annotation, the evaluation and curation of knowledge resources, the alignment of ontologies and data integration, as well as specific types of query processing and question answering

[1] Especially in tasks like data annotation or data quality assessment, which involve defining and encoding the meaning of the resources published on the Web or resolving semantic conflicts such as data ambiguity or inconsistency.

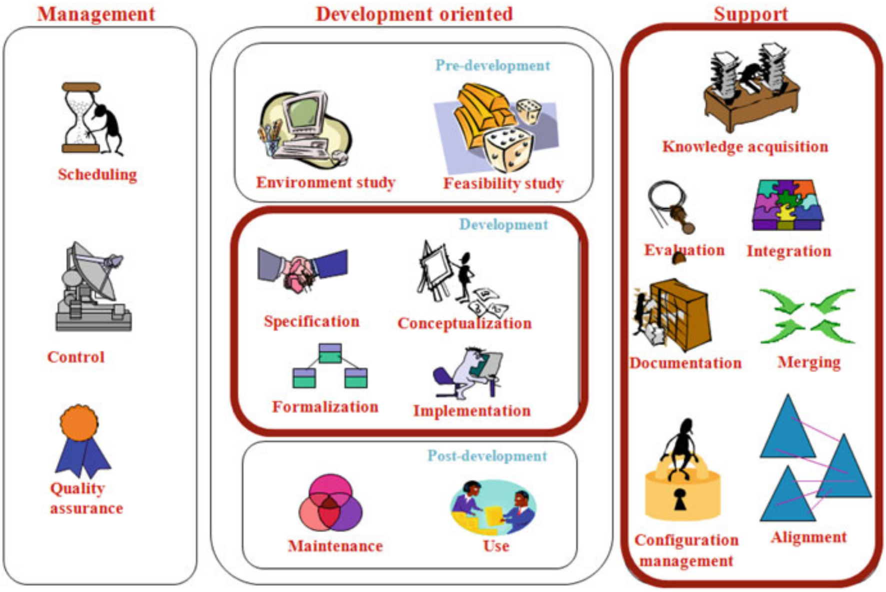


Fig. 1 Ontology engineering activities (Gómez-Pérez et al. 2004)

(Maribel Acosta et al. 2012). To an equal extent, it comprises almost everything that has to do with the creation of human-readable interfaces to such sources, in particular labeling, where human capabilities are indispensable to tackle those particular aspects that are acknowledged to be hardly approachable in a systematic, engineering-driven fashion; and also, though to a lesser extent, to the wide array of methods and techniques that have been proposed as an attempt to perform others automatically. In this second category, despite constant progress in improving the performance of the corresponding algorithms and the quality of their results, experiences show that human assistance is nevertheless required, even if it is just for the validation of algorithm outputs.

Figure 1 gives an overview of the knowledge engineering life cycle at the level of ontologies.[2] The activities we will look into are an essential part of knowledge (or ontology) engineering, but are not necessarily unique to this area. Nevertheless, compared to software engineering or relational data bases, it is primarily knowledge engineering, and in particular its use on the (Semantic) Web, that has been increasingly the subject to a crowdsourcing approach. This is due primarily to the (recent) strong Web orientation of knowledge engineering as a field, which led to a variety

[2] A similar process model applies to the creation, management and use of instance data. Management and pre-development activities cover the entire scope of the knowledge-engineering exercise. Development, post-development and support activities are equally relevant to both schema and data, though there might be differences in their actual realization. For example, instance data is typically lifted from existing sources into the newly created ontological schema, while a greater share of activities at the ontology level are carried out manually.

of knowledge base development projects and applications thereof being initiated and executed with the help of open Web communities, leveraging Web 2.0 participatory principles and tools. The high costs associated with creating and maintaining real-world knowledge bases in a classical work environment motivated experiments with alternative approaches that rely on the wisdom of open crowds, volunteer contributions, or services such as Amazon Mechanical Turk. Especially the latter is still an expanding field of research, with initial trials for various types of knowledge domains and tasks delivering very promising results. Nevertheless, in scenarios which are less open, both in terms of audiences addressed and the technologies they use, crowdsourcing methods need to take into account additional aspects to be effective, including the human and computational resources available, and how the results could be optimally acquired from, and integrated into, productive environments while avoiding to disrupt established workflows and practices.

The development life cycle in Fig. 1 distinguishes among management, development, and support activities. *Ontology management* refers primarily to scheduling, controlling and quality assurance. Scheduling is about coordinating and managing an ontology development project, including resource and time management. Controlling ensures that the scheduled tasks are accomplished as planned. Finally, quality assurance evaluates the quality of the outcomes of each activity, most notably of the implemented ontology. *Ontology development* can be split into three phases: pre-development, development, and post-development. As part of the pre-development phase, an environment study investigates the intended purpose and use of the ontology. Furthermore, a feasibility study ensures that the ontology can actually be built within the time and resources assigned to the project. These two activities are followed by the actual development, which includes first and foremost the requirements specification that eventually results in a conceptual model and its implementation in a given knowledge representation language. In the final, post-development phase, the ontology is updated and maintained as required; this phase also includes the reuse of the ontology in other application scenarios. *Support* stands for a wide range of different activities that can be performed in parallel or subsequent to the actual ontology development. The aim of these activities is to augment the results of the, typically manual, ontology development by automatizing parts of the process, providing auxiliary information sources that could be used to inform the conceptualization and implementation tasks, and evaluating and documenting intermediary results. Typical support activities include *knowledge acquisition*, *ontology evaluation*, *ontology alignment*, and *ontology learning* and *ontology population*. *Ontology population* is closely related to *semantic annotation*, by which information artifacts of various forms and flavors are described through instances of a given ontology. *Data interlinking* is closely related to the area of *ontology alignment*, and involves the definition of correspondences between entities located in different data sets, and the description of these correspondences through specific predicates (equivalence, related to, or domain-specific ones). The two activities not only share commonalities in terms of the types of basic (machine-driven) algorithms they make use of, but can also influence each other. Based on mappings at the schema level, one can identify potentially related instances; conversely, the availability of links between sets of entities may indicate similarities between classes.

In previous work of ours (Siorpaes and Simperl 2010) we surveyed methodologies, methods and tools covering each activity in order to learn about the types of processes knowledge and ontology engineering projects conform to, and the extent and reasons they might rely on human intervention. In the remainder of this section, we summarize the results of this analysis for a selection of activities: conceptual modeling as part of ontology development, alignment and interlinking as a prominent support activity in the engineering life cycle introduced earlier, and finally documentation, as a classical example of human-driven activity.

Developing Ontologies

Developing ontologies requires domain expertise and the ability to capture domain knowledge in a clean but purposeful conceptual model. An ontology describes the things that are important in a specific domain of interest, their properties, and the way they are interrelated. It defines a common vocabulary and the meaning of the terms used in the vocabulary. In the last 15 years, a wide array of ontology development methodologies have been proposed (Gómez-Pérez et al. 2004). Many suggest to start with the specification of the scope the ontology should cover and the requirements it should fulfil. This is often complemented by the informal and formal specification of competency questions. Based on that, relevant terms in the domain are then collected. Widely accepted ontology representation formalisms use classes, properties, instances and axioms as ontological primitives to describe domain knowledge. The overall process can be performed in a centralized (within a predefined team of knowledge engineers and domain experts) or a decentralized fashion (within a potentially open community of stakeholders, domain experts, and users).

The *conceptual modeling* process includes the definition of classes and the associated class hierarchy, as well as the definition of properties and additional axioms. Several automatic approaches have been proposed to discover specific types of relationships, in particular specialization and generalization extracted from natural language text, but human intervention is required for training the underlying algorithms, building the text corpus on which they operate, and validating their results (Bouquet et al. 2006; Buitelaar and Cimiano 2008). In addition, efforts need to be typically invested in post-processing the domain and ranges of individual properties, so that these are defined at the most appropriate level in the abstraction hierarchy. Defining axioms, on the other side, involves specifying precise, logics-based rules, such as cardinality constraints on certain properties and disjointness that apply to classes. Approaches for automatically specifying such axioms are very limited in their scope and require substantial training and validation (Völker et al. 2007).

As explained previously, the creation of instances is related to *semantic annotation*; we investigate it in more detail below. Relevant for the context of ontology development is the definition of so-called ‘fixed’ or ‘ontological’ instances which are the result of explicit modeling choices during the conceptualization phase. The distinction between classes and instances is very specific to the application setting, and we are not aware of any approaches aiming at automatizing this task.

There is a wide range of approaches that carry out *semi-automatic annotation of texts*: most of them make use of natural language processing and information extraction techniques. Even though they require training, a large share of the work can be automated (Reeve and Han 2005; Uren et al. 2006). The situation is slightly different with the *annotation of multimedia* content: approaches for the annotation of media, no matter if manual, semi-automatic or automatic, aim at closing the so-called "semantic gap", which is a term coined to describe the discrepancy between low-level technical features of multimedia, which can be automatically extracted to a great extent, and the high-level, meaning-bearing features a user is typically interested in and refers to when searching for content. Recent research in the area of semantic multimedia retrieval attempts to automatically derive meaning from low-level features, or other available basic metadata. This can so far be achieved to a very limited extent, i.e., by applying machine learning techniques with a vertical focus for a specific domain (such as face recognition), in turn for a substantial training and tuning, all undertaken with human intervention (Bloehdorn et al. 2005). The *annotation of Web services* is currently a manual task, but more research is needed in order to clearly determine whether this can be traced back to the nature of the task, or to the fact that the corresponding area is not mature enough to produce approaches that can offer reliable automatic results (Dimitrov et al. 2007; Kerrigan et al. 2008, 2007).

In Siorpaes and Simperl (2010) we analyzed various tools for text and media annotation which create semantic metadata with respect to the degree of automation they can support (nine tools in the first category, and six in the second one). In the case of textual resources, the main challenge is finding optimal ways to integrate human inputs (both workflow-wise and implementation-wise) with existing pre-computed results. On the contrary, multimedia annotation remains largely unsolved; there the typical scenario would use human computation as a main source of input for the creation of annotations, though specific optimizations of the process are nevertheless required. In Simperl et al. (2013) we embark on a broader discussion about how users could be motivated to engage with different types of participatory applications, including human computation ones, and on the principles and methods that could be applied to study and change user behavior to encourage engagement.

Supporting Ontology Development

Support activities accompany the development of ontologies. One prominent example thereof is the *alignment* of heterogeneous ontologies. Many of the existing ontology engineering environments provide means for the manual definition of mappings between ontologies. In addition, there is a wide range of algorithms that provide automatic support (Euzenat et al. 2007; Euzenat and Shvaiko 2007; Noy and Musen 2001, 2003), whilst it is generally accepted that the question of which ontological primitives match cannot (yet) be done fully automatically (Euzenat and Shvaiko 2007; Falconer and Storey 2007). This area is closely related to *data interlinking*, which we analyzed in more detail in (Simperl et al. 2012; Wölger et al. 2011).

Another support task is *documentation*, which contains two main components: the documentation of the overall process, and of the resulting knowledge base, in particular in terms of labels and commentaries associated to concepts, attributes, properties, axioms, and instances of the knowledge base. Either way, it remains human-driven, especially when it comes to recording modeling decisions and their rationales. Basic support for ontology documentation can be obtained by automatically creating entries for each ontological primitive which capture its core context in terms of labels and other annotations, as well as related classes, instances and properties. In this context, it is also worth mentioning the topic of *ontology localization*, which mainly refers to the translation of labels into different natural languages. Similarly to other areas in ontology engineering which employ natural language processing techniques for instance, ontology learning human input is required in order to solve translation questions which are highly context-specific, or to choose between different alternative translations.

We now turn to an analysis of how human computation could be applied to these activities in order to overcome the limitations of automatic techniques. For each activity, we will introduce examples of systems and applications such as games-with-a-purpose, microtask crowdsourcing projects, and community-driven collaborative initiatives that demonstrate the general feasibility of a hybrid human-machine approach.

Games-with-a-Purpose for Knowledge Engineering

Games-with-a-purpose is one of the most popular instances of social computing approaches to knowledge acquisition proposed in the last years. The game designer capitalizes on the appeal of key game properties such as fun, intellectual challenge, competition, and social status to turn the significant number of hours willingly spent playing by users through sophisticated algorithms into meaningful results that lead to qualitative improvements of information management technology. The concept is particularly useful to problems in knowledge engineering, an area which historically has targeted highly specialized audiences rather than casual Internet users. Tasks such identifying classes as groups of similar individuals, relating objects through properties, defining sub- and super-classes, validating whether two entities are the same or different, or labeling things in a given natural language are, though not always trivial to answer, much easier to tackle by humans than by machines.[3]

In the remainder of this section we will give a number of examples for this type of games illustrating the general concept.

[3] Exceptions include highly contextualized systems, which require extensive training and/or background knowledge. In these cases, the manual efforts shifts from the creation and maintenance of the knowledge base to the generation of training data sets and background corpora.

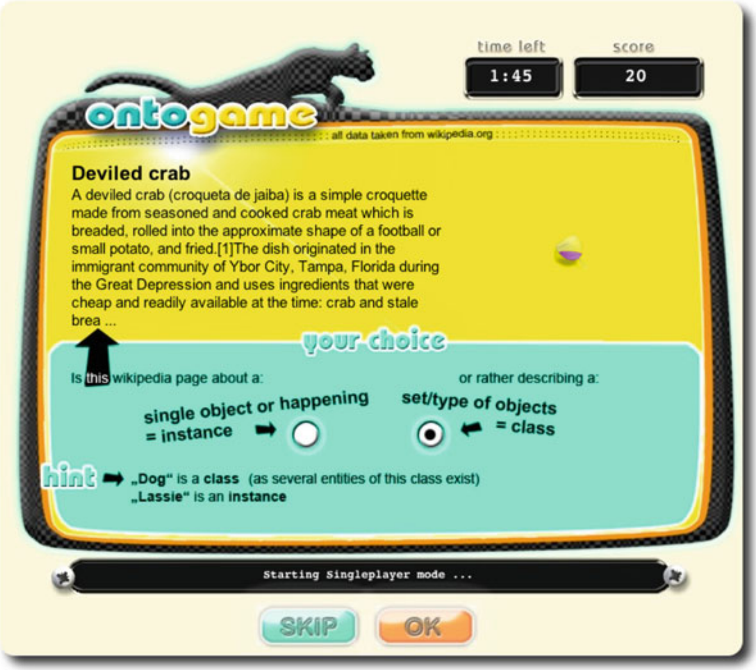


Fig. 2 OntoPronto: Expanding an existing ontology with Wikipedia concepts

Conceptual Modeling

OntoPronto

OntoPronto (Siorpaes and Hepp 2008) (see Fig. 2) is a real-time quiz game for the development and population of ontologies. The knowledge corpus used to generate challenges to be addressed by players is based on the English Wikipedia. Random Wikipedia articles are classified to the most specific class of an upper-level ontological structured called Proton (SEKT). The game can be played in a single- and two-players modus, where the former uses pre-recorded answers to simulate interaction. In the most general case, two players are randomly playing and can gain points by consensually answering two types of questions referring to the same Wikipedia article. In the first step, they are shown the first paragraph of an article and (if applicable) a picture, and are asked to agree whether the topic of the article stands for a class of similar objects or a concrete object. Once this issue has been settled, they enter the second step of the game, in which they navigate through the hierarchy of classes of the Proton ontology in order to identify the most specific level which will be extended through the topic represented by the Wikipedia article. The game back-end uses a number of standard means to validate the players' results.

Fig. 3 Virtual Pet: Creating a Chinese semantic network

Questions are subject to several game rounds and repeated, consensual answers are considered correct if they were authored by reliable players.

Virtual Pet and Rapport

Virtual Pet Game[4] (see Fig. 3) aims at constructing a semantic network that encodes common knowledge (a Chinese ConceptNet[5] equivalent). The game is built on top of PPT, a popular Chinese Bulletin Board System, which is accessible through a terminal interface. Each player has a pet which he should take care of in order to satisfy its needs, otherwise it could die. In order to take care of the pet (e.g., buy food), the player has to earn points by answering quiz-like questions that are relevant to the semantic network creation task at hand. The pet, in this game, is just a substitute for other players which receive the questions/answers and respond or validate them. Question and answers are provided by players using given templates (e.g., subject, predetermined relation, object). The validation of players' inputs is based on majority decision.

The purpose of the Rapport Game (Yen-ling Kuo et al. 2009) is very similar. Rapport Game, however, is built on top of Facebook (see Fig. 4) and uses direct interaction between players, rather than relying on the pet-mediated model

[4] http://agents.csie.ntu.edu.tw/commonsense/cate2_1_en.html

[5] http://conceptnet5.media.mit.edu/

Fig. 4 Rapport Game: Building a semantic network via Facebook

implemented by Virtual Pet. The players ask their opponents questions. These, in turn, answer them and the answers are evaluated by the user community. Points are granted for each type of action, from raising questions to answering and rating.

Guess What?!

Guess What?![6] (see Fig. 5) is a semantic game-with-a-purpose that creates formal domain ontologies from Linked Open Data.[7] Each game session is based on a seed concept that is chosen manually. In the back-end the application tries then to find a matching URI in a set of pre-defined ontologies of the Linking Open Data Cloud and gather additional information about the resources identified by the URI from interconnected Linked Data repositories. Additional information relies mainly on the adjacent graph in the Linking Open Data Cloud, including related classes and

[6] http://nitemaster.de/guesswhat/manual.html

[7] http://linkeddata.org/

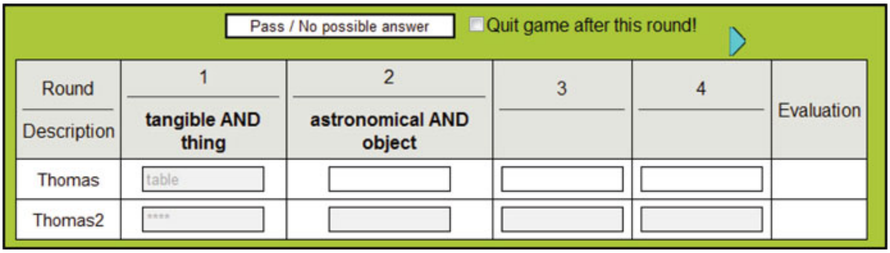


Fig. 5 Guess What?!: identifying complex concepts

entities, but also documentation such as labels. The resulting labels and URIs are analyzed using natural language processing techniques in order to identify expressions which can be translated into logical connectors such as 'AND'. Complex descriptions are broken down into smaller fragments, which are then weighed by a generality and confidence value. These fragments are used to generate the challenges solved in each game round. More specifically, a round starts with the most general fragment, and in the subsequent rounds a more specific one is connected to it through a logical operator. The goal of each round is for the player to guess the concept described by the interconnected fragments. For instance, in an initial round the fragment shown to the user contains the fragments 'fruit' AND 'yellow' AND 'oval', with solutions such as 'lemon' OR 'citrus'. Quality assurance is achieved through consensus and majority voting.

Alignment and Interlinking

WordHunger

WordHunger[8] (see Fig. 6) is a turn-based Web application that integrates among two large knowledge bases: WordNet and Freebase.[9] WordNet is a large lexical data base in which elements are grouped into synonym sets. Freebase is a structured knowledge base. Each game round consists of a WordNet term and up to three suggested possible Freebase concepts. The player then has to select the most fitting of those, or in case of insecurity, pass. Players may also select "no match" in case the articles are not related. After one of these possible choices the player proceeds to the next WordNet term. A player gets a point for each answer given. The data is validated through repeated answers.

[8] http://wordhunger.freebaseapps.com/

[9] WordNet: http://wordnet.princeton.edu/, Freebase: http://www.freebase.com/

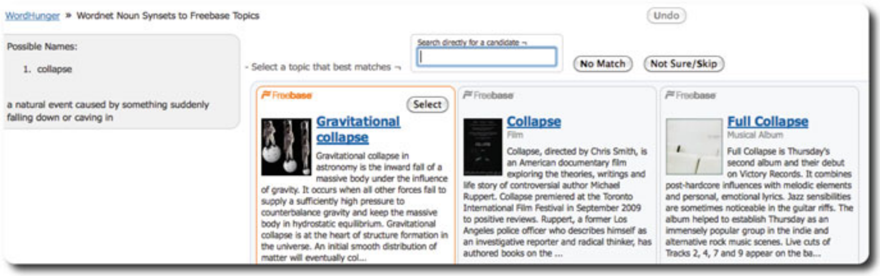


Fig. 6 WordHunger: Mapping WordNet to Freebase

Fig. 7 SpotTheLink: ontology alignment illustrated on DBpedia and Proton

SpotTheLink

SpotTheLink (Siorpaes and Hepp 2008) (see Fig. 7) is a real-time quiz-like game for ontology alignment. It aligns random concepts of the DBpedia ontology[10] to the Proton upper-level ontology that was already used in OntoPronto. Each game round

[10] http://wiki.dbpedia.org/Ontology

Fig. 8 UrbanMatch: connecting points of interest with image data sets

is centered around a randomly chosen DBpedia concept. In the first step of a game round, both players have to select a fitting concept from the Proton ontology. In case they choose the same concept they proceed with agreeing on a relationship between these concepts, either *is the same* or *is more specific*. They earn points for each consensual answer. After successfully matching a DBpedia class with a Proton class the players have to match the same DBpedia to the hierarchical next level of the Proton class. Otherwise, they play a new random DBpedia class. The validation of the results is based on consensus and majority voting.

UrbanMatch

UrbanMatch[11] is an application used to interlink Smart Cities data sources by exploiting games-with-a-purpose and gamification techniques (Fig. 8). It is built as a mobile, location-aware application in which players are expected to match points of interest related to a urban area to representative photos retrieved from the Web. To generate the challenges to be played, in each game round the application uses a mixture of trusted and less trusted online sources, including OpenStreeMap,[12] a geo-information repository, Flickr and Wikimedia Commons, the collection of images used by

[11] http://swa.cefriel.it/urbangames/urbanmatch/index.html

[12] http://www.openstreetmap.org

the Wikipedia encyclopedia. Candidate links are validated based on a metric taking into account the source of the image and of the answers. There are six difficulty levels and two game modes: in a 'race against time' players maximize the number of links founds between points of interest and pictures, thus optimizing recall; accuracy is addressed by a 'wise choice' option in which players have to identify the best possible links and submit their best-four selection without any time constraints.

Semantic Annotation

There is a wide array of games applied to tasks related to object identification and annotation of multimedia content. A selection of some of these games published in the human computation literature of the last five years can be found on the SemanticGames site.[13] They apply a large variety of games models (input agreement, output agreement, see GWAP),[14] and further distinguish themselves in the choice of game narrative, quality assurance (majority voting and beyond), and selection of challenges in each game round.

Microtask Crowdsourcing for Knowledge Engineering

In this section we introduce a number of approaches that have used microtask crowdsourcing to execute knowledge engineering tasks in a highly parallel fashion by using services of established crowdsourcing labor markets such as Amazon Mechanical Turk (AMT) and CrowdFlower.[15] For this purpose the actual task was first decomposed into small work units (denomintaed *microtasks*) and published on these platforms. The input collected from the crowds was incorporated into knowledge-based systems to be further consumed by the systems themselves, other automatic approaches, or even processed by human workers in more complex tasks.

Conceptual Modeling

CrowdSPARQL: Ontological Classification

CrowdSPARQL (Maribel Acosta et al. 2012) is a hybrid query engine for graph-based data which combines automatic query processing capabilities with microtask

---

[13] http://www.semanticgames.org

[14] http://www.gwap.com/

[15] Amazon Mechanical Turk: http://mturk.com, CrowdFlower: http://crowdflower.com

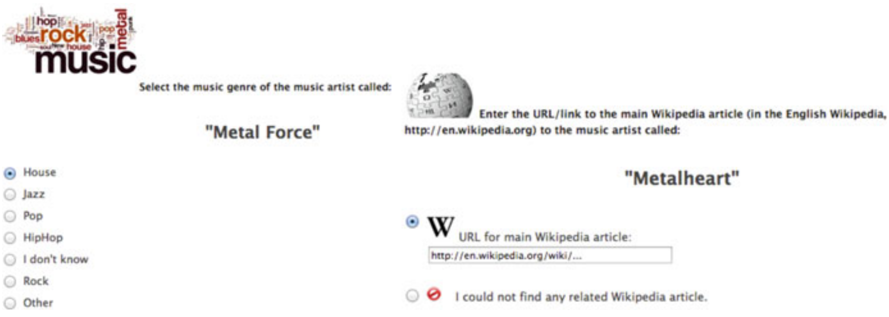


Fig. 9 CrowdSPARQL HIT interface: Ontological classification (*left*), entity resolution (*right*)

crowdsourcing. The aim is to produce enhanced results by evaluating the SPARQL queries against data stores and crowdsourcing parts of the query to discover relationships between Linked Data resources via a microtask platform such as AMT. The human tasks created by the engine are declaratively described in terms of input and output, which allows translating the results from the crowd directly into data that can be further processed by a conventional graph query engine.

From a knowledge engineering point of view, this hybrid engine will select specific patterns in the 'WHERE' clause of a SPARQL query that refer to tasks such as ontological classification and interlinking. Where such information is not available in the original data repositories, these patterns will be translated into microtasks (see Fig. 9). CrowdSPARQL implements several mechanisms for spam detection and quality assessment, including the creation of control questions within the microtasks where the correct answer is a priori known. The new relationships provided by the crowd are evaluated using majority voting (and some variations of this rule) and the consolidated answers are integrated into the Linked Data sets.

InPhO System: Conceptual Herarchies

The InPhO system (Niepert et al. 2007) attempts to dynamically generate a taxonomy of philosophical concepts defined in the Indiana Philosophy ontology.[16] The system relies on a user community composed of domain experts to construct and develop a philosophical hierarchy via asynchronous feedback, where the users (dis)confirm the existence of semantic relationships between the ontology concepts. The system follows a human-computation-based approach, where the feedback is collected and incorporated automatically into the taxonomy, evolving it and allowing new users' contributions.

Eckert et al. (2010) applied microtask crowdsourcing to populate the InPhO taxonomy via AMT, and compared the quality of the AMT workers' input with the

[16] https://inpho.cogs.indiana.edu/

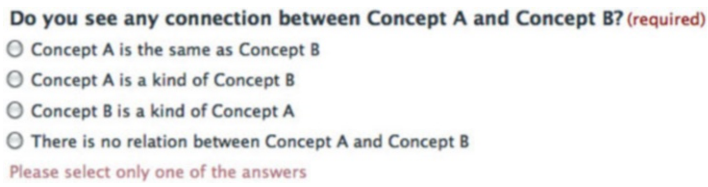


Fig. 10 CrowdMAP human task interface

feedback provided by the experts. The experiment involved the crowdsourcing of 1,154 pairs of philosophical concepts; each HIT submitted to AMT consisted of 12 questions where the users must first determine the relatedness ('unrelated' vs. 'highly related') of concept pairs and then select a predefined semantic relationship between these concepts. Each HIT was answered by five distinct workers. In addition, the authors implemented filtering mechanisms to detect low quality answers. By applying the right combination of these filters, the results suggest that it is possible to achieve high quality answers via crowdsourcing, as the feedback from the crowd and the experts is comparable.

Alignment and Interlinking

CrowdMAP: Ontology Alignment

CrowdMAP (Sarasua et al. 2012) introduces a human-loop in the ontology alignment process by crowdsourcing the possible mappings between ontologies as microtasks with individual alignment questions. The CrowdMAP architecture receives as input two ontologies to be aligned and an automatic algorithm to generate an initial mapping. Based on this information, CrowdMAP generates the human tasks (see Fig. 10) and submits them to CrowdFlower, where the workers suggest the type relationships between a pair of concepts ('same', 'subclass of', 'superclass of'). During the microtask generation, control questions where included in the tasks in order to facilitate the spam detection. In addition, the quality assurance and answer consolidation mechanisms supported by CrowdMap are those offered by the platform CrowdFlower. The experimental study in Sarasua et al. (2012) showed that CrowdMap on average is able to outperform automatic solutions, and the results suggest that the combination of ontology alignment algorithms with human-driven approaches may produce optimal results.

CrowdSPARQL: Entity Resolution

In Linked Data it is often the case that different data sets create their own resource identifier to refer to the same concepts. CrowdSPARQL (Maribel Acosta et al. 2012) is designed to handle entity resolution tasks via crowdsourcing when links between data sets are required while processing a SPARQL query. The current status of the engine allows the interlinking of Linked Data resources to DBpedia, which contains the RDF representations of knowledge extracted from Wikipedia. The workers perform the discovery of 'same as' correspondences providing the Wikipedia entry (URL) for a given Linked Data resource (see Fig. 9).

ZenCrowd: Entity Linking

ZenCrowd (Demartini et al. 2012) is a hybrid system that combines algorithmic and manual techniques in order to improve the quality of entity extraction on a corpus of news articles and linking them to Linked Data resources, by executing state-of-the-art solutions to find candidate matches and selecting the right one via microtask crowdsourcing. In each microtask, the workers have to select the correct Linked Data resource for a given entity. The results from the crowd are analyzed by ZenCrowd using a quality model to select the right answer based on probabilistic graphs, where entities, workers and candidate matches are represented as nodes, which are connected through factors. The experimental results showed that ZenCrowd is able to outperform automatic approaches by crowdsourcing entity linking, reflected as an improvement of the overall system accuracy.

Documentation

Mechanical Protégé: Ontology Documentation

Mechanical Protégé[17] is a plug-in for the open source Protégé[18] ontology editor tool and knowledge-base framework, which allows crowdsourcing ontology development activities such as creating classification hierarchies or labeling concepts and translating them into different languages. The ontology editor selects a task and the concepts within the ontology subject to crowdsourcing as illustrated in Fig. 11, and Mechanical Protégé creates and submits the human tasks to AMT. The types of tasks handled by Mechanical Protégé are considered complex tasks due to the variety of answers that may be retrieved from the crowd, therefore the ontology editor must perform the analysis and validation of the human input manually.

---

[17] http://people.aifb.kit.edu/mac/mechanicalProtege

[18] http://protege.stanford.edu/

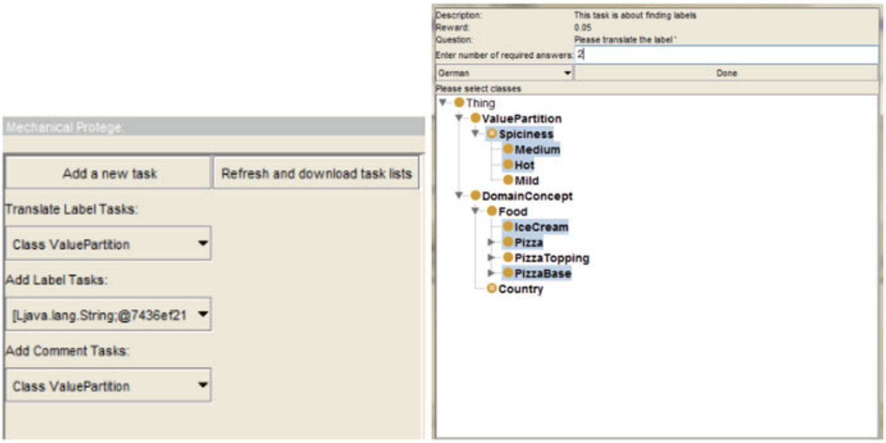


Fig. 11 Mechanical Protégé creation of microtasks: Selecting the type of task (*left*), selecting entities to crowdsource (*right*)

Other Approaches

Other human computation-based approaches rely on collaborative contributions from users to achieve their goals. One recent example comes from the Linked Data community for evaluating the well-known data set DBpedia. The DBpedia Evaluation Campaign,[19] aimed at detecting possible quality issues in the DBpedia data set; it was performed in two phases: first, a taxonomy of common quality issues was built by experts; then, Linked Data enthusiasts were invited to use the TripleCheckMate tool (see Fig. 12) in order to arbitrarily explore the data set resources and identify possible quality problems contemplated in the taxonomy. The second phase was performed as an open contest; the user submissions were analyzed and verified by experts, who selected a winner based on his contributions. Although the campaign has finished already, the information collected from the participants represents a valuable input to correct future versions of the data set and implement better (semi-)automatic data extractors on top of the Wikipedia mappings.[20]

While DBpedia is the attempt to extract structured, semantic data from the only partly ordered, enormous knowledge base that is Wikipedia, the project Wikidata[21] takes a different, more fundamental approach by letting the community directly build structured data relations to be then used by automated systems. This happens,

[19] http://nl.dbpedia.org:8080/TripleCheckMate/

[20] Wikipedia extractors. http://wiki.dbpedia.org/DeveloperDocumentation/Extractor

[21] http://www.wikidata.org/

Fig. 12 TripleCheckMate tool for exploring resources and selecting quality problems

e.g., through inline queries from Wikipedia, pulling for example up to date inhabitant numbers from Wikidata into info boxes of articles about cities. Community members, mostly Wikipedia editors, establish and maintain the data entries in a collaborative, open fashion. Data entries are stored as triples and can be accessed using Linked Data technologies. Wikidata is operated by the Wikimedia Foundation as the 'Data Layer' for its projects, much like Wikimedia Commons acts as its overall storage for media files. The project bears many similarities with initiatives such as Freebase, which applied a combination of volunteer and paid crowdsourcing to collaboratively create and maintain a structured knowledge base.[22]

Conclusions

In this chapter we gave an overview of how human computation methods such as paid microtasks and games-with-a-purpose could be used to advance the state of the art in knowledge engineering research, and develop and curate valuable (structured)

[22] http://www.freebase.com/

knowledge bases in different domains. Given the inherently human-driven nature of many knowledge engineering tasks, most notably knowledge acquisition and modeling, human computation has received great attention in this community as an alternative to costly, tedious expert-driven approaches followed in the past, with promising results. This resulted in an impressive number of systems, in particular casual games, tackling tasks as diverse as ontological classification, labeling, property elicitation, entity linking, ontology alignment or the annotation of different types of media. Besides these promising prospects, many of these projects still need to prove themselves in terms of sustainability and actual added value in the data they produce. More research is needed in order to enable the reuse of human-computation data, and even allow for different methods to be applied in combination. This would not only increase the quality of the crowd-engineering knowledge, which will be curated in time through various tools and platforms, but it would possibly facilitate the application of human-computation methods to different types of tasks and workflows, that are less amenable to parallelization.

References

Acosta M, Simperl E, Flöck F, Norton B (2012) A sparql engine for crowdsourcing query processing using microtasks – technical report. Institute AIFB, KIT, Karlsruhe

Berners-Lee T, Hendler J, Lassila O (2001) The semantic web. Sci Am 284(5):34–43

Bloehdorn S, Petridis K, Saathoff C, Simou N, Tzouvaras V, Avrithis Y, Handschuh S, Kompatsiaris Y, Staab S, Strintzis MG (2005) Semantic annotation of images and videos for multimedia analysis. LNCS. Springer The Semantic Web: Research and Applications, volume 3532 of Lecture Notes in Computer Science, pp 592–607. Springer Berlin Heidel-berg

Bouquet P, Scefor S, Serafini L, Zanobini S (2006) Booststrapping semantics on the web: meaning elicitation from schemas. In: 15th international conference on ACM, New York, NY, USA, http://doi.acm.org/10.1145/1135777.1135851, pp 505–512

Buitelaar P, Cimiano P (2008) Ontology learning and population: bridging the gap between text and knowledge. IOS-Press, Amsterdam

Celino, Irene, Simone Contessa, Marta Corubolo, Daniele Dell'Aglio, Emanuele Della Valle, Stefano Fumeo, Thorsten Krüger, and Thorsten Krüger. "UrbanMatch-linking and improving Smart Cities Data." In LDOW. 2012

Demartini G, Difallah DE, Cudré-Mauroux P (2012) Zencrowd: leveraging probabilistic reasoning and crowdsourcing techniques for large-scale entity linking. In: Proceedings of the 21st international conference on world wide web, WWW'12, Lyon. ACM, New York, pp 469–478

Dimitrov M, Simov A, Momtchev V, Konstantinov M (2007) Wsmo studio – a semantic web services modelling environment for wsmo (system description). In: European semantic web conference (ESWC 2007), Innsbruck

Eckert K, Niepert M, Niemann C, Buckner C, Allen C, Stuckenschmidt H (2010) Crowdsourcing the assembly of concept hierarchies. In: Proceedings of the 10th annual joint conference on digital libraries, JCDL '10, Gold Coast. ACM, New York, pp 139–148

Euzenat J, Shvaiko P (2007) Ontology matching. Springer, Berlin/New York

Euzenat J, Mocan A, Scharffe F (2007) Ontology alignments. Volume 6 of semantic web and beyond. Springer, p 350

Falconer SM, Storey M-A (2007) A cognitive support framework for ontology mapping. In: Asian semantic web conference (ASWC 2007), Busan

Fensel D (2001) Ontologies: a silver bullet for knowledge management and electronic commerce. Springer, Berlin/New York

Gómez-Pérez A, Fernández-Lopéz M, Corcho O (2004) Ontological engineering – with examples form the areas of knowledge management, e-Commerce and the semantic web. Advanced information and knowledge processing. Springer

Guarino N (1998) Formal ontology and information systems. In: Proceedings of the 1st international conference on formal ontologies in information systems FOIS1998, Amsterdam/ Washington. IOS-Press, pp 3–15

Kerrigan M, Mocan A, Tanler M, Fensel D (2007) The web service modeling toolkit – an integrated development environment for semantic web services (system description). In: European semantic web conference (ESWC 2007), Innsbruck

Kerrigan M, Mocan A, Simperl E, Fensel D (2008) Modeling semantic web services with the web service modeling toolkit. Technical report, Semantic Technology Institute (STI)

Kuo Y-L, Lee J-C, Chiang K-Y, Wang R, Shen E, Chan C-W, Hsu J-Y (2009) Community-based game design: experiments on social games for commonsense data collection. In: International conference on knowledge discovery and data mining, HCOMP'09, Paris. ACM, New York, pp 15–22

Neches R, Fikes RE, Finin T, Gruber TR, Senator T, Swartout WR (1991) Enabling technology for knowledge sharing. AI Mag 12(3):35–56

Niepert M, Buckner C, Allen C (2007) A dynamic ontology for a dynamic reference work. In: Proceedings of the 7th ACM/IEEE-CS joint conference on digital libraries, JCDL '07, Vancouver. ACM, New York, pp 288–297

Noy N, Musen M (2001) Anchor-prompt: using non-logical context for semantic matching. In: IJCAI workshop on ontologies and information sharing, Seattle, pp 63–70

Noy NF, Musen M (2003) The prompt suite: interactive tools for ontology merging and mapping. Int J Hum Comput Stud 59(6):983–1024

Reeve L, Han H (2005) Survey of semantic annotation platforms. ACM Press, New York, pp 1634–1638

Sarasua C, Simperl E, Noy NF (2012) Crowdmap: crowdsourcing ontology alignment with microtasks. In: International Semantic Web Conference (1)'12, Boston. pp 525–541

Schreiber G, Akkermans H, Anjewierden A, de Hoog R, Shadbolt N, Van de Velde W, Wielinga B (1999) Knowledge engineering and management: the CommonKADS methodology. MIT, Cambridge http://proton.semanticweb.org

SEKT Consortium. Proton ontology

Simperl E, Wölger S, Norton B, Thaler S, Bürger T (2012) Combining human and computational intelligence: the case of data interlinking tools. Int J Metadata Semant Ontol 7.2 (2012):77–92

Simperl E, Cuel R, Stein M (2013) Incentive-Centric semantic web application engineering. Synthesis lectures on the semantic web: theory and technology. Morgan & Claypool, San Rafael

Siorpaes K, Hepp M (2008) Ontogame: weaving the semantic web by online games The Semantic Web: Research and Applications, volume 5021 of Lecture Notes in Computer Science, pp 751–766. Springer Berlin Heidelberg, 2008

Siorpaes K, Simperl E (2010) Human intelligence in the process of semantic content creation. World Wide Web J 13(1):33–59

Studer R, Benjamins VR, Fensel D (1998) Knowledge engineering principles and methods. Data Knowl Eng 25(1/2):161–197

Uren V, Cimiano P, Iria J, Handschuh S, Vargas-Vera M, Motta E, Ciravegna F (2006) Semantic annotation for knowledge management: requirements and a survey of the state of the art. Web Semant Sci Serv Agents World Wide Web 4(1):14–28

Völker J, Vrandecic D, Sure Y, Hotho A (2007) Learning disjointness. In: Proceedings of the 4th european semantic web conference ESWC 2012, Innsbruck, pp 175–189

Wölger S, Siorpaes K, Bürger T, Simperl E, Thaler S, Hofer C (2011) Interlinking data – approaches and tools. Technical report, STI Innsbruck, University of Innsbruck

Yen-ling Kuo et al. (2009): In Proceedings of the ACM SIGKDD Workshop on Human Computation (HCOMP '09)

Human Computation in Citizen Science

Chris Lintott and Jason Reed

Early Development of Citizen Science Through Human Computation

Human computation is playing an increasingly important and interesting role in the scientific process through projects which variously label themselves as 'citizen science', 'scientific crowdsourcing' or simply 'public participation in scientific research' (PPSR). The willingness of large crowds of volunteers to give their time to projects that offer the promise of an authentic contribution to science, and the projects themselves are forming a critical piece of the response to the growing challenge of big data facing researchers in fields from astronomy to zoology.

In many ways, these distributed citizen science projects were a response to the success in the early years of distributed computing projects, particularly those such as SETI@Home (Anderson et al. 2002), which make use of the widely distributed BOINC platform and library (Anderson 2004). Projects such as ClimatePrediction.net (Massey et al. 2006), Einstein@home (Knipsel et al. 2010) and ROSETTA@home (Raman et al. 2008) have all produced significant scientific research. However, there are plenty of scientific tasks where human classification, transcription, intervention or computation is still superior to available machine learning solutions. In common with other fields in which human computation is deployed, many—most—of these problems are not impossible to attack with automated routines, but merely difficult. The lack of infinite resources for machine learning and computer vision mean that there are plenty of scientists and researchers classifying images, sorting through data and performing other repetitive tasks, and these are the problems that distributed citizen science projects can attempt to solve.

C. Lintott (✉)
University of Oxford, UK
e-mail: cjl@astro.ox.ac.uk

J. Reed
Adler Planetarium, Chicago, USA

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_14,

An early large attempt to develop a program such as this was NASA's clickworkers program, launched in 2000. Volunteers, known as clickworkers, were invited to look at images of the Martian surface from the Mars Global Surveyor spacecraft and record the positions of craters in order to determine the age of surface features. (Kanefsky et al. 2001) Clickworkers was primarily designed as a proof of concept, and it was successful, with click workers coming within a few pixels of the known catalogues (Barlow et al. 2000). The project team also noted user behaviours which were to become familiar to many running such projects; both a sudden rise in the number of classifiers and a division of labour between 'super-classifiers' and those who only visit once were noted in the project report (Kanefsky et al. 2001).

Although further iterations of the clickworkers project were developed, it did not become a general platform for citizen science, nor did it produce significant scientific results. Stardust@home, launched in 2006, fulfilled both of these goals and, more importantly, dispelled the notion that attractive images were necessary in order to engage a crowd (Mendez 2008). The task was to sort through images of dust grains returned by the Stardust probe from Comet 81P/Wild (Wild-2), with the goal of identifying interstellar dust grains trapped in the spacecraft's aerogel detector. A significantly challenging test was introduced to ensure classifiers were providing data of sufficient quality, and the incentive of having the right to informally name any verified interstellar dust grains found (!) was introduced. Tens of thousands of participants took part, and in 2010 two particles were announced as candidate interstellar grains. The team responsible for Stardust@home went on to build the Berkeley Open System for Skill Aggregation (BOSSA), the first scientific crowdsourcing platform, which more recently has, with the advent of pyBOSSA (http://crowdcrafting.org), been ported into Python.

The Advantages of Citizen Science

In this section, I focus on Galaxy Zoo, a project developed by one of the authors [CL] and an interdisciplinary team, directly inspired by the success of Stardust@home. The original aim of Galaxy Zoo was to provide morphological classifications of nearly one million galaxies imaged by the Sloan Digital Sky Survey (Lintott et al. 2008; York et al. 2000). Galaxy Zoo received strong support from a volunteer community from its launch in July 2007, reaching a peak rate of more than 70,000 classifications per hour (Lintott et al. 2008) and still receives classifications in a new version 6 years later. It has produced data used in many scientific papers (see zooniverse.org/publications for an updated list). This longevity and productivity make it ideal to illustrate the advantages of human computation in this scientific space.

The scientific results of Galaxy Zoo have come from two routes. The first is the designed route, resulting from the collection and combination of user responses delivered through the main interface, allowing the project to meet the key challenge of being able to provide the scale of effort required. In the case of Galaxy Zoo, users were asked to provide answers to questions presented in a decision tree next to an image of a galaxy. (The process of data reduction for citizen science is discussed below, in section "Citizen Science Motivations: Gaming the System").

However, two key design decisions enabled not only data analysis but serendipitous discovery. Firstly, classifiers could follow a link to the ‘professional’ view of their galaxy, which provided more than 100 items of metadata, spectra and the ability to view images at different wavelengths. In later iterations of the project, this link is only available after the image has been classified in order to guard against the possibility of biasing classifications. Secondly, a forum was added to the site, initially in response to the large numbers of emails received which quickly overwhelmed the team’s ability to respond.

Volunteers quickly adopted the forum as a place to discuss unusual finds, which ranged from galaxies shaped like letters of the alphabet (mygalaxies.co.uk) to more scientifically interesting finds. For initial discoveries, such as ‘Hanny’s Voorwerp’, a galaxy sized gas cloud heated to 50,000 K by a jet from a now quiescent supermassive black hole (Lintott et al. 2009), the forum and thus the communities’ role was limited to simple discovery (even if individual volunteers were included in discussions about follow-up observations and papers), but as the community matured more complex behaviours emerged.

In some cases, this was self directed. The Galaxy Zoo ‘peas’, a set of small, round and (in SDSS imagery) green galaxies which turn out to be the most efficient factories of stars in the local Universe were identified and investigated by approximately twenty volunteers who downloaded data, analysed spectra and even collaborated to produce their own crowdsourcing site to sort through candidate objects. (Cardamone et al. 2009) In this example, therefore, the community of volunteers was able to identify objects of interest and classify them according to a schema they devised, but also to carry out more advanced work, which contributed greatly to the final scientific paper. This mode of operation, in which small groups of volunteers are able to work on rare classes of objects can be enabled by interaction with the scientists; members of the Galaxy Zoo team who interacted with forum participants were able to arrange significant searches for multiple classes of such objects (Keel et al. 2013, 2012).

Serendipitous discovery is thus a key advantage of this form of human computation, in both directed and undirected modes, and projects built on the Zooniverse platform (Fortson et al. 2012) which grew from Galaxy Zoo now almost always explicitly allow for this in design. However, it became clear that the forum employed by Galaxy Zoo was not adequate; it used off the shelf software that could not easily be incorporated with the main classification task, and as the forum’s content grew both in size and complexity it became increasingly hard to navigate. (The Galaxy Zoo forum (www.galaxyzooforum.org) currently contains more than 600,000 posts in nearly 20,000 topics). The percentage of Galaxy Zoo users using the forum dropped steadily over time as a result, and scientists were increasingly unwilling to spend time looking for interesting conversations. This latter factor was a particular problem as a large part of the forum’s scientific output had come from collaboration between citizen and professional scientists.

In response to these challenges, a new object-oriented discussion system was developed and deployed on later Zooniverse projects. Known as ‘Talk’ (http://github.com/zooniverse/talk) it is linked from the main classification page, and allows users to quickly discuss interesting or curious subjects they have encountered while classifying. Forum-like boards enable longer discussions, and collections and hashtags increase the visibility of interesting objects. Talk has been successful, in particular on the Planet Hunters project (Schwamb et al. 2012) which looks for extrasolar planets in data provided by NASA's Kepler space telescope. More than 60 % of Planet Hunters users have used the system, and more than half of those have made comments (Fischer et al. 2012). Significant discoveries, including Planet Hunters 1b, the first planet in a four-star system (Schwamb et al. 2013) and more than forty planet candidates in the habitable zone of their parent stars (Ji et al. 2013), have come from efforts led by Talk users. Where active communities have grown up on projects using talk, they show an interesting propensity to become interested in topics which diverge from the interests of the science team; examples include the community of bird watchers on SnapshotSerengeti.org (an animal identification project which lumps almost all birds into a single category) and those planet hunters studying variable stars. Talk has not always succeeded, however, and early engagement with the community by the scientific team seems to be a critical factor. Nonetheless, it is clear that serendipitous discovery can not only be a crucial part of a successful citizen science project, but also can be designed for and encouraged.

A third key advantage is in the feedback between citizen science and machine learning approaches to the same problems. In many cases, the limiting factor in the performance of such algorithms is the lack of sufficiently large training sets; these can be provided by citizen science projects. (Smith et al. 2010; Bloom et al. 2012) In addition, the design of a typical project in which several classifiers independently work on the same subject allows systems to be trained not only with gold standard data but given an estimate of the uncertainty in the resulting classification; such data has been shown to improve the performance of neural networks significantly (Banerji et al. 2010). This ability to improve the performance of automated routines is key as datasets continue to grow in size, requiring an ever-higher proportion to be automatically processed even in cases where the need for human computation remains.

These are all advantages which are intrinsic to the process of research; reasons why the adoption of a citizen science approach might be of use for the researcher attempting to sort through a large dataset. Citizen science projects also often have extrinsic goals, whether or not there are explicitly stated, in serving as part of an educational or outreach effort whose goals are not limited to imparting the knowledge necessary to be an effective participant. Both formal (Raddick et al. 2013) and informal education programs have been developed which make use of such projects; research into the effectiveness of such efforts is in the early stages, but it is clear that the ability of hybrid systems which contain both citizen science activities and interventions designed for learning provides a rich laboratory for experiments in the field.

Citizen Science Motivations: Gaming the System

Discussion of learning through citizen science introduces the question of user motivation. A large survey of Galaxy Zoo volunteers suggested that the most common primary motivation lies in a desire to contribute to research, with an interest in the specific subject concerned second. (Raddick et al. 2013; von Ahn et al. 2008)

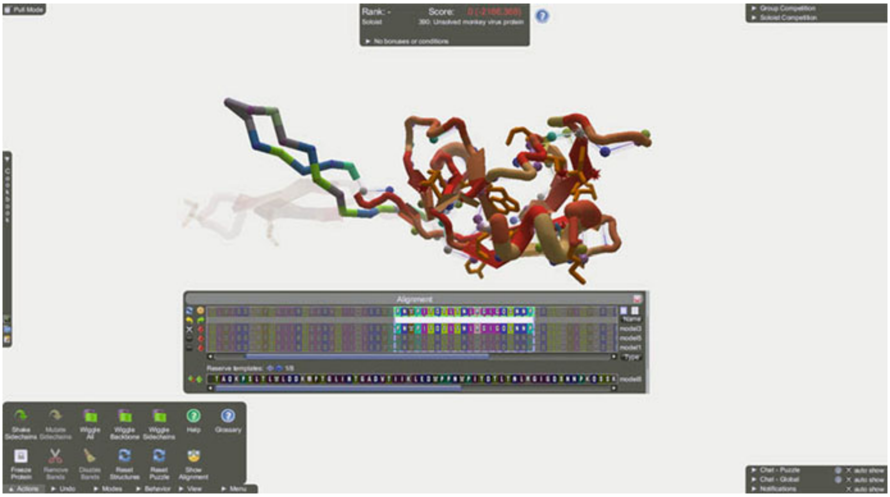


Fig. 1 A screenshot of the "Unsolved monkey virus protein" Foldit puzzle, showing the complexity of both task and interface. Fold.it players produced a structure which was accurate enough to enable the ultimate solution of the three dimensional structure for this molecule

For developers of projects, this is encouraging as it suggests that significant engagement need not be limited to subjects such as astronomy which have large public followings already.

However, the reach of citizen science is not limited to projects which appeal directly to these motivations; the success of the reCAPCHA project (von Ahn et al. 2008) in digitizing words entered by users as an anti-spam device suggests that the deployment of scientific crowdsourcing as a tool for distinguishing humans from machines might be worth trying. More relevantly, several projects have attempted to build systems which involve the user in playing a game which also happens to produce scientifically useful results.

The most successful of these attempts is Fold.it (Khatib et al. 2011a), a protein folding game. The new Fold.it user is presented with a series of puzzles (See Fig. 1) of increasing difficulty which teach the use of a variety of tools which can be used to manipulate the structure of a simulated protein. A score can be assigned such that the lowest energy state achieved scores most highly (in fact, the calculation is an approximation to the actual energy), incentivizing 'correct' behaviour. Once a user has progressed through the training levels, they can choose to work on a series of real problems where the correct answer is not known. The system design is both elegant, incorporating the necessary knowledge about how chemical bonds behave, for example, in tool behaviour, and successful, scoring highly in international protein folding competitions (Popovic 2008) and making discoveries worthy of follow-up (Khatib et al. 2011b). Fold.it allows collaboration on problems, with volunteers refining and adding to the solutions found by others, and an intriguing degree of specialization has grown up with some users, for example, priding themselves on specializing in the 'end game'—the final adjustments that can improve already good

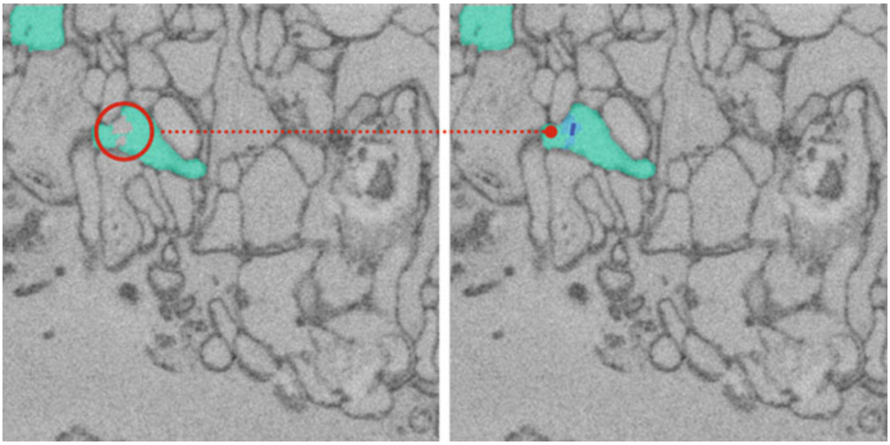


Fig. 2 This image shows an example of the data used in Eyewire, where the task is to assign regions of an image to neurons. On the left, the blue area represents an automatic attempt to solve the problem, but a small part is erroneously excluded. This mistake is corrected by human intervention (*right*)

solutions. Over time, the Fold.it team have added further complexity, with the introduction of macros (scripts with perform predetermined tasks automatically) allowing the community a means of recording and passing on knowledge, even as individual classifiers come and go.

Fold.it demonstrates the power of a gameified approach to citizen science projects, and has inspired a series of newer projects which take a similar approach. Critically, it has some differences from the classifier-based model described earlier. Most strikingly, rather than combining classifications from a large crowd of volunteers, it looks for individual solutions from a small number of classifiers (or a small number of small groups of classifiers). Whereas all participants in, for example, Galaxy Zoo are providing data which directly impacts the scientific results, most fold.it players are working their way through the training levels or failing to improve on existing solutions. This suggests that the motivation of such volunteers might be slightly different from those participating in classification citizen science projects.

The kind of game treatment deployed by fold.it will not be suitable for all scientific human classification problems. In particular, the protein folding problem benefited from an easily calculated proxy for the correct answer; in problems where the correct answer is unknown devising a reward scheme that encourages the right behaviour may prove impossible. The development costs associated with the deployment of a fully game-like system are likely to be high, the failure rate higher (building computer games which people want to play is hard, even without the constraint of producing useful results) and the results less generalizable.

These limitations suggest the development of a hybrid model in which game-like features are introduced to an otherwise standard classification system. This can be effective (a particularly comprehensive attempt has been made by EyeWire (see Fig. 2), a recent project which aims to map the 3d structure of neurons) but it risks conflicting motivations and thus having a detrimental effect on user behaviour. For example, the introduction of a points system in an early trial version of Galaxy Zoo led to the loss of both particularly poor and particularly good classifiers. The former makes sense, but we can hypothesize that the latter were no longer motivated by a desire to contribute but rather by a gaming motivation that resulted in stopping playing when the game was ‘won’ or no further improvement was possible. Even the addition of rewards in the form of achievement badges, a common simple game-like addition can be problematic—in SETI@home, an experimental Zooniverse project, the interruption of receiving a badge meant that people were likely to leave instead of being motivated to continue classifying. Gameification, then, is to be done carefully rather than forming a panacea to all problems.

Data Reduction and User Weighting

A discussion of the deployment of human computation through citizen science would not be complete without including discussion of the methods used to combine work from multiple classifiers. These have become increasingly sophisticated as projects seek increases in efficiency and accuracy. Moving on from simply taking a majority vote, for example, the Galaxy Zoo project weighted users who consistently agreed with the majority of users (Lintott et al. 2008), and measured and adjusted the bias inherently in classifying smaller, fainter or more distant galaxies (Land et al. 2008; Bamford et al. 2009; Willett et al. 2013). The Planet Hunters (see Fig. 3) project mentioned above used classifications of ‘gold standard’ data (in their case, classifications of both known and simulated planets were sought for verification purposes) (Schwamb et al. 2012). The Milky Way project, which analysed infrared images of the sky looking for ‘bubbles’ associated with the formation of stars used user behaviour as a proxy for ability, discounting classifications from those who did not use the full set of tools provided (Simpson et al. 2012).

However, much more sophisticated analyses are possible, and several researchers have used archived data from citizen science projects to demonstrate the effectiveness of a Bayesian approach to the problem. In this picture, each classification provides information not only about the image being classified but also about the volunteer doing the classification (see Fig. 4). This information can be used both to provide more accurate classifications, but also to direct attention through the efficient choice of the next classifier (Simpson et al. 2013; Kamar; Waterhouse 2013). Details of these systems are described elsewhere, but they typically achieve an increase in efficiency of between 30 % and 70 %. This has led to interest in deploying these systems, especially from those researchers who are working on systems, which interact not with a large archive of data but with a live stream of information, necessitating rapid response for follow-up or decisions about data storage.

There are substantial technical challenges to such implementation, primarily the time or computational resources needed to deal with a system of tens of thousands

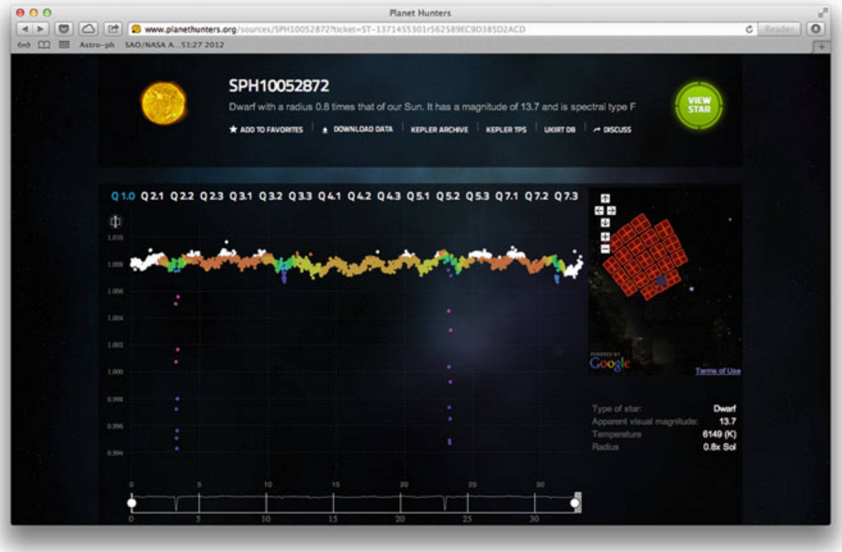


Fig. 3 Light curve of Planet Hunters 1b, the first planet in a four star system, which was discovered by citizen scientists. The x-axis is time in days, the y-axis normalized brightness. The large dips are due to transits of one star in front of another, whereas the small dip at 11 days is the transit of the planet. Planethunters demonstrates that beautiful images are not a prerequisite for a successful citizen science project

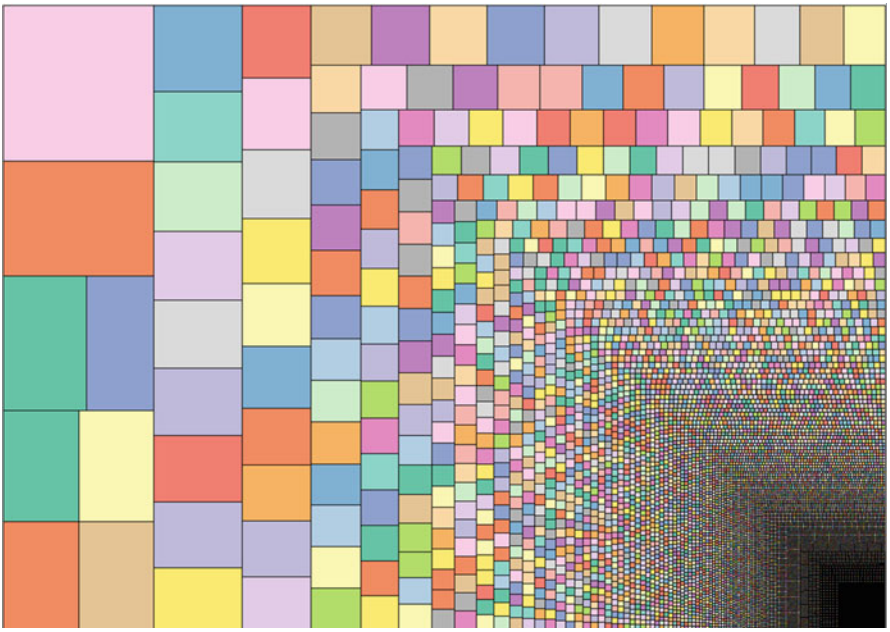


Fig. 4 Distribution of effort in the original Old Weather project, with each square corresponding to the contributions of a single transcriber. The general pattern—in which substantial effort is provided by both occasional users as well as those who are much more committed—is seen in many Zooniverse projects

of classifiers and hundreds of thousands of classifications in real time. Volunteer classifiers are unlikely to be tolerant of even small delays in deciding what they should see next. However, the real problem is that such systems make explicit the conflicts between motivations mentioned above. In systems which combine many imperfect automatic classifiers, and in much of the literature (including those papers cited in the previous paragraph) it is assumed that each classification provided has a fixed cost; whether the subject for classification is selected because we want to know more about that subject, because we want to know more about the performance of the classifier on subjects like it, or whether we are training users by asking them to classify it, the cost is usually assumed to be equal. The goal is then to minimize cost.

With a system containing volunteer classifications things are much more complex. Consider a simple system that shows the most difficult images (say, the faintest galaxies) to the best human classifiers (those that perform most strongly on a training set of faint galaxies). Now consider that in this system, all human classifiers are motivated by seeing varied images of galaxies—a perfectly plausible situation. Given that, unlike their automatically classifying robot counterparts, the volunteers are free to stop classifying at any time, this system would systematically drive away the best classifiers from the site due to the monotony of the visual experience. Without a proper understanding of motivation, therefore, attempts to improve the efficiency of citizen science projects are likely to fail.

This is a key example of the ability of citizen science to throw up interesting general problems in the field of human computation. There has been a rapid increase in the number of such projects in the last 5 years, as well as in their variety and scope. There is thus much need for collaboration between scientists making use of citizen science and experts in computational systems to move beyond simple 'clockwork' and toward dynamic and complex systems that balance learning with simply getting the job done.

References

Anderson DP (2004) BOINC: a system for public-resource computing and storage. In: 5th IEEE/ACM international workshop on Grid computing, Pittsburgh, 8 Nov 2004, pp 1–7

Anderson D, Cobb J, Korpela E, Lebofsky M, Werthimer D (2002) SETI@home: an experiment in public-resource computing. Commun ACM 45(11):56

Bamford S et al (2009) Galaxy Zoo: the dependence of morphology and colour on environment. MNRAS 393:1324

Banerji M et al (2010) Galaxy Zoo: reproducing galaxy morphologies via machine learning. Mon Not R Astron Soc 406:342

Barlow NG et al (2000) LPSC XXXI, Abstract #1475

Bloom J et al (2012) Automating discovery and classification of transients and variable stars in the synoptic survey era. Publ Astron Soc Pac124:1175

Cardamone C et al (2009) Galaxy Zoo Green Peas: discovery of a class of compact extremely star-forming galaxies, Mon Not R Astron Soc 399:1191

Fischer D et al (2012) Planet Hunters: the first two planet candidates identified by the public using the Kepler public archive data. Mon Not R Astron Soc 419:2900

Fortson L et al (2012) Galaxy Zoo: morphological classification and citizen science. In: Way MJ, Scargle JD, Ali K, Srivastava AN (eds) Advances in machine learning and data mining for astronomy. Chapman & Hall, Boca Raton, FL

Ji W et al (2013) Planet Hunters. V. A confirmed Jupiter-size planet in the habitable zone and 42 planet candidates from the Kepler archive data. Submitt Astrophys J. Accessible at arXiv/1301.1644

Kamar E, Hacker S, Horvitz E Combining human and machine intelligence in large-scale crowdsourcing. In: Proceedings of the 11th international conference on autonomous agents and multiagent systems, vol 1. p 467

Kanefsky B, Barlow NG, Gulick VC (2001) Can distributed volunteers accomplish massive data analysis tasks? Presented at 32nd annual Lunar & planetary science conference, Abstract #1272

Keel W et al (2012) The Galaxy Zoo survey for giant AGN-ionized clouds: past and present black hole accretion events. Mon Not R Astron Soc 420:878

Keel W et al (2013) Galaxy Zoo: a catalog of overlapping Galaxy Pairs for dust studies. Publ Astron Soc Pac 125:2

Khatib F et al (2011a) Algorithm discovery by protein folding game players. PNAS. doi 10.1073/pnas.1115898108

Khatib F et al (2011b) Crystal structure of a monomeric retroviral protease solved by protein folding game players. Nat Struct Mol Biol 18:1175

Knipsel B et al (2010) Pulsar discovery by global volunteer computing. Science 329:1305

Land K et al (2008) Galaxy Zoo: the large-scale spin statistics of spiral galaxies in the Sloan digital sky survey. MNRAS 388:1686

Lintott C et al (2008) Mon Not R Astron Soc 389:1179

Lintott C et al (2009) Mon Not R Astron Soc 339:129

Massey N, Aina T, Allen M, Christensen C, Frame D, Goodman D, Kettleborough J, Martin A, Pascoe S, Stainforth D (2006) Data access and analysis with distributed federated data servers in climateprediction.net. Adv Geosci 8:49–56

Mendez BJH (2008) In: Garmany C, Gibbs MG, Moody JW, (eds) ASP conference series, vol 389. EPO and a changing world: creating linkages and expanding partnerships. Astron Soc Pac, San Francisco, p 219

Popovic Z (2008) CASP8 results, Foldit Blog, http://fold.it/portal/node/729520. 17 Dec 2008

Raddick J et al (2013) Galaxy Zoo: motivations of citizen science volunteers. Astron Educ Rev (in press) http://arxiv.org/abs/1303.6886

Raman S, Baker D, Qian B, Walker RC (2008) Advances in Rosetta protein structure prediction on massively parallel systems. J Res Dev 52(1–2):7

Schwamb et al (2012) Planet Hunters: assessing the Kepler inventory of short-period planets. Astrophys J 754:129

Schwamb et al (2013) Planet Hunters: a transiting circumbinary planet in a quadruple star system. Astrophys J 768:127

See http://crowdcrafting.org

Simpson R et al (2012) The milky way project first data release: a bubblier galactic disc. Mon Not R Astron Soc 424:2442

Simpson E, Roberts S, Psorakis I, Smith A (2013) Dynamic Bayesian combination of multiple imperfect classifiers. Stud Comput Intell 474:1–35

Smith A et al (2010) Galaxy Zoo supernovae. Mon Not R Astron Soc 412:1309

von Ahn L, Maurer B, McMillen C, Abraham D, Blum M (2008) reCAPTCHA: human-based character recognition via web security measures. Science 321(5895):1465

Waterhouse T (2013) Pay by the bit: an information-theoretic metric for collective human judgment. In: Proceeding CSCW, ACM, New York, pp 623–638

Willett et al (2013) MNRAS, 435, 2835

York, D et al (2000) Astron J 120:1579

Human Computation as an Educational Opportunity

Carole R. Beal, Clayton T. Morrison, and Juan C. Villegas

Introduction: Citizen Science

Volunteers have long made important contributions to science. For example, Charles Darwin was a volunteer naturalist during his travels on the HMS Beagle (Silvertown 2009). However, in the last two decades the dramatic growth in technologies that support global-scale communication and data sharing have contributed to a movement in which large numbers of individual volunteers can participate in significant scientific research through collecting and processing data. Such "citizen science" projects represent an example of human computation in that powerful computational and communication tools now allow for the distribution of the work of science to hundreds, thousands or millions of volunteers (Clery 2011; Cohn 2008; Silvertown 2009). Citizen science projects offer the potential for individuals to cooperate to solve the major challenges facing humanity, from documenting the impact of climate change to early detection of earthquakes and discovering unique objects in the night sky. By providing novices with the tools to record, share and review data, scientists can crowd-source some of the data collection, making it possible to work at a much larger scale than is possible for a single research team.

One of the earliest examples of citizen participation in science involves bird counting, such as the annual Christmas Bird Count conducted by the Audubon Society for more than a century, and now expanded into year-round bird-monitoring activities directed by the Cornell Laboratory of Ornithology (2012). People who are

C.R. Beal (✉) • C.T. Morrison
School of Information: Science, Technology, and Arts, University of Arizona, Tucson, USA
e-mail: crbeal@email.arizona.edu

J.C. Villegas
Universidad de Antioquia, Medellín, Colombia

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_15,

interested in birds and knowledgeable about species characteristics and behavior voluntarily conduct systematic observations in their neighborhoods, and then contribute valuable information about the location and density of different species. The breadth of the observations supports large-scale mapping of ranges and changes in populations over time that would not otherwise have been possible. Early data contributions by bird observers were conducted primarily by mail but now involve online data entry via a Web portal. Related examples include projects involving searching for and marking sea turtle nest sites (Bradford and Israel 2004) and nesting bird sites (King and Lynch 1998). These are all excellent examples of citizen science capitalizing on natural interest ranging from curiosity to serious hobby, and existing expertise in the community can opportunistically provide useful observations in the field.

Astronomy is another area that has long utilized novices' observations of the night sky to move the field forward. Since 1911, amateur astronomers have made recordings of changes in the brightness of individual stars, contributing more than 20 million data points to a database used by professional astronomers (Williams 2001). In the Galaxy Zoo project, more than 200,000 people have logged into the Sloan Digital Sky Survey website where they view and classify images of galaxies (e.g., is the galaxy spiral in shape and if so, what is the direction of its rotation?). Their volunteer labor contributed to dozens of scholarly publications as well as independent research initiated by some of the volunteers themselves (Raddick et al. 2010). Here, citizen scientists perform a critical role in processing and analyzing data that has been collected but requires decision-making that is still difficult to automate yet takes relatively little training for humans to accurately manage.

More recent examples of citizen science projects involve a range of new technologies for data collection and contribution, such as the large-scale data collection about seismic activity contributed by users via accelerometers in their laptops (Cochran et al. 2009). The "Quake-Catcher Network" is improving the speed with which pending earthquakes are detected by contributing data from locations that are between established recording stations. Mobile devices support "BioBlitz" events in which volunteers use their phones to conduct rapid inventories of biological diversity in a specific area within 1 or 2 days (Lundmark 2003). Brightsmith et al. (2008) found that volunteers were even willing to pay for the opportunity to participate in a research project to photograph and document the behavior of Blue-headed Macaws in the rainforests of Peru, pointing to the potential of integrating citizen science with ecotourism. Such collaborations are being explored for global coral reef monitoring (Marshall et al. 2012) and preservation of the Galapagos Islands (Gibbs and Milstead 2012).

As these examples illustrate, citizen participation in large-scale science projects can benefit professional researchers through the accumulation of data at a scale that may not otherwise be possible (Silvertown 2009). Related work suggests that there are also benefits to the volunteer participants in terms of the opportunity to build scientific thinking skills (Trumbull et al. 2000). Citizen scientists may also benefit through a deeper understanding of the value of science and its impact on their daily lives (Fusco 2001).

Citizen Science and STEM Education

The power of citizen science has led an increasing number of scientists, agencies and professional research organizations to involve student volunteers in data collection, raising questions about whether volunteers are able to contribute high quality data consistent with professional standards (Cohn 2008). New methods for automatically calibrating the quality of volunteer-provided data and identifying suspicious entries are being developed (Yu et al. 2012). Some studies revealed the need for careful protocols and detailed training to ensure that the data collected by the volunteers were consistent with the work of professionals (Engel and Voshell 2002). Overall, though, it appears that volunteers are able to contribute valuable data. For example, in the Oregon White Oak Stand Survey, students were trained to identify and log the locations of the white oak, a native tree that provides valuable habit to endangered species in the region. More than 600 students were transported to the area by bus and assigned to a transect, where they reported the location and canopy shape of oaks versus other tree species. Comparisons with data provided by professionals indicated that the students' data were of generally high quality although the novices tended to over-report observations of rarer species (Galloway et al. 2006). Similar positive conclusions about the quality and value of volunteer-contributed data were noted in Delaney et al. (2008) with regard to novices' ability to recognize and document the presence of native and invasive crab species in the Eastern coastal region of the United States. In a project involving large-scale monitoring of mammal (e.g., badger) signs in the Oxfordshire forest, Newman et al. (2003) found that volunteers typically took longer to record observations but that the quality and utility of their data was equivalent to that of the professionals.

Case Study: The B2E2 Project

Taking into account these lessons learned in practical citizen science applications, the Biosphere 2 Evapotranspiration Experiment (B2E2) project explored the intersection of science, public participation, and education by combining a citizen science approach to expand research beyond the laboratory with a special emphasis on creating educational opportunities for middle and high school students.

Climate Change and the Water Cycle

The project was born out of the dissertation research of Dr. Juan Villegas at the University of Arizona and Biosphere 2 facility located outside of Tucson, AZ. Dr. Villegas's research focused on the study of evapotranspiration, the combined process of water loss from soil (evaporation) and from plants through their leaves (transpiration). Somewhat surprisingly, factors that affect water loss through these two processes are not fully understood; nor is it clear how environmental characteristics such as foliage density or woody versus non-woody plant types affect the ways in which water is lost into the atmosphere. These questions are increasingly urgent with the impact of climate change and long-standing drought on the Southwestern region of the United States, as well as in other drought-ravaged regions such as Australia (Adams et al. 2009).

The goal of Dr. Villegas's project was to learn more about how the changing vegetation in the desert might affect the water cycle, more specifically, the ratio of water loss into the atmosphere from evaporation from soil, and from transpiration from plants. Climate changes have led to the invasion of non-native species into the desert; non-natives tend to grow more closely together than native species and there is growing concern that greater density may alter the traditional seasonal weather patterns such as the summer monsoon rains around which the desert ecosystem revolves (Huxman et al. 2005).

The Biosphere 2 Research Laboratory

The Biosphere 2 facility provided an ideal laboratory to conduct Dr. Villegas's research. Biosphere 2 is a large (over three acres) glass-enclosed laboratory that includes five distinct biomes and its own internal ocean, and serves as a research facility for The University of Arizona. The unique design allows an unprecedented degree of control over climate factors such as humidity and temperature, and provides a venue for systematic, controlled environmental science research at a scale not possible in traditional laboratories or in the field. Taking advantage of this facility, the investigators set up experimental comparisons involving arrangements of mature trees planted in large boxes and interspersed with boxes that contained only soil. Measurements of evaporation from the soil and transpiration from the trees indicated that the relative ratio of foliage and soil did affect water loss patterns (Villegas et al. 2009).

The control afforded by the Biosphere 2 facility allowed Dr. Villegas to precisely monitor details of evapotranspiration processes that would be difficult to replicate on a large scale, such as sap flow monitoring, precise measurement of microclimate conditions that require specialized equipment, and the chosen study subjects of mature mesquite trees. At the same time, the landscape-scale: study using living trees spaced as they would occur in the desert pointed to the need for additional research to better understand the broad range of parameters affecting evapotranspiration. This led to considering alternate forms of the general evapotranspiration study design in which these parameters might be explored but using smaller scale materials and apparatus. Furthermore, it was recognized that there was an opportunity to turn these smaller scale studies into an educational opportunity in which the studies could be run in classrooms.

The "Tabletop" Experiment

A partnership was formed with middle schools in the area in which students conducted table-top versions of the experiment and contributed their data to the research team. The science curriculum included units on scientific inquiry, earth sciences, and the water cycle, making the citizen science activity well-aligned with the ongoing instruction.

The activity began with a pre-test of students' initial understanding of the water cycle and science methods that was conducted by their science teachers. The next day, a member of the research team visited each classroom and gave a 40 min presentation about the water cycle, the processes of evaporation and transpiration, and the study being conducted at the Biosphere 2, and invited the students to participate. Students appeared highly engaged by the idea of contributing to the research through their classroom activity.

The week after the pre-test and introduction, sets of small pots with soil and pots planted with snapdragon plants were delivered to the classrooms, along with instruction for arranging the pots in specific patterns designed to contrast relative density (i.e., how many pots out of 20 were planted versus just had soil) and arrangement (e.g., if there were five plants, they could be grouped together or spaced apart with bare soil between them). The protocol specified how much water to administer each day and the intervals at which the pots would be weighed to measure overall water loss. The difference between the weight change attributed to soil-only pots and planted pots was attributed to the contribution of evaporation and transpiration that make up the overall evapotranspiration factor.

Over the next 2 weeks, the students worked in teams and followed their specified protocol with regard to the amount of water to administer. They weighed their pots and recorded the data, and transmitted it to the research team. The data were entered by students into a spreadsheet that was displayed on a interactive smartboard. As the weights were logged into the appropriate cells, the calculations representing water loss were automatically updated along with the graph showing loss attributed to evaporation versus transpiration. At the end of the week-long activity students completed a post test regarding their knowledge of evapotranspiration and science inquiry.

Study Results and Extensions

Teachers reported that the students took the activity seriously and that there was very little "off-task" behavior during the week. In addition, students appeared concerned with guaranteeing the correctness of their measurements. Having the values displayed publicly in an easy to read fashion on the smartboard led to interesting classroom discussion and promoted an informal peer review-like process in which the students corrected errors immediately during the data collection and analysis phases of the project. In addition, the pre-post test comparisons showed a positive impact on the student participants with regard to their understanding of the water cycle.

After the activity, students were more likely to provide accurate definitions for key terms, to recognize the factors and conditions that influenced evapotranspiration, and to suggest reasonable strategies for calculating weight (Villegas et al. 2010).

Students were also able to follow the experimental protocols successfully and to provide valuable data. The success of the activity was somewhat surprising given that the process of weighing plants required considerable attention to relatively fine-grained measurements and the use of the metric scale—topics that are known to be challenging for middle school students in the United States (Slavin and Lake 2008). In addition, the activity was sensitive to the successful completion of multiple sequential steps: students had to weigh the plants, record the values in decimal format to the second decimal place, and then enter the data into the workbook on the smart board. Processes that involve multiple steps are often prone to error, especially with relatively young participants (Michaels et al. 2008; Hassard and Diaz 2008). However, the data generated by the students in the table-top experiments were considered credible and valuable by the research team, and were directly used to increase the breadth of experimental results of the larger scale Biosphere 2 study. Ultimately, the students' work led to the development of new conceptual frameworks about the effects of vegetation type and structure on the partitioning of evapotranspiration (Villegas et al. 2009).

The table-top experiment has now been administered several times, each time varying the plant type and other conditions. The initial table-top experiment was conducted in local Arizona schools in the spring of 2009 and 2010. In the summer of 2009, a group of high school students visiting the Biosphere 2 for a summer program carried out the experiment. Then, in the fall of 2009, another opportunity arose, this time partnering with two schools in New South Wales, Australia, the Dubbo Public School and Trangie School. In this variation the same experiment design was used, and video conferencing was used in which Dr. Villegas presented the introductory lecture and provided the same interaction that was used in the local studies. Feedback from the students and teachers at the remote sites was extremely positive, suggesting the potential to expand the activities via technology.

Conclusions

The possibilities for harnessing human computation on a large scale are far from exhausted but already the opportunities and benefits that result from engaging the public to participate in scientific research are being realized. The B2E2 project offers an example of one such opportunity: if an experiment can be presented at the right scale, there is opportunity to combine active science with appropriate STEM curricula goals, while at the same time also affording broader exploration of scientific questions, varying parameters and conditions as a kind of exploratory data collection. Our entry into the age of large and pervasive data is not only a boon for scientists, but will transform how we conduct science and how science communicates with and educates the public.

Acknowledgments We would like to thank our colleagues Matt Adamson, Javier Espeleta, and Katherine Gerst, as well as classroom teachers Nicole Tilecki and Cheryl Cook and their students for their enthusiastic participation in the project. Special thanks are also due to Daniel Espeleta for his important contributions.

References

Adams H, Guardiola-Claramonte M, Barron-Gafford G, Villegas JC, Breshears D, Zou C, Troch P, Huxman T (2009) Temperature sensitivity of drought-induced tree mortality portends increased regional die-off under global-change-type drought. Proc Natl Acad Sci 106:7063–7066

Bradford BM, Israel GD (2004) Evaluating volunteer motivation for sea turtle conservation in Florida. Retrieved August 12, 2005, from University of Florida, Institute of Food and Agricultural Sciences website: http://edis.ifas.ufl.edu

Brightsmith DJ, Stronza A, Holle K (2008) Ecotourism, conservation biology and volunteer tourism: a mutually beneficial triumvirate. Biol Conserv 141:2832–242

Clery D (2011) Galaxy zoo volunteers share pain and glory of research. Science 333:173

Cochran ES, Lawrence JF, Christensen C, Jakka RS (2009) The quake-catcher network: citizen science expanding seismic horizons. Seismol Res Lett 80:26–30

Cohn JP (2008) Citizen science: can volunteers do real research? Bioscience 58:192–197

Cornell Laboratory of Ornithology (2012) Annual report. Cornell University. Retrieved from https://secure3.birds.cornell.edu/Page.aspx?pid=2560. Accessed 5 June 2013

Delaney DG, Sperling C, Adams C, Leung B (2008) Marine invasive species: validation of citizen science and implications for national monitoring networks. Biol Invasions 10:117–128

Engel SR, Voshell JR Jr (2002) Volunteer biological monitoring: can it accurately assess the ecological condition of streams? Am Entomol 48:164–177

Fusco D (2001) Creating relevant science through urban planning and gardening. J Res Sci Teach 38(8):960–877

Galloway A, Tudor M, Vander Haegen WM (2006) The reliability of citizen science: a case study of Oregon white oak stand surveys. Wildl Soc Bull 34:1425–1429

Gibbs JP, Milstead B (2012) Monitoring the Galapagos ecosystem. In: Wolf M, Gardener M (eds) The role of science for conservation. Routledge, New York, NY, pp 105–118

Hassard J, Diaz M (2008) The art of teaching science. London: Routledge

Huxman TE, Wilcox BP, Breshears DD, Scott RL, Snyder KA, Small EE, Hultine K, Pockman WT, Jackson RB (2005) Ecohydrological implications of woody plant encroachment. Ecology 86:308–319

King C, Lynch CV (1998) The motivation of volunteers in the nature conservancy. J Vol Admin 16:5

Lundmark C (2003) BioBlitz: getting into backyard biodiversity. Bioscience 53:329

Marshall NJ, Kleine DA, Dean AJ (2012) CoralWatch: education, monitoring and sustainability through citizen science. Front Ecol Environ 10:332–334

Michaels S, Shouse AW, Schweingruber HA (2008) Ready, set, science. National Academies Press: Washington, DC

Newman C, Buesching CD, Macdonald DW (2003) Validating mammal monitoring methods and assessing the performance of volunteers in wildlife conservation. Biol Conserv 113:189–197

Raddick MJ, Bracey G, Gay PL, Lintott CJ, Murray P, Schawinksi K, Szalay AS, Vandenberg J (2010) Galaxy zoo: exploring the motivations of citizen science volunteers. Astron Educ Rev 9:1–18

Silvertown J (2009) A new dawn for citizen science. Trends Ecol Evol 24:467–471

Slavin R, Lake C (2008) Effective programs in elementary mathematics: A best-evidence synthesis. Rev Educ Res 78:427–515

Trumbull DJ, Bonney R, Bascom D, Cabral A (2000) Thinking scientifically during participation in a citizen-science project. Sci Educ 84:265

Villegas JC, Barron-Gafford G, Adams HA, Guardiola-Claramonte M, Sommer E, Wiede AL, Breshears DD, Zou CB, Huxman TE (2009) Evapotranspiration partitioning along gradients of tree cover: ecohydrological insights from experimental evaluation In the biosphere 2 glass-house facility. AGU Chapman conference on examining ecohydrological feedbacks of landscape change along elevation gradients in semiarid regions. Sun Valley, Id, Oct 5–9 2009

Villegas JC, Morrison CT, Gerst KL, Beal CR, Espeleta J, Adamson M (2010) Impact of an eco-hydrology classroom activity on middle school students' understanding of evapotranspiration. J Nat Res Life Sci Educ 39:150–156

Williams TR (2001) Reconsidering the history of the AAVSO. J Am Assoc Var Star Obs 29:132

Yu J, Kelling S, Gerbratcht J, Wong W (2012) Automated data verification in a large-scale citizen science project: a case study. In: Proceedings of the IEEE 8th international conference on E-Science, pp 1–8

Search and Discovery Through Human Computation

Albert Yu-Min Lin, Andrew Huynh, Luke Barrington, and Gert Lanckriet

Introduction

Machines are good at handling huge amounts of data, but they lack the flexibility and sensitivity of human perception when making decisions or observations. To understand human perception, we look toward what defines being human. To sense, observe, and make sense of the world around us, we combine our biological receptors (eyes, ears, etc.) with our cognitive faculties (memory, emotion, etc.). But the memory banks that we pull from to create comparative reasonings are unique from individual to individual. Thus, we each see things in slightly different ways, i.e. what

A.Y.M. Lin (✉)
California Institute for Telecommunications and Information Technology,
UC San Diego Division, 9500 Gilman Drive, La Jolla, CA 92093-0436, USA

National Geographic Society, Washington, D.C., USA
e-mail: a5lin@ucsd.edu

A. Huynh
California Institute for Telecommunications and Information Technology,
UC San Diego Division, 9500 Gilman Drive, La Jolla, CA 92093-0436, USA

Computer Science and Engineering Department, University of California at San Diego,
La Jolla, CA, USA
e-mail: a5huynh@cs.ucsd.edu

L. Barrington
Digital Globe Corporation, Longmont, Colorado, USA
e-mail: lukeinusa@gmail.com

G. Lanckriet
Electrical and Computer Engineering Department, University of California at San Diego,
9500 Gilman Dr, La Jolla, CA 92093, USA
e-mail: gert@ece.ucsd.edu

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_16,

is beautiful to one person may not be to another. However, there are trends that emerge among our collective human consciousness and efforts to tap a consensus of human perception, i.e. crowdsourcing, depend upon these trends to scale up analytical tasks through massively parallel networks of eyes and minds. This concept of crowd based computing has become an important approach to the inevitable "data avalanches" we face.

The Modern Age of Human Information Processing: More than one quarter of the world's population has access to the Internet (Internet World Stats 2009), and these individuals now enjoy unprecedented access to data. For example, there are over one trillion unique URLs indexed by Google (Google Blog 2008), three billion photographs on Flickr, over six billion videos viewed every month on YouTube (comScore 2009), and one billion users of Facebook, the most popular social networking site. This explosion in digital data and connectivity presents a new source of massive-scale human information processing capital. User generated content fills blogs, classifieds (www.craigslist.org), and encyclopedias (www.wikipedia.org). Human users moderate the most popular news (www.reddit.com), technology (www.slashdot.org), and dating (www.plentyoffish.com) sites. The power of the internet is the power of the people that compose it, and through it we are finding new ways to organize and connect networks of people to create increasingly powerful analytical engines.

Breaking up the Problem: To combine the large-scale strength of online data collection with the precision and reliability of human annotation, we take a creative approach that brings the data collection process close to humans, in a scalable way

Fig. 1 Ultra-high resolution imagery of Mongolia displayed on the HiperSpace visualization facility at UC San Diego

that can motivate the generation of high quality data. Human computation has emerged to leverage the vast human connectivity offered by the Internet to solve problems that are too large for individuals or too challenging for automatic methods. Human computation harnesses this online resource and motivates participants to contribute to a solution by creating enjoyable experiences, appealing to scientific altruism, or offering incentives such as payment or recognition. These systems have been applied to tackle problems such as image annotation (von Ahn and Dabbish 2004), galaxy classification (www.galaxyzoo.org), protein folding (Cooper et al. 2010), and text transcription (von Ahn et al. 2008). They have demonstrated that reliable analytics can produced in large scales through incremental contributions from parallel frameworks of human participantion.

One approach to human computation motivates participants by creating enjoyable, compelling, engaging games to produce reliable annotations of multimedia data. Markus Krause's chapter (in this book) on gamification provides a brilliant investigation of this specific topic. These "games with a purpose" (von Ahn 2006) have been applied to classify images (von Ahn and Dabbish 2004; von Ahn 2006), text (von Ahn et al. 2006) and music (Mandel and Ellis 2007; Barrington et al. 2012b; Law and vonAhn 2009). In general, these games reward players when they agree on labels for the data and, in turn, collect information that the consensus deems reliable. The goal of these games has been to collect data on such a massive scale that all the available images, text or music content could be manually annotated by humans. Although simple and approachable online games – "casual games" – have broadened the video gaming demographic (International Game Developers Association 2006), designing a human computation game that meets these data collection goals while being sufficiently attractive to players in massive volumes remains a challenge.

In this chapter we describe several efforts to produce game like frameworks that take on a needle-in-a-haystack problems, often when the needle is undefined. Specifically, we explore innovative networks of human computation to take on the ever expanding data challenges of satellite imagery analytics in search and discovery. We describe frameworks designed to facilitate peer directed training, security through the partitioning and randomization of data, and statistical validation through parallel consensus. In each case it is clear that careful architecture of information piping is a determinate in the success of parallel human computation. We begin with an overview of our initial efforts in satellite remote sensing for archaeology, followed by subsequent experiences in disaster assessment, and search and rescue.

Case Study: Archaeological Remote Sensing

In 2010 we launched "Expedition: Mongolia" as the satellite imagery analytics solution for the *Valley of the Khans Project* (VOTK), an international collaboration between UC San Diego, the National Geographic Society, and the International Association for Mongol Studies to perform a multidisciplinary non-invasive search for the tomb of Genghis Khan (*Chinggis Khaan*). We turned to massively parallel human computation out of frustration from the inability to effectively survey the vast quantity of imagery data through automated or individual means.

Since the invention of photography, aerial images have been utilized in archaeological research to provide greater understanding of the spatial context of ground features and a perspective that accentuates features which are not otherwise apparent (Riley 1987; Bewley 2003; Deuel 1969; Lyons 1977). Buried features can produce small changes in surface conditions such as slight differences in ground level, soil density and water retention, which in turn induce vegetation patterns (cropmarks), create variability in soil color (soilmarks) or even shadows (shadowmarks) that can be seen from above.

The introduction of earth sensing satellites has further contributed to the integration of remote sensing in archaeology (Fowler 1996; Parcak 2009). The ability of detecting features on the ground from space is largely dependent upon the ratio of feature size to data resolution. As sensor technologies have improved, the potential to utilize satellite imagery for landscape surveys has also improved (Wilkinson et al. 2006; Lasaponara and Masini 2006; Blom et al. 2000). In September of 2008 the GeoEye-1 ultra-high resolution earth observation satellite was launched by GeoEye Inc. to generate the world's highest resolution commercial earth-imaging (at the time of launch) (Madden 2009). Generating 41 cm panchromatic and 1.65 m multispectral data this sensor further expanded the potential of satellite based archaeological landscape surveys. However, the massive amount of data that is collected each day by these sensors has far exceeded the capacity of traditional analytical processes. Thus, we turn to the crowds to scale human computation towards a new age of exploration.

We construct a massive parallel sampling of human perception to seek and survey the undefined. Specifically, we aim to identify anomalies in vast quantities of ultra-high resolution satellite imagery that represent archaeological features on the ground. Because these features are unknown we are searching for something we cannot predefine. Our internet-based collaborative system is constructed such that individual impact is determined by independent agreement from the "crowd" (pool of other participants who have observed the same data). Furthermore, the only direction that is provided to a given participant comes from the feedback in the form of crowd generated data shown upon the completion of each input. Thus, a collective perception emerges around the definition of an "anomaly".

The Framework

Ultra-high resolution satellite imagery covering approximately 6,000 km^2 of landscape was tiled and presented to the public on a National Geographic website[1] through a platform that enabled detailed labeling of anomalies.

Within the data interface participants are asked to annotate features within five categories: "roads", "rivers", "modern structures", "ancient structures", and "other". For each image tile, participants were limited to create no more then five separate annotations. This limitation was designed to limit the influence that any single individual could have on a given section of imagery (see Fig. 2).

Image tiles (with georeference meta data removed) were distributed to participants in random order. By providing segmented data in random order a collection of participants (or participant with multiple registrations) could not coordinate a directed manipulation of any given location. This was designed to both secure the system against malicious data manipulation as well as to protect the location of potential sites from archaeological looters.

At the onset of the analysis, ground truth information did not exist to provide an administrative source of feedback of the accuracy of analysis to participants. Thus we depend upon peer feedback from data previously collected by other randomly and independent observers of that image tile to provide a consensus based reference to position ones input in relation to the "crowd" (see Fig. 3).

The semi-transparent feedback tags provide a reference to gauge one's input to the perceptive consensus of a crowd. This reference information cannot be used to change the input provided to that particular image tile, however is designed to influence the participant for the following image tiles. Basing training on an evolving

[1] http://exploration.nationalgeographic.com/mongolia

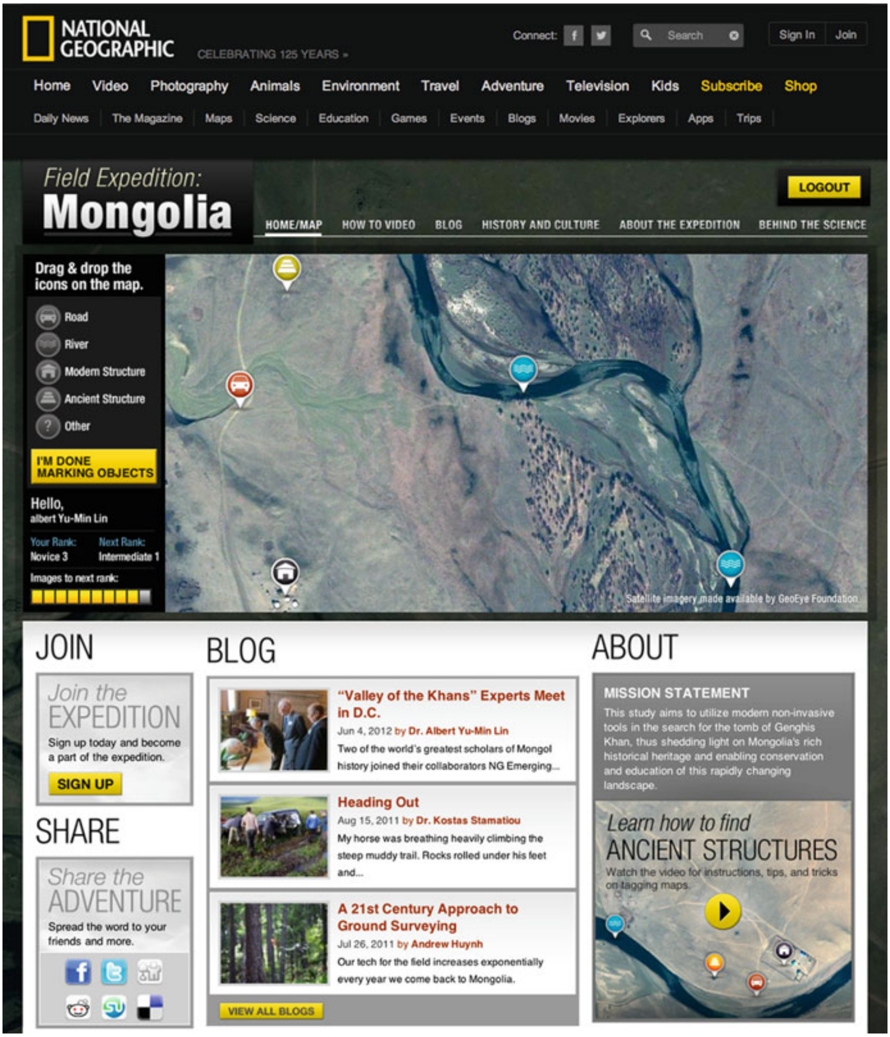


Fig. 2 User interface for online participants to identify anomalies within randomly presented sub-sectioned satellite imagery (Presented on http://exploration.nationalgeographic.com/mongolia)

peer generated data set we allow a form of emergent collective reasoning to determine the classifications, an important design element when searching for something that cannot be predefined.

The emergence of "hotspots" of human agreement also provide a form of validation through agreement among independent observers (a multiply parallel blind test). The mathematical quantification of agreement is the basis for extracting insight from the noisy human data. A detailed investigation of this framework and the role of collective reasoning will be reported in a forthcoming manuscript (Lin et al. 2013).

Fig. 3 Peer based feedback loop (Presented on http://exploration.nationalgeographic.com/mongolia)

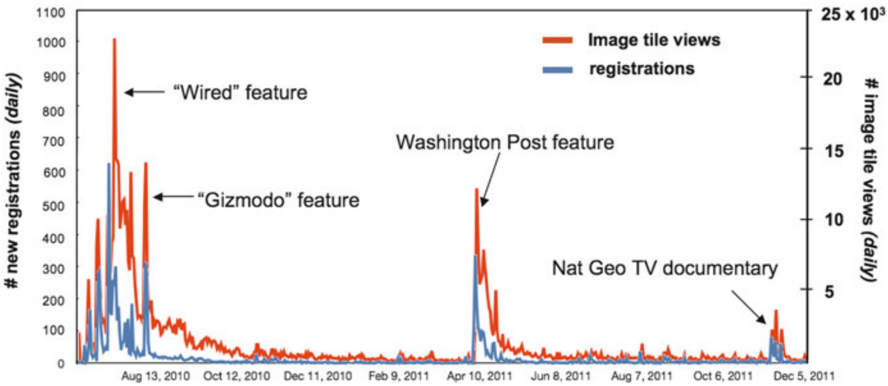


Fig. 4 Registration (*blue*) and image view (*red*) statistics across the duration of the experiment

Opening the Flood Gates

Since its launch over 2.3 million annotations from tens of thousands of registered participants were collected. Recruitment was facilitated through public media highlights, i.e. news articles and blogs. These highlighting events provide observable spikes of registration/participation, as seen in Fig. 4. We show this trend to articulate the importance of external communities to drive participation in crowdsourced initiatives.

Overlaying this huge volume of human inputs on top of satellite imagery creates a complex visualization challenge (Huynh et al. 2013) a subset of which is depicted in Fig. 5. While independently generated human inputs are inherently noisy, clusters

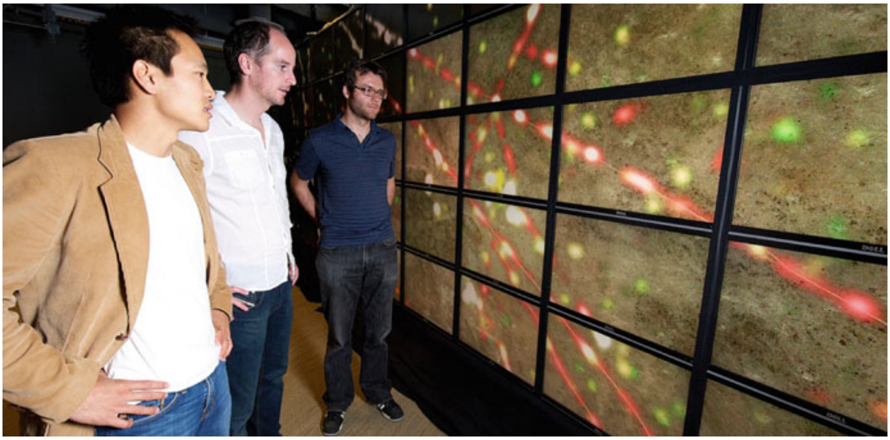


Fig. 5 Human generated tags overlaid on satellite imagery showing emergent agreement around features. Tag categories "road" and "ancient" are represented in *red* and *yellow*, respectively. We have explored methods of clustering to define linear features through tags (roads and rivers, Huynh and Lin (2012))

of non-random organization do emerge. Categorical filtering highlights road networks, rivers, and archeological anomalies, respectively.

Guided by this global knowledge of public consensus, we launched an expedition to Mongolia to explore and groundruthed locations of greatest convergence (defined mathematically through kernel density estimations). From three base camp locations along Mongolia's Onon River Valley we were restricted to a proximity boundary based upon 1 day's travel range and limitations associated with extreme inaccessibility. This created an available coverage distance of approximately 400 square miles. Within these physical boundaries we created and explored a priority list of the 100 highest crowd rated locations of archaeological anomalies. The team applied a combination of surface, subsurface geophysical (ground penetrating radar and magnetometry), and aerial (UAV based) survey to ground truth identified anomalies (Lin et al. 2011). Of those 100 locations, over 50 archaeological anomalies were confirmed ranging in origins from the Bronze age to the Mongol period (see example in Fig. 6).

Case Study: Christchurch Earthquake Damage Mapping

Born out of the success of "Expedition:Mongolia" Tomnod Inc. was formed in 2011 to explore broader application of human computation in remote sensing. While search targets varied, the computation challenge was consistent. The methodology of large scale human collaboration for earth satellite imagery analytics was quickly applied in the aftermath of a 6.3 magnitude earthquake that devastated the city of Christchurch, New Zealand in February 2011.

Fig. 6 Rectangular burial mound (identified through our human computation network) from early to late Bronze Age origins (Allard and Erdenebaatar 2005; Jacobson-Tepfer et al. 2010)

Once again, a website was developed to solicit the public's help in analyzing large amounts of high-resolution imagery: in this case 10 cm aerial imagery (Barrington et al. 2012a). Users were asked to compare imagery taken before and after the quake and to delineate building footprints of collapsed or very heavily damaged buildings. The interface was designed to be simple and intuitive to use, building on widespread public familiarity with web-mapping platforms (Google Maps, Google Earth, Bing Maps, etc.), so that more of the user's time is spent analyzing data versus learning how to use the interface. Using a simple interface that runs in a web browser, rather than an 'experts-only' geographic information system (GIS) platform, opens the initiative to a larger group of untrained analysts drawn from the general Internet public (Fig. 7)

After just a few days, thousands of polygons outline areas of damage were contributed by hundreds of users. The results are visualized in Fig. 8 below where areas of crowd consensus can be clearly identified by densely overlapping polygons. The crowd's results were validated by comparison to ground-truth field surveys conducted in the days immediately following the earthquake. The field surveys marked buildings with red (condemned), yellow (dangerous) or green (intact) tags, indicating the level of damage. Ninety-four percentage of the buildings tagged by the crowd were actually reported as damaged (red or yellow) by the field survey (Foulser-Piggott et al. 2012).

Fig. 7 Tomnod Disaster Mapper Interface in the Christchurch GEOCAN effort

Case Study: Peru Mountain Search & Rescue

The previous case studies demonstrated the capability of large networks of distributed human analysts to identify undefined features and apply visual analytics to remote sensing datasets on a massive scale. The final application of crowdsourced remote sensing we discuss highlights the timeliness that can be achieved when hundreds of humans help search through imagery and rapidly identify features of interest. On July 25, 2012, two climbers were reported to be lost in the Peruvian Andes. Missing in a remote, inaccessible region, the fastest way for their friends in the US to help find them was to search through satellite images. DigitalGlobe's WorldView-2 satellite captured a 50 cm resolution image and, once again, Tomnod launched a crowdsourcing website to facilitate large scale human collaboration. Friends, family and fellow climbers scoured the mountain that the climbers were believed to have been ascending. The crowd tagged features that looked like campsites, people, or footprints and, within hours, every pixel of the entire mountainside had been viewed by multiple people (Fig. 9).

One of the first features identified within just 15 min of launching the website showed the 3-man rescue team making their way up the glacier in search of the climbers. Over the next 8 h, consensus locations were validated by experienced

Fig. 8 Results of the crowd-contributed damage outlines and highlights of two destroyed buildings. *Red* = completely destroyed, *orange* = heavy damage, *yellow* = light damage

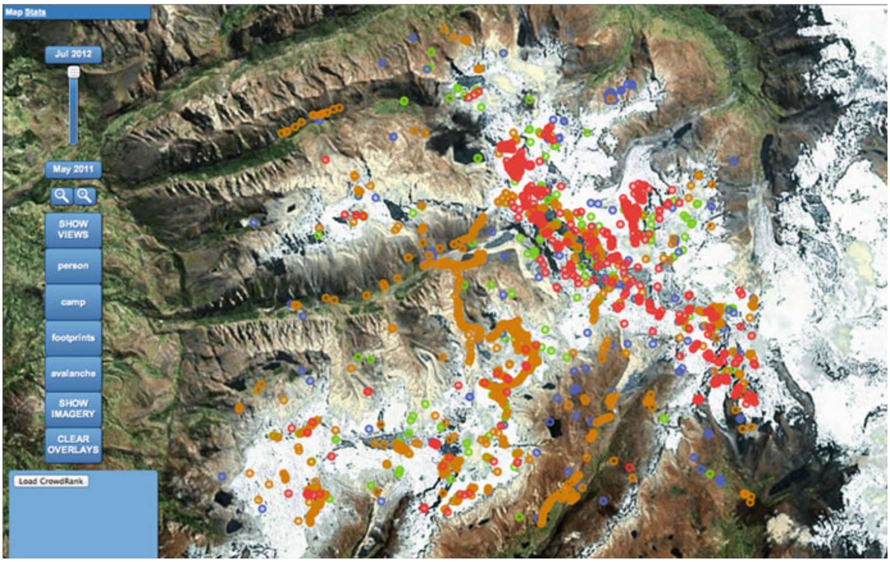


Fig. 9 Comprehensive crowdsourcing maps an entire mountain in just a few hours. The crowd identified possible footsteps (*orange*), people (*green*), campsites (*blue*) and avalanche regions (*red*)

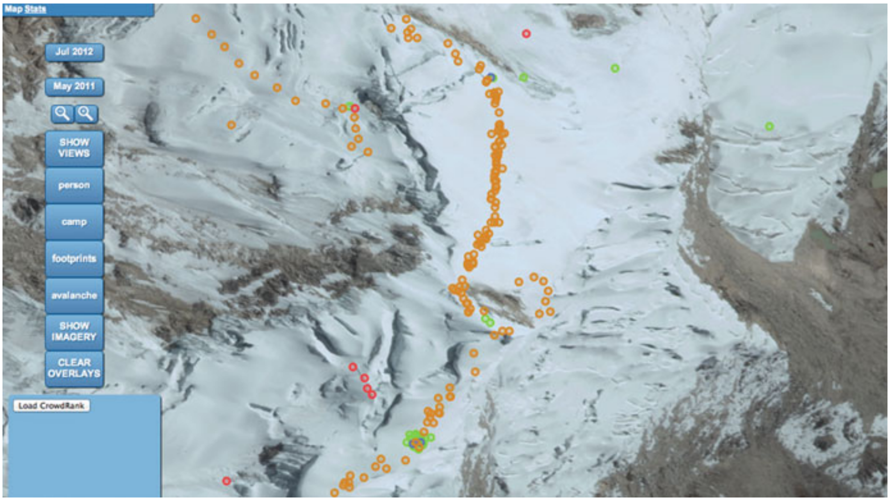


Fig. 10 Fresh foot tracks in the snow outlined through crowdsource analytics of near real time ultrahigh resolution satellite imagery

mountaineers and priority locations were sent to the rescue team on the ground (e.g., footprints in the snow, Fig. 10).

The search ended the next morning when the climbers bodies were discovered where they had fallen, immediately below the footprints identified by the crowd. While this case study has a tragic ending, the story highlights the power of human collaboration networks to search a huge area for subtle clues and, in just a few hours, go from image acquisition to insight. Furthermore, we observe that in times of need, humans want to help, and when channeled in appropriate collaborative pipelines can do so through computation.

Next Step: Collaborating with the Machine

While we have shown three examples of scalable human analytics, it would be a challenge for human computation alone to analyze every image on the web, every galaxy in the sky or every cell in the human body. However, human computation systems can produce well-labeled examples in sufficient volume to develop machine learning methods that can tackle such massive problems autonomously (Barrington et al. 2012b; Snow et al. 2008; Novotney and Callison-Burch 2010). By integrating machine intelligence systems with human computation, it is possible to both focus the human effort on areas of the problem that can not yet be understood by machines and also optimize the machine's learning by actively querying humans for labels of examples that currently confound the machine.

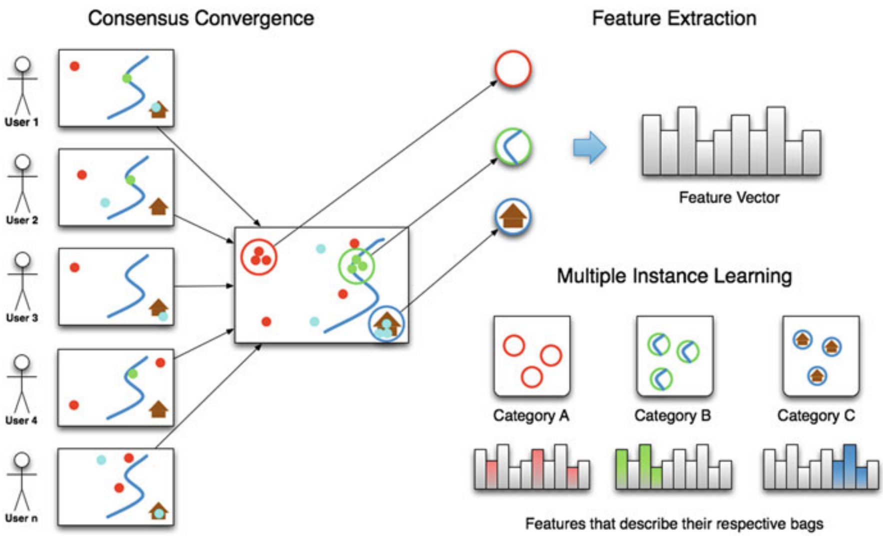


Fig. 11 Three phase approach to combine machine learning with search and discovery human computation: consensus convergence; feature extraction; and multiple instance learning

The detection of anomalies within an image is a difficult problem: we know that they may be located in regions of the image, but we don't know exactly where. We believe the application of multiple instance learning (Babenko et al. 2006; Maron and Lozano-Pérez 1998; Maron and Ratan 1998; Zhang et al. 2002) would be best suited for the problem at hand. Unlike the classical approach to learning, which is based on strict sets of positive and negative examples, multiple instance learning uses the concept of positive and negative bags to address the nature of fuzzy data. Each bag may contain many instances, but while a negative bag is comprised of only negative instances, a positive bag is comprised of many instances which are undetermined. While there may be negative examples in the positive bag due to noisy human input, the majority the positive examples will tend to lie in the same feature space, with negative examples spread all over. Multiple instance learning is able to rely on this insight to extrapolate a set of features that describes the positive bag. This is very appropriate for our data since a single image patch may contain many alternative feature vectors that describe it, and yet only some of those feature vectors may be responsible for the observed classification of the patch. A schematic of a proposed workflow for combining human computation and multiple instance learning (a machine based method) is outlined in Fig. 11.

If we are able to pool human perception to identify and categorize hard to define anomalies, we can begin applying this approach. From each of the many instances in a given category bag (i.e. ancient structure) we extract a set of image feature vectors. Since not every instance in the bag truly represents the labeled concept, some of these features will describe random image details, while others may be drawn from an actual ancient structure and will, for example, exhibit a certain rectangular shape. As we iterate through all the instances in multiple bags, the aim is that the features that describe an anomaly will become statistically significant. As the signal from multiple positive instances emerges from the uniformly distributed background noise, we can identify the features that best describe certain classes of anomaly. Thus even with multiple, noisy, weakly-labeled instances from our training set, applying multiple-instance learning will result in a set of features that describe each anomaly and which we can apply to new data to find anomalies therein.

Conclusions

The idea of collecting distributed inputs to tap the consensus of the crowd for decision making is as old as the democratic function of voting, but in this digital age, networks of individuals can be formed to perform increasingly complicated computational tasks. Here, we have described how the combined contribution of parallel human micro-inputs can quickly and accurately map landscapes and features through collective satellite imagery analytics.

In "Expedition:Mongolia" we designed a system of peer based feedback to define archaeological anomalies that have not been previously characterized, to leverage a collective human perception to determine normal from abnormal. Participants without pre-determined remote sensing training were able to independently agree upon image features based on human intuition, an approach avails of the flexibility and sensitivity of human perception that remains beyond the capability of automated systems. This was critical in our search for the "undefined needle in a haystack".

While this initial effort focused on an archaeological survey, applications of crowdsourced remote sensing exist across domains including search & rescue and disaster assessment. This was demonstrated through the efforts of Tomnod Inc., a group born out of the experiences in Mongolia to tackle the data challenges of the commercial satellite imaging industry through crowdsourced human computation. In the Christchurch disaster mapper effort we observe a remarkable 94 % accuracy to ground truth. This result opens new possibilities for human computation and remote sensing in the assessment and ultimately recovery of disaster events. The Peruvian Mountain search & rescue operation demonstrated the remarkable speed with which insight could be gained from pooling human effort for large scale data analytics, suggesting that a combination of networked human minds and fast data pipelines could actually save lives.

Each example demonstrates the potential of online communities to mine unbounded volumes of digital data and catalyze discovery through consensus-based analytics. We have shown how human perception can play a powerful role when seeking unexpected answers in noisy unbounded data.

However, while our approach depends upon emergent trends of agreement as the validating principle of actionable information, we observe this inherently does not capture the value of outliers (independent thinkers). Future work may identify mechanisms to reward "out of the box" thinking, possibly through a more detailed understanding and utilization of the individual human variables that contribute to a distributed human computation engine.

Finally, we observe that the natural next step in the evolution of human centered computation will be the collaboration between human and automated systems. This synergy will likely be required as we face the increasingly overwhelming data avalanches of the digital world.

Acknowledgements We thank N. Ricklin, S. Har-Noy of Tomnod Inc. as well as the entire Valley of the Khans (VOTK) project team; S. Bira, and T. Ishdorj of the International Association for Mongol Studies and F. Hiebert of the National Geographic Society for co-leadership in field expeditions; D. Vanoni, K. Ponto, D. Lomas, J. Lewis, V. deSa, F. Kuester, and S. Belongie for critical discussions and contributions; S. Poulton and A. Bucci of National Geographic Digital Media; the Digitaria team; Ron Eguchi and ImageCat Inc.; Digital Globe. This effort was made possible by the support of the National Geographic Society, the Waitt Institute for Discovery, the GeoEye Foundation, and the National Science Foundation EAGER ISS-1145291 and HCC IIS-1219138.

References

Allard F, Erdenebaatar D (2005) Khirigsuurs, ritual and mobility in the bronze age of mongolia. Antiquity 79:547–563

Babenko B, Dollár P, Belongie S (2006) Multiple instance learning with query bags, pp 1–9. vision.ucsd.edu

Barrington L, Ghosh S, Greene M, Har-Noy S, Berger J, Gill S, Lin AYM, Huyck C (2012) Crowdsourcing earthquake damage assessment using remote sensing imagery. Ann Geophys 54(6):680–687

Barrington L, Turnbull D, Lanckriet G (2012) Game-powered machine learning. Proc Natl Acad Sci 109(17):6411–6416

Bewley RH Aerial survey for archaeology. Photogramm Rec 18:273–292 (2003)

Blom RG, Chapman B, Podest E, Murowchick R (2000) Applications of remote sensing to archaeological studies of early Shang civilization in northern China. In: Proceedings IEEE 2000 international geoscience and remote sensing symposium, IGARSS 2000, vol 6, pp 2483–2485 Honolulu, Hawaii

comScore. http://www.comscore.com/Press_Events/Press_Releases/2009/3/YouTube_Surpasses_100_Million_US_Viewers, Mar 2009

Cooper S, Khatib F, Treuille A, Barbero J, Lee J, Beenen M, Leaver-Fay A, Baker D, Popović Z, Players F (2010) Predicting protein structures with a multiplayer online game. Nature 466(7307):756–760

Deuel L (1969) Flights into yesterday: the story of aerial archaeology. St. Martin's Press, New York

Foulser-Piggott R, Spence R, Brown D (2012). 15th World Conference on Earthquake Engineering (Lisbon) following February 2012 The use of remote sensing for building damage assessment following 22 nd February 2011 Christchurch earthquake: the GEOCAN study and its validation

Fowler MJF (1996) High-resolution satellite imagery in archaeological application: a Russian satellite photograph of Stonehenge region. Archaeological Prospection, 70; pp 667–671 Antiquity Portland Press

Google Blog. http://googleblog.blogspot.com/2008/07/we-knew-web-was-big.html, July 2008

Huynh A, Lin AYM (2012) Connecting the dots: triadic clustering of crowdsourced data to map dirt roads. In: Proceedings of 21st international conference on pattern recognition, Tsukuba
Huynh A, Ponto K, Lin A Yu-Min, Kuester F (2013) Visual analytics of inherently noisy crowdsourced data on ultra high resolution displays. Aerospace conference proceedings IEEE. 1--8
International game developers association (2006) Casual games white paper
Internet world stats. http://www.internetworldstats.com/stats.htm, Aug 2009
Jacobson-Tepfer E, Meacham JE, Tepfer G (2010) Archaeology and landscape in the Mongolian Altai: an Atlas. ESRI, Redlands
Lasaponara R, Masini N (2006) Identification of archaeological buried remains based on the normalized difference vegetation index (NDVI) from Quickbird satellite data. Geosci Remote Sens Lett IEEE 3:325–328
Law E, vonAhn L (2009) Input-agreement: a new mechanism for collecting data using human computation games. In: ACM CHI, Boston
Lin AY, Novo A, Har-Noy S, Ricklin ND, Stamatiou K (2011) Combining geoeye-1 satellite remote sensing, uav aerial imaging, and geophysical surveys in anomaly detection applied to archaeology. IEEE J Sel Top Appl Earth Obs Remote Sens 4(4):870–876
Lin AYM, Huynh A, Lanckriet G, Barrington L (2013) Crowdsourcing the Unknown: The Search for Genghis Khan (in preparation)
Lyons TR (1977) Remote sensing: a handbook for archeologists and cultural resource managers. Cultural Resources Management Division, National Park Service, US. Department of the Interior: For sale by the Supt. of Docs., U.S. Govt. Print. Off.
Madden M (2009) Geoeye-1, the world's highest resolution commercial satellite. In: Conference on lasers and electro-optics/international quantum electronics conference, OSA technical digest (CD). San Jose, CA, USA
Mandel M, Ellis D (2007) A web-based game for collecting music metadata. In: ISMIR, Vienna
Maron O, Lozano-Pérez T (1998) A framework for multiple-instance learning. In: Advances in neural information processing systems. Madison, Wisconsin, USA, pp 579–576 (Citeseer)
Maron O, Ratan AL (1998) Multiple-instance learning for natural scene classification. The fifteenth international conference on machine learning, San Francisco, pp 341–349, April 1998
Novotney S, Callison-Burch C (2010) Cheap, fast and good enough: automatic speech recognition with non-expert transcription. In: Human language technologies: 11th conference of the North American chapter of the association for computational linguistics (NAACL HLT), Los Angeles
Parcak SH (2009) Satellite remote sensing for archaeology. Routledge, London/New York
Riley DN (1987) Aerial photography in archaeology. University of Pennsylvania Press, Philadelphia, PA
Snow R, O'Connor B, Jurafsky D, Ng AY (2008) Cheap and fast—but is it good? Evaluating non-expert annotations for natural language tasks. In: 13th conference on empirical methods in natural language processing (EMNLP), Waikiki
von Ahn L (2006) Games with a purpose. IEEE Comput Mag 39(6):92–94
von Ahn L, Dabbish L (2004) Labeling images with a computer game. In: ACM CHI, Vienna
von Ahn L, Kedia M, Blum M (2006) Verbosity: a game for collecting common-sense facts. In: ACM CHI, Montral, Montréal, Canada
von Ahn L, Maurer B, McMillen C, Abraham D, Blum M (2008) reCAPTCHA: human-based character recognition via web security measures. Science 321(5895):1465–1468
Wilkinson KN, Beck AR, Philip G (2006) Satellite imagery as a resource in the prospection for archaeological sites in central Syria. Geoarchaeology 21:735–750
Zhang Q, Goldman SA, Yu W, Fritts JE (2002) Content-based image retrieval using multiple-instance learning. In: Machine learning-international workshop then conference- number 2, pp 682–689 (Citeseer). San Jose, CA, USA

Human Computation in Electronic Literature

Scott Rettberg

Introduction

Louis von Ahn (2009) has described human computation as "a paradigm for utilizing human processing power to solve problems that computers cannot solve." Quinn and Bederson (2011) further describe a consensus that what constitutes human computation are the problems that fit the general paradigm of computation, and as such might be solvable by computers; and in which the human participation is directed by the computational system or process. A typical example of human computation would be an Amazon Mechanical Turk process using the incremental labor of internet workers to verify that images of red shoes for sale in an online store actually match the description of the product's color advertised on the site.

Most forms of electronic literature can be considered to have some elements of human computation: the majority of works in this field consist of texts authored by humans which are then subject to some sort of computational process or algorithmic manipulation. Electronic literature is a field of literary and artistic practice that, according to the Electronic Literature Organization, involves "works with important literary aspects that take advantage of the capabilities and contexts provided by the stand-alone or networked computer." This encompasses a wide range of digital literary practices including hypertext fiction, kinetic poetry, chatbots, interactive fiction, interactive drama, generated poetry and narratives, narratives situated in networked communication technologies such as email, SMS, blogs, Twitter, and wikis, textual digital art installations, and many other practices. With electronic literature, human

S. Rettberg (✉)
The University of Bergen, Bergen, Norway
e-mail: scott.rettberg@lle.uib.no

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_17,

authors develop texts that involve computational processes—both texts that are themselves computer programs and texts that are the result of human interaction with algorithms—and human readers engage in reading practices that are technologically mediated.

Considering electronic literature from the standpoint of human computation is something of an inversion of the standard perspective. Scholars in this field more typically focus on how computers, networks, and computational processes can be useful in enabling humans to create new forms of literary expression, rather than beginning from the question of what roles humans play in a computational process. The challenges of creating a convincing and engaging narrative or producing a rich poetic use of language are still not generally solvable by computation alone. Even in the case of successful story or poetry generation, aspects of human writing are deeply integrated into the development of the system.

Hayles (2008) refers to the relationship between humans and computers evident in many works of electronic literature in terms of symbiotic loops: "Humans engineer computers and computers reengineer humans in systems bound together by recursive feedback and feedforward loops, with emergent complexities catalyzed by leaps between different media substrates and levels of complexity." Likewise, the relationship between the system and the human participants/authors in works of electronic literature is often more complexly layered than a single iteration of enlisting humans to perform tasks the system cannot provide without human input. There are examples of works of electronic literature where human authorship is directed by computational processes. We encounter systems that are first developed—by humans—as literary platforms, which then computationally direct, arrange, or integrate contributions by other humans.[1] The system may or may not be altered in response, in a recursive cycle that can continue.

After briefly discussing architectures of participation in collective narratives, I will focus herein on three types of human computation relevant to electronic literature:

1. Digital art projects involving human computation which offer some lessons for human-computation-driven electronic literature;
2. Poetry engines that use human contributions or human judgment to produce or refine combinatory or generate poetry;
3. Literary projects that are self-consciously engaged in a meta-level critique of the role that large-scale systems of human computation—for examples Google's global-scale harvesting of search queries—play in reconstructing contemporary human culture and social practices.

[1] The *ePluribus Solver* project (Greene et al. 2014) provides an example from the domain of collective journalism. Working with small fragments of a story in pictures using only a few characters or words, team members cast into descriptive and evaluative roles worked together to develop a collective narrative of the given situation.

Architectures of Participation: Frameworks for Collaboration

A literary project involving human computation should be understood to have an architecture of participation, a system that affords and constrains human participation. This architecture can be understood both as a platform in the sense of a computational system and a stage on which human interaction with the text, the system, and other authors and editors takes place.

Human computation in electronic literature is relatively uncharted territory. Paul Rohwer's "A Note on Human Computation Limits" (2010) considers two projects: *A Million Penguins*, a crowdsourced wiki novel produced by De Montfort University and Penguin Books in 2007, and two audio books produced by BBC Audiobooks America, that harvested Twitter responses to the first line of a story in "an iterative progression, singular integration model" to result in a collective fiction. The wiki novel project was an experiment in using the collaborative wiki platform—in which any user may edit any other user's text at any time (though those changes may be reverted)—to create collectively written novel. In their "A Million Penguins Research Report" (Mason and Thomas 2008) produced after the conclusion of the project, project organizers concluded that the result was ultimately less interesting as a novel than it was as a cultural text or performance. Penguin Digital Publisher Jeremy Ettinghausen reports "as the project evolved, I stopped thinking about it as literary experiment and starting thinking about more as a social experiment." Other critics and co-authors of the project recorded similar responses. The lightly controlled chaos of the wiki, it appears, served as a compelling arena for textual performance, but not for the development of a cohesive narrative.

Rohwer contrasts this project with one he considers successful, *Hearts, Keys, and Puppetry* by Neil Gaiman and the Twitterverse (2010). The story began with one tweet by Neil Gaiman, and readers then contributed Tweet-long continuations of the story. A single editor reviewed these tweets and selected the next line that would be included in the canonical version of the story, one line at a time. Rohwer argues that the "single real-time editor may be the natural requirement to achieve a sufficiently coherent narrative." While it is problematic to suggest that there is any "natural" requirement for coherent narrative—there are certainly many examples of multi-authored texts that did not have a single editor—it is clear that the two projects had different architectures of participation and control. The problem with narrative cohesion in *A Million Penguins* may have simply been that this architecture was not established as a system in which contributory and control roles were clearly defined and functional.

In a previous article focused on collective narratives (Rettberg 2011), I discussed a number of different online literary narrative projects that involved collaborative methods. These range from collaboration in small groups of authors, such as in the hypertext novel (Gillespie et al. 1998) to the attempt in the early 1980s by the Seattle writing group The Invisibles to use questionnaires and an early form of literary computer database to gather material for a novel, *Invisible Seattle* (1987), written by the whole city of Seattle, to projects such as Barbara Campbell's *1001 Nights Cast* (2005)—a durational performance in which Campbell daily solicited individual texts from internet participants in response to a prompt which changed each day, and then performed a reading of one texts each night 1001 nights in a row. Surveying collective narrative projects, I identified three different types of participation:

Conscious participation: Contributors are fully conscious of explicit constraints, of the nature of the project, and of how their contribution to it might be utilized.
Contributory participation: Contributors may not be aware of how their contribution fits into the overall architecture of the project, or even of the nature of the project itself, but they do take conscious steps to make their contribution available to the project.
Unwitting participation: Texts utilized in the collective narrative are gathered by the text-machine itself, and contributors have no conscious involvement in the process of gathering the material.

Human-computation-driven literary projects might involve any of these three different types of participation. People might be consciously participating as co-authors (for example by writing or editing a chapter of a wiki-based novel), may simply provide some text or information that will then be integrated by editors or by a computational system into a larger literary structure (for example respondents in the *Invisible Seattle* project who answered questions like "What is the best restaurant in Seattle to go for a break-up dinner?" and thus provided settings for the novel), or could be participating in a completely unwitting way (I will later discuss of Twitter haiku projects which harvest unwitting haiku from a general Twitter stream).

Digital Artworks Based on Human Computation

Electronic literature and digital art practices are deeply intertwined, so before moving to further specifically literary examples, it is useful to consider some notable examples of non-linguistic digital art that involve human computation. Aaron Kolbin's "The Sheep Market" (2006) is a project that involved the production of 10,000 sheep by workers on Amazon's Mechanical Turk. The workers were paid $.02 for each sheep they produced. Kolbin developed a Processing-based drawing tool, which recorded the drawing of each sheep. Each worker was instructed to "draw a sheep facing left." The results of the project included installations with prints of all 10,000 of the sheep, and animations, which reproduce the process of each sheep being drawn. Kolbin reports that the average wage paid to each worker was $.69 per hour, and the average time spent on drawing each sheep 105 seconds.

One might reasonably ask what the point of such an experiment might be, or where we should locate the "art" in a project which is based very much on the idea of "amateur" production (albeit "professional" in the sense that each of the workers was paid). Certainly on some level there is an embedded critique of the labor dynamics of human computation. Paying someone $.69 an hour for labor of any sort is unconscionable by

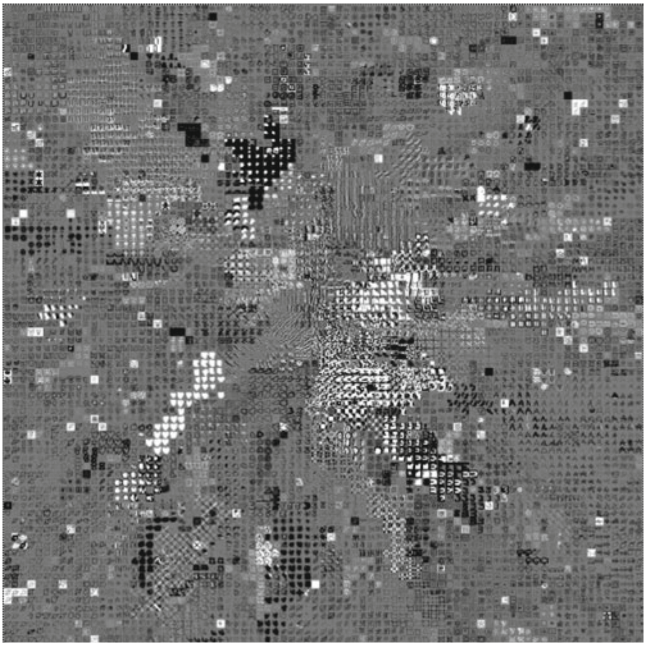


Fig. 1 Overview of "Seed Drawing 52" by Clement Valla (Reproduced from the artist's website)

the standards of most developed nations.[2] It calls into question other projects that use Mechanical Turk and similar platforms—is human computation simply a way of lowering labor costs to avoid paying human workers a reasonable minimum wage? And of course, the project also mirrors some more general global labor issues: Western consumers would not have access to such a plethora of affordable and wondrous consumer electronics without laborers in the East who are paid subsistence wages in poor working conditions to perform repetitive tasks. So on one level, the work can be understood as being about the political economy of contemporary consumer markets.

On the other hand, the process of human computation here also reveals tremendous creativity and diversity in a generalized class of human producers. Even in a simple rectangular black-and-white drawing environment, we encounter a diverse variety of approaches to producing a drawing of a barnyard animal. Like snowflakes, each of the 10,000 sheep in the market is in some way distinct from the others. The most fascinating aspect of watching the animations of the sheep drawings is seeing a human decision-making process unfold, as the workers draw, hesitate, make half-starts and scratch-outs. The drawings themselves are not nearly as affective as these ghostly presences, these invisible hands (Fig. 1).

[2] In his contribution to this volume, "Labor Standards," Alek Felstiner (2014) begins to unpack some of the thorny conceptual and jurisdictional issues involved in utilizing a globally distributed casual labor pool for crowdsourced human-computation-based labor.

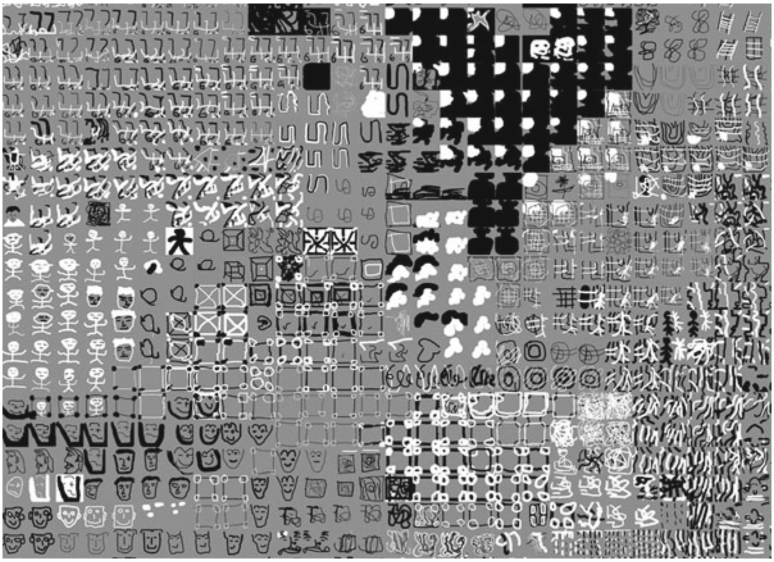


Fig. 2 Detail of “Seed Drawing 52” by Clement Valla (Reproduced from the artist’s website)

Clement Valla’s “Seed Drawings” series (2011) likewise uses Mechanical Turk as an engine for a collective art practice. In this case, rather than being provided with a written instruction of what to draw, each online worker is provided with a “seed drawing”—a pattern—and instructed to reproduce it using a simple drawing tool. The results, the artist notes, are much like a game of “telephone.” The first drawing is placed on the center of a grid, and the drawings based on it appear adjacent to it as they are produced. As each worker produces a new drawing based on another copy, the variability also increases dramatically. So what, in “Seed Drawing 52,” for example, is seeded as a simple black-and-white line pattern might, several generations later, evolve into an image of a face, or a coffee mug, or a letter, or a fish, or a star. As the original “message” is interpolated, its content changes significantly. One particularly interesting aspect of the drawings in “Seed Drawing 52” is that as the drawings are interpreted by different human actors, they generally appear to move from abstraction towards representation—at the center of the image we see abstract drawings but as we move to the outer parts of the grid, many more of the drawings are of recognizable objects or symbols. When charged with the pure task of mechanical reproduction, it seems the workers could not simply engage in automatic reproduction of the previous image, but were instead driven first towards interpretation. While a simple computer program could have replicated the seed drawing accurately in all 6,560 squares, the human workers first reflected on *what they thought it was*, reproducing not the image but an idea of the object it signified, even if it may have originally signified nothing (Fig. 2).

Kolbin, Valla, and a number of artists have continued to explore this type of collective, human-computation-driven methodology in subsequent works. From the perspective of narrative generation, Kolbin and Chris Milk’s recent 2012–2013 project "This Exquisite Forest (2013)" is perhaps the most intriguing. In this case, each work begins with a seed animation: for example of a stick figure falling down at the beginning of "A Bad Day." A HTML5 web-browser-based tool then allows successive users to add new frames to the new animation. They might continue to build from the seed narrative, or they might build upon any of the resulting branches. The branching tree structure can be used in a number of different narrative or thematic ways. In some cases the trees are clearly based on continuing established narratives and taking a story to a new turn or diverted path, while in other examples the continuities are limited to those of visual style.

We can note common features in each of the three art projects discussed above that provide lessons for the production of successful literary works based on human computation:

1. In each case, the artists provide users with a simple tool and platform for developing their contributions;
2. Contributors are also provided with a clear and concise *constraint*;
3. While the constraint or instruction is explicit, the interaction of the user with the constraint is also the point at which *play* takes place in the system, as it involves a moment of interpretation and decision on the part of the contributor;
4. The essential element of what makes each work appreciable, as an aggregate, collective work of art is not the *accuracy* of the human response to instructions, but the *variability* of the human responses to the given constraints recognizable in the aggregate.

Online Haiku Generators Involving Human Computation

Many of the early experiments of net.art involved the aggregation of contributed texts by a number of different anonymous human actors. *The World's First Collaborative Sentence,* launched by Douglas Davis in 1994 is one simple example of this. When reader/contributors open TWFCS in a web browser, they encounter a long unbroken stream of text, and a link to a web form which they can use to contribute to the work-in-forever-progress. The primary goal of the project appears to have been open performance on a global network—the instructions encouraged contributors to "WRITE, PERFORM, OR SING ANYTHING YOU WISH TO ADD IN WHATEVER LANGUAGE YOU LOVE TO THIS COLLABORATIVE WORK, JOINING HANDS AND MINDS WITH YOUR SISTERS AND BROTHERS OF WHATEVER RACE, REGION, OR BELIEF ANYWHERE IN THE WORLD…" Contributors were encouraged to add not only text but also "PHOTOGRAPHS, VIDEO, SOUND." The only constraint was that the contribution could not include a period and therefore the sentence could theoretically go on forever.[3]

[3] Davis's work was live until the early 2000s when the scripts driving the project became nonfunctional in the context of the contemporary Web. In 2012, the Whitney Museum restored the digital work, releasing both a "restored" historical version and a fully functional live version which allows for new contributions.

Like many net.art projects, TWFCS was largely about the early idealistic exuberance and utopianism with which many people took the Web as they first encountered it as a new medium for human expression. The possibilities of instantaneous publication with nearly global reach and the ability to share texts and collaborate with thousands of other people, without the intrusion of institutional gatekeepers, were still very new in 1994. The focus is largely on the novelty of the device and the medium itself. The project was successful insofar as its aim was to simply be a large-scale participatory text—more than 200,000 contributions were made to TWFCS between 1994 and 2000. But it would be difficult to assess its interest or merit as a literary work. When the goal of the project is unstructured participation, it is no surprise that the result was rambling and largely incoherent.

From the standpoint of human computation, more compelling examples of digital literature involve participatory structures that use human contributions in more specific ways, driven by constraints and processes intended to result in a coherent reading experience. These often involve the use of literary forms that are themselves constrained. Let us consider for example three projects that enlist human participation in the generation of online haiku.

Though the structure of the traditional Japanese haiku is more refined, in its English incarnation, haiku is generally understood to be a form of three lines in a 5 -7 -5 syllable structure. Haiku are often imagistic, and typically deal with two aspects of nature that when juxtaposed, can serve to startle the reader or bring about some sense of recognition. Given the comparative simplicity of the form in its English incarnation compared say to a Shakespearian sonnet, it is no surprise that it has been the subject of many experiments with combinatory, generative, or collective poetry. Haiku were in fact among the forms of some of the earliest experiments with poetry generation—in 1967 John Morris published "How to Write Poems with a Computer" describing his haiku generation program developed at Michigan State University. Morris both described his actualized program and conceptualized a better one that would balance an algorithmic process with elements of randomness, though, he confessed that he found the most affective poetry to be "…communication from a particular human being. And this is precisely what a computer is not."

Nanette Wylde's *haikU* (2001) is a project based on principles of user participation and on the use of a randomizing function to produce haiku that startle in the sense of producing *unintended* juxtapositions—no single author has determined which lines will appear together. The reading interface is a simple, spare web page. Every time a reader reloads the page, a new haiku is produced. Following a link to "Write haiku" individuals can submit their own haiku in three lines, each of which has its own button to post the line to bins of first, middle, and last lines. The poems delivered on each reload of the site are not the individual haiku as submitted by readers, but recombinations of these first, middle, and last lines of haiku pulled together in a variable way. Two reloads of the page produced for example "working round the world/the oven melting fire/brushed by a warm hand" and "under the rainbow/dew softly lays upon grass/hot sex in the night." Reloading the page 20 times or so, it is remarkable how many of the poems read as if they have been individually intended by a human intelligence. Most of the haiku, perhaps 80 %, cohere quite well as poetry (Fig. 3).

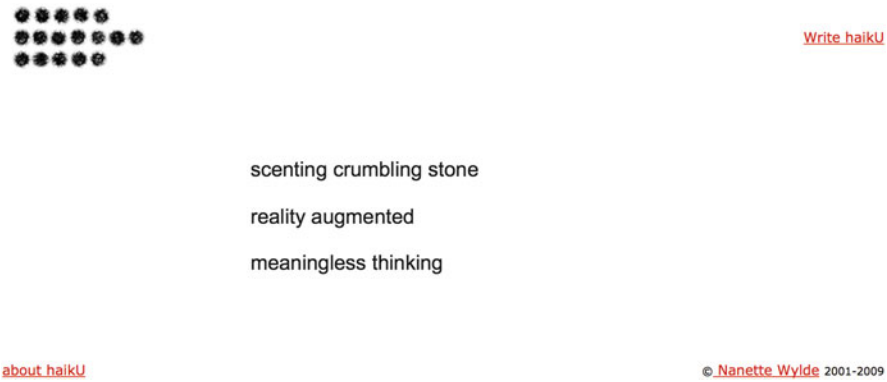


Fig. 3 Example of a haikU (Reproduced from the project site)

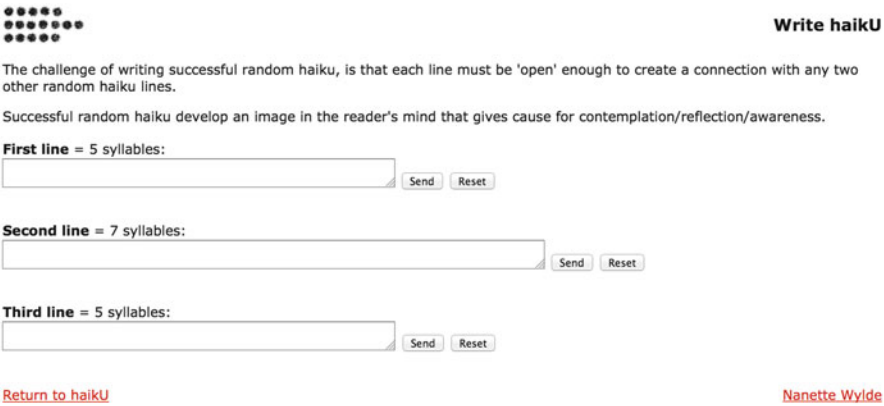


Fig. 4 *HaikU* writing interface (Reproduced from the project site)

Wylde provides two opportunities for instructions to contributors. The first is on the brief "about haikU" page where she explains not just what the project is but what Haiku are: "Haiku traditionally reference a season and are generally observations of everyday life" and she attests that the "challenge of writing successful random haiku is that each line must be 'open' enough to create a connection with any two other random haiku lines. Successful random haiku develop an image in the reader's mind that gives cause for contemplation/reflection/awareness." She reiterates these last two instructions on the "write haikU" page (Fig. 4).

In *haikU*, the combinatory form and structure of the project, in concert with the form and structure of the poetic form, and the fairly subtle instructions to contributors, lead to the production of a poetic database that works fairly well. While extremely simple in concept and execution, the combination of human-written lines and arbitrary structure results in new poetry neither completely determined by any human nor free of authorial intention.

Another online haiku generator project produced during the early 2000s, *HaikuTree.org*, (Goodwin 2000) attempted to bring human judgment to computer-generated haiku. Web readers would place generated haiku on a virtual tree. The haiku would be ranked by all these readers and would further "weather" over time. Only the most popular haiku would survive this process. In theory—though the project and its source code are no longer online—these selections would inform the process of refining the generator itself, to "help computers write better poetry." It is unclear from the remaining project documentation whether by this the project developer meant that human judgment was directly informing and training the system via a machine learning approach or simply informing the human developer as he refined the system itself. In any case, poetry or story generators that are trained by human response to output are certainly conceivable as a branch of further research.

A number of more recent online haiku generator projects harvest human-written texts from the Internet, scan them for 17 syllable count and appropriate word-breaks, break them into lines, and redisplay them as haiku. One example of this is John Berger's @HaikuD2 Twitter account (Berger 2013). In this case all of the text is human-produced but none of it is necessarily intended as haiku. It is only when Berger's bot provides line breaks and a #haiku tag that it becomes recognizable as such. The Twitter bot approach, at least in this iteration, may be more limited than Wylde's simpler system, which involves more intentionality on the part of the contributors. While some of the resulting haiku are clever or amusing in the way that they formalize language that is otherwise colloquial or banal, most of them simply read as tweets with line breaks, and not necessarily as particularly good poetry.

Based on a similar process to that of the Twitter haiku bots but generally producing more compelling results is *Times Haiku* (2013). Developed by the software architecture staff of *The New York Times*, *Times Haiku* is driven by an algorithm that scans the text of articles published on the *Times* home page for potential haikus using a syllable count dictionary. The dictionary is regularly updated and modified by the *Times'* staff "with words like 'Rhianna' and 'terroir' to keep pace with the broad vocabulary of The Times" (Harris 2013). The algorithm discards haiku "if they are awkwardly constructed" (presumably meaning they don't break lines properly) and do not scan articles "covering sensitive topics" (presumably to avoid the production of deeply offensive haiku). Staff of *The Times* then read the haiku found by the algorithm. Human journalists who find a haiku "beautiful or funny or just a gem of a haiku" then select them for posting to a Tumblr blog. Selected haiku are posted by the system as an image file on the blog, and from there readers can share them on a number of social network sites. Each posting also includes a link to the original *Times* story. If the haiku produced by this process are not often imagistic or concerned with nature, they are often timely and amusing in their relation to contemporary culture. A couple of choice examples of haiku resulting from this process during June 2013 include: "There are horses who/can uplift, cause a chuckle, / spur a memory." (from June 11, 2013 story "Philotimo: A Horse Rescue Story") and "Young skin is spandex; / older is linen and needs/loving attention." (from June 4, 2013 story "'Counterclockwise' and 'Up'—In Pursuit of Longevity").

Consider the relationships between computer and human involved in the production of *Times Haiku*:

1. Human journalists write stories including lines which (presumably unwittingly) have the syllabic count of a haiku;
2. These are automatically fed into an algorithm which flags them as haiku;
3. The program's syllabic vocabulary is further modified by human actors;
4. Human curators then interact with a feed of texts that meet the basic formal requirements of haiku;
5. Selected haiku are then formatted by the system as image files and posted on a Tumblr blog;
6. Human readers then choose to share and propagate their favorite haiku.

Times Haiku provides a superb case of a recursive literary use of human computation. Without the computational system, the majority of the texts from *The Times* would never be recognized as haiku. Without the unwitting participation of human contributors, the texts would not exist at all. Without the conscious participation of human curators, the system would have a more limited vocabulary and would provide less aesthetically satisfying results.

Literary Meta-critique of Human Computation

During recent years several e-lit authors have produced works that engage critically with human computation as an aspect of the contemporary network environment. In this case, the authors are not concerned as much with using human computation to develop collectively produced narratives or poetry, but instead with the systems of large international corporations such as Google and Facebook that regularly harvest and monetize information about their users and their behaviors on the network. Human computation is of course occurring on a large scale in these cases, as every time a user posts on Facebook or searches on Google, another contribution is made to a very large graph of extremely marketable information both about that specific user and about the broader contexts of human language and society. A group of authors loosely centered on the Digital Language Arts program at Brown University have this taken as a particular concern and derived literary art from it (Fig. 5).

Mimi Cabell and Jason Huff's *American Psycho* (2010) is a work that provides a context for considering how Google's different feedback mechanisms shape and control human experiences on the Internet. With this project, Cabell and Huff focused in particular on the Google Mail platform. They note "Google reads our mails, garners information from our personal messages, and uses that profiling strategy to select 'relevant' ads. It then displays those ads on the screen next to the very emails from which they were initially taken." In order to test the behaviors of this system, the authors chose to send the entire text of Brett Easton Ellis's novel *American Psycho* through Gmail one page at a time. They then collected the links

Fig. 5 *American Psycho* recontextualized (Photo reproduced from the project site)

that Google displayed, and printed a book, in which they left intact Ellis's chapter titles but eliminated the text of Ellis's novel, leaving only footnotes that recorded the links Google had provided for each page of the novel. They report that some of the ads Google returned were directly relevant to the text from the novel—a scene in the novel involving the brutal stabbing of a dog and a man generated ads for knives and knife sharpeners—if at other times completely irrelevant to the context of the novel. Sections of the novel including racist language did not return any ads at all, indicating that Google's technology has at least some censoring in place. Ads for Crest Whitestrips coupons were the most frequent single item to appear. The project might be described as a work of conceptual writing focused on revealing and foregrounding processes of human computation that we might take for granted in the course of everyday interactions on the network that simultaneously take advantage of us and make marginal but significant alterations to our communications environments (Fig. 6).

Complex questions of who has—and who should have—access to shared literary heritage and linguistic data are at play in John Cayley and Daniel Howe's *How It Is in Common Tongues* project (2012). They describe the overall project of *Common Tongues* as remediating "practices of and processes of reading" and critically addressing "the commodification of reading itself, and the proprietary enclosure of a growing portion of our linguistic cultural commons." In particular the project addresses the fact that on the Internet many texts are now first read, processed, recomposed, and "multimediated" by computers in "pages that precede and predetermine any further or deeper 'human' reading." The project, installed at the ELMCIP Remediating the Social exhibition at the Inspace Gallery in Edinburgh in

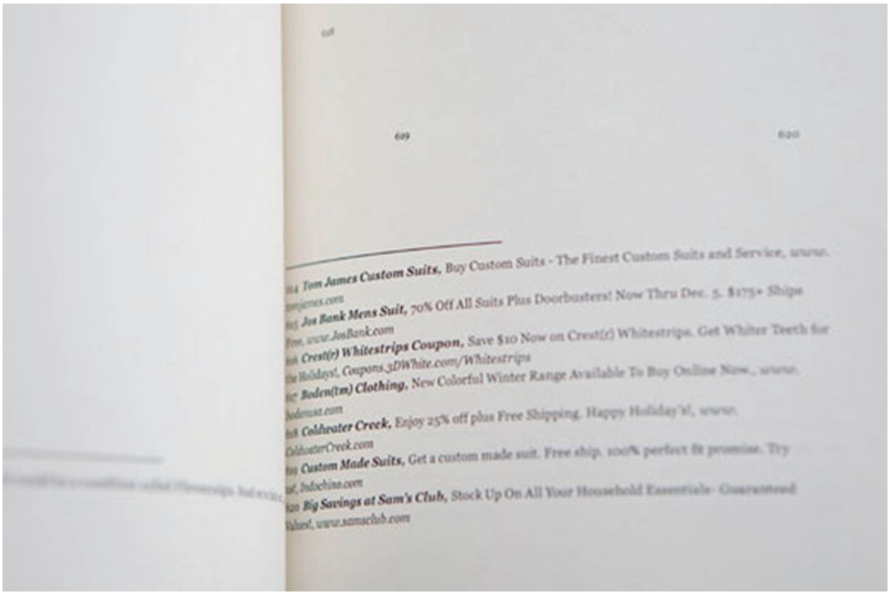


Fig. 6 A page of Cabell's and Huff's *American Psycho* showing only references to advertising URLs generated by sending Brett Easton Ellis's novel through Google Mail (Photo reproduced from the project site)

November 2012, had a number of different digitally mediated text components that engage in different ways with the Google search engine, practices of reading, ownership of language, and Samuel Beckett's work *How It Is*.

One aspect of Cayley and Howe's installation notable for its engagement with copyright was a printed copy of Samuel Beckett's text *How It Is*. While the text of the literary work printed in the book is identical on a word-for-word basis to Beckett's text, every phrase in the book is footnoted with a URL. This URL corresponds to a non-Beckett use of the phrase found as a result of using a search engine. In his description of the project in the ELMCIP Knowledge Base, Cayley notes that all of the words in the book "are quoted from a portion of the commons of language that happens to have been indexed by a universally accessible engine." Samuel Beckett's estate, notorious for their enforcement of copyright, would doubtless have some issues with this citation practice. The work however makes the point that the text here is doubly enclosed: once in Beckett's text by a copyright system that makes texts unavailable for reuse and adaptation until long after the authors are dead, and again as the texts that appear as search results by Google's indexing system, which harvests texts written on the Internet by humans and machines and immediately begins making use of those texts everywhere it encounters them.

Samantha Gorman's *Completely Automated* (2011) is an "exploration of how our written histories are forged through the interplay between human and machine editing." The project engages critically with the human-computation-based archival project reCAPTCHA—the system developed by Louis von Ahn which serves both as a spam blocker—by using language recognition to test whether a user is human—and as aid in the process of digital archiving of scanned texts—by using human responses to images of individual words in scanned archival texts to verify optical character recognition. Gorman produced a short film (2012) enacting a fictional scenario in which she can first be seen typing a text, “Pronouncement Against Domestick Production of Fraudlent Coinage as Decreed by Sovereign Law and Writ” by John Cartwright, into a page layout program, making modifications such as changing the name of the author, as she goes. She prints the modified text, outlines over the printed letters with painted ink, stains the paper with tea, giving it an aged appearance, before scanning the text into a university library’s archive system, and then planting it in a folder in the rare books room. The video concludes with other Internet users scrutinizing individual semi-obscured words of the fraudulent text, as these fragments are approved one at a time.

Gorman explains the crux of her issue with the reCAPTCHA process on the project site: “Essentially, even a slight deviation from the original may escape the loop’s filters and be preserved digitally as a final authoritative text: our cultural heritage. Meanwhile, the original print is less conveniently accessible than the digital version and begins to lose authority within its physical library archive.” Gorman further suggests that, in privileging human language recognition, the reCAPTCHA system suggests that these processes are what “define us as human and… best distinguish human cognition from that of a machine.” So Gorman’s project raises conceptual issues with both the inherent uncertainty involved in integrating humans into computational processes—humans might not only make errors but conceivably could purposefully subvert the system—and with the effect human computation might have on the role and function of human cognition. Furthermore, in integrating steps of human cognition into processes that are controlled by machines, are we in effect subordinating human cognition, treating humans as superior sensory apparatuses, but lesser cognizers, than the machines they serve?

As the three projects discussed above reveal, the relationship between electronic literature and human computation is not simply procedural. While electronic literature authors may design architectures of participation to develop more effective collectively produced narratives, or new ways of harvesting poetry from streams of network discourse, they also have a role to play in critiquing the technological apparatus in which humans are increasingly embedded as actors, if not ghosts, in the machine.

Conclusion and Potential for Further Research

This chapter has considered human computation in a number of different aesthetic contexts: in the development of collective narratives, in massively crowdsourced visual and conceptual art, in haiku generators that automatically harvest and represent poetry from a Twitter stream or the news of the day. It has also considered how authors and artists are responding to a context in which their agency as creators or co-creators is resituated in relation to networked systems that are increasingly harvesting and interpreting human communications, reading and reformulating texts, and composing and determining narratives. The relationship of contemporary digital literary practice to human computation is neither entirely symbiotic nor essentially adversarial.

The field of electronic literature by nature experimental: practices from a number of different fields including writing, computation, visual arts, performance, communication, and design meet in this sphere. If there is a general commonality to the various practices and artifacts grouped under the rubric, it is that they all share an interest in exploring the relationships between literature and computation. It is important to emphasize that this a reciprocal set of concerns: we explore both the ways in which new possibilities for literature are afforded and constrained by computational processes and the networked environment and, in turn, the new possibilities for computation and the networked environment afforded by literary practice.

In the specific area of human computation and network-based collective writing projects, although there is a rich and growing body of experimental work in the area, a great deal of practical research remains to be done. Detailed analytic case studies are necessary to better understand how collective writing systems can best be harnessed to establish a level of aesthetic control and structure that would result in a sufficiently coherent reader experience while allowing for a degree of play, variability of response, and diversity of collective knowledge that could usefully enhance these sorts of projects and distinguish them from single-author literary endeavors. Our understanding of these practices would also be furthered by greater research collaboration between writers and artists working in electronic literature and digital art with computer scientists working in human computation, machine language learning, and other areas.

Given world enough and time, this chapter could have detailed many other extant experimental works of collective writing. It is a growing area of interest. Projects such as Judd Morrissey, Mark Jeffrey and the Goat Island Collective's 2007–2010 project *The Last Performance* (Morrissey et al. 2007), for instance, involved a collective narrative contributed to by more than 100 other writers, all responding to the same provided constraints. The short narrative and poetic texts they produced were then machine-interpreted, thematically cross-linked, and visualized in a number of different configurations. This deconstructed/reconstructed narrative architecture further served as a text and context for live performance.[4] Projects such as Brendan Howell's *Exquisite Code* bring algorithmic processes even more deeply into the writing process. In that project, a group of writers sit together in rooms writing for extended periods of time in response to prompts that they and system generate. The texts that they write are then periodically subject to "select/mangle" processes by the system. Each performance of this project so far has resulted in a book-length text which could be said to have been written both by the participating authors and by the machine itself, in what Howell refers to as a "c[ad]aver[n]ous exquisite_code life-work" (Howell et al. 2008).

[4] See Rettberg (2010) for further discussion of this work and strategies for reading *The Last Performance* as text and collective performance.

There are many questions we have only begun to address: of how to best make use of human computation strategies to develop compelling collectively written narratives, of how to integrate algorithmic procedures into writing processes in ways that produce aesthetically satisfying results, of how to productively integrate the artistic research strategies of electronic literature with the experimental methodologies of computer science, and indeed of how the function of literary writing in general changes in an environment in which networked systems are constantly harvesting and reframing texts of all kinds. We can only be certain that when confronted with technological opportunity, writers will continue to invent new literary forms and that contemporary literary works will continue to offer opportunities for reflection on the communication technologies, languages, and cultures of the era in which they are produced.

References

Ahn L (2009) Human computation. In: DAC'09 proceedings of the 46th annual design automation conference. ACM, New York, pp 418–419

Berger J @HaikuD2 (2013) Twitter stream. Accessed 11 June 2013.

Cabell M, Huff J (2010) American psycho. http://www.mimicabell.com/gmail.html. Accessed 17 June 2013

Campell B (2005) 1001 nights cast. http://1001.net.au/. Accessed 18 June 2013

Cayley J, Howe D (2012) How it is in common tongues. Documentation on the ELMCIP electronic literature knowledge base. http://elmcip.net/node/5194. Accessed 17 June 2013

Davis D (1994) The world's first collaborative sentence. http://whitney.org/Exhibitions/Artport/DouglasDavis. Accessed 11 June 2013

Felstiner A (2014) Labor standards. In: Michelucci P (ed) The handbook of human computation. Springer, New York

Gaiman N, Twitterverse (2010) Hearts, keys, and puppetry. Audiobook. BBC Audiobooks America, North Kingstown

Gillespie W, Rettberg S, Stratton D, Marquardt F (1998) The unknown. http://unknownhypertext.com. Accessed 18 June 1023

Gorman S (2011) Completely automated. Project site. http://samanthagorman.net/Completely-Automated. Accessed 17 June 2013

Gorman S (2012) Completely automated. Video on Vimeo. http://vimeo.com/33118528. Accessed 17 June 2013

Goodwin D (2000) HaikuTree.org. Documentation of project at Rhizome.org. http://rhizome.org/artbase/artwork/2178/. Accessed 11 June 2013

Greene K, Thomson D, Michelucci P (2014) Explorations in massively collaborative problem solving. In: Michelucci P (ed) The handbook of human computation. Springer, New York

Hayles NK (2008) Electronic literature: new horizons for the literary. University of Notre Dame, Notre Dame

Harris J (2013) About Times haiku. http://haiku.nytimes.com/about. Accessed 12 June 2013

Howell B et al (2008) Exquisite code. http://exquisite-code.com/. Accessed 21 June 2013

Invisible Seattle (1987) The novel of Seattle, by Seattle. Function Industries Press, Seattle

Kolbin A (2006) The sheep market. http://www.aaronkoblin.com/work/thesheepmarket/index.html. Accessed 11 June 2013

Kolbin A, Milk C (2012) This exquisite forest. http://www.exquisiteforest.com. Accessed 11 June 2013

Mason B, Thomas S (2008) A million penguins research report. http://www.ioct.dmu.ac.uk/documents/amillionpenguinsreport.pdf. Accessed 17 June 2013

Morris J (1967) How to write poems with a computer. Michigan Q Rev 6.1: 17–20

Morrissey J et al (2007) The last performance [dot org]. http://thelastperformance.org. Accessed 21 June 2013

Quinn AJ, Bederson BB (2011) Human computation: a survey and taxonomy of a growing field. In: CHI'11: proceedings of the SIGCHI conference on human factors in computing systems. ACM, New York, pp 1403–1412

Rettberg S (2010) Performative reading: attending the last performance [dot org]. Dichtung digital 40. http://dichtung-digital.mewi.unibas.ch/2010/rettberg/rettberg.htm. Accessed 21 June 2013

Rettberg S (2011) All together now: hypertext, collective narratives, and online collective knowledge communities. In: Page R, Thomas B (eds) New narratives: stories and storytelling in the digital age. University of Nebraska Press, Lincoln

Rohwer P (2010) HCOMP '10: proceedings of the ACM SIGKDD workshop on human computation. ACM, New York, pp 38–40

Times Haiku Generator (2013) http://haiku.nytimes.com/. Accessed 12 June 2013

Valla C (2011) Seed drawings. http://clementvalla.com/category/work/seed-drawings/. Accessed 11 June 2013

Wylde N (2001) haikU. http://preneo.org/nwylde/haikU/index.shtml. Accessed 11 June 2013

Human Computation for Information Retrieval

Christopher G. Harris and Padmini Srinivasan

Introduction

Information Retrieval (IR) involves locating documents that match an information need. Although searching documents is a core component of any IR system, few user information needs are satisfied by the initial query. In studies of Web searches, which parallel document searches, more than half of all queries are subsequently reformulated by users after results are returned from an initial query (Spink et al. 2002). Query refinement is often necessary due to the presence of over- or under-specified search terms, inappropriate terms retrieving non-relevant documents, and typos. Thus, query refinement is an important step and a core area of study in information retrieval. It is widely acknowledged that an initial query refined using a reasonable strategy will yield better results than the initial query. The basis of the refinement may be human-assessed feedback or pseudo relevance feedback[1] derived from the documents retrieved by the initial query.

Two recent human computation developments, crowdsourcing and games with a purpose (GWAP), charge us to return to query design research. Crowdsourcing is a framework whereby tasks may be accomplished quickly and cheaply by soliciting from a largely anonymous pool of participants. GWAP interfaces are similar except that these devices are also games meant to entertain, reward with scores, be

[1] Pseudo relevance feedback, also known as blind relevance feedback, automates the manual part of relevance feedback through local document analysis. The pseudo relevance feedback method is to perform normal retrieval to find an initial set of most relevant documents, assume that the top "k" ranked documents are relevant, and then perform relevance feedback techniques as before under this assumption. Evidence suggests that this method tends to work better than global document analysis (Xu and Croft 1996).

C.G. Harris (✉)
SUNY Oswego, Oswego, New York 13126, USA
e-mail: christopher.harris@oswego.edu

P. Srinivasan
Iowa City, Iowa 52242, USA

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_18,

interactive, and provide the look and feel of a game. Crowdsourcing has gained widespread attention as illustrated by recent conferences and workshops in the IR context (Alonso and Lease 2011; Lease and Yilmaz 2012). GWAP interfaces, while harder to implement in IR, have nonetheless obtained a fair amount of interest though to a lesser extent than crowdsourcing. These recent developments motivate our goal, which is to assess the use of human intelligence for both for initial query design and for query refinement in document retrieval. These methods provide us with the beginnings of a new approach for assisting searchers with query design (Harris and Srinivasan 2013). The use of human computation mechanisms in query formulation may be invoked when a query is difficult and the information need has longevity (e.g., in topic detection and tracking (Allan et al. 1998)) or where some latency in the returned results can be tolerated.

We study the value of using participants via crowdsourcing in our query design; this includes both initial query formulation and query refinement given some relevance feedback. We study this approach in game (GWAP) and non-game settings and compare performance with a machine algorithm baseline. We compare retrieval results obtained using these query design methods applied to a common set of topics and by running the resulting queries with the same retrieval algorithms against the same collection. We ask the following three research questions:

1. Does retrieval performance differ when the initial query is designed by humans versus the machine?
2. Does retrieval performance differ when feedback-based query refinement is done by humans versus the machine?
3. Does retrieval performance differ for humans using the non-game (a basic web interface) versus the game interface? (Note this question is asked both for initial query design and for query refinement with feedback).

The remainder of this paper is organized as follows. In the next section, we briefly discuss the background of our approaches. In section "Experimental Methods", we provide a description of our experimental methods. In section "Results", we provide our results. This is followed by a discussion of our general findings in section "Analysis and Discussion". We conclude and briefly discuss future directions of our work in section "Conclusion".

Background and Motivation

Crowdsourcing-Based Approaches

To date, most crowdsourcing studies in IR have examined relevance assessment. Several studies, such as (Alonso and Mizzaro 2012; McKibbon et al. 1990) have compared the crowd to experts in document assessment, concluding there is little difference in quality, particularly when multiple assessors are used. Few evaluations have been conducted to compare crowd-based and lab-based participants on search performance.

One study compared crowd and lab participants on multimedia search results in (Harris 2012), concluding that the two groups were indistinguishable in quality.

Integrating the crowd is becoming more commonplace for the difficult searches, perhaps indicating the crowd represents a nice tradeoff between speed, cost, and quality. A study by Yan et al. (2010) described a mobile search application in (Xu and Croft 1996); claiming a search precision of 95 %. Ageev et al. conducted an experiment to evaluate crowd search techniques in (Ageev et al. 2011). Harris and Srinivasan conduct a study to evaluate queries using crowdsourcing participants in (Harris and Srinivasan 2012). Each of these studies illustrate that the crowd can be used effectively to deliver search results with reasonable precision.

Game-Based Approaches

Only a few games with a purpose (GWAP) have been constructed to address initial query and query reformulation effectiveness. Search War (Law et al. 2009) is a game used to obtain data on search relevance and intent for a user-provided query. Players are paired and each given a unique search query and the objective of guessing their opponent's search query first. The design relies on the premise that players will select the least relevant webpage w.r.t. the search query, to provide to their opponent as hints, which implicitly provides a relevance judgment. Thumbs-up (Dasdan et al. 2009) is another GWAP that uses output-agreement mechanism to gather relevance data. This game asks players to evaluate search terms and attempt to independently determine the most relevant document to a given query. Another game, Koru (Milne et al. 2008), allows users to assess their search skills relative to other searchers and evaluate how their own searches might be improved; however, it is limited to a small document collection from a single source. The aforementioned Harris and Srinivasan study (Harris and Srinivasan 2012) evaluated query refinement in a news collection and found that the game format had higher average precision than the non-game version.

Machine-Based Approaches

A number of studies have examined interactive query expansion versus automatic query expansion and reformulation. Interactive query expansion and reformulation can be used as an effective means of improving a search. Efthimiadis (2000) found system-provided terms, on average, when selected, improved retrieval performance. Ruthven (2003) demonstrated that human searchers are less likely than machine-based systems to make good reformulation decisions. Anick (2003) found that users rarely used machine-suggested terms to expand and refine their queries, but when they did it improved retrieval performance. Thus, there are mixed performance results from machine-provided query reformulation and these approaches have not been adequately evaluated against human computation-based methods.

Experimental Methods

We evaluated performance on three treatments: two different *query types* (initial queries and queries refined based on feedback) and three different *approaches* (crowdsourcing using a game interface, crowdsourcing using a web (non-game) interface and machine).

Datasets

We randomly selected 10 topics used in the OHSUMED test collection used in the TREC-9 filtering task. The 10 topic numbers chosen were: 3, 4, 9, 13, 20, 28, 30, 36, and 41. These topics were presented to each user in the same order. We used the relevance judgments provided by OHSUMED assessors as our gold standard. Since the OHSUMED collection was assessed on a three-point relevance scale (0 = non-relevant, 1 = partially relevant, 2 = definitely relevant), we take the approach consistent with the assessors and consider "partially relevant" and "definitely relevant" documents as "relevant". The number of relevant documents per topic ranged from 12 (for topic 4) to 172 (for topic 30), with an average of 68.8 relevant documents per topic.

Query Design Approaches

Web Interface

Initial Query Formulation. Users were provided with the title and description for each of the 10 topics. Participants were given a large text box to input their query, with a pop-up help screen available to them throughout the task. We provided detailed instructions and examples of how to construct queries using terms and simple operators (AND, OR and NOT), and provided the following objective to participants: "The objective of this task is to construct queries that will bring back as many relevant documents as possible while excluding non-relevant documents". For example, the information provided and information need request for topic 4 is given as:

`Title:` `57year old male with hypercalcemia secondary to carcinoma.`
`Description:` `Effectiveness of gallium therapy for hypercalcemia.`
`Information Need:` `Find documents that describe the effectiveness of gallium therapy for hypercalcemia.`

Query Reformulation with Feedback. Once a user had provided initial input for each of the 10 topics, they were instructed to return after 2 hours to allow us time to run the provided queries against our document collection and to provide the recall and precision for each query for the second round. The user's original search terms were pre-loaded in the input text boxes for each topic, allowing easy modification to their original query. Also, in the second round, we provided users with the highest-ranked relevant and non-relevant document from the collection to aid them in their query refinement.

Game Interface

Some users invited to participate in this exercise were randomly selected to use a PHP-based game instead of the standard web interface.

Initial Query Formulation. Users selected to use game interface were given a different URL and were presented with the same initial screen outlining the game's objectives, instructions on term and operator rules as the web interface participants. Participants were asked to enter the initial query. The game instructions also had the following additions. First, there was a time-based constraint that required search terms to be entered within 45 seconds, with a point bonus awarded on a sliding scale based on how quickly the query was entered. Second, scoring was provided instantly (explained soon). Third, participants had musical sound effects to enhance the interface's game-like feel. Last, a leaderboard and badges, or icons, were awarded for superior game performance.

Query Reformulation with Feedback. Unlike the web non-game interface, the game interface did not provide users with precision and recall information from their initial round as they began their second round. This was because the calculation of this information was not integrated into the game interface and would take away from the feeling of engagement. Instead once a user entered a set of terms for a topic, these terms were parsed to remove stop-words, stemmed, and compared against a weighted list of stemmed terms obtained from documents judged relevant for that topic. A pop-up screen provided scoring and bonus information to each player after they submitted their query. A higher score was awarded for the use of relevant terms not commonly used by other participants. This score was immediately calculated and issued to the user, along with a time-based bonus for completing the search quickly. Once a user completed the first round, they could begin the query refinement round without delay. Users were instructed to refine their initial query based on their score and a relevant and non-relevant document provided to them to aid their refinement, subject to the same 45 seconds time restriction. Stars were awarded to users who scored above a certain threshold. Badges were given to users having the highest overall score, and a leaderboard was shown to the users, providing the option for top scorers to add their names for "bragging rights".

Algorithmic Baseline

Initial Query Formulation. The machine-based queries used the title and the description, as provided from the OHSUMED topics data in TREC-9 Filtering task (Hersh et al. 1994). Similar to the web and game interfaces, this input had stop-words removed using the same stop-word list and were stemmed using the Porter stemmer. These terms were transformed into a query written in the INQUERY query formulation language and run against an Indri created index. The ranked list returned by Indri was evaluated against our gold standard dataset.

Query Reformulation with Feedback. Using the ranked list returned by Indri (Strohman et al. 2005), we selected the highest-ranked relevant document from the results of the initial query. If no relevant documents were returned, we randomly selected a relevant document to use. We appended the terms contained within the title and Medical Subject Heading (MeSH) terms of the relevant document as additional inputs to the initial query, applied the stemming and stop-word list to the added terms. This became our refined query.

Participants

Crowdsourcing participants (N = 40) were recruited from Amazon Mechanical Turk (MTurk) participants and oDesk participants. MTurk participants were paid $0.20 to complete both rounds, whether they were assigned to the game or the non-game interface. We structured the task such that, to receive any compensation, these crowd participants would have to complete both rounds of initial query design and query refinement. We discarded the inputs for those participants who did not complete all 10 topics in both rounds.

Assigning Participants to Interfaces

Crowd participants were assigned randomly to either the web or the non-game interface or the game interface. Twenty-four of the MTurk participants failed to complete both rounds; those participants who did not complete both rounds and the two surveys had their inputs removed from our dataset and replaced by another participant. Participants were divided equally between game and non-game treatments.

Retrieval Algorithms

We used an Okapi retrieval algorithm (Robertson et al. 1995), which has been shown in a previous study on query formulation (Harris and Srinivasan 2012) to outperform a more commonly-used *tf.idf* approach (Jones 1972). The Okapi algorithm was implemented using the Indri (Strohman et al. 2005) system. We used parameter values k1 = 0.75, b = 0.75, and k3 = 7.

Results

The results from our study, comparing different human-based approaches and interfaces to the machine algorithm baseline, are summarized in Tables 1 and 2. We also conducted tests to examine each of our research questions, which are provided in Table 3. For each research question, we provide two-tailed t-tests at the $p<0.05$ level of significance for results.

Table 1 Overall results for the initial query comparing human computation approaches to the machine baseline

	Initial query		
Approach	P@10	MAP	Recall
Crowd – Non-game (N = 20)	0.25	0.161	0.170
Crowd – Game (N = 20)	0.22	0.161	0.171
Algorithm (Okapi)	0.07	0.108	0.123

Table 2 Overall results for the initial query comparing human computation approaches to the machine baseline

	Query reformulation w/feedback		
Approach	P@10	MAP	Recall
Crowd – Non-game (N = 20)	0.30	0.237	0.243
Crowd – Game (N = 20)	0.32	0.244	0.241
Algorithm (Okapi)	0.22	0.201	0.202

Table 3 Summary of findings for the research questions that examine (a) the initial query, and (b) query reformulation with feedback. Standard deviation is given in parentheses next to each mean value. An asterisk indicates it is statistically significant at $p < 0.05$

	Initial query		
Research question	P@10	MAP	Recall
RQ1: Machine (A) vs. Humans (B)	A: 0.070 (0.082) B: 0.235 (0.088) $p < 0.001^*$	A: 0.108 (0.059) B: 0.156 (0.031) $p = 0.017^*$	A: 0.123 (0.065) B: 0.171 (0.047) $p = 0.004^*$
RQ3: Non-game (A) vs. Game (B)	A: 0.250 (0.085) B: 0.220 (0.092) $p = 0.081$	A: 0.156 (0.029) B: 0.156 (0.035) $p = 0.962$	A: 0.170 (0.046) B: 0.171 (0.049) $p = 0.327$
	Query reformulation w/Feedback		
Research question	P@10	MAP	Recall
RQ2: Machine (A) vs. Humans (B)	A: 0.220 (0.132) B: 0.310 (0.079) $p = 0.014^*$	A: 0.201 (0.074) B: 0.241 (0.063) $p = 0.001^*$	A: 0.202 (0.064) B: 0.242 (0.086) $p < 0.001^*$
RQ3: Non-game (A) vs. Game (B)	A: 0.300 (0.113) B: 0.320 (0.097) $p = 0.343$	A: 0.237 (0.064) B: 0.244 (0.065) $p = 0.042^*$	A: 0.243 (0.094) B: 0.241 (0.083) $p = 0.759$

Our first two research questions compared human-based and machine approaches on mean average precision (MAP), precision across the top 10 documents (p@10) and recall for both the initial query formulation and the query refinement across all 10 topics (See Buckley and Voorhees 2004) for further discussion of these parameters. We found a significant difference for both initial query and query refinements between the crowd and machine approaches. Thus the crowd-supplied queries outperformed the machine algorithm queries.

For our third research question, we conducted a two-tailed t-test to compare the game and non-game interfaces for the crowd on MAP, p@10 and recall across all 10 topics. The only significant difference found was for MAP in query reformulation, with better performance provided by the game over the non-game interface.

Analysis and Discussion

Our findings on game vs. non-game performance partially supports the findings of a previous study by Harris and Srinivasan (Harris and Srinivasan 2012), which found that games did provide a better MAP, while non-games provided a better p@10. We did observe the better MAP by games, but not the better p@10 for non-game interfaces.

Crowdsourcing participants supplied fewer terms than machine approaches (4.2 terms vs. 6.1 terms for initial query; 7.1 terms vs. 32.3 terms for the query refinement). Game participants supplied fewer terms than non-game participants (3.7 terms for game vs. 4.7 terms for non-game in the initial query; 5.5 terms vs. 6.7 terms for the query refinement). Understandably, using the correct terms affected recall and precision more than simply supplying a larger number of terms.

Consistent with numerous earlier findings on Web searches, all of our treatments improved as a result of the query refinement with feedback. Given that the collection searched contained medical text documents, the opportunity for users to expand their queries through the use of synonyms or additional terms to describe the information need more accurately. The algorithm, with its access to the MeSH terms from a relevant document, improved the most between the initial query phase and the query refinement phase, indicating the power of using a taxonomical approach to document search. A post-hoc evaluation found that few crowd participants made use of this information.

Conclusion

We have illustrated how human computation mechanisms, including crowdsourcing and GWAP can be applied to document searches, a key area of IR. Although query design, term expansion strategies, methods for reformulating term weights etc., have been studied extensively, crowdsourcing and GWAP have motivated a new investigation of query design research. We conduct a study to evaluate different how these developments may impact precision in initial query construction and feedback-based query refinement. Using identical retrieval algorithms, this study examines how human-based query approaches compare with machine-based approaches on 10 OHSUMED topics, concluding that the human computation approach we evaluated provides better MAP, p@10 and recall compared a machine algorithm approach. We also evaluate these same three metrics to compare a web-based interface and a game interface, discovering that games provide a higher MAP score for reformulated queries. Experiments that apply human computation mechanisms to new domains are still relatively new and there is considerable room for novel human computation techniques to be applied to well-studied areas such as IR.

References

Ageev M, Guo Q, Lagun D, Agichtein E (2011) Find it if you can: a game for modeling different types of web search success using interaction data. In: Proceedings of SIGIR'11. ACM, New York, pp 345–354

Allan J, Papka R, Lavrenko V (1998) On-line new event detection and tracking. In: Proceedings of SIGIR'98. ACM, New York, pp 37–45

Alonso O, Lease M (2011) Crowdsourcing for information retrieval: principles, methods, and applications. In: Proceedings of SIGIR'11. ACM, New York, pp 1299–1300

Alonso O, Mizzaro S (2012) Using crowdsourcing for TREC relevance assessment. Inf Process Manage 48(6):1053–1066

Anick P (2003) Using terminological feedback for web search refinement: a log-based study. In: Proceedings of SIGIR'03. ACM, New York, pp 88–95

Buckley C, Voorhees EM (2004) Retrieval evaluation with incomplete information. In: Proceedings of SIGIR'04. ACM, New York, pp 25–32

Dasdan A, Drome C, Kolay S, Alpern M, Han A, Chi T, Hoover J, Davtchev I, Verma S (2009) Thumbs-Up: a game for playing to rank search results. In: Proceedings of the ACM SIGKDD workshop on human computation, Paris. ACM, New York, pp 36–37

Efthimiadis EN (2000) Interactive query expansion: a user-based evaluation in a relevance feedback environment. J Am Soc Inf Sci 51(11):989–1003

Harris CG (2012) An evaluation of search strategies for user-generated video content. In: Proceedings of the WWW Workshop on Crowdsourcing Web Search (Lyon, France), pp 48–53

Harris CG, Srinivasan P (2012) Applying human computation mechanisms to information retrieval. Proc Am Soc Inf Sci Technol 49(1):1–10

Harris CG, Srinivasan P (2013) Comparing crowd-based, game-based, and machine-based approaches in initial query and query refinement tasks. In: Advances in information retrieval. Springer, Berlin Heidelberg, pp 495–506

Hersh W, Buckley C, Leone T, Hickam D (1994) OHSUMED: an interactive retrieval evaluation and new large test collection for research. In: Proceedings of SIGIR'94. Springer, London, pp 192–201

Jones KS (1972) A statistical interpretation of term specificity and its application in retrieval. J Doc 28(1):11–21

Law E, Ahn L von, Mitchell T (2009) Search war: a game for improving web search. In: Proceedings of the ACM SIGKDD workshop on human computation, Paris. ACM, New York, pp 31–31

Lease M, Yilmaz E (2012) Crowdsourcing for information retrieval. ACM, New York, SIGIR Forum 45:2 (January 2012), pp 66–75

McKibbon KA, Haynes RB, Walker Dilks CJ, Ramsden MF, Ryan NC, Baker L, Flemming T, Fitzgerald D (1990) How good are clinical MEDLINE searches? A comparative study of clinical end-user and librarian searches. Comput Biomed Res 23(6):583–593

Milne D, Nichols DM, Witten IH (2008) A competitive environment for exploratory query expansion. In: Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries (JCDL'08). ACM, New York, pp 197–200

Robertson SE, Walker S, Jones S, Hancock-Beaulieu MM, Gatford M (1995) Okapi at TREC-3. NIST Special Publication SP-1995. Gaithersburg, Maryland, USA, pp 109–121

Ruthven I (2003) Re-examining the potential effectiveness of interactive query expansion. In: Proceedings of SIGIR'03. ACM, New York, pp 213–220

Spink A, Jansen BJ, Wolfram D, Saracevic T (2002) From e-sex to e-commerce: web search changes. Computer 35(3):107–109

Strohman T, Metzler D, Turtle H, Croft WB (2005) Indri: a language model-based search engine for complex queries. In: Proceedings of the international conference on intelligence analysis, McLean, VA. Poster, 2–6 May 2005

Xu J, Croft WB (1996) Query expansion using local and global document analysis. In: Proceedings of SIGIR'96. ACM, New York, pp 4–11

Yan T, Kumar V, Ganesan D (2010) CrowdSearch: exploiting crowds for accurate real-time image search on mobile phones. In: Proceedings of MobiSys'10. ACM, New York, pp 77–90

Human Computation-Enabled Network Analysis for a Systemic Credit Risk Rating

François Bry

Introduction

This chapter proposes a novel approach to credit risk rating on financial markets based upon Network Analysis and Human Computation and consisting in a dual-purpose participatory mechanism (Quinn and Bederson 2009).

Credit risk rating is an important activity for participants in financial markets which has become difficult with the advent of financial contracts called derivatives and structured notes and of credit risk management techniques called securitization. A wide-spread improper rating of credit risk, especially of the risk associated with derivative and securitization instruments, has been recognized as a major cause of the financial crisis of 2007–2009 (Soros 2008; Caccioli et al. 2009; Sarkar 2009; Gregory 2010; Simkovic 2009, 2010; National Commission on the Causes of the Financial and Economic Crisis in the United States 2011; Haldane and May 2011; Hull 2012; Simkovic 2013; Arora et al. 2012) which sparked a great recession, the European Sovereign-Debt Crisis (Haidar 2012) and recessions which, after half a decade, are still going on in many countries.

An improper credit risk rating could wide-spread because derivatives, structured notes, and securitization challenge the methods used in current credit risk rating. The disregard of counter-party risk, which is absent in conventional contracts but inherent to derivatives, undoubtedly played a role in the financial crisis of 2007–2009 (Gregory 2010) (but has been largely irrelevant to the Subprime Mortgage Crisis which led to that crisis (Hull 2012)). Therefore, deficiencies of current credit risk rating methods, or of the current credit risk rating practice, can be seen as core reasons for the improper credit risk rating which has been a major cause of the financial crisis of 2007–2009.

F. Bry (✉)
Institute for Informatics, Ludwig-Maximilian University of Munich,
Oettingenstr. 67, 80538 München, Germany
e-mail: bry@lmu.de

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_19,

The approach to credit risk rating proposed in this chapter radically departs from current credit risk rating in four aspects. First, it collects credit risks assessments from the debtors and not, as usual, from the creditors. Second, it propagates debtors' risk estimates through the risk dependency graph induced by credit contracts, derivative contracts, and currencies by aggregating as eigenvector centralities the agents' contributions in the global market's risk. Third, it is not based upon stochastic methods and statistical data. As a consequence, it keeps its relevance in exceptional situations such as rare crises or bubbles. Fourth, its principle promises much earlier warnings of an increasing credit risk than possible with current credit risk rating methods.

Since it combines human computed credit risk assessments and a machine computed eigenvector in which these human inputs are aggregated, the proposed method is a Human Computation algorithm (Law and von Ahn 2011, Chap. 2, p. 15). An essential part of this Human Computation algorithm is an incentive, the "Grace Period Reward" (GPR), for an actual or potential debtor to compute, constantly actualize, and disclose to a system running the proposed Human Computation algorithm estimates of the risk that, in the future, she will fail to honor her debts. The approach to credit risk rating proposed in this chapter is a dual-purpose system (Quinn and Bederson 2009): On the one hand, the GPR gives actual or potential debtors a reason to assess and to disclose to the system the risk of their own defaulting; on the other hand the system provides the market participants with a systemic credit risk rating. Since the reason for an agent to contribute to the Human Computation system, namely her use of the GPR, is not the primary purpose of the system, one can call it a passive Human Computation system.[1]

The novel credit risk rating proposed in this chapter is systemic because of its global assessment of credit risk by Network Analysis as eigenvector centralities. This distinguishes it from current credit risk rating performed locally by financial agents for themselves, or by credit rating agencies (such as Standard & Poor's, Moody's Investor Service, and Fitch Ratings) for financial agents, and which do not at all, or only to a very limited extent, propagate credit risk estimates between agents bound by financial contracts.

Implementing the approach proposed would require and induce changes on financial markets that are briefly discussed in this chapter.

Human Computation systems (whether they are Crowdsourcing marketplaces such as Amazon Turk, online job marketplaces such as oDesk, prediction markets (Pennock et al. 2001; Servan-Schreiber et al. 2004; Gjerstad 2005; Wolfers and Zit zewitz 2006; Hubbard 2007; Snowberg et al. 2007; Berg et al. 2008; Arrow et al. 2008), decision markets (Leutenmayr and Bry 2011), or games with a purpose (von Ahn 2006)) on the one hand and markets on the other hand have much in common. These commonalities are finally investigated. This chapter argues that markets can be seen as Human Computation systems *avant la lettre*. This chapter also argues that, as markets become global and transactions get faster, markets' good

[1] This denomination has been suggested by Pietro Michelucci.

functioning will require Human Computation-enabled network analyses of the kind proposed in this chapter for financial markets.

This chapter is based upon the research report (Bry 2012) which it extends.

The contributions of this chapter are as follows:

- A Human Computation algorithm for a systemic credit risk rating.
- A discussion of the practicability of this Human Computation algorithm and of implications of its deployment.
- The thesis that the good functioning of many markets and Human Computation systems will, in the future, benefit from Human Computation-enabled network analyses of the kind proposed in this chapter.

This chapter is structured as follows. Section "Introduction" is this introduction. Section "Credit Risk Rating Challenged by Derivatives, Structured Notes, and Securitization" briefly introduces into credit risk rating, derivatives, structured notes, and securitization explaining why these financial instruments and techniques challenge current credit risk rating. Section "Human Computation: Potential Debtors Assess Their Own Risk of Defaulting" proposes an incentive, the "Grace Period Reward" or GPR, for debtors to determine, constantly actualize, and disclose estimates of the risk that they fail to honor their debts. Section "Network Analysis: Aggregating Human Estimates into a Systemic Credit Risk Rating" proposes a systemic credit risk rating as eigenvector centralities. Section "Discussion" discusses the practicability of the Human Computation algorithm proposed and a few perspectives its deployment on financial markets would open. Section "Human Computation and Markets" compares Human Computation systems and markets. Section "Conclusion" is a conclusion.

Credit Risk Rating Challenged by Derivatives, Structured Notes, and Securitization

Importance of Credit Risk Rating

Credit risk is the name given to the risk that a financial agent (like a bank) will not recover the money it is owed according to financial contracts (like mortgages).

Credit risk is essential on financial markets for several reasons. First the values of financial assets depend on the risks associated with these assets. Second, taking too much risk can lead to bankruptcy. Third, financial institutions are expected to reduce credit risk, that is, to convey to their creditors less credit risk than they themselves face (Bhattacharya et al. 1998). For this reason, depository financial institutions (like banks) have to enforce risk-based capital guidelines or "capital requirements" (such as those issued by the Board of Governors of the U.S. Federal Reserve System and the Basel Accords commonly referred to as Basel I, BaseI II and Basel III) that specify a "capital adequacy" ensuring that a depositary financial institution holds enough capital to both, sustain possible losses and honor withdrawals (Hull 2010). The minimum of capital required by risk-based capital guidelines is called "regulatory capital" (Bhattacharya et al. 1998; Hull 2010).

Derivatives, structured notes, securitization and a considerable speed differential between financial transactions and credit risk rating, as it is currently performed, have been challenging credit risk rating since at least four decades (Buffet 2002). This challenge, which so far has not been met, is one of the acknowledged causes of the financial crisis of 2007–2009 (Soros 2008; Caccioli et al. 2009; Sarkar 2009; Gregory 2010; Simkovic 2009; National Commission on the Causes of the Financial and Economic Crisis in the United States 2011; Simkovic 2010; Haldane and May 2011; Hull 2012; Simkovic 2013; Arora et al. 2012) which sparked worldwide recessions. In Simkovic (2010, p. 1), Michael Simkovic stresses as follows the role of an improper credit risk rating in igniting the financial crisis of 2007–2009:

> "One of the most important contributors to the financial crisis of 2008 was the proliferation of opaque and complex financial instruments that effectively withheld key information from market participants. Without detailed, reliable information about debtors' off-balance-sheet debts and conditional liabilities such as derivatives exposures, creditors cannot accurately evaluate the creditworthiness of debtors and the markets cannot appropriately price risk."

Current Credit Risk Rating

On financial markets, not only credit contracts of various kinds (loans, mortgages, bills, and bonds) are traded with but also derivative contracts of many kinds (futures, forwards, options, warrants, swaps among others credit default swaps (CDS) and contingent credit default swaps (CCDS), structured notes and securitization instruments) (Chance 2008; Hull 2012).

Technically, with a derivative there are no creditors and no debtors because, when the derivative contract is entered and during most of the contract's lifetime the direction of money flows between contract parties is left unspecified.This is a fundamental difference between credit and derivative contracts: A credit contract fully specifies the flows of money between the contract parties, a derivative doesn't.

The payments specified in derivatives, like those specified in credits, may not be honored. Thus, while with a credit only the creditor assumes a risk, with a derivative both parties in the derivative assume a risk (Chance 2008; Duffie and Singleton 2003; Gregory 2010; Hull 2012; Arora et al. 2012). The risk induced by both, derivatives and credits, is called credit risk. The credit risk associated with derivatives is often called counterparty risk (Gregory 2010; Arora et al. 2012) reflecting that both parties in a derivative assume a risk. Counterparty risk as well as the credit risk assumed by holders of structured notes and securitized instruments is difficult to assess and, so far, is often improperly assessed (Chance 2008; Gregory 2010; Arora et al. 2012; Hull 2012).

Abusing the terms, we shall call debtor (creditor, respectively) a party in a credit or derivative contract which has, or may have, to perform (receive, respectively) a payment.

We shall use the phrase "actual or potential debtor" for stressing this abuse of terminology.

The credit risk faced by an agent *i* is measured as a weighted sum (Treacy and Carey 1998; Boucheaud and Potters 2000, 2003; Duffie and Singleton 2003; Gregory 2010; Arora et al. 2012): It is the sum over all debtors of *i* of the likelihood that this debtor fails to serve and/or reimburse her debt weighted by the loss this failure would entail for *i*. If, for example, an agent *i* is creditor of three agents *a*, *b*, and *c* for the following sums *a*: 10$, *b*, 20$, and *c*: 60$ and if the risks that these agents will fail to reimburse their debts to *i* are *a*: 25 %, *b*: 5 %, and *c*: 90 %, then the credit risk faced by *i* is $(10 \times 25\%) + (20 \times 5\%) + (60 \times 90\%)$.

Creditor and parties in derivatives use sophisticated stochastic models and statistical methods for assessing the credit risk induced by credits and derivatives (Treacy and Carey 1998; Altman and Saunders 2009; Boucheaud and Potters 2000, 2003; Metz and Cantor 2006; Board of Governors 2007; Lando 2009; Gregory 2010; Kothari 2012; Arora et al. 2012; Kothari 2012). This is called credit risk rating or credit risk assessment. Some of these methods are codified in national and international regulations such as Basel I, II and III. In spite of a large number of models, mathematical methods, procedures and regulations, credit risk rating remains awkward and is far from being reliable (Jarrow and Turnbull 1995; Beaver et al. 2006; Gregory 2010; Haldane and May 2011; National Commission on the Causes of the Financial and Economic Crisis in the United States 2011; Arora et al. 2012).

Derivatives, Structured Notes, and Securitization

This subsection is a brief introduction into derivatives, structured notes, and securitization. See Chance (2008) and Hull (2012) for detailed presentations. This section can be skipped by readers familiar with derivatives, structured notes, securitization, and current credit risk rating and who are aware of the limitations, and criticisms, of current credit risk rating.

Motivating Example

The following example may help to understand derivatives and, indirectly, structured notes and securitization that, though different from derivatives, are used for similar reasons.

Assume that a family lets a small apartment in Vienna and that a child of this family goes to study to Heidelberg. Finding affordable accommodations at predictable costs is a major challenge for students in Europe in general and in Heidelberg in particular. The family could enter a contract over the duration of its child's studies granting the owner of an apartment in Heidelberg the Vienna appartment's rent for the use of her apartment. With such a contract, the family would make a loss if housing rents raise more in Vienna than in Heidelberg and a gain if the rent differential evolves in the opposite direction. However, with such a contract, the family does not need to concern itself any longer with its child's accommodation and entering such a contract is much cheaper, especially as taxes are concerned, than selling the Vienna apartment for buying an apartment in Heidelberg. For the owner of the Heidelberg apartment, the contract may have advantages as well like low-cost, especially low-taxes, income diversification, and securing a tenant for several years. A derivative is, basically, such a contract. Reasons for parties to enter derivatives, structured note, and securitized instruments are, basically, like in this example.

Derivatives

A derivative contract, short derivative, is a contract between two parties whose value derives from the value of an underlying asset, reference rate, or index. A derivative serves to transfer at low costs the risk associated with its underlying financial instrument or asset from one party to another. Since the end of the 1970s of the twentieth century, the use of derivatives has grown considerably. Economics' "law of comparative advantages" (Ricardo 1817; Jones 1961), that is, the ability of an agent to produce a good or service at a lower marginal and opportunity cost than another, explain why a transfer of risk between parties may make sense.

There are different types of derivatives: futures, forwards, options (among others swaptions), warrants, swaps (among others credit default swaps (CDS) and contingent credit default swaps (CCDS)).

A *future* is contract to buy or sell an asset on, or before, a future date at a price specified at contract entering time. Futures have no entering costs, are exchange-traded and standardized. Futures are written (that is, guaranteed) by a clearing house: The clearing house becomes the buyer to a future's seller, and the seller to a future's buyer, so that if a party defaults, then the clearing house assumes the loss. To reduce the credit risk incurred by the clearing house, each party in a future must post a margin (that is, provide an initial amount of cash or a performance bond), usually 5–15 % of the future's price. The margin is adjusted daily in a process called "marking to market".

Forwards are like a future except that they are not traded on an exchange (they are "off-exchange" or "traded over-the-counter (OTC)"), they induce no interim payments (they require no "marking to market"), and they are not standardized.

An *option* is a contract giving its owner the right, but not the obligation, to buy or sell an asset (commonly a stock, a bond, a currency or a future) at some future time. Options can be both, "exchange-traded" or "traded over-the-counter". Exchange-traded options are, like futures, standardized. A *swaption* or *swapion* is an option on a swap –see below.

A *warrant* is a long-dated option, that is, a contract similar to an option but having a maturity period of more than 1 year. Warrants are mostly, but not only, "traded over-the-counter" and not standardized.

A *swap* is a contract to exchange over a period of time, usually up to 15 years, the cash flows of one party's financial instrument for those of the other party's financial instrument. Most swaps are "traded over-the-counter" and not standardized.

With *credit default swaps (CDS)* and *contingent credit default swaps (CDS)* the exchange of cash flows depend on "credit events" (such as capital restructuring, bankruptcy, if an agent's credit rating is downgraded) independent of the two financial instruments the swap is based upon. A CDS is comparable to an insurance because in return for a premium the buyer receives from the seller a sum of money if one of the credit events specified in the contract occur. Unlike an insurance, however, a CDS may, and usually does, cover an asset not owned by its buyer. Not being called "insurances", CDSs escape the (state and federal) regulations insurances are subject to in the USA. A *contingent credit default swap (CCDS)* is like a CDS except that the notional amount of protection is also referenced to an additional "credit event", usually a change in a market or another variable. Thus, the credit risk induced by a CDS to each of its parties depends upon a third party, the party responsible of the credit event the CDS refers to. The credit risk induced by a CCDS to each of its parties depends in addition on a further party, the party which is responsible of the contingent credit event the CCDS refers to.

Finally, derivatives may be squared, that is, a derivative may be derived from … a derivative.

Except futures, that are guaranteed by a clearing house, all derivatives induce a counterparty risk (that is, a risk for both parties in a derivative) making their credit risk rating more complex than that of credits. Rating the credit risk of CDSs and CCDSs is especially challenging because of CDSs' and CCDSs' "credit events" referring to assets usually not owned by a party in the CDSs or CCDSs.

Derivatives that are guaranteed by a clearing house or exchange-traded are usually standardized, other derivatives are usually not standardized. The reason is that standardization makes possible current credit risk rating, which is based on statistics. The need for standardization, which restricts derivatives, is often mentioned against proposals to regulate the derivative market by requiring all derivatives to be guaranteed by clearing houses and/or to be traded on exchanges.

The approach to credit risk rating proposed in the following requires an institution keeping track of, or "list", trades in credits, derivatives and securitization instruments. It does not require, however, credits, derivatives or securitization instruments to be standardized.

Structured Notes

Structured notes are debt securities (like mortgages, government and corporate bonds) and therefore no derivatives. Like derivatives, however, the interest on a structured note depends on another security, or on price moves, or on a rate (like the London Interbank Offered Rate known as LIBOR). The formula specifying this dependency may be complex.

Thus, like a derivative, a structured note induces a credit risk for both of its parties which depend on the structured note's security of reference. The credit risk assumed by a structured note's creditor, depends on the note's security of reference as well as on the debtor of the structured note. The formula specifying how a structured note's interest refers to its security of reference complicates, often significantly, the rating of the credit risk incurred by the structured notes' parties (Chance 2008).

Securitization

Securitization consists in building portfolios of debt securities (like mortgages and government or corporate bonds) called securitized instruments or securitized assets and in issuing new securities with claims on the portfolio called "tranches". The payments of interest and principal by the debtors in the debt securities underlying a securitized instrument are allocated to the tranches. The tranches are served by decreasing seniority, the tranche with smallest seniority, called equity tranche, receiving what remains. Thus, for investors, with decreasing tranches' seniorities the credit risk increases.

Securitized instruments are often built from debt securities like home or loan mortgages for which prepayments are possible. As a consequence, the credit risk of all tranches, including the most senior one, of a securitized instrument also depend on the prepayment risk, that is, the risk that debtors in the debt securities underlying the securitized instrument prepay all or part of their debts prior to their debts' maturity. Prepayments happen when interest rates on the credit market fall sufficiently what makes the credit risk assumed by the holders of some securitized instruments dependent on the interest rates. Prepayment risk is often underestimated, or even ignored, by investor estimating the credit risk of securitized instruments (Hull 2012). Since a contractual prepayment is no defaulting, technically, prepayment risk is no credit risk. However, the aim of credit risk rating –that is, assessing the likelihood that contractual flows of payments may stop in the future– makes it appropriate to consider prepayment risk as credit risk.

There are several types of securitized instruments: Mortgage-backed securities (MBS) among other Agency MBS, collateralized mortgage obligations (CMO), and collateralized debt obligations (CDO). How and when their tranches are served is specified in the instrument's contract which may be several hundred pages long and quite complicated what, in turn, may make credit risk rating difficult.

Agency MBS are MBS, the principal and interest of their underlying mortgages are guaranteed by US government entities or government-sponsored enterprises (like the Government National Mortgage Association, GNMA, also known as Ginnie Mae, the Federal Home Loan Mortgage Corporation, FHLMC also known as known as Freddie Mac, and the home Loans Banks). Holders of Agency MBS nonetheless assume a risk because of the afore-mentioned prepayment risk. In the past, most holders of Agency MBS have ignored this risk. This was one of the causes of the Subprime Mortgage Crisis (Hull 2012).

With a CMO, the various tranches assume different prepayment risk and other credit risk. With some simple CMOs, some tranches receive only interest and therefore assume only prepayment risk, while other tranches receive only principal payments and therefore only assume the credit risk of the debt securities underlying the CMO. Thus, the credit risk assumed by holders of CMOs depends on the tranches and in turn on the CMO contract which may be complicated. Furthermore, CMOs are issued as follows by financial entities, the financial health of which is in general difficult to assess. A financial institution creates a legal entity called in the USA special purpose entities (SPE), outside the USA special purpose vehicle (SPV), and transfer mortgages, the "collateral", to this SPE which use them for issuing mortgage-backed securities. SPE isolate the financial institution which create them from the risk of the CMOs the SPE has been created to issue.

Securitized instruments are also built from debt securities such as commercial mortgages, car loans and credit card debt obligations. Such securitzed instruments are called collateralized debt obligations (CDO). The credit risk induced by CDO depend on many debtors, and therefore on many economical variables, what makes it difficult to assess.

Finally, securitized instruments can, like derivatives, be squared, that is, securitized instruments can be built from … securitized instruments. With such constructions, the credit risk even of the most senior tranches of a squared securitized instrument can increase considerably and the credit risk can become, even for large financial institutions, extremely difficult to assess. One of the acknowledged causes of the Subprime Mortgage Crisis is that credit rating agencies and investors have under-estimated the credit risk induced by such constructions (National Commission on the Causes of the Financial and Economic Crisis in the United States 2011; Simkovic 2010; Hull 2012).

Limitations of Current Credit Risk Rating

Credit risk rating, as it is currently performed, has been criticized for empirical and methodological reasons.

As of empirical criticisms, it is acknowledged that an inaccurate assessment of credit risk has been instrumental in the Subprime Mortgage Crisis and the Financial Crisis of 2007–2009 (Soros 2008; Caccioli et al. 2009; Gregory 2010; Simkovic 2009; National Commission on the Causes of the Financial and Economic Crisis in the United States 2011; Haldane and May 2011; Hull 2012; Simkovic 2013) and that the wide-spread disregard, or under-estimation, of the credit risk induced by derivatives, structured notes, and securitization has been one of the major causes of the Financial Crisis of 2007–2009 (National Commission on the Causes of the Financial and Economic Crisis in the United States 2011; Gregory 2010; Simkovic 2013). Assessing the credit risk derivatives induce is considered rather complex (Boucheaud and Potters 2000, 2003; Buffet 2002; Gregory 2010; Simkovic 2009; Caccioli et al. 2009; Haldane and May 2011).

A further empirical criticism of current credit risk rating is that it mostly fails when applied to securitization instruments. Securitization limits an investor's ability to assess the risk associated with mortgage-backed securities and CMOs. An improper credit risk rating of securitization instruments is seen as a cause of the US Subprime Mortgage Crisis (Hull 2012, Chap. 8) which sparked the Financial Crisis of 2007–2009. Off balance sheet securitization, which is based on a transfer of unqualified risk, is believed to have played a significant role in the high leverage level of US financial institutions before the financial crisis, and the need for bailouts after the outbreak of the financial crisis of 2007–2009 (Simkovic 2013). No credit rating agencies, for example, downgraded the investment bank Bear Stearn, which had issued large amounts of asset-backed securities, before its collapse in 2008.

As of methodological criticisms of current credit risk rating, some, prominently Benoit Mandelbrot and Nassim Nicholas Taleb, have argued that, since current credit risk rating is based on stochastic methods and statistical data, it is inherently inaccurate in exceptional situations such as market crises and bubbles (Mandelbrot and Hudson 2004; Malevergne and Sornette 2005; Taleb 2007, 2010).

Current credit risk rating is necessarily inaccurate during market bubbles because current credit risk rating is performed by the parties assuming the risk, not by those causing it, and because bubbles, from the Tulip Mania Bubble in seventeenth century Holland to the Dot-com Bubble in twenty-first century USA, always result from a loss of sense of assumed risk (Blanchard and Watson 1982): As a bubble booms, that is, some prices keep raising more and over longer periods of time than usual, more and more traders get seduced by the perspective of unexpected gains, loose their sense of risk and join in the frenzy, buying because they expect to later sell at higher prices, thus contributing to keep the price raising up until enough traders come to reason, what causes the bust.

A further wide-spread methodological criticism of current credit risk rating concerns biases. As mentioned in Treacy and Carey (1998, p. 921) biased views, whatever their causes, often result in an inaccurate credit risk rating.

A further wide-spread methodological criticism of current credit risk rating is that, being performed mostly by banks to display evidence of their financial health and by credit rating agencies on behalf, and often at the expenses, of debtors that need good ratings for being granted credits at good conditions, current credit risk rating is not free from moral hazard. The charts of Treacy and Carey (1998, p. 917) for example report on much more optimistic credit risk ratings at banks than at credit rating agencies long before the Subprime Mortgage Crisis and the Financial Crisis of 2007–2008. The article (Božović et al. 2011) cautiously states that

> "the way the current rating market is organized may provide [rating] agencies with intrinsic disincentives to accurately report credit risk of securities they rate."

A further problem with current credit risk rating, which, admitedly, is rarely mentioned, is the considerable speed differential between financial transactions and credit risk rating. Algorithmic trading (Joyce 2008; Gomolka 2011), in particular "high-frequency trading", automatically reacts to index variations in fractions of seconds, much faster than humans can react to observations they make. In contrast, credit risk rating is computed by humans working mostly in committees delivering their updates at best weekly (for example for the home mortgages of a region), usually every couple of weeks, at worst every quarter of a year (for example for government bonds) (Treacy and Carey 1998; Altman and Saunders 2009; Metz and Cantor 2006; Board of Governors 2007).

The Human Computation-enabled network analysis for credit risk rating described in the following addresses the afore mentioned limitations of current credit risk rating: It is affected neither by the nature nor by the complexity of financial instruments traded with on a market, it is not based on stochastic methods and statistical data what makes it reliable also in exceptional situations, the rating it delivers is neither impaired by investors loosing their sense of risk, nor by financial institutions eager to demonstrate a good financial health, and it significantly reduces the speed differential between transactions and credit risk rating. And, importantly, it does not require a standardization of financial instruments.

Human Computation: Potential Debtors Assess Their Own Risk of Defaulting

We propose to collect from the market's agents assessments of the risk that, in the future, they fail to honor their debts. Collecting such assessments from actual or potential debtors has three advantages:

- It provides earlier estimates than current credit risk rating performed by creditors or credit rating agencies. Indeed, debtors suspect their possible defaulting earlier, usually much earlier, than their creditors.
- It complements current credit risk rating performed by creditors and credit rating agencies.
- It is not subject to the moral hazard of current credit risk rating (mentioned above in section "Current Credit Risk Rating").

The Incentive: The Grace Period Reward (GPR)

For an agent facing its possible defaulting, time is extremely precious. Time makes it possible to recover outstanding debts or to take a credit and thus, in some cases, to prevent one's defaulting and, possibly, bankruptcy. We exploit this in devising an incentive, the "Grace Period Reward" (GPR), for an actual or potential debtor to compute, constantly actualize, and disclose its own estimates of the risk of defaulting.

The GPR functions like a credit default insurance but, importantly, only for a limited period of time of a few weeks to a few months, the "grace period" and at costs that are the same for all agents on the financial market. As a consequence, the GPR is not a credit default insurance and does not yield moral hazard as do credit default insurances.

The GPR can be activated by (actual or potential) debtors at any time t so as to begin at any future time $t_1 \geq t$ and for any coverage (that is, percentage) of actual or possible outstanding payments. Once activated by an agent for an actual or potential debt, the GPR can be deactivated at any time by this agent.

An activation by an agent of the GPR at time t beginning at a later time t_1 for $x\%$ of an actual and possible debt expresses the opinion at time t of this agent that, at time t_1 or later, it may default to pay the principal or the interest of this debt. Increasing (decreasing, respectively) values of x reflect an increase (a decrease, respectively) of the agent's belief in its own defaulting. An activated GPR comes at a cost for the agents, what incites them only to activate the GPR when they see a need. The costs of an activated GPR are proportional to both, the outstanding payments and the activation duration, making GPR activations reliable estimates of how likely debtors hold their own defaulting.

The costs of an activated GPR are covered from a compulsory GPR deposit to be made by actual or potential debtors when entering a credit or derivative contract. The hight of this compulsory deposit depends on the credit or derivative contract. The GPR deposit is lost (to the creditor or the agency running the GPR) by the debtor if it defaults while the GPR is not activated and otherwise refunded at the end of the credit or derivative contract up to the costs resulting from, possibly temporary, activations of the GPR. The possible loss of the GPR deposit incites debtors to activate the GPR accordingly to the risk of defaulting they perceive.

Furthermore, it would make sense not to grant the GPR's grace period, or to grant it only to a limited extent, to defaulting agents that have activated the GPR much later than when they acquired knowledge of events motivating their activating the GPR.

Whether a debtor activates the GPR or not is not disclosed within the agent's community. This ensures that no moral hazard impairs the risk assessments deduced from GPR activations.

Finally, the GPR could come at a low cost so as to cover its management costs as well as the costs of the network analysis described in the next section.

Calibrating the GPR

The GPR requires calibration. The costs of an activated GPR must be set according to insurances' good practices, the duration of the grace period must be defined (most likely depending on the type of agents), the types of credits, and types of derivatives, and the value of the GPR deposit must be appropriately set (most likely depending on the types of agents and contracts), etc.

Part of the GPR's calibration might consist in "socio-cultural adjustments" of the following kind: If a social and/or economical group of agents is known to overestimate (or underestimate) their own credit risk, than this could be accounted for with adjustment factors reducing (or enhancing) the credit risk estimates they express through the GPR. Furthermore, such socio-cultural adjustments could be democratically agreed upon in the community of all agents, possibly using ad hoc Human Computation systems.

Calibrating the GPR requires further investigations and is out of the scope of this chapter.

Assessing Prepayment Risk

As mentioned in section "Securitization", it is appropriate to consider prepayment risk as credit risk. The question therefore arises, whether the GPR could contribute to an early assessment of prepayment risk.

This seems to be the case. The GPR would incite debtors to inform early of possible prepayment if, while activating the GPR, a debtor could limit the activation duration to the date of an expected prepayment and if the GPR costs would be reduced by early activations.

Like calibrating the GPR, tuning the GPR towards assessing prepayment risk requires further investigations and is out of the scope of this chapter.

Social Control

The GPR gives room to social control. If some agents, say some banks, feel that other agents, say home mortgage debtors, over-estimate, or under-estimate the likelihood of their defaulting, then the first agents can trigger a debate on the issue what, eventually, can lead to the other agents changing their assessments of their risk of defaulting.

The GPR and Traditional Credit Risk Rating

The GPR complements traditional credit risk rating. It neither replaces it nor conflicts with it. Indeed, in deciding whether or not to activate the GPR, agents are well advised to make use of all information and all risk rating methods at their disposal.

The estimates the GPR would collect differ from those obtained with current credit risk rating in several essential aspects. First, the GPR returns estimates by actual or potential debtors of the likelihood of their own defaulting while current risk rating is performed by the actual or potential creditors. Arguably, estimates collected by the GPR from debtors are less biased than current credit risk rating. Second, the estimates collected with the GPR can be expected to be updated at least daily and, in case of algorithmic trading, much more often. Indeed, algorithmic trading calls for algorithmic GPR activations. Third, in contrast to traditional credit risk rating, the GPR promises estimates differentiated after different time points in the future. Fourth, estimates collected by the GPR are not based on stochastic methods and statistical data. This makes GPR-based estimates reliable in crisis times (such as bubbles) and when new financial instruments are introduced for which no statistics are available.

The GPR and Moral Hazard

Since the grace period is limited to a short period of time, the GPR is not subject to the moral hazard of a credit default insurance which may induce debtors to take risks that, without insurance, they would not take.

The GPR as a Transaction Tax

The GPR can be expected to act like a financial transaction tax, or Tobin tax, the objective of which is to hinder short-term speculative financial "round-trip transactions" (Tobin 1978). Indeed, adequately activating the GPR requires human work and come at a cost while not activating it may result in losses. Algorithmic GPR activations would not reduce the value of the credit risk rating based on the GPR activations because of both, the costs of activating the GPR and the time-limited safety the GPR provides to both, the GPR activators and the market as a whole.

A fundamental difference, however, is that in contrast to a financial transaction tax, the GPR collects information which, as described in the next section, is used for a systemic credit risk rating. Thus, the GPR can be called an "informationally productive" financial transaction tax, while the standard financial transaction tax, or Tobin tax, how effective it might be, can be seen as "informationally unproductive".

Network Analysis: Aggregating Human Estimates into a Systemic Credit Risk Rating

This section describes how the estimates collected by the GPR –see above section "Human Computation: Potential Debtors Assess Their Own Risk of Defaulting"– can be aggregated as eigenvectors expressing a systemic credit risk rating.

Eigenvectors are solutions of systems of linear equations. They are commonly used for expressing the stability of physical systems and the relative importance, so-called centralities, of the nodes of a network (Bonacich 2001). It is this second usage which is relevant here.

Formalizing an Agent's Credit Risk

Using the estimates of defaulting likelihood provided by the GPR, the credit risk $\mathrm{CR}^t(i)$ of an agent i at any future time t can be estimated as follows:

$$\mathrm{CR}^t(i) = \sum_j (w^t_{ji} \times c^t_{ji}) \tag{1}$$

where:

- $w^t_{ji} \in [0,1]$ is the estimate collected by the GPR of the likelihood that agent j defaults to agent i at time t
- c^t_{ji} is the payment surely or possibly due by agent j to agent i at time t

Variations of $\mathrm{CR}^t(i)$, or of a conveniently chosen aggregation $\Phi_{i\in G}\mathrm{CR}^t(i)$ for a group G of agents like, for example, the home mortgage debtors of a region, are useful indicators of agent i's, or of the group's, financial health. A constant increase of $\mathrm{CR}^t(i)$, or of $\Phi_{i\in G}\mathrm{CR}^t(i)$, over a period of time can help in restoring early enough agent i's, or group G's, credit strength so as to prevent cascading defaulting.

If agent i has a large number of debtors, for example, if i is a sufficiently large financial institution, then an appropriate aggregation $\Phi_{i\in D(i)}\mathrm{CR}^t(i)$ over the set $D(i)$ of (actual or potential) debtors of agent i can be disclosed to this agent without disclosing the estimates w^t_{ji} what, as discussed in section "Human Computation: Potential Debtors Assess Their Own Risk of Defaulting" would compromise the good working of the GPR as a collector of reliable estimates of defaulting likelihood. The variations over time of conveniently chosen aggregates would be valuable credit risk indicators for a financial institution that would complement its own credit risk rating.

Credit Risk Flow

Considering how credit risk flows from agent to agent on a financial market suggests a systemic credit risk rating, that is, a rating reflecting the flows of credit risk on the financial market. Credit risk flows are first discussed.

A first observation is that nowadays on financial markets creditors are also debtors. Indeed, cash (that is, assets that can be realized immediately or almost immediately) is marginal in backing the credits that financial institutions (be they depositary institution, investment institutions, or insurances or pension funds) grant. Indeed, it would not make much sense to take a credit and to back it with cash! For the same reason, cash is also marginal in backing future payments a company may face due to a derivative contract.

A second observation is that, as of credit risk, governments are like other agents on financial markets both debtors and creditors. They are creditors, since taxpayers owe governments taxes, and debtors, since governments issue securities, government bonds, and bills.

A further observation is that some financial institutions borrow and lend money from central banks, that is, are debtor and creditor of central banks. Furthermore, central banks differ from other financial institutions inasmuch that they are not expected to reduce credit risk. While reducing credit risk (that is, to convey to their creditors less credit risk than they receive from their debtors) belongs to the *raisons d'être* of depositary banks, investment institutes, insurances and pension funds (Bhattacharya et al. 1998), the *raison d'être* of a central bank is to control the monetary base (that is, to create money and control its quantity), to control the interest rates, and to be a lender of last resort (Bordo 2007).

A last observation is that money and more generally a currency, too, conveys credit risk (Mitchell-Innes 1914; Graeber 2011). Indeed, money is a "claim upon society" (Simmel 1978). Credit risk, flowing from debtors to creditors, reaches central banks from which it flows back to all agents of the financial market. The European Sovereign-Debt Crisis since 2009 gives ample evidence of this backflow through a currency of real, as well as perceived, credit risk (Haidar 2012).

A Central Bank's Contribution to Credit Risk

What, in Eq. (1) of section "Formalizing an Agent's Credit Risk" above defining the credit risk $\mathrm{CR}^t(i)$ of agent i (at time t), should be the contribution of the central bank b to the credit risk of i (at time t)? Since, as observed above, a central bank b does not reduce credit risk, we propose to define this contribution as the amount of credit as well as cash or cash equivalents (that is, assets readily convertible into cash) owned by agent i (at time t). Thus, Eq. (1) is refined as follows so as to take the central bank into account:

$$\mathrm{CR}^t(i) = C_i^t + \sum_j c_{ji}^t + \sum_j (w_{ji}^t \times c_{ji}^t) \tag{2}$$

where C_i^t denotes the cash and cash equivalents owned by agent i at time t.

For financial institutions i subject to risk-based capital guidelines, the so-called capital requirements, C_i^t can easily be known. Indeed, it is disclosed by i to the agencies controlling the enforcement of capital requirements. For other agents i on the financial market, C_i^t can either be neglected as very small in comparison to $\sum_j c_{ji}^t$ or estimated from statistics.

All agents i should disclose c_{ji}^t when entering the corresponding credit or derivative contract. Knowledge of c_{ji}^t contract could be given by agents i and j to the GPR (or any other agency) upon entering a credit or derivative contract even if the GPR is not immediately activated, what, most likely, should be the most frequent case.

Thus, $\mathrm{CR}^t(i)$, as defined in Eq. 2, can be considered a known value for all agents i.

Credit Risk Graph

Call Credit Risk Graph (CRG) of a financial market the directed graph the nodes of which are the market agents (including government(s) and the central bank) and the labelled edges of which are defined as follows:

- There is an edge

$$j \xrightarrow[w_{ji}^t \times c_{ji}^t]{t} i$$

  from each agent j which is not the central bank and which is an actual or potential debtor of an agent i expressing the contribution of agent j to the credit risk $\mathrm{CR}^t(i)$ of agent i.
- There is an edge

$$b \xrightarrow[C_i^t + \sum_k c_{ki}^t]{t} i$$

  from the central bank b to each agent i.

The CRG of a financial market is strongly connected. Indeed, as observed above, every agent is a debtor of a financial institution. Every financial institution is directly or indirectly connected by a directed path, or “debtor path”, to a central bank. Indeed, this is because of such paths that central banks can control the monetary base. Thus, in the CRG, there is a directed path from every agent j to a central bank. Since there is a directed edge from a central bank to every agent i, there is, in the CRG, a directed path from every agent j to every other agent j, that is, the CRG is strongly connected.

Credit Risk Rating as Eigenvector Centralities

The facts that credit risk is defined as a weighted sum, that is, as a linear combination, and that the CRG of a financial market is strongly connected suggests that eigenvector centralities (Seeley 1949; Katz 1953; Gould 1967; Bonacich 1972, 2001, 2007; Koschützki et al. 2005; Vigna 2009) are appropriate as credit risk ratings. The following elaborates on this intuition and shows that it is adequate.

Let $A^t = (a_{ji}^t)$ denote the adjacency matrix of the labelled graph CRG at time t. Beware that the superscript t denotes time, not matrix transposition. Matrix transposition is denoted, as usual, by the superscript T.

$A^t = (a^t_{ji})$ is defined by

$$a^t_{ji} = \begin{cases} w^t_{ji} \times c^t_{ji} \; if \; j \, is \, not \, the \, central \, bank \, and \, is \, an \, actual \, or \, potential \\ \qquad debtor \, of \, i \, at \, time \, t \\ C^t_i + \sum_k c^t_{ki} \; if \; j \, is \, the \, central \, bank \\ 0 \text{ otherwise} \end{cases} \tag{3}$$

Note that matrix A^t is real and non-negative and the diagonal elements of A are all 0. If I is the identity matrix of same size $n \times n$ as A^t, $A^t + I$ is a real and non-negative matrix the diagonal elements of which are all 1, that is, non-negative. Considering $A^t + I$ instead of A^t is needed for the Perron vector considered below to exist.

Let $B^t = (b_{j\,i})$ be the column-stochastic normalization of the transposed $(A^t + I)^T$ of $A^t + I$ defined as follows:

$$b^t_{ji} = \begin{cases} \dfrac{a^t_{ij}}{\sum_j a^t_{ij}} \; if \; j \neq i \\ \dfrac{1}{\sum_j a^t_{ij}} \; if \; j = i \end{cases} \tag{4}$$

Like matrix A^t, matrix B^t is real and non-negative and its diagonal elements b^t_{ii} are all positive. Matrix B^t is, by definition, column-stochastic. The CRG being, as observed above, at any time t strongly connected, B^t is irreducible.

While the entry a^t_{ji} of A^t expresses the *absolute* contribution of agent j to the credit risk of i (at time t), the entry b^t_{ji} with $j \neq i$ of B^t expresses the *relative* contribution of agent i to the credit risk of j (at time t).

Equation (2) can be re-expressed as follows:

$$(\overrightarrow{cr^t})^T = \vec{1}^T A \tag{5}$$

where:

- $\overrightarrow{cr^t}$ is the credit risk (column) vector at time t, the ith element of which is $CR^t(i)$
- $\vec{1}$ is the unity (column) vector of dimension n (n being the number of agents on the financial market), all components of which are 1

Disregarding the diagonal elements of B, a relative version of (5) and therefore of (2) is:

$$\overrightarrow{cr^t} = B\vec{1} \tag{6}$$

Equation (6) suggests to define a systemic risk rating as follows. Assume $CRR^t(i)$ is a rating, or index, expressing the "credit risk strength" of agent i at time t. The credit risk strength of i (at time t) should surely be seen as proportional to the average of the credit risk strengths of its debtors j (at time t) weighted by the proportions b^t_{ji} to which debtors j contribute to the credit risk of i (at time t):

$$\mathrm{CRR}^t(i) = \frac{1}{\lambda} \times \sum_j (b^t_{ij} \times \mathrm{CRR}^t(j)) \tag{7}$$

where λ is a positive real scalar. Equation 7 can be expressed as:

$$\lambda \overrightarrow{crr^t} = B^t \overrightarrow{crr^t} \tag{8}$$

where $\overrightarrow{crr^t}$ is the credit risk rating vector (at time t) the ith component of which is $\mathrm{CRR}^t(i)$, the credit risk rating of agent i (at time t).

Equation 8 specifies a credit risk rating as eigenvector centrality. The rest of this section, which shows that Eq. 8 is an acceptable definition, is common knowledge (Seeley 1949; Katz 1953; Gould 1967; Bonacich 1972, 2001, 2007; Koschützki et al. 2005; Vigna 2009). It is included here for the sake of completeness.

B^t being, as observed above, irreducible, real, non-negative, column-stochastic and each diagonal element of B^t being positive, it follows from celebrated theorems by Perron and Frobenius (Perron 1907; Frobenius 1912; Wielandt 1950; Langville and Meyer 2006) that λ = 1 is a simple and strictly dominant eigenvalue of B^t the eigenvector associated with is called Perron vector of B^t.

Since λ = 1 is a simple eigenvalue of B^t, the Perron vector of B^t is, up to a scalar, the unique real non-zero solution of Eq. 8. This makes the Perron vector B^t with norm 1 an acceptable credit risk rating vector. Indeed, if Eq. 8 had several real solutions with norm 1, then there would be no reasons to choose the one instead of another as a credit risk rating vector.

Finally, if a vector $\vec{u}$ is not orthogonal to the Perron vector of B^t, then normalized power sequences of B^t and $\vec{u}$ converge to the Perron vector of B^t (von Mises and Pollaczek-Geiringer 1929; Langville and Meyer 2006). This makes it possible to apply the power iterations (von Mises and Pollaczek-Geiringer 1929; Langville and Meyer 2006) to compute the Perron vector of B^t expressing a credit risk rating. If, by chance, a vector $\vec{u}$ normalized power sequences would start with would be orthogonal to the Perron vector of B^t, then rounding errors would nonetheless ensure convergence to the Perron vector of B^t.

Discussion

This section first questions the model underlying the systemic credit risk rating proposed in the former section. Then, it discusses the technical practicability of running the Human Computation algorithm consisting of the Grace Period Reward (GPR) of section "Human Computation: Potential Debtors Assess Their Own Risk of Defaulting" and the computation of the Perron vector defined by Eq. (8) of section "Network Analysis: Aggregating Human Estimates into a Systemic Credit Risk Rating". It also discusses whether financial markets could accept the constraints imposed by the GPR. Further, it stresses that the Human Computation algorithm proposed in this chapter would provide the information necessary for more accurate, dynamic, settings of the regulatory capital (Bhattacharya et al. 1998; Hull 2010). Finally, it explains why the credit risk rating proposed in this article is a systemic rating which has the potential of detecting systemic risk.

Questioning the Model

Even though there are good arguments for assuming that all agents on a financial market are debtors, the reality is sometimes more complex than one expects.

If some agents are, at some time t no debtors, then the Credit Risk Graph (CRG) of section "Network Analysis: Aggregating Human Estimates into a Systemic Credit Risk Rating" would not always be strongly connected and the matrix B^t (obtained from the adjacency matrix A^t of the CRG) would not fulfill the conditions ensuring the existence of the Perron vector. In such a case, the matrix A^t could be modified so as to express that an agent, who in reality is no debtor, being in equal proportions the debtor of all other agents on the financial market. PageRank (Page et al. 1999; Langville and Meyer 2006) is based upon a similar transformation of the Hyperlink matrix. This transformation does not impair PageRank's adequacy as a ranking.

In section "Network Analysis: Aggregating Human Estimates into a Systemic Credit Risk Rating", a single central bank is considered. The CRG, and therefore the systemic credit risk rating specified in that section, can easily be extended to several central banks. Such an extension requires to consider different currencies. To this aim, instead of the matrices A^t (B^t, respectively), 3-rank tensors need being considered each consisting, for each currency of a matrix A^t (B^t, respectively). Exchange rates between the currencies also need to be considered, if, as one may expect, some agents are parties in credit or derivative contracts in different currencies.

The model is based upon the view that a currency in general, and money in particular, conveys credit risk. This view implies that the value of a currency (and of money) is two dimensional, one dimension reflecting the prosperity of, the other the systemic credit risk in, the area where the currency is a unit of account. The European Sovereign-Debt Crisis since 2009 gives ample evidence that this interpretation is appropriate (Haidar 2012).

Practicability

Could, from a computing viewpoint, the GPR of section "Human Computation: Potential Debtors Assess Their Own Risk of Defaulting" be deployed and repeated computations of the Perron vector defined by Eq. (8) of section "Network Analysis: Aggregating Human Estimates into a Systemic Credit Risk Rating" be performed?

The GPR requires a transaction for each new financial contract and each time a contract party activates, de-activate, or modifies an activation of the GPR. Except in crisis times, most contracts can be expected to result in no or a very small number of activations of the GPR. Hence, the IT support of the GPR would, in normal times, be no challenge for the IT infrastructure of a financial market. If in crisis times the GPR requires too high an IT support, then this would be a sufficient reason to slow down the market activity and, as a consequence, the GPR usage. Indeed, a "GPR frenzy" would be a clear sign of a market getting out of control.

Repeated computations of the Perron vector of Eq. (8) is a real challenge which would require a significant extension of the IT infrastructure of a financial market and which deserves more investigations. Since it refers to time, Eq. (8) in fact refers to a 3-rank tensor A, one of the ranks of which is the time line. The time considered is, of course, discrete (the time unit being the (bank) day) and finite (the latest time point being the latest credit event referred to in an activation of the GPR). This tensor is likely to be very sparse, because only a few different degrees of activations of the GPR for a same contract by a same agent can be expected. This sparsity provides a hook for efficient power iterations which, admittedly, remain to be fully worked out. This tensor is also sparse because,

- At each time t, the matrix A^t, and therefore the matrix B^t, are sparse. Indeed, most agents are no financial institutions and therefore enter financial contracts with a limited number of agents,
- There are only a few financial institutions, that at any time are bound by contracts with large numbers of agents.

Repeated updates of the Perron vector defined by (8) could be performed, like the Google search engine does, by considering only those parts of the matrices B^t that have been updated since the last computation.

Thus, computing the Perron vector defined by (8) would be rather similar, in the techniques and computing effort needed, to computing the structure ranking of a Web search engine.

Acceptance

Deploying on a financial market the systemic credit risk rating proposed in this chapter would require significant changes on this market. First, all financial transactions would have to be registered. Second, computations similar to that of a Web search engine would have to be performed.

In Simkovic (2010) Michael Simkovic makes a plea for "recordation", that is, for creditors to make

> a full and complete disclosure [of creditor-debtors-liens] in return for payment priority

in case of a debtor's defaulting. Disclosure of creditor-debtors-liens to the system running the GPR and computing the systemic credit ranking is all the systemic credit risk rating proposed in this chapter requires, that is, less than Michael Simkovic considers necessary for other reasons.

One might expect that the inertia natural to individuals and organizations will prevent such changes. The promise of a financial market better and earlier foreseeing a wide-spreading of credit risk should be sufficient an incentive to changes, if not for the financial markets themselves, for the executives and legislatives responsible for the bailouts of financial institutions deemed "to big to fail" that, as the financial crisis of 2007–2009 has shown, may result from today's credit risk rating.

The GPR should be appealing to financial markets, regulators of these markets, and the society because its functioning can be seen for three reasons as "bailouts in the small": First, the GPR makes it possible to bail out debtors before too high a credit risk has been concentrated in financial institutions deemed "too big to fail". Second, the GPR is limited in time and, in contrast to a credit risk insurance, does not bail out defaulting debtors but instead only give them, and the society, time to financially recover. Third, the GPR's costs are covered from the markets' agents, the GPR fee acting like a transaction tax.

Feedback Loop: Dynamic Regulatory Capital

Currently, the regulatory capital (Bhattacharya et al. 1998; Hull 2010) a depository financial institution must hold depends on the credits it has given but not on its credit risk. This is consistent with the fact that, so far, there is no systemic credit risk rating and that assessing credit risk is, in spite of regulations, to a large extent "one's own affair".

The systemic credit risk rating proposed in this chapter would make it possible to specify for each depository institution an amount of regulatory capital depending on the systemic credit risk rating of this institution.

The systemic credit risk rating proposed in this chapter would also make it possible to foresee changes in the regulatory capital of a depositary financial institution implied by changes of the institution's systemic credit risk rating.

Systemic Risk

The rating of credit risk proposed in this article is "systemic" because it is computed as an equilibrium property of the credit risk graph induced by credit contracts, derivative contracts, and currencies. Can this systemic rating of credit risk also serve as a rating of systemic risk?

There is an abundant research literature on systemic risk, a large part of which has been published after, and because of, the Asian Crisis of 1997–1998. Following the Financial Crisis of 2007–2009 and the European Sovereign-Debt Crisis going on since 2009, the Board of Governors of the Federal Reserve System in the US, the boards of banks in the US and in Europe, and public agencies such as the European Securities and Markets Authority (ESMA) have launched research groups commissioned to propose systemic risk measurements. This triggered many more publications on the subject. The recent survey (Bisias et al. 2012) the authors of which state (on p. 4) "we do not attempt to be exhaustive in our breadth" lists no less than 31 proposals for measuring systemic risk! In spite, or because, of this intense research activity, systemic risk remains a concept without widely accepted definition (Dow 2000; Bisias et al. 2012; Hansen 2013) as the article (Bisias et al. 2012) stresses in its introduction:

> The truism that 'one cannot manage what one does not measure' is especially compelling for financial stability since policymakers, regulators, academics, and practitioners have yet to reach a consensus on how to define 'systemic risk'. While regulators sometimes apply Justice Potter Stewarts definition of pornography, i.e., systemic risk may be hard to define but they know it when they see it, such a vague and subjective approach is not particularly useful for measurement and analysis, a pre-requisite for addressing threats to financial stability.

Informally, "systemic risk" refers to the risk of breakdown of a financial market as a consequence of cascading insolvabilities of market agents bound to each other by financial contracts. The article Hansen (2013, Sect. 2.1) points to three different acceptations of "systemic risk" in the research literature: (1) "a modern-day counterpart to a bank run triggered by liquidity concerns", (2) "the vulnerability of a financial network in which adverse consequences of internal shocks can spread and even magnify within the network", and (3) one of the former senses extended so as to "include the potential insolvency of a major player in or component of the financial system."

The purpose of the systemic credit risk rating proposed in this article is an early detection of systemic risk in the first of the above mentioned senses. Most likely, it would also help in detecting some "vulnerabilities in a financial network". There are no reasons to believe, though, that it could help in detecting all of them. It should also be useful for an early detection of potential insolvencies of components of a financial market. Indeed, its principle, an eigenvector computation, makes it easily adaptable to compute the centralities of groups of market agents. It would provide with an estimate of the systemic risk assumed by a currency, that is, by the community of the currency area. The matrices A and B of section "Credit Risk Rating as Eigenvector Centralities" would be useful for many kinds of systemic risk investigations.

In contrast to most systemic risk measures proposed so far (see Bisias et al. (2012) for a survey) the rating proposed in this article is neither based on the assumption that systemic risk arises endogenously within a financial market nor does it rely on stochastic methods and statistical data. This makes it a plausible risk indicator in times of exogenous shocks and of crises that rarely happen.

The articles May et al. (2008); Haldane and May (2011); Kyriakopoulos et al. (2009) deserve a special mention here because, like the present proposal and unlike most of the credit risk and systemic risk literature, they propose approaches based on Network Analysis and eigenvector computations. The article Haldane and May (2011) develops models of financial networks inspired from ecological food webs and from networks within which infectious diseases spread. The article Kyriakopoulos et al. (2009) which reports on an empirical time series analysis of the financial transactions over 1 year between the major financial agents in Austria uses eigenvalue spectra.

Human Computation and Markets

In conceiving well-working Human Computation systems (such as Crowdsourcing marketplaces (Law and von Ahn 2011, Sect. 5.1, p. 45), prediction markets (Hubbard 2007; Wolfers and Zitzewitz 2006; Pennock et al. 2001; Servan-Schreiber et al. 2004; Gjerstad 2005; Snowberg et al. 2007; Berg et al. 2008; Arrow et al. 2008), decision markets (Leutenmayr and Bry 2011), immediate response Crowdsourcing systems (Munro 2010; Starbird 2011), or games with a purpose (GWAPs) (von Ahn 2006; Steinmayr et al. 2011; Bry and Wieser 2012; Kneissl and Bry 2012), the following issues are worth considering: The incentives provided for humans to contribute to the systems (Papaioannou and Stamoulis 2005; Harris 2011; Findley et al. 2012; Zhang and van der Schaar 2012); whether the systems are capable of growing (in the sense of attracting more human participation) (Kim 2000; Preece 2000; Powazek 2002; O'Keefe 2008; Kraut and Resnick 2011); whether the systems are self-sufficient (inasmuch that they generate all data needed for their proper working);[2] and whether they are efficient (in the sense of achieving maximum productivity with minimum wasted human and machine effort or expense) (Jain and Parkes 2008; Archak and Sundararajan 2009; Ghosh and McAfee 2011; Cavallo and Jain 2012).

It is striking that these issues are economical characteristics. This is not by chance. The amount of human computation a Human Computation system makes possible can be seen as the "wealth" it generates and economics is the field concerned with the production, consumption, and transfer of wealth (Jewell et al. 2005). Furthermore, markets can be seen as Human Computation systems *avant la lettre*, that is, before the term had been coined. Indeed, on markets traders perform the following "computations" (even though in the past without computer support): Interpreting information on the goods traded with and adjusting the trading prices. On markets so-called "market-makers" also perform "computations" (in the past without computer support) when they ensure the markets' liquidity by selling (purchasing, respectively) at prices lower (higher, respectively) than the current sale (purchase, respectively) prices and possibly speculating on the prices' evolution for making a profit (Grossman and Miller 1988).

The Efficient Market Hypothesis (Samuelson 1965; Fama 1970; O'Sullivan and Sheffrin 2007) according to which prices on financial markets reflect all information[3] on the assets traded with makes sense because of these "computations" performed by humans, that is, the human activity necessary for the timely wide-spreading of the information the traders and market-makers rely on for their price adjustments. Thus, Adam Smith's "invisible hand" (Smith 1776; O'Sullivan and Sheffrin 2007),

[2] Surprisingly, self-sufficiency of Human Computation systems does not seem to have, so far, attracted much attention within the research community. The system's self-sufficiency has been one of the author's concerns in building the Human Computation platform metropolitalia.org (Kneissl and Bry 2012).

[3] Past, present and even hidden information, depending on which of the Weak, Semi-Strong and Strong Efficient Market Hypotheses is considered.

a metaphor expressing markets' capability of self-regulation, can be seen as a Human Computation *avant la lettre.*

A worthwhile question is thus not only whether today's financial markets are efficient, a question which has been much debated (Nicholson 1968; Basu 1977; R osenberg et al. 1985; Beechey et al. 2000), but also what can be done for ensuring, or improving, financial markets' efficiency. This chapter is a contribution to ensuring the efficiency of financial markets. Its thesis is, that Human Computation provides with novel means, that so far were unthinkable, for ensuring markets' efficiency. Since markets and certain forms of human computation are related, more insights into markets' efficiency may also have implications on the design of efficient Human Computation systems. In other words, approaches similar to the Human Computation-enabled systemic credit risk rating proposed in this chapter could be conceived for regulating Human Computation systems.

Like credit risk on financial markets, reputation is a primary concern among the human contributors to many Human Computation systems (like Crowdsourcing marketplaces and immediate response Crowdsourcing systems) and of course among traders on online auction systems (like eBay). It is rather natural to think that a network analysis similar to the one proposed in this chapter for credit risk rating would be promising for assessing reputation. This view has recently been shown to be accurate (Gkorou et al. 2012; Chiluka et al. 2012).

Prices are as important for the good working of a Human Computation labor market (like Amazon Turk) as a proper credit risk rating is for the good working of a financial market. So far, Human Computation labor markets are far from being efficient both, in the sense of Human Computation system efficiency mentioned at the beginning of this section, and in the sense of the Efficient Market Hypothesis (Samuelson 1965; Fama 1970; O'Sullivan and Sheffrin 2007). Indeed, a major criticism of Mechanical Turk concerns its pricing of labor (Ross et al. 2010; Silberman et al. 2010a,b; Felstiner 2011). This criticism is summarized as follows in Law and von Ahn (2011, p. 74):

> […] there exists power asymmetry in Mechanical Turk, where requesters can reject work without providing justification, thereby not only forbidding payment but hurting the future chances of work by damaging the workers reputation.

On a labor market like Amazon Turk, a Human Computation-enabled network analysis of the type described in this chapter holds the promise of a better labor pricing, either through the detection of unfair requesters by a ranking of their reputations, through a "bipartite ranking" considering both, requesters' and workers' (or turkers') reputations, or through a ranking combining prices and reputation. What such a ranking may be, is an open issue. Undoubtedly, Network Analysis holds the promise of such rankings, that is, of making Human Computation labor markets more efficient in both senses mentioned above.

The author believes that Human Computation-enabled network analyses of the kind proposed in this chapter will, in the future, contribute to the good-working of many Human Computations systems and markets. In an age of progressing globalization (IMF Staff 2000), only a systematic collecting of information and its timely aggregation can ensure the efficiency of both, Human Computation systems and markets. Human Computation is needed for collecting all the information needed by humans contributing to Human Computation systems and markets because of the variety and complexity of goods, services, or tasks dealt with on most Human Computation systems and markets.

Conclusion

This chapter has proposed a novel approach to credit risk rating based upon Network Analysis and enabled by Human Computation. The approach proposed is a dual-purpose participatory mechanism.

Section "Credit Risk Rating Challenged by Derivatives, Structured Notes, and Securitization", an introduction into credit risk rating, derivatives, structured notes, and securitization, has stressed the need for a novel, systemic, assessment of credit risk on financial markets.

The system proposed in this chapter for computing a systemic credit risk rating consists of an incentive, the Grace Period Reward (GPR) introduced in section "Human Computation: Potential Debtors Assess Their Own Risk of Defaulting", for humans to contribute with inputs and a network analysis for aggregating these human inputs. It is therefore a Human Computation algorithm.

The GPR, which resembles a credit risk insurance over a limited period of time, provides with an incentive for debtors to assess the risk of their own defaulting and to disclose these assessments to a system. The aggregation of the human inputs has been specified in section "Network Analysis: Aggregating Human Estimates into a Systemic Credit Risk Rating" as eigenvector centralities expressing a systemic rating of the credit risk faced by the market's agents.

This systemic credit risk rating has been shown in sections "Human Computation: Potential Debtors Assess Their Own Risk of Defaulting" and "Discussion" to hold the promise of overcoming many deficiencies of current credit risk rating. The GPR has been described in section "The GPR as a Transaction Tax" as an informationally productive transaction (or Tobin) tax and as making "bailouts in the small" possible.

The credit risk rating proposed in this chapter is unusual in several ways: It collects credit risks assessments from the debtors and not from the creditors; it propagates debtors' risk estimates through the risk dependency graph by computing eigenvector centralities; it is not based upon stochastic methods and statistical data; and it promises to deliver warnings of an increase in credit risk much earlier than current credit risk rating methods. Furthermore, it gives a means to specify the regulatory capital of a depository financial institution depending on its "credit risk centrality" in the credit risk graph of the financial market.

Importantly, the credit risk rating proposed in this chapter does not require credits, derivatives or securitization instruments to be standardized. Thus, it imposes no constraints on the kind of financial instruments dealt with on the financial market and can accommodate new, so far not thought of, financial instruments.

The approach proposed in this chapter seems to have several advantages over current credit risk rating, which has been briefly introduced in section "Credit Risk Rating Challenged by Derivatives, Structured Notes, and Securitization": it is systemic in the sense that it not only considers the credit risk assumed by one agent but instead aggregates the credit risk along the debtor-creditor-liens, it is not affected by the complexity of financial instruments traded with, it is not based on statistics what makes it appropriate in exceptional situations too, it is not subject to the moral hazard of current credit risk rating, it provides time-dependent ratings, and it significantly reduces the speed differential between transactions and credit risk rating.

Deploying the approach on a financial market would, as it is discussed in section "Discussion", require further investigations, especially a calibration of the GPR.

Whether the proposed method would, in practice, hold its promises, is an open issue which is out of the scope of this chapter.

The main obstacle to a deployment of the systemic credit risk rating proposed in this chapter should be of cultural and political nature. Is the time ripe for financial markets and policy makers to understand and accept a network analysis enabled by Human Computation as a means to credit risk rating? If not, how many more financial crises will be needed for convincing of the value of techniques that, in other fields, are already well established?

Finally, this chapter has argued in section "Human Computation and Markets" that Human Computation systems and markets share many commonalities. In that section, the view is expressed that Human Computation-enabled network analyses of the kinds proposed in this chapter will contribute, in the future, to the good-working of both, Human Computations systems and markets.

Acknowledgements The author is thankful to Norbert Eisinger (from the Institute for Informatics of the Ludwig-Maximilians University of Munich) for useful suggestions; to the team of the Play4Science research project for stimulating discussions on incentives in Human Computation systems; and to Pietro Michelucci, editor-in-chief of the "Handbook of Human Computation", and Haym Hirsh, editor of Section B "Application Domains" of that handbook, for useful advices on how to better present the research reported about in this chapter.

The work reported about in this chapter has benefited much from the knowledge and the experience the author gained through his contributions to the Human Computation platforms ARTigo.org and metropolitalia.org developed within the project Play4Science founded in part by the German Foundation for Research (DFG, project number 578416).

References

Altman EI, Saunders A (1998) Credit risk measurement: developments over the last 20 years. J Bank Finance 21:1721–1742

Archak N, Sundararajan A (2009) Optimal design of crowdsourcing contests. In: Proceedings of the 30th international conference on information systems (ICIS), Phoenix

Arora N, Gandhi P, Longstaff FA (2012) Counterparty credit risk and the credit default swap market. J Financ Econ 103(2):280–293

Arrow KJ, Forsythe R, Gorham M, Hahn R, Hanson R, Ledyard JO, Levmore S, Litan R, Milgrom P, Nelson FD, Neumann GR, Ottaviani M, Schelling TC, Shiller RJ, Smith VL, Snowberg E,

Sunstein CR, Tetlock PC, Tetlock PE, Varian HR, Wolfers J, Zitzewitz E (2008) The promise of prediction markets. Science 320:877–878
Basu S (1977) Investment performance of common stocks in relation to their price-earnings ratios: a test of the efficient markets hypothesis. J Finance 32:663–682
Beaver WH, Shakespeare C, Soliman MT (2003) Differential properties in the ratings of certified vs. non-certified bond rating agencies. Research report, Ross School of Business, University of Michigan, revised 2006
Beechey M, Gruen D, Vickery J (2000) The efficient markets hypothesis: a survey. RBA research discussion paper rdp2000-01, Economic Research Department, Reserve Bank of Australia
Berg JE, Nelson FD, Rietz TA (2008) Prediction market accuracy in the long run. Int J Forecast 24(2):285–300
Bhattacharya S, Boot AWA, Thakor AV (1998) The economics of bank regulation. J Money Creit Bank 30(4):745–770,
Bisias D, Flood MD, Lo AW, Valavanis S (2012) A survey of systemic risk analytics. Report 0001, US Department of Treasury, Office of Financial Research
Blanchard OJ, Watson MW (1982) Bubbles, rational expectations and financial markets. Working paper 945, National Bureau of Economic Research
Board of Governors (2007) Report to the congress on credit scoring and its effects on the availability and affordability of credit. Report, Federal Reserve System
Bonacich P (1972) Factoring and weighting approaches to clique identification. J Math Sociol 2:113–120
Bonacich P (2001) Eigenvector-like measures of centrality for asymmetric relations. Soc Netw 23:191–201
Bonacich P (2007) Some unique properties of eigenvector centrality. Soc Netw 29:555–564
Bordo MD (2007) A brief history of central banks. Economic Commentary of the Federal Reserve Bank of Cleveland
Boucheaud J-P, Potters M (2000, 2003) Theory of financial risk and derivative pricing – from statistical physics to risk management. Cambridge University Press, Cambridge
Božović M, Urošević B, Živković B (2011) Credit rating agencies and moral hazard. Panoeconomicus 58:221–227
Bry F (2012) Human computation and economics. Research report, Institute for Informatics, Ludwig-Maximilian University of Munich
Bry F, Wieser C. (2012) Squaring and scripting the ESP game. In: Proceedings of the 4th human computation workshop (HC), Toronto
Buffet WE (2002) Chairmans letter. Annual report, Berkshire Hathaway Inc.
Caccioli F, Marsili M, Vivo P (2009) Eroding market stability by proliferation of financial instruments. Eur Phys J B 71:467–479
Cavallo R, Jain S (2012) Eficient crowdsourcing contests. In: Proceedings of the 11th international conference on autonomous agents and multiagent systems (AAMAS), Valencia
Chance DM (2008) Essays in derivatives. Wiley, Hoboken
Chiluka N, Andrade N, Gkorou D, Pouwelse JA (2012) Personalizing eigentrust in the face of communities and centrality attack. In: Proceedings of the IEEE 26th international conference on advanced information networking and applications (AINA), Fukuoka, pp 503–510
Dow J (2000) What is systemic risk? Moral hazard, initial shocks, and propagation. Monetray and economic studies, Tokyo, pp 1–24
Duffie D, Singleton KJ (2003) Credit risk – pricing, measurement, and management. Princeton series in finances. Princeton University Press, Princeton and Oxford
Fama EF (1970) Efficient capital market: a review of theory and empirical work. J Financ 25(2); 383-417
Felstiner A (2011) Working the crowd: employment and labor law in the crowdsourcing industry. Berkeley J Employ Labor Law 32(1):143–204
Findley MG, Gleave MC, Morello RN, Nielson DL (2012) Extrinsic, intrinsic, and social incentives for crowdsourcing development information in Uganda: a field experiment. Technical report, Brigham Young University

Frobenius G (1912) Über Matrizen aus nicht negativen Elementen. In: Sitzungsbericht der königlich-preußichen Akademie der Wissenschaften. On Matrices with non-negative elements (in German)
Ghosh A, McAfee P (2011) Incentivizing high quality user generated content. In: Proceedings of the 20th international conference on world wide web (WWW), Hyderabad, pp 137–146
Gjerstad S (2005) Risk aversion, beliefs, and prediction market equilibrium. Research report, Economic Science Laboratory, University of Arizona
Gkorou D, Vinkó T, Chiluka N, Pouwelse JA, Epema DHJ (2012) Reducing the history in decentralized interaction-based reputation systems. In: Networking (2) – proceedings of the 11th international IFIP TC 6 networking conference, Prague, pp 238–251
Gomolka J (2011) Algorithmic Trading – Analyse von computergesteuerten Prozessen im Wertpapierhandel unter Verwendung der Multifaktorenregression. Doctoral dissertation, Wirtschafts- und Sozialwissenschaftliche Fakultät, University of Postdam. Analysis of computer-based processes in securities trading using multiple regression (in German)
Gould PR (1967) On the geographical interpretation of eigenvalues. Trans Inst Br Geogr 42:53–86
Graeber D (2011) Debt: The first 5000 years. Allen Lane. Melville House Publishing, New York
Gregory J (2010) Counterparty credit risk: the new challenge for global financial markets. Wiley, Chichester
Grossman SJ, Miller MH (1988) Liquidity and market structure. J Financ 63(3):617–633
Haidar JI (2012) Sovereign credit risk in the eurozone. World Econ 13(1):123–136
Haldane AG, May RM (2011) Systemic risk in banking ecosystems. Nature 469:351–355
Hansen LP (2013) Challenges in identifying and measuring systemic risk. Working paper 18505, National Bureau of Economic Research
Harris CG (2011) You're Hired! An examination of crowdsourcing incentive models in human resource tasks. In: Proceedings of the workshop on crowdsourcing for search and data mining (CSDM), Hong Kong, pp 15–18
Hubbard DW (2007) How to measure anything: finding the value of intangibles in business. Wiley, Hoboken
Hull JC (2010) Risk management and financial institutions, 2nd edn. Prentice Hall, Boston
Hull JC (2012) Options, futures, and other derivatives. Pearson, Boston/Munich
IMF Staff (2000) Globalization: threats or opportunity? Brief 00/012, International Monetray Fund, Washington, D.C.
Jain S, Parkes DC (2008) A game-theoretic analysis of games with a purpose. In: Proceedings of the 4th international workshop on internet and network economics (WINE), Shanghai, pp 342–350
Jarrow RA, Turnbull SM (1995) Pricing derivatives on financial securities subject to credit risk. J Financ 50(1):53–85
Jewell EJ, Abate F, McKean E (2005) The new Oxford american dictionary, 2nd edn. Oxford University Press, New York
Jones R (1961) Comparative advantage and the theory tariffs: a multi-country, multi-commodity model. Rev Econ Stud 77:161–175
Joyce TM (2008) 2008 U.S. Market Structure Update. White Paper, Knight Capital Group, Inc.
Katz L (1953) A new index derived from sociometric data analysis. Psychometrika 18:39–43
Kim AJ (2000) Community building on the web: secret strategies for successful online communities. Peachpit, Berkeley
Kneissl F, Bry F (2012) MetropolItalia: a crowdsourcing platform for linguistic field research. In: Proceedings of the IADIS international conference WWW/internet
Koschützki D, Lehmann KA, Peeters L, Richter S, Tenfelde-Podehl D, Zlotowski O (2005) Network Analysis: Methodological Foundations. Lecture notes in computer science. Springer-Verlag, Springer, Berlin, Heidelberg, New York, pp 16–61
Kothari V (2012) Credit Derivatives and structured credit trading, chapter approaches to quantification of credit risk. Wiley, Singapore

Kraut RE, Resnick P (2011) Evidence-based social design: mining the social ssciences to build online communities. MIT, Cambridge

Kyriakopoulos F, Thurner S, Puhr C, Schmitz SW (2009) Network and eigenvalue analysis of financial transaction networks. Eur Phys J B 71(4):523–531

Lando D (2009) Handbook of financial time series, chapter credit risk modeling. Springer, Berlin/London, pp 787–798

Langville AN, Meyer CD (2006) Google's PageRank and beyond: the science of search engine rankings. Princeton University Press, Princeton

Law E, von Ahn L (2011) Human computation. Synthesis lectures on artificial intelligence and machine learning. Morgan & Claypool, San Rafael

Leutenmayr S, Bry F (2011) Liquid Decision Making: an exploratory study. In: Proceedings of the 13th international conference on information integration and web-based applications and services (iiWAS), Ho Chi Minh City, pp 391–394. ACM

Malevergne Y, Sornette D (2005) Extreme financial risks: from dependence to risk management. Springer finance. Springer, Berlin/London

Mandelbrot B, Hudson RL (2004) The misbehavior of markets: a fractal view of financial turbulence. Basic Books, New York

May RM, Levin SA, Sugihara G (2008) Complex systems: ecology for bankers. Nature 451:893–895

Metz A, Cantor R (2006) Moody's credit rating prediction model. Moody's special comment, Moody

Mitchell-Innes A (1914) The credit theory of money. Bank Law J 31:151–168

Munro R (2010) Crowdsourced translation for emergency response in Haiti: the global collaboration of local knowledge. In: Proceedings of the AMTA workshop on collaborative crowdsourcing for translation, Denver

National commission on the causes of the financial and economic crisis in the United States (2011) The financial crisis inquiry report. Public Affairs, Perseus Books Group

Nicholson F (1968) Price-earnings ratios in relation to investment results. Financ Anal J 24(1):105–109

O'Keefe P (2008) Managing online forums: everything you need to know to create and run successful community discussion boards. AMACOM, New York

O'Sullivan A, Sheffrin SM (2007) Economics: principles in action. Prentice Hall, Boston

Page L, Brin S, Motwani R, Winograd T (1999) The PageRank citation ranking: bringing order to the Web. Technical report, Stanford InfoLab, Stanford University

Papaioannou TG, Stamoulis GD (2005) An incentives' mechanism promoting truthful feedback in peer-to-peer systems. In: Proceedings of the fifth IEEE international symposium on cluster computing and the grid (CCGRID), Cardiff

Pennock DM, Lawrence S, Giles CL, Nielsen F (2001) The real power of artificial markets. Science 291(5506):987–988

Perron O (1907) Zur theorie der Matrices. Mathematische Annalen 64(2):248–263. On matrix theory (in German)

Powazek DM (2002) Design for community: the art of connecting real people in virtual places. Pearson Technology, New Riders Publishing, Indianapolis

Preece J (2000) Online communities: designing usability and supporting sociability. Wiley, New York

Quinn AJ, Bederson BB (2009) A taxonomy of distributed human computation. Technical report HCIL-2009-23, University of Maryland

Ricardo D (1817) On the principles of political economy and taxation. John Murray, London

Rosenberg B, Reid K, Lanstein R (1985) Persuasive evidence of market inefficiency. J Portf Manag 13:9–17

Ross J, Irani L, Silberman MS, Zaldivar A, Tomlinson B (2010) Who are the crowdworkers? shifting demographics in mechanical turk. In: Proceedings of the 28th international conference on human factors in computing systems (CHI) – extended abstracts, Atlanta, pp 2863–2872

Samuelson P (1965) Proof that properly anticipated prices fluctuate randomly. Ind Manag Rev 6:41–49

Sarkar A (2009) Liquidity risk, credit risk, and the federal reserve's responses to the crisis. Financ Mark Portf Manag 23:335–348
Seeley JR (1949) The net of reciprocal influence: a problem in treating sociometric data. Can J Psychol 3:234–240
Servan-Schreiber E, Wolfers J, Pennock DM, Galebach B (2004) Prediction markets: does money matter? Electron Mark 14(3):243-251
Silberman MS, Irani L, Ross J (2010a) Ethics and tactics of professional crowdwork. ACM Crossroads 17(2):39–43
Silberman MS, Ross J, Irani L, Tomlinson B (2010b) Sellers' problems in human computation. In: Proceedings of the workshop on human computation (HC), Washington, DC
Simkovic M (2009) Secret liens and the financial crisis of 2008. Am Bankruptcy Law J 83:253–296, Washington, D.C.
Simkovic M (2010) Paving the way for the next financial crisis. Bank Financ Serv Policy Rep 29(3):1–20
Simkovic M (2013) Competition and crisis in mortgage securitization. Indiana Law J 88:213–271
Simmel G (1978) Philosophy of money. Routledge and Kegan Paul, Boston
Smith A (1776) An inquiry into the nature and causes of the wealth of nations. W. Strahan and T. Cadell, London
Snowberg E, Wolfers J, Zitzewitz E (2007) Partisan impacts on the economy: evidence from prediction markets and close elections. Q J Econ 122(2):807–829
Soros G (2008) The new paradigm for financial markets – the credit crisis of 2008 and what it means. PublicAffairs™, Perseus Books Group, Philadelphia
Starbird K (2011) Digital volunteerism during disaster: crowdsourcing information processing. In: Proceedings of the CHI wokshop on crowdsourcing and human computation, Vancouver
Steinmayr B, Wieser C, Kneißl F, Bry F (2011) Karido: a GWAP for telling artworks apart. In: Proceedings of the 16th international conference on computer games (CGAMES), Louisville, July 2011
Taleb NN (2007, 2010) The black swan – the impact of the highly improbable, 2nd edn. Random House, New York
Tobin J (1978) A proposal for international monetary reform. East Econ J 4:153–159
Treacy WF, Carey MS (1998) Credit risk rating at large U.S. banks. Fed Reserve Bull 84(11)897–921
Vigna S (2009) Spectral ranking. CoRR, abs/0912.0238, http://arxiv.org/abs/0912.0238
von Ahn L (2006) Games with a purpose. Computer 29(6):92–94
von Mises R, Pollaczek-Geiringer H (1929) Praktische Verfahren der Gleichungsauflösung. Zeitschrift für Angewandte Mathematik und Mechanik 9:152–164. Practical methods for solving equations (in German)
Wielandt H (1950) Unzerlegbare, nicht negative Matrizen – Herrn Oskar Perron zum 70. Geburtstag am 7. Mai 1950 gewidmet. Mathematische Zeitschrift 52(1):642–648. Irreducible non-negative matrices – Dedicated to Mr Oskar Perron for his 70th birthday on the 7th of May 1950 in German
Wolfers J, Zitzewitz E (2006) Interpreting prediction market prices as probabilities. Working paper 12200, NBER
Zhang Y, van der Schaar M (2012) Reputation-based incentive protocols in crowdsourcing applications. Technical report, Department of Electrical Engineering, University of California,

Innovation via Human Computation

Lisa Purvis and Manas Hardas

Introduction

This chapter considers the means by which many people can work together to generate new ideas that have practical value. A familiar example of such a process is "brainstorming", where people build off of each other's ideas. Network technology and social collaboration have allowed us to improve traditional brainstorming so that more people can contribute ideas and work together more effectively irrespective of time asynchronicity or geographical distance. This chapter describes the techniques we have found to be instrumental for achieving innovation on an organizational scale.

Innovation, the introduction of new methods, solutions, products, is important to today's global business growth. Innovative products enjoy 70 % higher margins than 'me-too' products (Aberdeen Group 2009). However, successful innovation at scale in today's enterprise environments is a complex and elusive process, especially as one tries to capture productive innovation in a large enterprise, where the internal culture, behaviors, and goals interact in sometimes unpredictable and changeable ways.

The collective wisdom of a crowd through the aggregation of information in groups has long been recognized as a powerful decision-making approach (Surowiecki 2004). Applying this approach to maximize outcomes from the innovation process makes intuitive sense. However, doing so productively requires intelligent approaches to understanding the innovation network and its interactions, the engagement with innovation that is occurring in order to help maximize outcomes, and to ensure that a healthy mix of collaboration and participation is happening.

L. Purvis (✉) • M. Hardas
Spigit Inc, Pleasanton, CA, USA
e-mail: lpurvis@spigit.com; mhardas@spigit.com

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_20,

Spigit is an Enterprise Innovation Platform that is used by the world's leading brands to invent new products and services, reduce costs and increase employee and customer engagement. Leveraging crowdsourcing, purpose driven social collaboration, game mechanics and big data analytics, Spigit helps companies identify and execute transformative ideas from their employees and customers at scale to drive business outcomes.

We have found that there are several key elements to a healthy innovation network, that all must be modeled and nurtured in a collaborative system in order to ensure optimal results. In this chapter we examine several of these elements in turn, and provide insight into current and future models and approaches for encompassing these elements in a collaborative innovation system.

In section "Modeling the Innovation Network" we examine the innovation network itself—what it consists of, its typical characteristics that build the foundation for an intelligent collaborative system around it. In section "Innovation Network Engagement" we discuss the role of engagement in innovation, and how to model and measure it. In section "Social Recognition and Rewards in Crowd Innovation" we examine the role of social recognition in crowd innovation—how it manifests and what behaviors and structures in the network support it. We continue in section "Future Work on Innovation Network Optimality Factors" by examining future aspects we are exploring to maximize the process of innovation through intelligent systems, and then conclude in section "Summary of Human Computation in Innovation" with a summary.

Please note that the mathematical expositions in this chapter are provided in order to reinforce the concepts for math-literate readers, but math is not required to understand the concepts herein. Readers may skip the formal expositions and still be confident about being able to follow conceptually.

Modeling the Innovation Network

An innovation network is a complex network of people and ideas, which can be represented as a graph consisting of vertices/nodes and edges/links. A vertex represents a person or idea in the network, and an edge represents some sort of connectivity between two vertices. One can think of an innovation network as a social graph with an additional layer of idea nodes integrated into it. In essence, the network consists of people, ideas, and the interactions between them.

Let the crowd 'C' be represented by a graph $G = (V,E)$ where $V = \{v_1, v_2, \ldots, v_n\}$ and $E = \{e_1, e_2, \ldots, e_m\}$ where n is the number of nodes and m is the number of edges.

In a typical *social* graph, people are the nodes and edges represent relationships between nodes, with the graph depicting the structure of how people are 'related' to one another through their relationships.

In an innovation network, nodes are not only people, but also ideas, and edges are not only the typical *person-to-person* connections, but also *person-to-idea* connections, capturing explicitly the additional interactions that occur on an idea by a person: an up vote, a down vote, comment, review, market trade, view, sharing with a friend, testimonial, workflow task, etc.

In this way in an innovation network, the ideas serve as connection hubs—connecting people nodes that otherwise would not be interacting or connected in a typical social network. What is interesting is that both traditional social networks as well as innovation networks exhibit small world properties—clustering and short paths (Watts and Strogatz 1998; Hardas 2013). The difference is that in a social network the clustering occurs between people, and in an innovation network the clustering happens around ideas. The idea layer brings the network together, with idea clusters that act as magnets to bring people together across the far reaches of the network.

Adding ideas as nodes to the network is a key aspect to scaling innovation, as otherwise the typical social clusters tend to magnify existing silos (groups of people that interact only with one another), and (act as silo magnifiers) as a network forms and operates in an enterprise—counteracting the diversity desired for healthy collaborative innovation. Next let's examine how to assess the health and engagement of the innovation network.

Innovation Network Engagement

For the innovation process to yield maximal results, the innovation network must be engaged and active. In an innovation network, it is not enough for the network to simply grow in number of nodes (ideas and people) to signal a healthy network as Metcalfe's law supposes for standard communication networks (http://en.wikipedia.org/wiki/Metcalfe's_law). Metcalfe's law makes a variety of assumptions on the type of structures in the network. Most real world complex networks are not homogeneously linked by similar types of edges. In an innovation network, the actual activity across nodes—in this case edge formation around idea nodes—must emerge and remain strong for optimal innovation results.

But how does one assess the engagement and activity growth or decline in an innovation network? Surely it ebbs and flows and changes over time. Thus having a computation that can serve as the engagement thermometer enables insight into the relative productiveness of the crowd at each moment in time.

We have created a model for measuring engagement in terms of the entropy of the innovation network (Hardas 2013). In our case, we define entropy as a measure of message activity flux. Entropy is calculated as a function of the probability frequency distributions of the incoming and outgoing messages, which represents the entropy in the activity over the network. And in this way we can translate the activity occurring over a network in terms of message exchange as a measure and prediction of the ongoing engagement of the network.

The engagement of a node is calculated in terms of the incoming and outgoing message entropy, which is the entropy of the incoming and outgoing probability distributions associated with a particular node. The measure of uncertainty is actually the information content in a distribution. Thus the entropy of an incoming and outgoing message probability distribution measures the information content in these distributions. The cumulative incoming and outgoing message entropies of a network are calculated as the summation of all the individual incoming and outgoing node entropies. Thus, the incoming entropy (1) and outgoing entropy (2):

$$H^{in} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} * x_{ij} \log\left(\frac{1}{x_{ij}}\right) \tag{1}$$

$$H^{out} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} * y_{ij} \log\left(\frac{1}{y_{ij}}\right) \tag{2}$$

Where x_{ij} is the incoming message distribution, y_{ij} is the outgoing message distribution, and $a_{ij}=0$ if i and j are not connected, and $a_{ij}=1$ if i and j are connected.

Finally, the total value of a network is calculated as a weighted measure of the incoming and outgoing entropies of the network. Thus, the value V is represented as a function of weighting variable α as in Eq. 3. The variable α allows for weighing the incoming and outgoing network entropies.

$$V(\alpha) = \alpha H^{in} + (1-\alpha)H^{out} \tag{3}$$

As it turns out, some of the structural features of a network indicate a disposition towards good engagement/cumulative entropy (Hardas 2013): high total number of active links, high clustering coefficient, low average shortest path length, and many connected components. Thus with this model, we can quickly assess not only the current engagement level of an innovation network, but also whether that network is predisposed towards engagement growth or decline—and thus recommend adjustments to optimize.

One of the aspects that plays strongly towards engagement growth is social recognition. In section "Social Recognition and Rewards in Crowd Innovation" we now examine a model for recognition in an innovation network, and how it interplays with engagement.

Social Recognition and Rewards in Crowd Innovation

One of the key behavioral aspects that drives good engagement in an innovation network is *social recognition*. This is due to the fact that social recognition is in reality an important motivator. Going back to Maslow's hierarchy of needs (http://en.wikipedia.org/wiki/Maslow's_hierarchy_of_needs) we can see that the 'esteem' layer of needs encompasses accomplishment, social status, attention, recognition needs.

Thus a key element to achieving good engagement in an innovation network is some way to model and externalize this esteem layer for social recognition. We have developed a model for reputation in an innovation network that captures in essence what the crowd thinks of an individual's contributions and interactions. We base this on the 'reaction' edges (votes and comment sentiment) to that person's idea nodes and comment edges in the innovation network.

However, it is not enough to simply look for volume of these edges signaling popularity—this would be a very shallow measure that does not over time scale to provide the motivational behavior desired, as one can quickly understand that simply soliciting volume of positive votes and comments on your ideas provides high reputation, artificially bringing everyone with any sort of activity on their contributions to the same level. We must be more intelligent about those who are positively reacting to an individual's contributions and factor in other measures as we incorporate their reactions into a reputation measure.

Key factors that we have found crucial to a more accurate and effective reputation measure are: decay over time, slower movement at the tails of the permissible range for the reputation score, continued engagement, reputation of the rater, and ability of the rater to discern good ideas from bad (Hardas 2012).

The first key aspect to this model is that an individual's actual rating on an idea or comment (R^A) is not taken at face value, but an effective rating (R^E) is computed from it. The modulation factors of reputation of the rater (m_1), discernment of the rater (m_2), how recently the rating happened (m_3) all contribute to the effective rating as in Eq. 4:

$$R^E = m_1 * m_2 * m_3 * R^A \tag{4}$$

The discernment of the rater is a measure of how skilled the rater is at finding the best ideas. For this, we use the Wisdom of the Crowds (WoC) principle (Surowiecki 2004) that tells us that the aggregated judgment of a number of individuals is closer to the answer than any of the 'best' individual estimates. And thus with this principle the crowd always comes up with the true value of an innovation. Thus, in this model, we predict the discernment of the rater based on: the history of the voter to side with the crowd, and the evidence about the idea in terms of what the crowd thinks.

These two are defined by Event C (the hypothesis) = rate with the crowd. Event I (the data/evidence about the value of the idea) = the cumulative crowd sentiment about the idea. Then the overall sentiment about an idea is computed as follows:

$$P(event\ I) = \frac{upRatings - downRatings}{totalRatings} \tag{5}$$

Now we can compute the probability of the hypothesis i.e. rating with the crowd given what the crowd is thinking about the idea. This probability is modeled using Bayesian inference as in Eq. 6,

$$p = P(C \mid I) = \frac{P(C) * P(I \mid C)}{P(I)} \tag{6}$$

where *P(C|I)* is the posterior probability which we are calculating. *P(C)* is the prior probability—the probability of the person voting with the crowd, i.e. the person's history of voting for the good idea. *P(I|C)* is the likelihood—the probability of the idea being a good/bad idea given the rater rates with the crowd. P(I) is the data/evidence about the idea, the probability that the idea is good/bad given the voter votes with/against the crowd respectively, given by Eq. 7:

$$P(I) = P(I \mid C) * P(C) + P(I \mid \sim C) * P(\sim C) \tag{7}$$

The overall probability p thus becomes the discernment of the rater—how probable is it that this rater tends to rate with the wisdom of the crowd, and becomes m_2 in our effective rating computation, to modulate the rating according to how discerning this rater is.

The sum of the effective ratings is then coupled with continued engagement e of the individual, a decay factor d over time and a tail velocity v that slows movement at the tails to define the overall reputation of an individual at time t:

$$R_t = d * R_{t-1} + \sum_{1}^{n} R_n^E * v * e \tag{8}$$

And in this way, we have a reputation measure that serves as a social recognition metric that goes much beyond simple volume of positive ratings to provide a reputation measure that elicits the engagement behaviors desired, and is robust to gaming as the innovation network grows.

Future Work on Innovation Network Optimality Factors

Several additional factors are key to ensuring productive and effective scaling and operation of an innovation network. Modeling and incorporating *trust* is one such factor. Finding particular *innovation personae* in the network and the optimal groupings and makeup of these personae in well-functioning innovation networks is another. And *emergent cooperation* and the factors and conditions necessary in the network to ensure cooperation grows rather than shrinks is a third important future area for exploration.

Network trust has been studied in social network and general collaboration contexts (Thirunarayan et al. 2010; Beckett and Jones 2012). We have also observed various trust behaviors that impact innovation network engagement, such as more readily-given trust in smaller innovation networks than in larger ones. In future work we will examine whether the traditional trust behaviors such as preference similarity (Liu et al. 2011) and frequent and regular communication (Abrams et al. 2003) are also the same factors that engender trust in an innovation network, and/or what additional trust metrics exist in the innovation context.

We have also observed that typical healthy innovation networks have a mixture of behavioral personae (innovation behavior types), for example innovators (skilled at creating good ideas), and discerners (skilled at finding good ideas of others). And that the mix of innovators and discerners remains fairly constant as the healthy innovation network scales. In future work we will explore additional personae that contribute to an especially healthy innovation network, such as creative collaborators, action-takers, etc., and in what contexts are they most effective and necessary, in order to be able to recommend innovation team groupings for optimality.

And finally, studies about emergent cooperation in social networks show that certain conditions favor cooperators over defectors in a network, ensuring that cooperation grows rather than shrinks (Ohtsuki et al. 2006). We will also examine as part of future work whether these conditions are also sufficient and necessary in innovation networks, and how the cost/benefit scenario can best be modeled to achieve emergent cooperation in innovation.

Summary of Human Computation in Innovation

Successful innovation has a distinct and crucial human element. Scaling innovation successfully requires encompassing and harnessing the knowledge of groups of individuals to provide outcomes greater than any one individual could achieve on their own. To provide a functioning computational system for true innovation at scale requires approaches for modeling and incorporating people's behaviors, trust, emergent crowd wisdom, social ties, rewards.

In this chapter we have examined how to model the innovation network as a social network layered with ideas that brings the social clustering across the network rather than in social silos. We have examined the role of engagement in an innovation network, and how it can be modeled and measured with a network entropy approach, as well as how to find structural indicators of whether a network is predisposed towards high or low engagement. We have also examined the crucial role of social recognition in an innovation network, and how to model reputation in such a way that it elicits desirable interaction behaviors, is robust to gaming, and scales as the network grows. And finally we examined the additional aspects of trust, emerging cooperation, and personae that are additional future key elements to incorporate into optimal innovation network models and approaches.

By modeling and incorporating these key factors into an intelligent human-machine computation system, the innovation network is positioned best to harness the knowledge of the group, and produce optimal innovation results.

References

Aberdeen Group (2009) Managing the innovation portfolio. In: Enabling engineering success to boost profits, Boston MA

Abrams LC, Lesser E, Levin DZ (2003) Nurturing interpersonal trust in knowledge-sharing networks. Acad Manage Exec 17(4):64–77

Beckett R, Jones M (2012) Collaborative network success and the variable nature of trust. Prod Plann Control 23(4):240–251

Hardas P (2012) Bayesian vote weighting in crowdSourcing systems. In: Computational collective intelligence, technologies and applications. Springer-Verlag Berlin, Heidelberg, pp 194–203

Hardas P (2013) Computing the value of a crowd. In: Social computing, behavioral-cultural modeling and prediction. Springer-Verlag Berlin, Heidelberg, pp 248–255

http://en.wikipedia.org/wiki/Maslow's_hierarchy_of_needs

http://en.wikipedia.org/wiki/Metcalfe's_law

Liu G, Wang Y, Orgun MA (2011) Trust transitivity in complex social networks. In: Proceedings of the 25th AAAI conference on artificial intelligence. San Francisco, CA

Ohtsuki H, Hauert C, Lieverman E, Nowak MA (2006) A simple rule for the evolution of cooperation on graphs and social networks. Nature 441(7092):502–505

Surowiecki J (2004) The wisdom of crowds: why the many are smarter than the few and how collective intelligence shapes business, economies, societies, and nations. Brown Publishing, Doubleday, New York

Thirunarayan K, Anantharam P, Henson CA, Sheth AP (2010) Some trust issues in social networks and sensor networks. In: International symposium on collaborative technologies and systems. Chicago, IL

Watts D, Strogatz S (1998) Collective dynamics of small-world networks. Chicago, IL, Nature 393:440

Human Computation for Organizations: Socializing Business Process Management

Marco Brambilla and Piero Fraternali

Introduction

The success of social networks, which have become the most intensely used applications on the Web, has demonstrated the centrality of communities of practice, whereby users can interact with the service providers and among themselves, to be informed, share experience, and express their opinion on the quality of products and services. This socialization of the user's online experience, for customers, citizens, or employees, will not be confined to the personal activities, but will carry over to the practices in the workplace, changing the interaction between organizations, employees, and customers. This shift will impact the way in which organization define and run their processes, which will more and more evolve from *closed* to *open and social*.

Social Business Process Management (also known as *Social BPM* or *SBPM*) is the approach that studies the integration of social interactions and business processes: it fuses business process management practices with social networking, with the aim of enhancing the enterprise performance by means of a controlled participation of external stakeholders to process design and enactment (Dengler et al. 2010; Erol et al. 2010; Johannesson et al. 2009; Koschmider et al. 2009; Schmidt et al. 2010). Notice that SBPM is not primarily concerned with human computation activity that simply produces some input for a business process; SPBM is mainly focused on socially enacted processes and social interactions intertwined with business processes.

In classical BPM, processes are defined centrally by the organization and deployed for execution by *internal performers*, i.e., actors formally entitled to execute the activities and directly advances a process case. This closed-world approach can be opened to social actors with different levels of control (Brambilla et al. 2011): from employees not normally entitled to participate to the process to unknown users

M. Brambilla (✉) • P. Fraternali
Politecnico di Milano, Dipartimento di Elettronica, Informazione e Bioingegneria, Via Ponzio 34/5, Milano, Italy
e-mail: marco.brambilla@polimi.it; piero.fraternali@polimi.it

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_21,

who contribute by acting on public social media platforms. Due to this variety of potential contributions, central to the successful implementation of SBPM is the possibility of controlling the tradeoff between centralized control (typical of BPM systems) and openness (typical of social applications) *specifically* for the business context of the target organization.

To respond to the demand for SBPM methodologies that allow a controlled integration of social interaction features within business processes, this chapter discusses a model-driven approach to participatory and social enactment of business processes. The key idea is to identify the goals of existing social interactions within the organization, turn these goals into social extensions of a business process model that supports the organization in achieving the stated goals, and provide powerful code generation and integration tools for the rapid prototyping and evolutive maintenance of the social business processes, so to allow the early trial and quick reconfiguration of the modelled social interactions. The proposed method is being implemented within the BPM4People project,[1] a Seventh Framework Programme project funded under the SME Capacities program of the Research Executive Agency.[2]

The Goals of Social BPM

- The ultimate goal of SBPM is improving a process by opening it to the contribution of more people. This general objective can be regarded as a process optimization phase, in which the organization seeks efficiency by extending the reach of a business process to a broader class of actors. This general motivation can be articulated more precisely into different sub-goals:**Exploitation of weak ties and implicit knowledge**: the goal is discovering and exploiting informal knowledge and relationships to improve activity execution.
- **Transparency**: the goal is making the decision procedures internal to the process more visible to the affected stakeholders.
- **Participation**: the goal is engaging a broader community to raise the awareness about, or the acceptance of, the process outcome.
- **Activity distribution**: the goal is assigning an activity to a broader set of performers or to find appropriate contributors for its execution.
- **Decision distribution**: the goal is eliciting opinions that contribute to the taking a decision.
- **Social feedback**: the goal is acquiring feedback from a broader set of stakeholders, for driving process improvement.
- **Knowledge sharing**: the goal is disseminating knowledge to improve task execution; at an extreme, this could entail fostering mutual support among users to avoid performing costly activities (e.g., technical support).

[1] http://www/bpm4people.org
[2] http://ec.europa.eu/rea/

Not all the aforementioned motivations apply to a specific business scenario. For example, a company may wish to exploit customers' opinions to decide among alternative product features in its product planning and design cycle (thus focusing only on decision distribution); another one may want to improve the process for staffing a new project by fostering the emergence of hidden internal competencies (thus requiring the exploitation of weak ties and implicit knowledge). As these examples show, quite different forms of social interactions should be modeled, implemented, and deployed, to meet the goals of specific organizational and business contexts. Several conceptual and technical ingredients are needed to support such a flexible approach: a methodology for deriving SBPM requirements from business goals, a notation for specifying the social aspects of processes, and tools for enabling social BPM rapid prototyping, deployment, and execution, directly from process models. SBPM is a perfect candidate for applying OODA (Observe, Orient, Decide, and Act) techniques; in particular, the Orientation phase that determines the way to observe, decide, and act, is crucial in SBPM because it considers the repository of our genetic heritage, cultural tradition, and previous experiences.

The SBPM practice in the industry is still in its infancy: even the most important BPM players only provide restricted forms of social interactions, which can be attached rather artificially to their standard BPM systems.

The BPM4People Approach

The BPM4People approach aims at applying Web engineering and model-driven development techniques to Social BPM. Its contributions can be summarized as follows:

- The identification of the main factors that drive the socialization of a business process (the socialization goals mentioned in section "The Goals of Social BPM").
- An extension of BPMN 2.0 enabling the specification of social roles, activities, and events (Social BPMN).
- A gallery of design patterns, expressed in Social BPMN, that represent archetypal solutions to recurrent process socialization problems (Social Process Patterns). Social patterns are referred to the goals they contribute to solve, and in this way help the construction of process models from business requirements.
- A technical framework for generating Social BPM applications from Social BPMN models, based on model transformations and on a runtime architecture integrating business process execution and social task enactment, implemented in a commercial tool suite called WebRatio (Brambilla et al. 2010).

A quite unique feature of the BPM4People framework is *one-click social process prototyping*. The model-driven approach and default code generation rules allow the business analyst to create a running prototype of the business process with all the social interaction functions fully implemented and integrated with a number of different platforms (including, public social networks like Twitter, Facebook, G+ and

LinkedIn, as well as services for ad hoc online activities, like Doodle polls and Google doc cooperative document editing). This functionality enables the early evaluation of the process activities offered to social actors, a key capability for assessing alternative social network interactions and for reducing the number of re-design cycles.

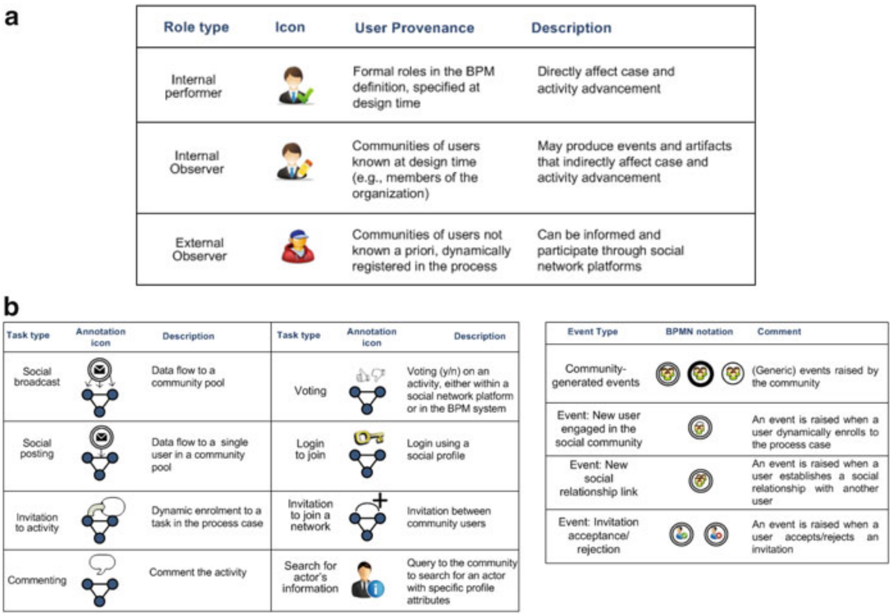
The Social BPMN Notation

Process design benefits from visual languages that convey the process structure and constraints in an clear way, immediately communicable also to non-technical stakeholders. Social process design should preserve the intuitiveness and expressivity of state-of-the-practice visual languages and build on standard notations already in use. To this end, social extensions of business processes can be conveyed using BPMN 2.0,[3] which incorporates a native extension mechanism. By enriching the existing BPMN concepts with a social meaning, it is possible to achieve a visual language that is both familiar to BPMN practitioners and possess enough expressive power to convey social behaviors. The Social BPMN extension proposed by BPM4People expands BPMN collaboration diagrams with social roles, events and activities (Brambilla et al. 2011).

The dimensions along which we have extended the BPMN metamodel include the specification of process roles with different degree of control, social task and event types representing the monitoring of the outcome of activities performed in social networks, the publication of activity to be executed in social networks by social actors, the publication of social content, and the definition of the platform supporting the social interactions. Figure 1a illustrates the notation for representing the different classes of roles that can participate to a social business process. Figure 1b exemplifies the notation employed for expressing the Social Task Type and the Event Type extension of the BPMN Task and Event concepts: a BPMN task or event can be visually stereotyped to denote that the activity is executed or the event is raised by social interactions, like voting, commenting, inviting more people to execute the task, etc.

An example of model expressed with Social BPMN is shown in Figure 2; it represents a social process started by the supervisor of a local government office, who selects and publishes some quality indicators on a public social network, for the citizens to express their opinions in the form of votes in a quality scale for a specific indicator or with free format comments. The voting and commenting activities take place in a *social pool* (denoted by a different icon), which means that execution is deployed onto a social platform; social interaction results produced by the social pool are then collected by a social monitoring task in the process pool denoting the PA office; after that step, the supervisor analyzes the data and broadcasts a summary of the results back to the social pool.

[3] http://www.bpmn.org/



a

Role type	Icon	User Provenance	Description
Internal performer		Formal roles in the BPM definition, specified at design time	Directly affect case and activity advancement
Internal Observer		Communities of users known at design time (e.g., members of the organization)	May produce events and artifacts that indirectly affect case and activity advancement
External Observer		Communities of users not known a priori, dynamically registered in the process	Can be informed and participate through social network platforms

b

Task type	Annotation icon	Description	Task type	Annotation icon	Description
Social broadcast		Data flow to a community pool	Voting		Voting (y/n) on an activity, either within a social network platform or in the BPM system
Social posting		Data flow to a single user in a community pool	Login to join		Login using a social profile
Invitation to activity		Dynamic enrolment to a task in the process case	Invitation to join a network		Invitation between community users
Commenting		Comment the activity	Search for actor's information		Query to the community to search for an actor with specific profile attributes

Event Type	BPMN notation	Comment
Community-generated events		(Generic) events raised by the community
Event: New user engaged in the social community		An event is raised when a user dynamically enrolls to the process case
Event: New social relationship link		An event is raised when a user establishes a social relationship with another user
Event: Invitation acceptance/ rejection		An event is raised when a user accepts/rejects an invitation

Fig. 1 (**a**) Social BPMN notation for expressing different roles in a social business process; (**b**) Example of Social BPM notation for some of the social task and event types

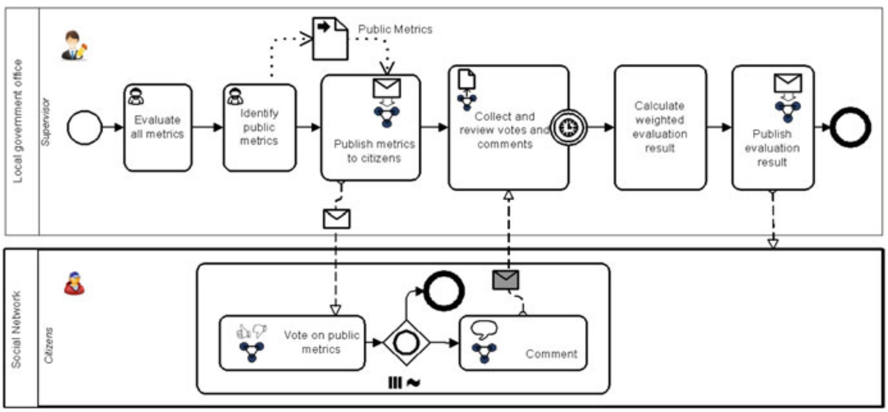


Fig. 2 Social BPMN model of a public administration process where citizens are requested to evaluate service performance. The example features social behaviour tasks (e.g., publish metrics to citizens), a social pool and lane (social network and citizens), and social monitoring tasks (e.g., collect and review votes and comments)

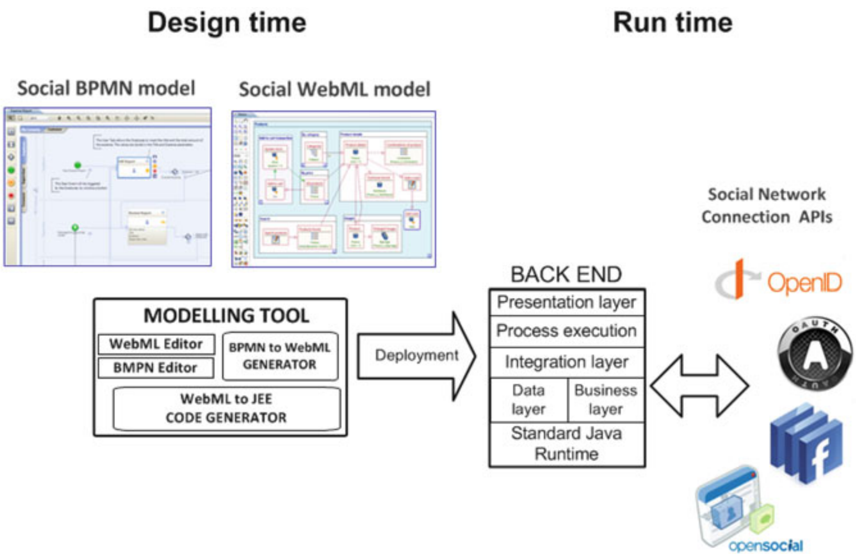


Fig. 3 Architecture of the BPM4People social BPM system

Rapid Prototyping of SBPM Solutions

To support the design, implementation, deployment, and monitoring of Social BPM solutions, BPM4People has constructed the technical architecture and tools shown in Figure 3. The architecture consists of a **design time** part, comprising a Model-Driven Integrated Development Environment, and of a **run time** part, where automatically generated applications are run on a standard JEE platforms connected via Web Service APIs to one or more Social Networking Platforms.

The distinguishing feature of the BPM4People approach is a **two-level Model Driven Development approach**, which exploits two modelling languages and two transformations: the two modeling levels are represented by **Social BPMN models** at the CIM level and by **WebML models** at the PIM level. WebML (Ceri et al. 2002) is a platform independent language for modeling interactive applications, which provides a native extension mechanism that has been used to incorporate new **social components** into the language. These social WebML components implement atomic social interaction functions (e.g., voting, commenting, polls, friend invitation, etc.) and support the connection to Social Networking Platforms via API invocations. Then, two model transformations enable the BPM4People quick prototyping approach:

- The model-to-model transformation from Social BPMN to WebML automatically maps Social BPMN activities into WebML application models for performing such activities, also in the case in which their execution requires the interaction with a designated social media platform.
- A model-to-code transformation automatically produces standard JEE code from WebML.

Implementation Experience

The social components that describe at PIM level the basic social interactions have been implemented and included in the WebRatio tool. They are available online in the WebRatio store (reachable at: http://www.webratio.com/store). Furthermore, some demonstrative applications have been implemented to showcase the BPM4People approach. A simple demonstration scenario for the social creation and execution of a poll about a meeting date is published on the BPM4People site at: http://www.bpm4people.org/cms/content/en/demos. A video shows the modeling approach step-by-step and describes the running application generated automatically and integrated with LinkedIn (for retrieving the users contacts) and with Doodle (for performing the poll).[4] Figure 4 shows a screenshot of the generated application, which displays the interface of a social task that requires the communication with people retrieved and selected from the user's LinkedIn contacts.

Additional applications have been developed that support: the fusion of crowdsourced input with enterprise knowledge management and information seeking tools (Bozzon et al. 2012); the embedding of social interfaces and enterprise processes in the domain of restaurant reservation and tourism services; the integrated management of corporate ideas crowdsourcing, microblogging, expert finding, and team building; the cloud-based publishing, sharing, and renting of extra-capacity resources of organizations (e.g., cars, office space, equipment, and personnel).

As an example, we provide some details of the extra-capacity social sharing application.

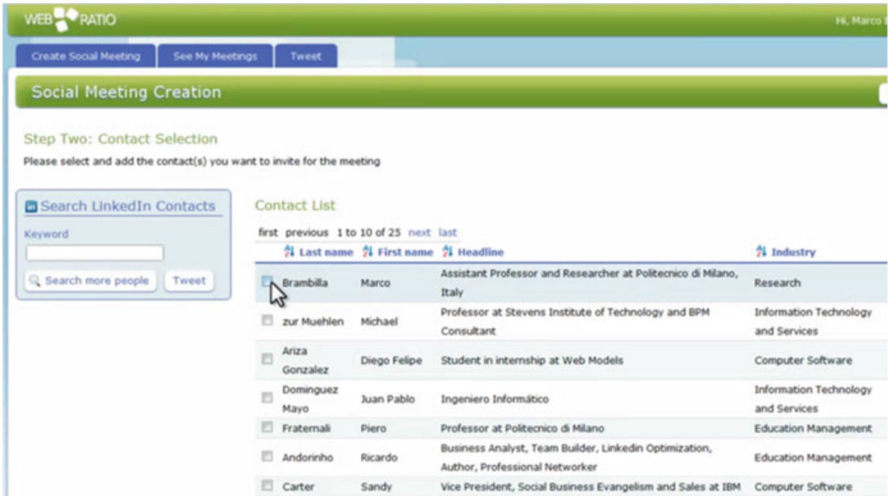


Fig. 4 Screenshot of social meeting application, generated automatically from a social WebML model

[4] The video is also available on YouTube at the following URL:http://www.youtube.com/watch?v=7qNV1Iw1oiA

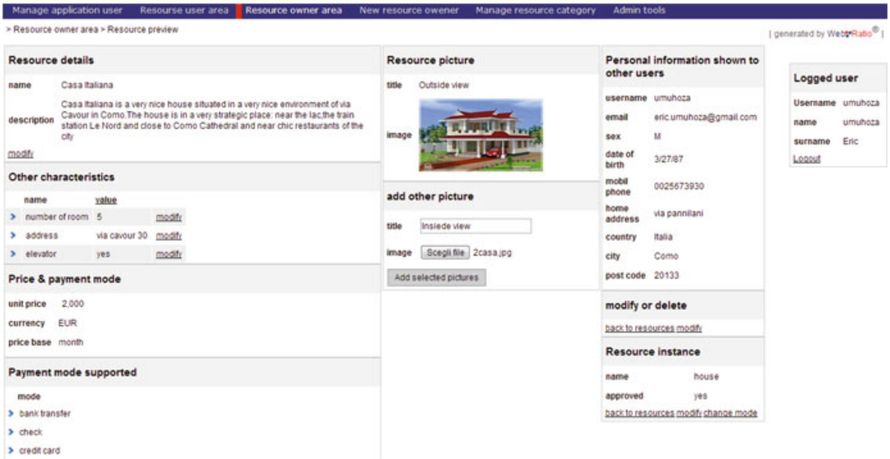


Fig. 5 Screenshot of social sharing application, showing the creation of a new resource (a house)

The Social Capacity Sharing Application

The Social Capacity Sharing application allows enterprises to exploit their extra capacity by sharing, renting or leasing assets that would otherwise remain unused for a certain amount of time. The social side of the application is that the company's extra capacity is offered on both a dedicated portal and on public social networks, with resources of different types automatically extracted from the enterprise systems based on their scheduled availability. Posting on social networks is used to improve the visibility of rentable resources, by encouraging people outside the company to spread the word to their acquaintances. The adoption of SBPM supports the integration in the same process of enterprise and social activities (e.g., by allowing automatic sharing of resources when these appear as free in the enterprise ERP system).

The application supports a workflow for the definition of resource types (e.g., cars, office space, houses), the publication of available resources of those types, and the reception of resources requests from customers. Resource availability can be advertised on multiple social networks, so as to spread the word and to collect feedback and declarations of interest.

Three roles can participate to the process: administrators, resource owners, and resource consumers.

An administrator can define the types of resources that can be shared; approve/reject new resource types proposed by other users; and manage the users subscriptions. The resource owner can publish his resources on the social networks of choice; disseminate such a publication to his friends/followers on enterprise or public social networks; monitor how the community reacts to his offers through a social activity monitor; and approve or reject bids from social network users. Figure 5 shows the result of the creation of a new resource (in this example, a house) to be shared.

Its properties have been filled in by the house owner through a blueprint dynamically designed at runtime by the administrator.

The consumer can search available resources; evaluate resources and owners from their community ranking, produced from people's votes; make a bid for a resource; post a comment on the application or on a connected social network about resources or providers; and propose new resource types. The social sharing application calculates the ranking of resources and owners according to criteria such as number of comments (likes or dislikes), etc. This ranking is made available to all users and may help resource owners improve the quality of their service, as in traditional marketplaces.

Conclusions

In this chapter we have defined an emerging area of Human Computation devoted to the integration of social interactions within the business processes of organizations, called *Social BPM*. We have then presented the BPM4People approach to Social BPM, which consists of identifying the goals of social BPMN within the organization, expressing the required social business processes by mapping the business goals to social BPM process patterns that support their realization, formalized using a notation called Social BPMN. Social BPMN extends BPMN 2.0 with social pools, tasks, and events, and is accompanied by a technical framework that allows enterprises to rapidly implement social business processes in a flexible and semiautomated way, on top of multiple social platforms.

Acknowledgements BPM4People (http://www.bpm4people.org) is partially funded by the Research Executive Agency of the European Commission, within the SME Capacities Program of the 7th FP. We wish to thank all the project partners: University of Trento, Universidad de Extremadura, Enterprise Concept, Homeria Open Solutions, and Nexture.

References

Bozzon A, Brambilla M, Ceri S (2012) Answering search queries with CrowdSearcher. In: World wide web conference (WWW), Lyon. ACM, pp 1009–1018

Brambilla M, Butti S, Fraternali P (2010) Webratio BPM: a tool for designing and deploying business processes on the web. In: 10th international conference on web engineering (ICWE), Vienna, pp 415–429

Brambilla M, Fraternali P, Vaca C (2011) A model-driven approach to social BPM applications. In: Social BPM. Future strategies Inc. and WfMC, pp 95–112. http://www.futstrat.com

Brambilla M, Fraternali P, Vaca C (2011) A notation for supporting social business process modeling, In: 3rd international workshop on BPMN, vol 95, Lucerne. LNBIP. Springer, pp 88–102

Ceri S, Fraternali P, Bongio A, Brambilla M, Comai S, Matera M (2002) Designing data-intensive web applications. Morgan Kaufmann, San Francisco/Calif

Dengler F, Koschmider A, Oberweis A, Zhang H (2010) Social software for coordination of collaborative process activities. In: Third workshop on business process management and social software, Hoboken, Sept 2010. LNBIP, pp 396–407

Erol S, Granitzer M, Happ S, Jantunen S, Jennings B, Johannesson P, Koschmider A, Nurcan S, Rossi D, Schmidt R (2010) Combining BPM and social software: contradiction or chance? J Softw Maint Evol 22:449–476

Johannesson P, Andersson B, Wohed P (2009) Business process management with social software systems-a new paradigm for work organization. In: BPM 2008 International Workshops, Milano, Italy, September 1–4, 2008. Revised Papers, LNBIP. Springer, Berlin/Heidelberg, pp 659–665

Koschmider A, Song M, Reijers HA (2009) Social software for modeling business processes. In: First workshop on BPM and social software, Milan. LNBIP

Schmidt R, Dengler F, Kieninger A (2010) Co-creation of value in IT service processes using semantic mediawiki. In: BPM Workshops, Hoboken. LNBIP

Solving Wicked Problems

Dan Thomsen

Human computation with massive crowds has the potential to solve problems never solved before. People can clearly define many unsolved problems, such as a human mission to Mars. However, a class of unsolved problems eludes an easy definition. Termed “wicked” these problems have a dynamic ill-defined nature that introduces difficulties to the solution process especially when trying to get thousands of humans to work together on the same problem. Wicked problems present such difficulty, it may be difficult to tell if a specific solution solves the problem or not.

The first examples of wicked problems arose from social policy planning, where any solution changes the problem specification and as a result the solution may be insufficient. Often wicked problems arise with multiple parties with different success criteria. As a result, no solution can maximize “happiness” across the entire population. The Middle East peace process provides an example of a wicked problem. Many challenges from wicked problems arise from people having different, incompatible, worldviews. Until the population shifts to compatible worldviews the wicked problem remains challenging to solve.

What does the existence of wicked problems mean for human computation? For the most part human computation assumes that the humans involved are working toward the same goal. When attempting to solve a wicked problem the human solvers could easily work at cross-purposes due to their different worldviews. Even non-wicked problems suffer from human solvers with different worldviews. For example, a sub-population of solvers could attempt to suppress potential solutions that do not fit their worldview.

Technology cannot pick the correct worldview or even average worldviews, however technology may causes people to shift their worldview. Let’s assume the problem sponsors, the people who wanted to solve the problem in the first place, truly want a solution found regardless of worldview. The sponsors set up the human

D. Thomsen (✉)
SIFT, LLC, Minneapolis, MN, USA
e-mail: dthomsen@sift.net

P. Michelucci (ed.), *Handbook of Human Computation*,

DOI 10.1007/978-1-4614-8806-4_22,

computation environment and set the initial goals. With a wicked problem the goals will most likely change. The sponsor must let the goals change in the hopes of finding a workable solution. However, one goal remains constant, educating the human solvers on different world viewpoints.

A cooperative human computation environment can provide not only a forum to solve problems, but a platform to educate the human solvers. Education can include technical information as well as information about different world viewpoints. As a rule people resist changes to their worldview, but given time and enough exposure they may at the very least be able to understand different worldviews allowing a solution to progress.

The problem sponsors must realize this and allow the problem and the solution to evolve over time. Since most people resist changes to their worldview the sponsors cannot recruit solvers by saying it will change they way they look at their world. Instead they must only indicate that the solution requires a great deal of education and knowledge. Before human solvers can participate in certain events and votes on solutions elements they must earn the right to vote by completing a number of educational requirements.

The educational requirements can have the people read through different histories told from different viewpoints and educate solvers on different viewpoints. Solvers must demonstrate an understanding of different viewpoints to contribute to some aspects of the problem. Forcing solvers to self educate, may seem like a stringent requirement, but remember these problems represent some of the most intractable problems known to humanity. You cannot solve things you cannot understand.

A human computation environment to solve wicked problems needs the following characteristics:

- Manage experts as a rare resource
- Acknowledge solutions to unsolved problems operate at the edge of human understanding
- Solving wicked problems require continuous learning on the part of human solvers
- Acknowledge that ideas evolve, this includes problem specifications
- Over come local maxima in a solution
- Manufacture Serendipity
- Motivate human solvers

Experts are a rare resource—Hard problems require specialized expertise. Many problem domains simply do not have thousands of experts available. All participants have different expertise and the computation environment must focus each individual where their skill does the most good. This means creating tasks to harvest the large number of novices to simplify tasks for the few experts. Solutions that create more experts also increase the chance of finding a solution.

Continuous Learning –To solve unsolved problems human solvers must learn continuously to advance the solution. Knowledge provides no value unless it allows a human to make a better decision. Overcoming both technical and wicked problems requires a continuous learning environment that continuously stretches the knowledge and worldview of solvers.

The Edge of Human Understanding—The computation environment must accept that it operates on the edge of human understanding, which means no one knows how to solve the problem. Existing curriculum doesn't directly cover the knowledge necessary to solve the problem so there is no clearly defined path to the solution. The system must build up a body of innovations and unique insights on the problem to make progress, basically developing a dynamic exploratory curriculum. The computation environment must aid humans to systematically explore a multitude of alternative approaches by directing the crowd to try many possibilities. Only one participant needs to find the key insight to solve the problem.

Manufacture Serendipity—Many key scientific discoveries arose from serendipitous combination of new ideas. Using the power of the crowd, the human computation environment can manufacture serendipity by systematically combining new concepts in a problem domain for the crowd to evaluate.

Breaking Functional Fixedness—People have a tough time seeing new uses for familiar objects, a problem called functional fixedness (Duncker 1945). As a result they get stuck on a local maxima for a solution. Technology can help present and combine concepts in new ways to help break functional fixedness. Locked in worldviews represent a particularly challenging form of functional fixedness to overcome.

Ideas Evolve—The human computation environment must create an environment where ideas evolve. In the case of wicked problems, even the problem specification must evolve. Ideas with more potential must spread, where ideas with less potential should be spread less. Understanding idea potential remains a huge challenge for wicked problems.

Motivation in the Face of Failure—Most science involves trying things that did not work. When tackling problems at the edge of human understanding, failure must be expected. The human computation environment must have a flexible rewards engine that creates a sense of accomplishment even in the face of long-term failures. Games and contests present one possibility of providing a feeling of incremental success while working toward a long-term goal.

This chapter explores each of these requirements for a human computation environment if more detail.

Experts as a Rare Resource

Any organization attempts to make good use of its experts to solve its most challenging problems. However, in an organization the experts are known, or at least suspected by the amount of salary they earn. In a human computation environment experts again must play a key role, however most environments will have no idea

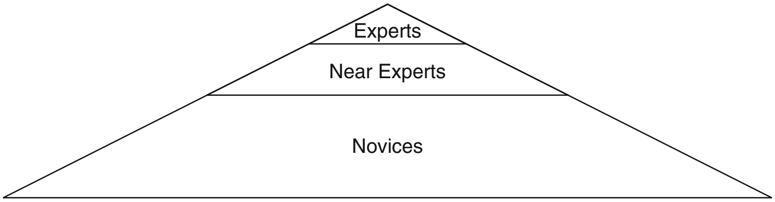


Fig. 1 Solving wicked problems requires using expert and near expert brain power wisely

who qualifies as an expert. Experts come with their own biases, and when tackling a wicked problem those biases could make the expert more of a liability than an asset. So the system must cultivate open minded experts.

Each human computation environment must develop its own metric for finding, testing and then utilizing experts effectively. In areas of established science real world reputation becomes an excellent metric for identifying experts. In the fractionated world of wicked problems, experts with a variety of viewpoints must contribute, but if experts were only picked on real world reputation it almost guarantees failure since the problem hasn't been solved previously.

A population with the skills, but none of the deeply ingrained biases offers the best chance of success. These "near experts" have experience in related fields, or simply a passion for the problem space and a capability to learn. By coming from outside the establishment they possess a variety of viewpoints, and will benefit their own reputation by solving a hard problem in a new domain.

The bulk of the human solvers will be novices that simply have a passion for the problem space. These novices represent the largest part of the labor force, see Fig. 1. Part of the environment will invariable include a number of tasks that track and maintain data, or evaluate ideas. The novices will provide most of this effort, and so the approach must use its experts and near-experts to vet ideas before or after the novices develop them. Because of expert biases, experts cannot dominate the hunt for a solution, but they can prune already explored avenues of research, focusing the human solvers on new ground.

Finding the how to use the various classes of solvers, remains a challenge for technical problems, for wicked problems it becomes even more critical. Clearly one of the best approaches for dealing with the relatively few number of experts is to create more experts, which is addressed next.

Continuous Learning

Both solving technical problems in science and solving wicked problems benefit from continuous education. In the case of wicked problems education can help expose solvers to the worldviews of the entire population affected by the solution; hopefully creating insight to determine more encompassing solutions.

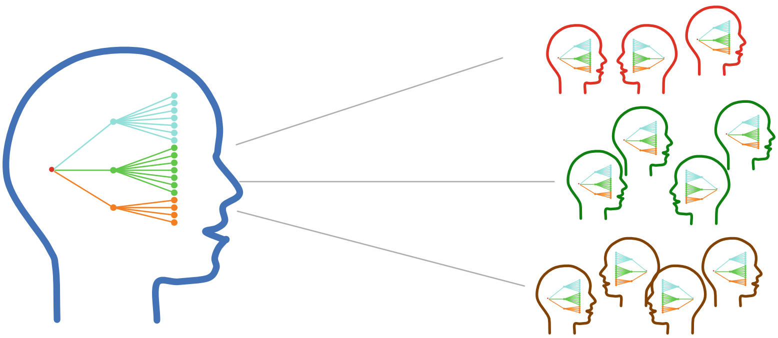


Fig. 2 A human computation environment manages idea exposure to the individual, sub-populations and the crowd as a whole

There are many on-line and computer based educations systems. However, all of these systems require someone to develop a curriculum. When working with wicked and unsolved problems on the edge of human understanding no one has created a curriculum. Often people will even have different opinions if known facts apply to the problem space or not. By taking advantage of the larger crowd many divergent curriculums can be developed and explored at the same time. An ideal human computation environment will understand the various sub-populations that make up the crowd and manage idea exposure.

Idea exposure means keeping track of what concepts have been shown to each member of the crowd and recording if the idea has "stuck". The system does not know if the idea, represents something useful or not by itself, but it can track the spread of the idea. If the human solvers provide some sort of metric on the usefulness of the idea the system can manage the spread of the idea exposing more useful ideas to more people.

As the system learns more about the population of human solvers it can also learn what viewpoints the various sub-populations hold. In science and technology key insights often come from combining science from different domains. The system can group the population by the technology and ideas they favor and then expose them to other viewpoints. The system must remain viewpoint neutral to avoid seeming "preachy" to the people. But human solvers trying to solve wicked problems must realize they have to understand different worldviews if they are going to make progress, see Fig. 2.

A continuous learning system produces a dynamic curriculum that changes based on new discoveries and insights. Solving problems on the edge of human understanding requires such a dynamic curriculum because invariably the wrong knowledge will creep into the curriculum and must be removed. Also if the problem remains unsolved, clearly the existing curriculum doesn't help solve the problem so the solution requires more new material.

A dynamic curriculum resembles building a bridge across an unknown ocean. You don't know if you are taking the shortest path, to a new continent or heading toward a deep trench you may never cross. The problem breaks down into known unknowns, and unknown unknowns (Morris 2010). An example of known unknown might you might have no idea how to build a bridge to withstand ocean storms, but you know at least storms exist. An unknown unknown, might be the existence of a miles deep ocean trench, filled with crustaceans that eat steel. No amount of planning and risk mitigation can address the problem. You must simply try and fail, before you can try and succeed.

For wicked problem unknown unknowns make up most of the solution space. Often evaluating a potential solution to a wicked problem involves implementing it and determining if it made the situation better or worse. Analyzing failures will invariably turn up complex issues that were simply unknown or poorly understood when the solution was implemented. Even the failure may only shed a little light on understanding the problem space.

Ideas Evolve

How can you create an organization and environment to solve a problem if the problem definition keeps changing? Surely such fluctuations will slow progress, but in a wicked problem you may not have understood the problem enough in the beginning to ask the right question.

Wicked problems mean you don't know where you will end up, but computer systems often lack flexibility. The human computation environment must allow ideas to evolve freely, or with minimal direction. Hopefully gentle nudges from time to time can keep to human solvers on track to produce a workable solution. Human computation system must address the question, "Does the system need to set bounds on how far the problem specification can stray?" Without bounds, the system could wind up solving already solved problems, or problems in a totally different domain. If humans set up the bounds on evolution of ideas, the bounds could easily incorporate their own biases and prevent the proper space from being explored.

Human solvers also must learn to let their own ideas and goals evolve. For example, suppose a solver joined a human computation to cure cancer. However, in the course of the exploration they found they could greatly limit the impact of cancer by changing people's diet; that efforts to change diet would provide immediate benefits. Would the human solvers want to switch to the new goal of better nutrition education?

Humans hold on to their biases much tighter than any computer program, actively resisting information and situations that would call those biases into account. A good human computation environment must educate its population of human solvers that biases exist and that they can present roadblocks to effective solutions. Such education can slowly chip away at biases hopefully leading to a solution.

Creating the Eureka Moment

In 1945 Duncker discovered that if an object was being used for one purpose people had a hard time seeing it used for a new purpose (Duncker 1945). In his experiment people had to affix a candle to a wall and light it without burning the wall. They were given a box of tacks, a book of matches and a candle, see Fig. 3.

The solution is to empty the box of tacks and tack the box to the wall and put the candle in the box. If the tacks were in the box it took the participants twice as long to solve the problem. Duncker called this problem functional fixedness. People see the box as a container to hold tacks. An empty box leads people to see it as a potential shelf in half the time.

Technology can help remove the impact of functional fixedness by organizing knowledge into its basic components and having the system randomly combine the components, leaving out key facts or relationships. With a large crowd people can look at the different sets of facts and see if they can draw different conclusions. Injecting new random facts into the description can extend the same approach at the cost of solvers learning incorrect knowledge. Human solvers provide the mental computation to understand the situation, while the computation environment can manage the many possibilities ensuring that people keep seeing unique combinations of facts. If a fact turns out to be wrong, the system knows which solvers were exposed to the fact and it can reeducate that individual.

Another way to look at controlling the facts that a solver, could be as manufacturing serendipity. Imagine the story of Archimedes in his bath discovering the displacement principle to calculate volume. He was tasked with solving a hard problem; figuring the volume of an irregular solid. The full bath combined with the problem in his brain allowed him to leap to a solution. This eureka moment arose because of the unique combination of circumstances; a full bath, a dirty philosopher and a problem that needed solving. Take away any one element of the situation and the insight would not have happened.

But what if you had a thousand Archimedes? Task them all with solving the volume problem and put them into different situations; some walking, some

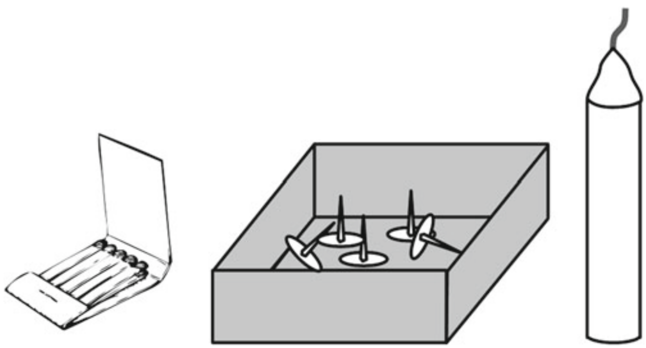


Fig. 3 The candle problem illustrates functional fixedness

skydiving, etc. How long until one of them jumps into a full bath, or a full rain barrel and has the eureka moment? Even if a potential Archimedes jumps in a full bath and doesn't have the insight, another one eventually will.

Almost all difficult problems require some insight, which starts the investigation on the correct path. To date most scientists solve hard problems by systematically looking at alternatives, taking educated guesses to narrow the search space down to something they can investigate. A human computation environment can combine the facts for humans to evaluate. Most of these combinations will have zero value, but if the time investment for a single individual is small, you can explore hundreds of thousands of combinations in a few weeks, greatly increasing the chance of at least one human solver experiencing the eureka moment.

Motivation in the Face of Failure

A wicked problem will most likely take years to solve. Since the problem has no good evaluation criteria, from day-to-day the human solvers will not know if they are closer to a solution or not. After years of perceived failure the human solvers will tend to drop out, unless they get feedback that their contributions are doing some good.

Solving this problem requires rewarding human solvers repeatedly to keep them engaged and motivated to continue working on the problem. However, determining near term rewards remains challenging because it will be impossible to tell if an individual's efforts have got the crowd closer to a solution. The system must instead reward solvers for quality participation and the amount of "ground" covered, not on if that effort produced a solution.

The human computation environment needs to create a knowledge economy. Basically people get rewarded for creating knowledge by gaining access to more knowledge, which in turn allows them to generate more knowledge. Science already works this way, with some concessions to allow scientists to get paid to have a roof over their head and a car in the garage. Of course many scientists strive to have the unique insight so they can capitalize on it and make more money. Such selfish behavior would be an anathema to a human computation environment because it would encourage information hiding.

Instead the computation environment needs a number of non-monetary rewards that blend together to reward a variety of contributor without requiring continuous capital investment by the sponsors. A working knowledge economy can be built using four different coins: altruism, recognition, competition and money.

Altruism

Many projects have failed in the past, even though they consisted of highly motivated, altruistic individuals. When an individual donates their time for the common good, they must understand their impact to the overall goal. Too often an individual feels their efforts are lost in bureaucratic red tape. The system must allow participants to quickly make a timely contribution that gives them a sense of accomplishment, and shows how their contribution has helped.

Human solvers need a visualization that shows an individual's contributions so that they can see they are having an impact. For example, a visualization could show the rate at which participants solve sub-problems and tasks on the project management dashboard. Thus participants can track the problem solving progress, even as the problem evolves.

Recognition

A human computation environment cannot generate altruism or money, but it can recognize the contributions of individuals. Intrinsic motivation works on the human desire to be acknowledged and recognized similar to social status (Mohtashemi and Mui 2003). Wikipedia succeeded in motivating participants to contributing their time to create knowledge about topics they felt passionate about, but contributors and their expertise remain nearly anonymous, providing no recognition reward. The successful Stack Overflow website answers technical computer questions and rewards people who provide answers by increasing their status (Anon 2009). As participants put in more effort, the system learns to trust them and allocates more power to them to administer the system.

Wicked problems require more time to solve than simply answering questions, and so the system must avoid point inflation, where the human contributor has earned so many badges and points that earning more provides no motivation. By having different domains and rewarding different aspects of problem solving the system can improve its population of solvers to get better at coordinating solutions, decomposing problems, deducing facts, fact checking, resolving conflicts, adding links, adding tags, creating solutions to test and integrating solutions. Each action could earn immediate and delayed recognition points. The delayed recognition points payout when the human solvers choice to invest their time in specific area yields results, even if the particular solver did not have the specific insight.

Competition

Competition provides great motivation for many individuals, often costing the sponsor nothing but recognizing the winner. To harness this motivation factor, a human computation system simply needs to rank contributors by a variety of metrics. Individuals will simply strive to be first or be on the top of a list of contributors (Rosner 2013). More structured contests can have solvers form teams or have specific contest goals. Contests will also provide recruitment to get people involved as their friends enlist them to help them win. The variety of metrics becomes critical otherwise a leaderboard demotivates those not within striking distance of first place. Different metrics create many leaderboards, to engage more of the population.

Monetary Rewards

Money becomes an important tool in the sponsor's arsenal when altruism, and recognition rewards do not provide enough motivation. Money cannot form the basis of motivation for a large crowd of solvers simply because conducting fair distribution of monetary rewards would be very time consuming and subject to fraud.

However, Amazon's mechanical Turk has shown small payments can motivate participation to complete specific tasks where other rewards may not be sufficient (Anon n.d.). Some wicked problems may require doing tasks that people will only do for money. For example, suppose a solution required thousands of genetic tests. Volunteers will probably not perform repetitive tests requiring expensive equipment for free. In these cases the sponsors should use their funds to hire people to complete the tasks.

Sponsors can use monetary rewards to shift focus to any task where the other reward factors do not provide enough motivation. However, once human solvers do get paid for a task they are unlikely to do it for free again in the future. Rather than base the knowledge economy on money the sponsors can use money as "solution lubricant" greasing the wheels when the solvers seem stuck or when no forward progress is being made.

Summary

Wicked problems provide the most challenging class of problem to solve in a human computation environment. However, often solutions to these very problems would create a tremendous benefit to human quality of life across the globe. Overall a human computation environment attempting to solve wicked problems must continuously flex to avoid getting stuck in the bias of the crowd of human solvers. The biases of the people posing the problem in the first place could prevent even a highly motivated crowd from finding a solution. Human solvers and sponsors alike must be prepared for a long battle to overcome these challenges.

References

Anon, Amazon Mechanical Turk—Welcome. Available at. https://www.mturk.com/mturk/welcome. Accessed 10 Aug 2010

Anon (2009) Stack overflow: a language-independent collaboratively edited question and answer site for programmers. Available at. http://www.stackoverflow.com/about

Duncker K (1945) On problem solving. Psychol Monogr 58(5) (Whole No. 270)

Mohtashemi M, Mui L (2003) Evolution of indirect reciprocity by social information: the role of trust and reputation in evolution of altrusim. J Theor Biol 223(4):523–531

Morris E (2010) The anosognosic's dilemma: something's wrong but you'll never know what it is (Part 1)—Opinionator blog—NYTimes.com. Available at. http://opinionator.blogs.nytimes.com/2010/06/20/the-anosognosics-dilemma-1/?hp. Accessed 25 June 2010
Rosner H (2013) Public participation in research back in vogue with ascent of "Citizen science": scientific American. Scientific American. Available at. http://www.scientificamerican.com/article.cfm?id=public-participation-research-back-in-vogue-ascent-citizen-science. Accessed 9 Mar 2013

Part III
Techniques and Modalities

Introduction to Techniques and Modalities

Kshanti A. Greene

Defined by Luis von Ahn as "systems that combine humans and computers to solve large-scale problems that neither can solve alone,[1]" human computational systems have attracted enough attention by researchers and developers to form a loosely connected community. This handbook aims to form stronger relationships between these researchers, many of whom happened into the field through artificial intelligence (AI) or human computer interfaces (HCI). In this section, several of these pioneers share "words of wisdom" from their own experiences with human computation. This section will be of considerable interest to anyone who is curious about how HC systems could be implemented, and those who would like to enhance existing systems. *Techniques and Modalities* highlights reusable techniques and approaches that can be applied to address common problems in human computation. Some recommendations are borrowed from other fields, such as human computer interfaces and biology. Others address issues that are unique to human computation, such as motivating contributors, incorporating the contributor's personal context, and aggregating multiple perspectives.

Trained computer scientists typically learn that it can be efficient to reuse software, architectures and algorithms, and to design for reuse when possible. Software systems can be described using different layers of abstraction and software engineering may address re-use at each of these layers. For example, an algorithm is an abstraction of a procedure for accomplishing a task in computer code. Architectures are another level of abstraction describing the structure and relationships between components of a system. *Patterns* are at a higher level than algorithms, but more process oriented than architectures. These are templates for solving a particular kind of problem. Many computer scientists have been exposed to the book,

[1] Luis von Ahn's website: https://www.cs.cmu.edu/~biglou/

K.A. Greene (✉)
Social Logic Institute, New Mexico, USA
e-mail: kshanti@gmail.com

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_23,

Design Patterns: Abstraction and Reuse of Object-Oriented Design, by Gamma et al. The book describes how to apply solutions to common problems in object-oriented design at a conceptual level that can be adapted to a specific implementation. Similarly, "Techniques and Modalities" investigates what kind of reusable patterns, techniques and processes are relevant to human computation.

The chapters in this section can be grouped into three loose categories; those that focus on tasks that are uniquely accomplished using human computation mechanisms, those that focus on how to engage human contributors in the process, and those that discuss how to effect certain emergent behaviors from human computational systems.

The following group of chapters focuses on tasks that are uniquely accomplished using human computation mechanisms.

Yolanda Gil's chapter, entitled "Collaborative Knowledge Collection" provides a survey of approaches that enable human populations to build knowledge bases with increasing semantic structure. Structuring contributions allows diverse information to be aggregated in order to more effectively answer complex queries. Existing systems require the contributors to manage both the content needs and structure of the data. They would be enhanced by automated mechanisms to support the humans.

Irene Celino's "Location-based Games for Citizen Computation" discusses how merging entertaining human computation tasks with mobile devices can create a rich texture of sensed and interpreted information about our spatial environments. This collected information can inform others seeking to maximize awareness of their surroundings and even encourage *citizen computation-* to collaboratively improve the world. Dr. Celino's chapter characterizes existing location based "games with a purpose" (GWAP) and provides guidelines for designing this class of HC application.

Mark Billinghurst's chapter, entitled "Augmented Reality Interfaces in Human Computation Systems" surveys applications near the boundary of mobile augmented reality games and human computation, and presents some novel applications that combine elements from both fields. Interfaces that merge digital information provided by human contributors with the user's physical environment can more effectively support distributed collaboration.

Joel Ross' "Pervasive Human Computing" represents another dimension of mobile human computation in which the focus is on harnessing information from the contributor's *situatedness,* incorporating the person's local and social context. While similar to location-based applications, the distinction is that pervasive computing focuses on the sensory aspects of the environment.

Kshanti Greene's "Building blocks for collective problem solving" describes a model that can be used by human contributors to collaboratively seek solutions to problems. Humans mark relationships between factors in a problem in a manner similar to ants marking pheromone trails– seeking paths indicating solutions through a problem space. The chapter shows that many existing problem solving approaches can be accomplished collaboratively using these building blocks.

Anamaria Barea's chapter, entitled "Adaptive Agents in Combinatorial Predictive Markets" discusses how humans and software agents can work together to improve prediction using a market-based approach. Instead of "betting" on singular events, combinatorial predictive markets evaluate the likelihood of combinations of events. Dr. Barea shows results from a study in which combinatorial prediction using human estimates is more accurate than non-combinatorial and agent-only approaches.

The following chapters discuss various mechanisms to engage human contributors in HC applications.

Jesse Chandler, Gabriele Paolacci and Pam Mueller's chapter, entitled "Risks and Rewards of Crowdsourcing Marketplaces" discusses the environments that match people with a task need (requesters) with humans willing to take on those tasks (workers). The most significant of these marketplaces is Amazon's Mechanical Turk. The authors discuss the benefits of these marketplaces and provide suggestions on how to overcome their limitations.

Markus Krause wrote "Designing Systems with Homo Ludens in the Loop", which introduces the concept of homo ludens (playful humans) as the "computers" in a human computational system. This chapter looks specifically at mechanisms that increase the entertainment aspect of an HC task. System developers are encouraged to take advantage of intrinsic motivation but are cautioned about how easy it is to go overboard. Dr. Krause discusses how he incorporates ludic elements in his prototype GWAP called "Empathy."

Stuart Reeves' chapter, entitled "Human-Computer Interaction issues in Human Computation" makes important parallels between the more mature field of human-computer interaction (HCI) and human computation (HC). In particular, he encourages HC system developers to engage the human element of human computation systems, in particular to consider how the aspect of being human frames the computational problem. Following from the field of HCI, HC systems should be informed by the user experience.

Authors in last group of papers discuss how to encourage certain emergent behaviors from a human computation system. Typically these require more collaboration and information exchange than the approaches focused on creating micro-tasks that can be completed independently of the actions of other human contributors.

Jasminko Novak's "Collective Action and Human Computation" asks how human computational systems can be used to create a participatory environment in which human contributors are working to benefit their shared community, not just their individual situation. In these environments it is particularly important that all data be non-excludable- in other words accessible to all contributors. The challenge then is to how to represent shared artifacts that incorporate and relay many different perspectives.

Liane Gabora's "Cultural Evolution as Distributed Computation" outlines human computational methods for modeling the emergence and evolution of cultural behavior and artifacts in a community of interacting agents or individuals.

Key features of the model include *restorative restructuring* (triggering a new idea from an existing one by considering it from a different perspective) and *communal exchange* (horizontal transfer of information), a concept she draws on in a manner inspired by metabolism-first theories of the origin of life.

Winter Mason's chapter, entitled "Collective Search as Human Computation" suggests that turning search into a collaborative process could enable collectives to answer more complex questions. The chapter describes interesting discoveries about the behavior of collectives working to search a complex problem space for an optimal solution. Success seems to be dependent on the amount of information sharing in a network, the motivation to explore versus exploiting known solutions, and the topology of the problem space.

Pietro Michelucci's chapter on Organismic Computing considers the synergies that might occur if mechanisms for shared sensing, collective reasoning, and coordinated action are incorporated into an augmented reality environment. In most large-sale collaborations, there are diminishing returns as more contributors are added. Dr. Michelucci proposes that, with the right HC framework, there is no such thing as "too many cooks", and a group's efficacy can actually increase as more people are added.

A few chapters in other sections of this handbook deserve mention for their introduction of techniques and modalities for human computation.

Jeffrey Nickerson's chapter in *Algorithms*, entitled "Human-based Evolutionary Computing" describes the process by which ideas are exchanged and modified over time by including the diverse perspectives of the crowd. This is similar to evolution in that ideas mutate and combine. This process can be augmented using memetic algorithms that infuse domain knowledge into the process. Such an infusion of knowledge comes naturally when crowds are organized to create and evaluate ideas.

Brambilla and Fraternali's chapter in *Application Domains*, entitled "Human Computation for Organizations: Socializing Business Process Management" seeks to enable socialization of business processes. They present a visual language for expressing the interrelated roles and behaviors of organization members. This language can be used to model any organization or community seeking to increase communication and leverage its member's perspectives.

Ido Guy's "Algorithms for Recommendation" (in *Algorithms*) analyzes a number of algorithms designed to provide recommendations to individuals seeking information on a wide variety of subjects. A pattern emerges from these algorithms that indicates the importance of providing explanations for recommendations. These explanations increase the likelihood that the receiver will take the recommendations to heart. We could extrapolate that explanations for any socially provided information, whether it be recommendations or solutions to a problem, should be provided.

"Methods for Engaging and Evaluating Users of Human Computation Systems," by Jon Chamberlain, Udo Kruschwitz and Massimo Poesio (in *Participation*) addresses the challenges of motivating human contributors to take on tasks and to encourage high quality results, based in part on their experiences with a system called Phrase Detectives. The authors provide effective strategies on how to design tasks, how to solicit workers, and how to evaluate and ensure quality.

The Techniques and Modalities section presents 13 unique perspectives from the emerging human computation field. Each article presents thoughts on how to best move forward. Based on these observations, the following challenges remain open and present opportunities for further consideration, research, and development.

- How to maximize the overall quality of results when using humans as "computers."
- How to grow a large, engaged community of human contributors.
- How to ensure fairness and ethics for crowdsourced workers as the number of people getting paid for these activities grows.
- How to move beyond the humans as computers aspect and encourage system developers to design for human motivations and incorporate human personalities.

It is interesting to note that the majority of the authors in this book were not aware of each other prior to its inception. How did so many people working in other fields begin to ask what humans and computers could accomplish together, instead of following the path initiated by artificial intelligence pioneers? Did we all collectively stop and ask whether we really wanted a world dominated by machines? Or did we realize that we were too intimate with and emotionally invested in these issues to let them be objectively pursued by a purely mathematical mind. On the other hand, it may have been a natural extension of distributed intelligence, but we are finally able to "scale up" conceptually to incorporate more complex agents into the mix.

Regardless of our individual motivations, let us continue to ask these questions while seeking solutions to the immediate challenges. Perhaps they will be stepping stones towards addressing the archetypal global problems that many of us hope to solve (or avoid!) such as war, genocide, hunger, disease, economic and environmental decline, and nature's caprice.

Social Knowledge Collection

Yolanda Gil

Introduction

In the early days of the Web, people contributed content in their individual Web pages and sites for the benefit of all. The turn of the millennium saw an emergence of social content collection sites as a new way to share information for the benefit of others. Social content collection sites range from wikis to blogs, and cover topics as broad as encyclopedias (http://www.wikipedia.org), health (http://www.healthnet.org), and how to do things (http://www.wikihow.com). What characterizes social content collection? First, these are social sites where many individual contributors collaboratively synthesize a body of content. There may be different kinds of contributions, some simply suggesting extensions and others with actual content and updates to the shared collection. Another feature is that there is some degree of coordination among the contributors. It can be very light coordination, for example a simple set of rules to organize the content. Alternatively, it can be very process-heavy where a complex editorial process is in place and contributors play specific roles with different oversight and responsibilities. For example, in its first year the English Wikipedia had fewer than 300 project pages (i.e., pages devoted to describing editorial processes and conventions) to organize the contributions of 21,000 topics, and as of September 2010 it reported 582,000 project pages and 7.9 M topics (http://stats.wikimedia.org/EN/TablesWikipediaEN.htm), quantifying the growth of bureaucracy in the editorial process from 1:70 to 1:13. Third, social content collection is organized around a coherent theme. For example, a wiki may be devoted to the theme of "how to do things". Finally, the content has a nascent structure. For example, wikis are organized so each page is devoted to a topic and may be related explicitly to other topics through hyperlinks. For example, a page about how to go camping could be linked to a page about how to set up a camping tent.

Y. Gil (✉)
University of Southern California Information Sciences Institute, Marina del Rey, USA
e-mail: gil@isi.edu

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_24,

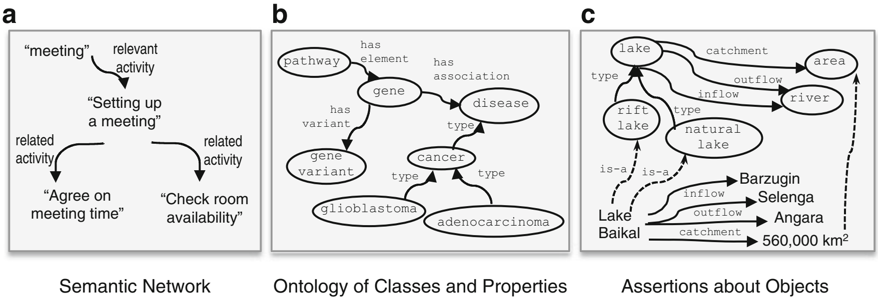


Fig. 1 Social knowledge collection can target different kinds of internal representations, which have implications on the kinds of users that can contribute and the kinds of reasoning that the system can do about its knowledge: (**a**) semantic networks of semi-structured knowledge, (**b**) ontologies of classes and their properties, (**c**) assertions about objects

Social content collection sites are incredibly popular. A search for "Powered by MediaWiki", the wiki software underlying Wikipedia that was developed by the Wikimedia Foundation and distributed under a Creative Commons license (Barrett 2008), showed 87 M hits in September 2010 and 150 M in March 2013. Myriads of other sites use other wiki software or different frameworks for web content management. Masses of volunteers are collaborating daily to create millions of formidable resources. They contribute content, play well-defined editorial roles, and organize the content around useful topic pages and categories.

Despite their popularity, social content collection sites have important limitations for search and query answering. Because the content has very little structure, they cannot aggregate information to answer simple queries. For example, Wikipedia content is well organized, but it is not structured to answer simple queries such as "What US Congress representatives own a business?", "What major cities in Europe have soccer teams that play in a national league?", or "What are all the versions to date of the Android software for cell phones?".

In recent years, new approaches for social content collection have emerged that are more focused on structuring contributions. These approaches support social knowledge collection, representing content in such a way that it can be aggregated in meaningful ways to answer reasonably complex questions. They share the characteristics discussed above for social content collection sites: many individual contributors, there is some coordination among contributors, and contributions revolve around a theme. A unique feature of social knowledge collection is that the content is structured. Figure 1 illustrates some useful distinctions in the way that knowledge can be structured, using different approaches to knowledge representation (Brachman and Levesque 2004). One possibility, shown in Fig. 1a, is to use semantic networks to link abstract concepts, but no reasoning is possible since the links and the concepts are not related to similar ones. In the figure, "Agree on meeting time" is a concept that has no relation to temporal representations of what time is, and therefore the system cannot answer questions about duration for example.

Another possibility, shown in Fig. 1b, is to structure knowledge by defining ontologies, where classes of objects are created as well as properties of the objects in each class. In the figure, the class "gene" has a property of having an association with another class "disease", which in turn has several subclasses such as "cancer" and "glioblastoma". A third kind of knowledge concerns assertions about objects. Shown in Fig. 1c are several assertions about the object Lake Baikal, for example that it has inflow from Barzugin and Selenga and has an area of catchment of 560,000 km^2. Note that these assertions can be linked to ontologies, in this example there is an ontology of classes of lakes and their properties. The choice of knowledge structures determine the kinds of automated reasoning that can be performed on the knowledge collected, and therefore the kinds of questions that the system can answer about its knowledge. For example, since the inflow and outflow of lakes are to rivers, the system can infer that Barzugin, Selenga, and Angara are all rivers. It can then answer questions about rivers that flow into Lake Baikal.

Although the acquisition of structured knowledge has been an active area of research in artificial intelligence, the advent of the Web and the opportunity for collaborative knowledge capture presents new challenges (Gil 2011). How should the interface be designed to guide contributors appropriately? What would be appropriate internal representation of the knowledge? What are successful approaches to attract and incentivize a healthy community of contributors? How can the quality of the knowledge collected be improved?

This chapter gives an overview of research to date and future challenges in social knowledge collection. Three major approaches are presented. The next section describes approaches to collect semi-structured repositories focused on common sense knowledge. The following section describes semantic wikis, extensions of traditional wikis that allow contributors to give more structure to topic pages and the links among them. After that, collaborative ontology editors are discussed as approaches to collect structured definitions of classes and properties. The chapter closes with a discussion of the research challenges ahead in this still nascent research area.

Collecting Semi-structured Knowledge Repositories

An interesting area of research in social knowledge collection targets the creation of semi-structured repositories of knowledge. The knowledge is organized as semantic networks that, as we mentioned above, relate concepts that have no formal definitions and that do not fully support reasoning. Creating semi-formal repositories is easier for contributors with no expertise in logic or knowledge engineering, because they provide simple English statements that the system then tries to organize into more formal knowledge structures. The research in this area has focused on the collection of common knowledge, including common sense knowledge about world objects as well as daily and routine activities that require no particular expertise and are known by everyone but are not known to computers.

An analogy-based approach to collect knowledge about common objects was used in LEARNER (Chklovski 2003a, b). LEARNER prompted volunteers for common objects, and upon an entry such as "newspaper" LEARNER would ask for useful things to know about newspapers. Contributors would respond with short sentences, for example "a newspaper is made of paper," "you can read a newspaper," and "you can carry a newspaper in your briefcase". LEARNER used simple natural language processing techniques to create a semantic network that made connections among the statements. LEARNER also used a novel analogical reasoning algorithm to detect commonalities among objects. So if a user entered "magazine" and said "a magazine is made of paper" and "you can read a magazine" then LEARNER would detect that magazines and newspapers seemed to have some things in common, and would ask whether "you can carry a magazine in your briefcase" was true along with other things it already knew about newspapers.

LEARNER2 (Chklovski 2005) was an extension of LEARNER focused on the collection of specific types of knowledge, originally designed to assist users with to-do lists (Gil et al. 2012). LEARNER2 toured for several years as an interactive kiosk at a science museum as part of a traveling exhibit called "Robots and Us" to raise public awareness of the challenges of teaching common sense to computers. It collected more than 600,000 raw entries concerning task-oriented knowledge, such as objects relevant to a task, repairing task failures, descriptions of tasks in natural language, and decompositions of tasks into subtasks.

A detailed analysis of the statements collected with LEARNER2 revealed important findings (Chklovski and Gil 2005a). First, redundancy of contributions helps identify high quality statements, so that if several contributors enter the same statement it is more likely to be correct. However, some of the statements also have overly high redundancy, drawing contributor effort away from areas where increasing coverage and increasing redundancy are more needed. That is, a large amount of contributors will think of entering the most common statements that are likely been already collected. This has consequences for the design of the user interfaces, so that contributors are enticed to make novel statements to the system (Chklovski and Gil 2005b, c). Figure 2 illustrates key aspects of the design of the user interface. The user was asked follow up questions using templates designed to collect additional knowledge piecemeal. The user would get guidance on the type of input that the system was expecting, and would tend to enter simple statements. The knowledge entered was analyzed with simple natural language techniques to discard unusable statements that would not conform to the simple structure expected. The knowledge was also aggregated and shown back to the user for confirmation, and as a way to detect whether the user had understood what was expected. Finally, the statements acquired were shown to other contributors for validation. These user interface features can significantly improve the quality and coverage of the knowledge collected.

The Cyc FACTory (Matuszek et al. 2005) allowed contributors to add facts to the Cyc knowledge base (Lenat and Guha 1990), which was designed to contain encyclopedic knowledge including common sense knowledge. Like LEARNER2, contributors were prompted with a template to fill, in this case a pre-defined schema based on the contents of the Cyc ontologies.

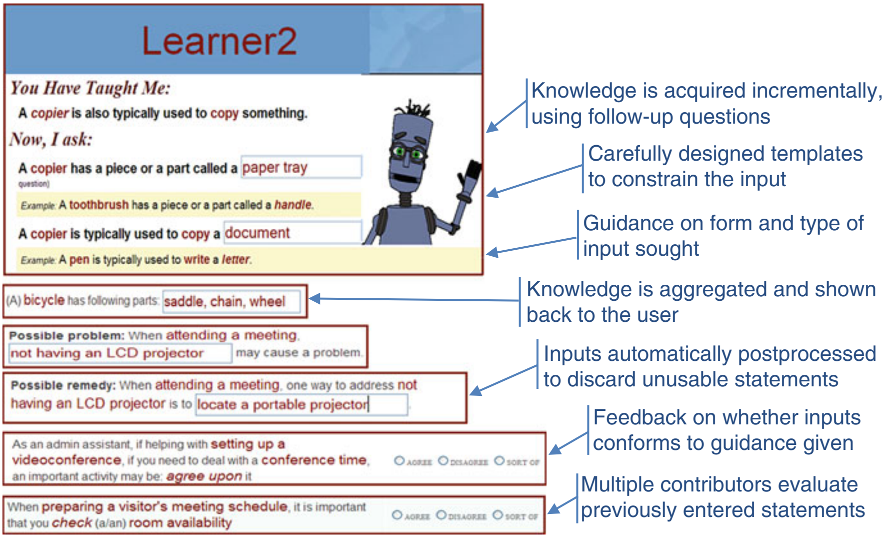


Fig. 2 Learner2 collected semi-structured statements from volunteers about common objects and tasks. Its user interface was designed to guide contributors and improve the quality and breadth of the knowledge collected

The Common Sense Computing Initiative (http://csc.media.mit.edu/) constellation of projects has been collecting common sense knowledge to create structured repositories (Havasi et al. 2007). Volunteers are prompted with objects that are mentioned in the contributions of others. A novel feature-based clustering technique was used to organize the contents collected (Speer et al. 2008). Specific collection efforts have been set up to collect knowledge about indoor objects to help with robot navigation (Gupta and Kochenderfer 2004), about common tasks and events (Lieberman et al. 2007), and about common objects and their properties (Havasi et al. 2007). These repositories have been used in a variety of contexts to assist users with tasks such as organizing pictures (Lieberman et al. 2004) and personal task management (Smith and Lieberman 2010). To date, the site has collected over a million sentences from over 15,000 contributors.

Semantic Wikis

Semantic wikis are wikis with extensions that support the creation of structured content, and have reasoning capabilities that exploit that structure to organize the wiki's knowledge. Traditional wikis support some ways to structure content, for example by assigning categories to topic pages. Wikipedia has infoboxes for athletes, politicians, and countries. Infoboxes are essentially just a form for users to organize content, and are often used to extract knowledge bases from wikis (notably from Wikipedia) (Auer et al. 2007; Weld et al. 2008). However, the system cannot reason about their content to answer questions, such as what rift lakes are in Russia. In contrast, a semantic wiki allows users to organize topic page categories as classes (or concepts) in a taxonomy, and to define properties that apply to each class. For example, the Wikipedia page for Lake Baikal would be linked to the page for Russia through a regular hyperlink such as Lake Baikal is in [[Russia]], while in a semantic wiki the hyperlink would be Lake Baikal is in [[country Russia]] where country is a property. This enables the system to answer questions about lakes in Russia. Semantic wikis allow users to constrain properties by the range of values that they can take, which are called structured properties. As content is added using these structured properties, the semantic wiki can use reasoning and inference. Users can then query the content to generate dynamic content for wiki pages. Visualizations can be created automatically by overlaying semantic information in maps or charts.

An important feature of semantic wikis is their integration with semantic web standards. Each assertion is turned into a triple of the form <object property value> that can be expressed in the Resource Description Framework (RDF) standard (Brickley and Guha 2004). This makes the knowledge collected through semantic wikis compatible with the data already captured in many billions of interlinked RDF triples that are accessible on the Web and are known as the Web of Data or Linked Open Data (http://www.w3.org/standards/semanticweb/data) (Heath and Bizer 2011; Auer et al. 2007).

Semantic wikis are becoming very popular, as they offer the simplicity of a wiki with additional capabilities to help contributors organize content. There are several implementations of semantic wikis. Semantic MediaWiki (Krotzsch et al. 2007) is a diverse set of extensions for the popular MediaWiki wiki platform, and allows users to easily create new concepts and structured properties without enforcing consistency up front. OntoWiki (Auer et al. 2006) is another semantic wiki that requires that a schema be defined before users enter content to populate it through a form-based web-interface. AceWiki (Kuhn 2009) provides a more powerful knowledge representation formalism than most other semantic wikis, with the cost of requiring the contributors to learn and use a semi-formal logical language designed for them by the wiki developers/administrators. (Bry et al. 2012) give a detailed overview of semantic wikis and a thorough comparison of semantic wiki frameworks. Perhaps because of its more permissive and organic approach to structuring knowledge collaboratively, Semantic MediaWiki has been adopted by hundreds of disparate communities for a variety of purposes such as science (e.g., organizing genomic knowledge), engineering (e.g., coding software), and hobbies (e.g., organizing gardening tips). A notable semantic wiki is Wikidata (http://www.wikidata.org), a project by the Wikimedia Foundation to build a comprehensive multilingual collection of facts that would complement their Wikipedia effort. Wikidata is built with Semantic MediaWiki, which extends the MediaWiki platform used by Wikipedia.

Shortipedia is a semantic wiki designed to collect structured knowledge about objects (Vrandečić et al. 2011). It is based on Semantic MediaWiki, and extends it to allow users to add new properties and values together with their provenance.

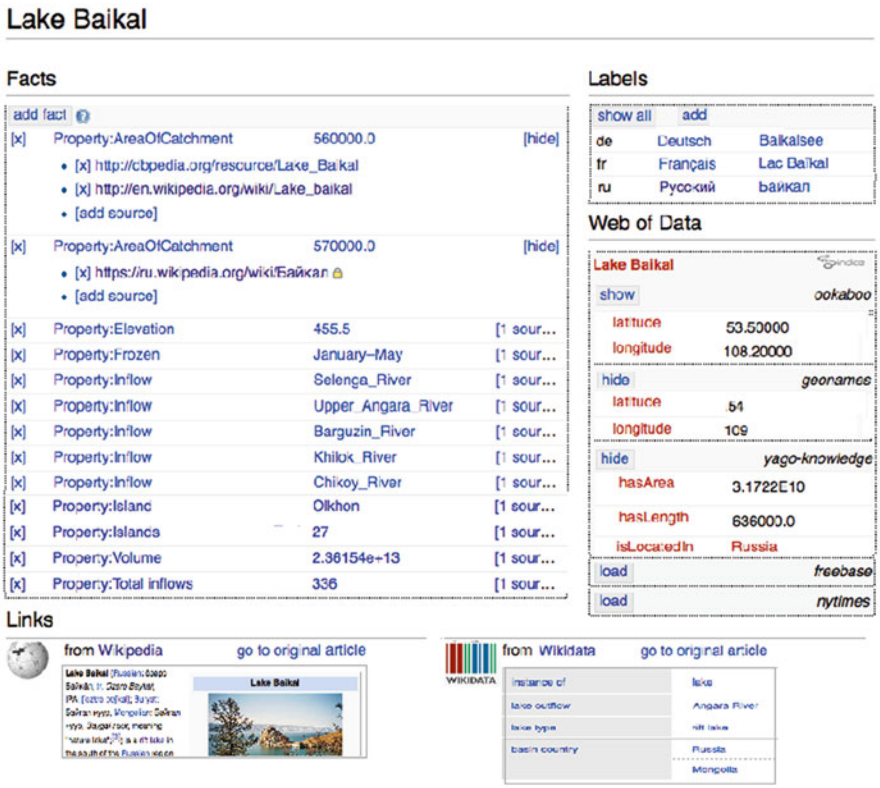


Fig. 3 Shortipedia was designed to collect structured knowledge about objects. On the top left, a page for Lake Baikal is shown, including properties such as its area of catchment, elevation, inflow and outflow. Note that each assertion is annotated with sources that support it. The figure shows that the area of catchment is different in the Russian and the English Wikipedia pages for the lake. On the *top right*, multilingual labels are shown. On the *middle right*, other known assertions on the Web are retrieved and shown to the user. Here, the latitude and longitude are different depending on the source. At the *bottom*, the original Wikipedia page is shown for reference, as well as the properties that appear in Wikidata

Figure 3 illustrates its user interface. On the top left, a page for Lake Baikal is shown, including properties such as its area of catchment, elevation, inflow and outflow, volume, and islands. Users can add new properties, together with the sources that support them. When the user adds a property, the system uses a command completion search to find existing properties that match what the user is typing. This encourages reuse and normalization of properties across contributors. Another feature of Shortipedia is that it allows contributors to state alternative values for a property. For example the area of catchment is different in the Russian and the English Wikipedia pages for Lake Baikal, so users can add both values with their respective sources. Shortipedia also enables users to add multilingual labels that allow the system to map assertions in different languages, shown on the top right in the figure. Shortipedia also allows users to easily include other known assertions on the Web of Data, by automatically retrieving them and allowing the user to select them as shown on the middle right of the figure. Users are also shown the original Wikipedia page for reference, and the properties that are contained in Wikidata so they can be included as well as shown in the bottom of the figure.

In order to understand how the semantic aspects of the wiki are used to structure the contributions, (Gil and Ratnakar 2013) carried out an analysis of more than 200 semantic wikis. The analysis showed the concepts and properties created in each wiki, and the amount of editors involved in creating them compared to the total amount of editors of the wiki. We found that concepts are not defined very often. In contrast, properties are very widely used. Large numbers of property assertions are used in almost every wiki. We also found that very small numbers of users edit properties. An important challenge is to understand the limited use of some semantic features of the wiki, such as concept definitions, as well as why there are relatively small amounts of users who create any definitions. One hypothesis is that this is due to the lack of support to the contributors in coordinating semantic edits, although further research is needed to understand this. In addition, semantic wiki communities might benefit from additional capabilities that make the system more proactive in making suggestions to contributors regarding the creation of new concepts, encouraging the reuse of properties created by other contributors, and resolving inconsistencies and missing knowledge.

Collaborative Ontology Development

For many years, ontology editors were used only by knowledge engineers, enabling them to create sophisticated ontologies of classes and properties either individually or in small well-orchestrated teams. Recently, ontology editors have been augmented to support the collaborative development of ontologies with contributors lacking prior training or prior knowledge about which specific areas each might be able to contribute to. Collaborative ontology development requires a framework that solicits and organizes contributions from people who might have different expertise and different views on the subject matter.

Collaborative Protégé is a framework for collaborative ontology development based on the widely used Protégé ontology editor (Tudorache et al. 2011). It has been used to develop biomedical ontologies of thousands of terms with dozens of contributors (Tudorache and Musen 2011), including the International Classification of Disease revision 11 (ICD-11) and the National Cancer Institute's Thesaurus (NCI Thesaurus). The system enables users to add new subclasses and properties, but it also allows them to override specific contributions made by others and post notes explaining disagreements that need to be discussed.

Understanding the processes or workflows that arise from different ontology editing patterns is helpful for developing new techniques that can support common patterns. For example, a recent analysis found a strong correlation between the amount of changes that a given contributor makes and the amount of notes that the contributor posts (Strohmaier et al. 2013). To provide a global view of the status of the ontology, visualization tools enable monitoring progress over time, expose areas of major disagreements, and measure the quality of the contributions (Walk et al. 2013). This exposes the breadth of expertise of specific contributors, and the most heavily edited areas of the ontology over time.

Further research is needed for supporting different editing patterns, different contributor skills, and managing the dynamic evolution of the ontology and its user community over time.

Research Challenges in Social Knowledge Collection

Social knowledge collection approaches have been demonstrated to create useful repositories of knowledge for a variety of purposes. However, further research is needed in designing systems that take a more active role in guiding the acquisition process, manage the knowledge collected, and coordinate contributions from different users. Research challenges in social knowledge collection include:

- **User interface design**: How can people detect errors and misconceptions in the system and fix them? How can contributors enter knowledge with minimal burden or prior training?
- **User feedback and prompting**: How can the system generate follow up questions that complement knowledge that users contribute on their own accord? How can users be assigned follow up questions based on their demonstrated expertise?
- **Coordination among contributors**: What are the most effective editorial processes to organize contributors? How can systems learn from several people who are providing overlapping and perhaps incompatible or even contradictory information?
- **Incentives**: What are successful ways to reach and recruit potential contributors to maintain a reasonable community over time? What are the right incentives and rewards to retain contributors?
- **Provenance**: How can users document the knowledge they enter so that the system can justify the sources of its knowledge to other users and be trusted?
- **Quality of the knowledge**: What mechanisms can be used to validate contributions?
- **Purpose**: What kinds of knowledge can we collect effectively through crowdsourcing approaches? What are appropriate knowledge acquisition tasks that contributors can handle?
- **Nature of knowledge collected**: What kinds of knowledge can be collected through volunteer contributors? What are appropriate uses of the knowledge collected? What knowledge formalism is adequate for a given use and kind of knowledge targeted?

- **Managing updates over time**: What are appropriate mechanisms to manage updates and changes, particularly when other systems may have been designed to use the knowledge being collected?
- **Combining interactive and automatic extraction**: How can we combine volunteer contributions with automatic extraction of knowledge from text? Can volunteers validate and extend knowledge automatically extracted that with varying accuracy?

Some of these issues have been studied in social content collection frameworks, notably Wikipedia (Adler and de Alfaro 2007; Almeida et al. 2007; Benson et al. 2010; Erickson 2008; Hoffmann et al. 2009; Hsieh et al. 2010; Kittur et al. 2008, 2009; Kittur and Kraut 2008, 2010; Lam et al. 2010; Leskovec et al. 2010; Panciera et al. 2010; Raban et al. 2010; Spinellis and Louridas 2008). However, the applicability of these results for social knowledge collection should be carefully considered. In addition, social knowledge collection presents its own set of challenges that need to be addressed.

We foresee in the not too distant future that knowledge repositories created through social knowledge collection could be interlinked through semantic web infrastructure, enabling knowledge sharing across communities of contributors. For example, a repository of genomics knowledge and a repository of biodiversity knowledge could be interconnected to relate genomic information to specific species. The provenance of knowledge sources will be crucial to propagate updates throughout the knowledge bases and to assess trust and resolve conflicting views.

Acknowledgments We gratefully acknowledge support from the National Science Foundation with grant IIS-1117281. We also thank Kshanti Greene and Pietro Michelucci for their comments on earlier drafts of this chapter.

References

Adler BT, de Alfaro L (2007) A content-driven reputation system for the wikipedia. In: WWW '07: proceedings of the 16th international conference on the world wide web (WWW)
Almeida R, Mozafari B, Cho J (2007) On the evolution of wikipedia. In: Proceedings of the international conference on weblogs and social media (ICWSM)
Auer S, Dietzold S, Riechert T (2006) OntoWiki—a tool for social, semantic collaboration. In: 5th international semantic web conference
Auer S, Bizer C, Kobilarov G, Lehmann J, Cyganiak R, Ives Z (2007) DBpedia: a nucleus for a web of open data. In: Proceedings of the 6th international semantic web conference
Barrett DJ (2008) MediaWiki, O'Reilly Media
Benson E, Marcus A, Howahl F, Karger D (2010) Talking about data: sharing richly structured information through blogs and wikis. In: Proceedings of the 19th international conference on the world wide web (WWW)
Brachman RJ, Levesque HJ (2004) Knowledge representation and reasoning. Elsevier
Brickley D, Guha RV (2004) RDF vocabulary description language 1.0: RDF schema. World wide web consortium. Available from http://www.w3.org/TR/rdf-schema

Bry F, Schaffert S, Vrandečić D, Weiand K (2012) Semantic wikis: approaches, applications, and perspectives. Lecture notes in computer science, reasoning web. Semantic technologies for advanced query answering, vol 7487
Chklovski T (2003a)“Using analogy to acquire commonsense knowledge from human contributors. Ph.D. thesis, department of computer science. MIT Artificial intelligence lab technical report AITR-2003-002. Massachusetts Institute of Technology
Chklovski T (2003b) LEARNER: a system for acquiring commonsense knowledge by analogy. In: Proceedings of 2nd international conference on knowledge capture (K-CAP)
Chklovski T (2005) Designing interfaces for guided collection of knowledge about everyday objects from volunteers. In: Proceedings of the ACM international conference on intelligent user interfaces (IUI)
Chklovski T, Gil Y (2005a) An analysis of knowledge collected from volunteer contributors. In: Proceedings of the 20th national conference on artificial intelligence (AAAI), Pittsburgh, pp 564–571
Chklovski T, Gil Y (2005b) Towards managing knowledge collection from volunteer contributors. In: Proceedings of the AAAI spring symposium on knowledge collection from volunteer contributors (KCVC), Stanford
Chklovski T, Gil Y (2005c) Improving the design of intelligent acquisition interfaces for collecting world knowledge from web contributors. In: Proceedings of the 3rd international conference on knowledge capture (K-CAP), Banff
Erickson T (2008) ‘Social systems’: designing digital systems that support social intelligence. Journal AI & Society—Special Issue: social intelligence design: a junction between engineering and social sciences 23(2)
Gil Y (2011) Interactive knowledge capture in the new millennium: how the semantic web changed everything. Knowl Eng Rev 26(1)
Gil Y, Ratnakar V (2013) Knowledge capture in the wild: a perspective from semantic wiki communities. In: Proceedings of the international conference on knowledge capture (K-CAP), Banff
Gil Y, Ratnakar V, Chklovski T, Groth P, Vrandecic D (2012) Capturing common knowledge about tasks: intelligent assistance for to do lists. ACM Trans Interact Intell Syst 2(3)
Gupta R, Kochenderfer M (2004) Common sense data acquisition for indoor mobile robots. In: Proceedings of the 19th national conference on artificial intelligence (AAAI-04)
Havasi C, Speer R, Alonso J (2007) ConceptNet 3: a flexible, multilingual semantic network for common sense knowledge. In: Proceedings of the conference on recent advances in natural language processing (RANLP)
Heath T, Bizer C (2011) Linked data: evolving the web into a global data space. Synthesis lectures on the semantic web. Morgan and Claypool Publishers
Hoffmann R, Amershi S, Patel K, Wu F, Fogarty J, Weld DS (2009) Amplifying community content creation with mixed initiative information extraction. In: Proceedings of the 27th international conference on human factors in computing systems (CHI)
Hsieh G, Kraut RE, Hudson SE (2010) Why pay?: exploring how financial incentives are used for question & answer. In: Proceedings of the 28th international conference on human factors in computing systems (CHI)
Kittur A, Kraut RE (2008) Harnessing the wisdom of crowds in wikipedia: quality through coordination. In: Proceedings of the ACM conference on computer supported cooperative work
Kittur A, Kraut RE (2010) Beyond wikipedia: coordination and conflict in online production groups. In: Proceedings of the ACM conference on computer supported cooperative work (CSCW)
Kittur A, Suh B, Chi Ed H (2008) Can you ever trust a wiki? Impacting perceived trustworthiness in wikipedia. In: Proceedings of the ACM conference on computer supported cooperative work (CSCW)
Kittur A, Lee B, Kraut RE (2009) Coordination in collective intelligence: the role of team structure and task interdependence. In: Proceedings of the 27th international conference on human factors in computing systems (CHI)

Krötzsch M, Vrandecic D, Völkel M, Haller H, Studer R (2007) Semantic wikipedia. J Web Semantics 5(4)

Kuhn T (2009) AceWiki: a natural and expressive semantic wiki. In: Proceedings of the 5th international workshop on semantic web user interaction (SWUI), CEUR workshop proceedings, vol 543

Lam S (Tony) K, Karim J, Riedl J (2010) The effects of group composition on decision quality in a social production community. In: Proceedings of the 16th ACM international conference on supporting group work (GROUP)

Lenat DB, Guha RV (1990) Building large knowledge-based systems: representation and inference in the Cyc project. Addison Wesley

Leskovec J, Huttenlocher D, Kleinberg J (2010) Governance in social media: a case study of the wikipedia promotion process. In: Proceedings of the AAAI international conference on weblogs and social media (ICWSM)

Lieberman H, Liu H, Singh P, Barry B (2004) Beating some common sense into interactive applications. AI Mag 25(4)

Lieberman H, Smith D, Teeters A (2007) Common consensus: a web-based game for collecting commonsense goals. In: Intelligent user interfaces (IUI-07), Honolulu

Matuszek C, Witbrock MJ, Kahlert RC, Cabral J, Schneider D, Shah P, Lenat DB (2005) Searching for common sense: populating cyc from the web. In: Proceedings of the 20th national Conference on artificial intelligence (AAAI)

Panciera K, Priedhorsky R, Erickson T, Terveen L (2010) Lurking? Cyclopaths? A quantitative lifecyle analysis of user behavior in a geowiki. In: Proceedings of the 28th international conference on human factors in computing systems (CHI)

Raban DR, Moldovan M, Jones Q (2010) An empirical study of critical mass and online community survival. In: Proceedings of the ACM conference on computer supported cooperative work (CSCW)

Smith DA, Lieberman H (2010) The why UI: using goal networks to improve user interfaces. In: Proceedings of the international conference on intelligent user interfaces (IUI), Hong Kong

Speer R, Havasi C, Lieberman H (2008) AnalogySpace: reducing the dimensionality of common sense knowledge. In: Proceedings of the conference of the association for the advancement of artificial intelligence (AAAI)

Spinellis D, Louridas P (2008) The collaborative organization of knowledge. Commun ACM 51(8)

Strohmaier M, Walk S, Poeschko J, Lamprecht D, Tudorache T, Nyulas C, Musen M, Noy NF (2013) How ontologies are made: evaluation of the hidden social dynamics behind collaborative ontology engineering projects. J Web Semantics 20

Tudorache T, Musen MA (2011) Collaborative development of large-scale biomedical ontologies. In: Ekins S, Hupcey MAZ, Williams AJ (eds) Collaborative computational technologies for biomedical research. Wiley, Hoboken

Tudorache T, Nyulas CI, Musen MA, Noy NF (2011) WebProtégé: a collaborative ontology editor and knowledge acquisition tool for the web. Semant Web J 4(1):165

Vrandečić D, Ratnakar V, Krötzsch M, Gil Y (2011) Shortipedia: aggregating and curating semantic web data. J Web Semantics 9(3)

Walk S, Pöschko J, Strohmaier M, Andrews K, Tudorache T, Noy NF, Nyulas CI, Musen MA (2013) PragmatiX: an interactive tool for visualizing the creation process behind collaboratively engineered ontologies. International Journal on Semantic Web and Information Systems, IJSWIS, Special issue on visualization of and interaction with semantic web data

Weld DS, Wu F, Adar E, Amershi S, Fogarty J, Hoffmann R, Patel K, Skinner M (2008) Intelligence in wikipedia. In: Proceedings of the 23rd conference of the association for the advancement of artificial intelligence (AAAI)

Location-Based Games for Citizen Computation

Irene Celino

Introduction

Since the turn of the century, several Web-based initiatives sprung out to exploit on-line human knowledge and people's willingness to contribute in a proactive way, the most renowned case being Wikipedia. But in the latest years, we are witnessing a transition of those phenomena from the virtual to the physical world. An exemplary case is constituted by OpenStreetMap,[1] the Wikipedia of maps. With the advent of Volunteered Geography (Goodchild 2007), wiki initiatives arose to exploit the Web to collect data about the physical space: OpenStreetMap is now the most renowned editable map of the world, born to provide free geographic data and mapping to overcome the legal or technical restrictions on the use of most commercial maps.

Additionally, the new generation of mobile phones equips a large part of the population (Nielsen Mobile Insights 2013) with smart devices enriched with sensors, not alone GPS and other positioning systems. People moving in urban environment—both inhabitants and occasional visitors—can therefore turn to their smartphones to get useful knowledge about the surrounding space (Kamvar and Baluja 2006; Dou et al. 2007): maps give directions, specialized apps and websites provide recommendations and local information.

Finally, the generation of "digital natives" finds it natural to share bits of their lives on-line: through social networks and location-based applications, they reveal information about where they are, what they are doing, and with whom. Web-mediated communication has become a usual means to maintain friendships and to organize the activities in the "physical world", also thanks to always-on mobile Internet connections.

[1] Cf. http://www.openstreetmap.org/

I. Celino (✉)
CEFRIEL – ICT Institute – Politecnico di Milano, via Fucini 2, 20133 Milano, Italy
e-mail: irene.celino@cefriel.it

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_25, © Springer Science+Business Media New York 2013

In this context, we are witnessing the rise of a virtual community that is joined by the spatial dimension. Location-based mobile apps can embody Human Computation principles and techniques (Law and Von Ahn 2011) to engage this on-site workforce to solve tasks and achieve purposes. In particular, the large-scale success of mobile gaming applications (Nielsen Mobile Insights 2013) shows a potential opportunity for location-based Games with a Purpose (GWAPs) (Von Ahn 2006).

The remainder of this chapter is structured as follows. Section "The Rise of Location-Based Applications" presents an analysis of successful location-based mobile apps: after a survey of different categories of apps (section "Survey of Location-Based Applications"), we present a comparative examination to highlight and understand distinctive features (section "Comparative Analysis of Location-Based Applications"). On the basis of this analysis, in section "A Reference Methodology for Location-Based Games with a Purpose" we propose a model and a methodology to design successful location-based GWAPs. We discuss the role of this kind of applications in section "Towards Citizen Computation", referring to the broader and rising discipline that we name Citizen Computation, at the crossroads of Human Computation and Citizen Science (Irwin 1995).

The Rise of Location-Based Applications

We have collected information about location-based applications related to our areas of interest, in particular those with a gaming flavour and those aimed at involving the users in some intelligent task. We analysed them in order to identify similarities and differences. In this section, after presenting a set of relevant applications, we provide some dimensional axes to interpret and compare them.

Survey of Location-Based Applications

Without the claim to be comprehensive, in this section, we present a number of mobile apps that leverage smart phones capabilities, with special regard to location-based features. The presented apps are very heterogeneous: from pure games to simple personal apps, from efforts oriented to advertising/marketing to applications designed with a crowdsourcing/human computation purpose. We cluster the apps in some categories to better highlight their distinctive features.

Social Networking Apps

The popularity of social networks on the Web led to an obvious rise of social networking apps for smart phones. Apart from the mobile phone access to Web social

Fig. 1 Social Networking mobile apps; from *top-left*, clockwise: *FullCircle*, Skout, face2face and Blendr (Source: respective websites)

networks (e.g. Facebook app for iPhone and Android,[2]) the location-based capabilities of phones allowed for a new generation of apps.

Some good examples are *FullCircle*[3]—that is "all about bringing people together based on their location in real time"—and *Skout*[4]—"the global network for meeting new people". Both apps find potentially interesting people, willing to be contacted, in the surrounding of the user's current location.

Other apps focus more explicitly on a specific social networking purpose. Some examples are *face2face*,[5] which is targeted to a business-oriented networking with a faceted search that lets the user specify details like people's employer or school affiliation, or *Blendr*,[6] which is expressly aimed at dating, with friends selection on the basis of photos and interests (Fig. 1).

---

[2] Cf. https://itunes.apple.com/en/app/facebook/id284882215?mt=8 and https://play.google.com/store/apps/details?id=com.facebook.katana

[3] Cf. http://www.fullcircle.net/

[4] Cf. http://www.skout.com/

[5] Cf. http://www.face2face.co/

[6] Cf. http://blendr.com/

Pure Gaming Apps

The success of mobile games is unquestionable: millions of gaming apps are today available for the most common mobile platforms. According to Entertainment Software Association (2012), 33 % of gamers play games on their smartphone, and 25 % play on their hand-held device.

An increasing number of those games have a location-based flavour thanks to the exploitation of smart phones' positioning sensors. This characteristic makes the games more engaging and immersive for players, using the surrounding environment as an incentive to play.

Some originate from games played without the support of a phone. This is the case of *Geocaching*,[7] an outdoor treasure hunting game for GPS-enabled devices. A "geocache" is an object of little value put into a small container, such as a glass bottle, that is hidden in a location; on the Geocaching website, the GPS coordinates of the existing geocaches are listed together with some clues. Players with a GPS-enabled device have to find the geocache; once they find it, they have to report their hunting experience by entering date, comments or photos about the found geocache on the website; they can also take the geocache and hide another similar object in its place or they can add geocaches elsewhere. The players win competitions or championships accordingly to the number of geocaches they find.

Other games take inspiration from board games, like *Gowar*,[8] a location-aware game inspired by the popular strategic game Risk. Players have to build their empires conquering POIs from Facebook Places, to collect points, new soldiers and rewards. The mechanism to conquer a place requires the player to be physically near the target; at this point the player can use his army (composed by virtual soldiers) to try to defeat the current owner's defence. If he wins, he conquers the POI. Subsequently users can unlock and buy new categories of weapons as soon as they have reached a specific score threshold. In this game the player can directly interact with his friends or opponents through Facebook (Fig. 2).

An app similar to those above and aimed at "blending" the boundary between the real and the virtual world is *GoblinsNGold*,[9] a location based strategic game where all actions depend on player's position within the game area (in this case, the campus area of Danmarks Tekniske Universitet—DTU). Player can investigate the area, harvest different resources, craft new resources and use them to hire and train creatures, which are then used to conquer the area and gain authority.

Gbanga[10] is a Swiss game studio that works in the area of *mixed-reality games*, i.e. apps that adapt to the player context: the real-world surrounding is reflected in the game, the real-weather has an impact on gameplay, real news also have side-effects in the game and the mobile phone sensors are used to control the game. An example is "Gbanga Famiglia", a strategic game to conquer neighbourhoods, which

[7] Cf. http://www.geocaching.com

[8] Cf. http://www.gowar.com

[9] Cf. https://play.google.com/store/apps/details?id=uadk.dtu.wdnsd.gng

[10] Cf. http://gbanga.com/

Fig. 2 Pure gaming apps: Geocaching and Gowar (Source: respective websites)

employs a mafia-like storyboard and enriches the game experience with the purchase of virtual goods (Fig. 3).

The last arrival in this category is *Google Ingress*,[11] an augmented-reality fighting game—in closed beta at the time of writing—which "represents a big step towards ubiquitous, accurate augmented reality (AR), in which real-world objects are annotated with a virtual layer of information that is displayed on a smartphone's camera" (Hodson 2012). Ingress in not completely a pure gaming app, in that Google aims at employing it to get information about new places, to generate more interesting search results, focused on what local people say are interesting (Fig. 4). For more information about Augmented Reality and its adoption in combination with Human Computation, interested readers should refer to Billinghurst (2013).

Commerce and Marketing Apps

The opportunity to attract customers based on their location is a strong incentive to create apps that ease this process. Thus, a local business can either provide a custom app or can join those location-based apps aimed to provide challenges and rewards based on a generic location .

The most popular and successful application of this kind is *foursquare*.[12] By employing a gamification (Deterding et al. 2011) approach, foursquare allows users to "check-in" in the surrounding locations, gaining points and badges to level up in the leaderboard of friends. The statistics of frequency and timing of check-in actions gives a powerful marketing insight on the popularity of places; moreover, local businesses can offer deals in relation to users' check-in actions (e.g., discounts and special offers), in order to create or reinforce a fidelity relationship.

[11] Cf. http://www.ingress.com/

[12] Cf. https://foursquare.com/

Fig. 3 Pure gaming apps: GoblinsNGold and Gbanga Famiglia (Source: respective websites)

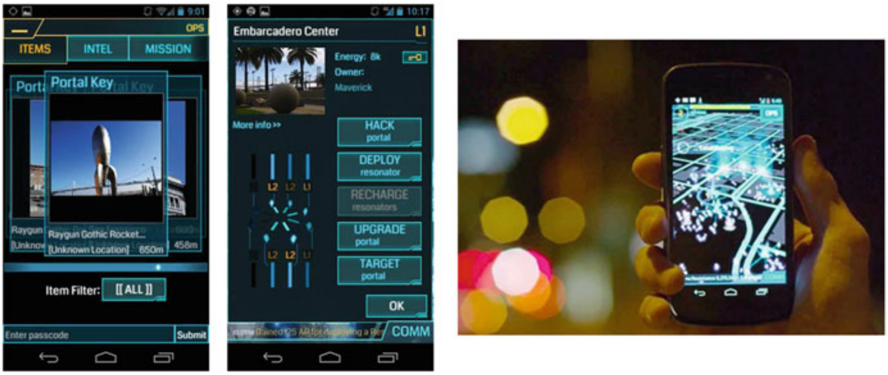


Fig. 4 Pure gaming apps: Google ingress (Source: www.ingress.com)

A number of competitors of foursquare emerged, applying the same principles with different flavours and specific features. *SCVNGR*[13] is a game that asks its players to face challenges in the surrounding locations—taken from Facebook Places—to gain points and rewards. It also allows users to create their own challenges and "treks", guided tours across different places.

Similarly but focusing on restaurants and food chain businesses, *Foodspotting*[14] centres the user experience on dishes instead of places, letting users discover good food through a visual guide. Users post pictures of dishes and rate them, thus allowing the app to provide recommendations and suggestions. Also in the case of Foodspotting, deals and discounts are possible for business owners to attract customers (Fig. 5).

The data collected through the applications just described are helpful to understand the current status and the customers' perception of an environment. But apps

[13] Cf. http://scvngr.com/

[14] Cf. http://www.foodspotting.com/

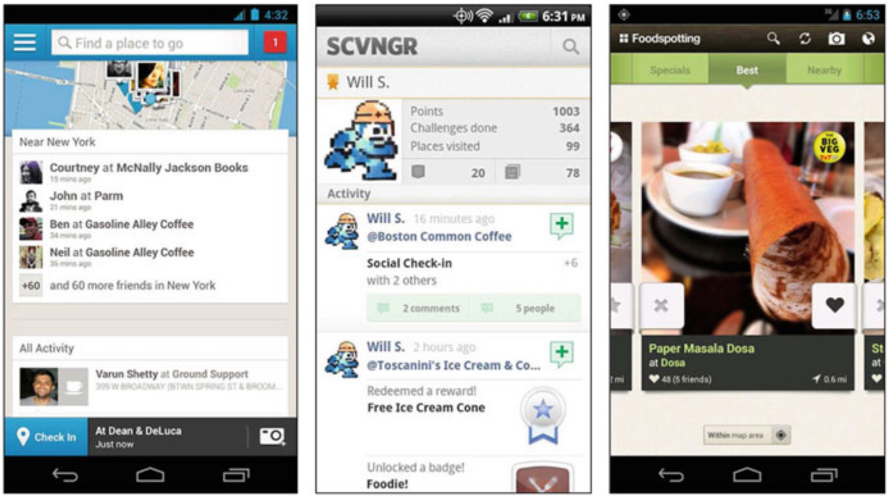


Fig. 5 Commerce and marketing apps: foursquare, SCVNGR and foodspotting (Source: respective websites)

can be used also to "design" the space, by letting imaginative users play to create a new or an improved world. This is the case of the *MyTown* app series developed by the Booyah company[15]: the player builds "the city of his dreams" with his favourite places from the real world. The company—founded with the meme "where you play matters"—has partnered with leading consumer brands to allow and foster the connection between gamers and products, thus revealing the marketing/advertising purpose of the games (Fig. 6).

Experience Sharing Apps

Beside the marketing or advertising objective, which is present in the majority of mobile apps, location-based application often aim at engaging the user in a long-lasting experience, by providing the expedient or the motivation to frequently launch the app. This happens for example in those apps that let the users record their memories and their itineraries.

Dopplr[16] is a travel planning app. Travellers register their upcoming trips, share their memories and advices with friends and contacts, and enjoy recommendations and suggestions during the journeys. In a very similar way, *Trippy*[17] ties into social sites like Facebook to find friends travelling to the same place, based on the intuition that travel experiences of friends are more valuable than random strangers's suggestions (Fig. 7).

[15] Cf. http://www.booyah.com/

[16] Cf. http://www.dopplr.com/

[17] Cf. http://www.trippy.com/

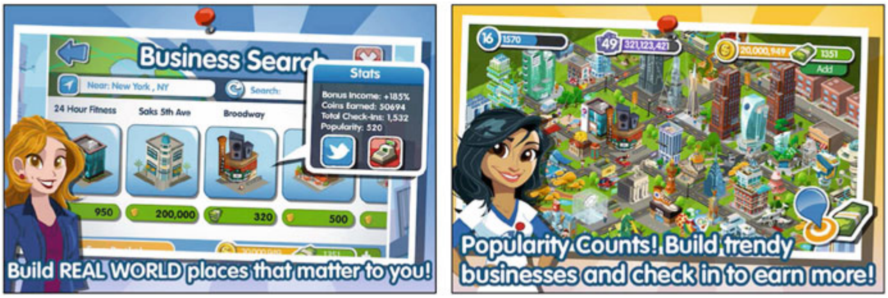


Fig. 6 Commerce and marketing apps: MyTown (Source: www.booyah.com)

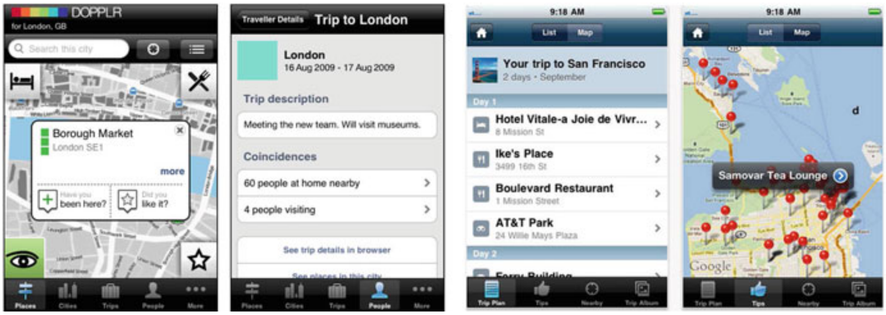


Fig. 7 Experience sharing apps: Dopplr and Trippy (Source: respective websites)

Hotlist[18] is a location-based social planning platform, which—unlike check-in based apps like foursquare—is centred on what the user is going to do in the future: the app helps to discover fun events and provides customized and personalized event recommendations, e.g. on the basis of events that Facebook friends are planning to attend.

MapHook[19] allow users to create geo-tagged digital memories about events, locations, and activities. These geo-tagged and user-created "hooks" are then shared and published, together with useful information about the points of interest from Wikipedia or Groupon offers (Fig. 8).

Health and Well-Being Apps

Another category of applications deals more closely with the user's personal sphere. In this case, the location-based features are an addition to the basic app functionalities.

[18] Cf. http://www.hotlist.com/

[19] Cf. http://www.maphook.com

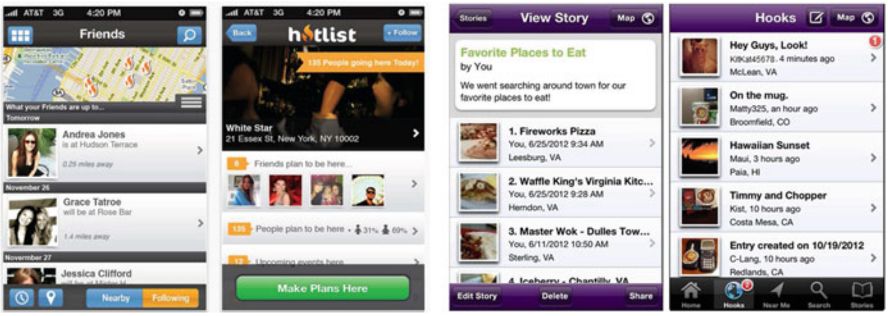


Fig. 8 Experience sharing apps: Hotlist and MapHook (Source: respective websites)

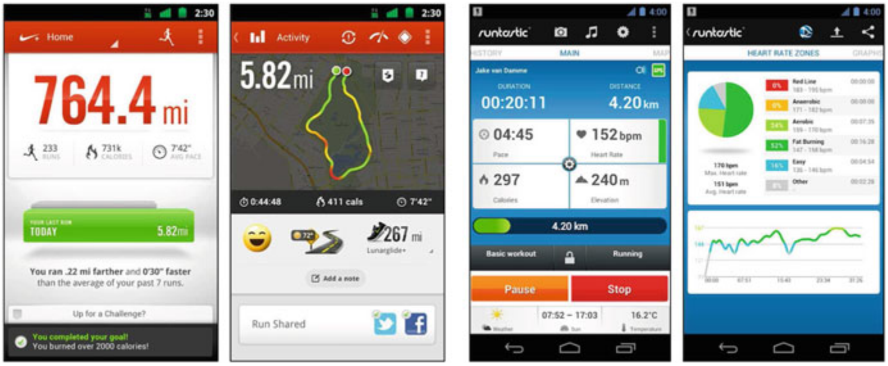


Fig. 9 Health and well-being apps: Nike+ and Runtastic (Source: respective websites)

In this field, an enormous success was gained by fitness-related apps, like *Nike+*[20] or *Runtastic*.[21] Thanks to the phone's GPS and accelerometer, the user can track his runs (in terms of time, distance, elevation change, speed, calories) and his training progress; the social and gaming flavour of those apps also provides the motivation to reach the user's goals. The integration with other sensor information (e.g., heart-beat) provide additional monitoring features to improve or personalize the user's training (Fig. 9).

Alongside those ones, other apps leverage the user's location to a more limited extent: calories' tracking apps suggest local food businesses or restaurants; health monitoring apps suggest the closest physician, hospital or pharmacy; food-specific apps provide personalized suggestions of suitable shops and restaurant, for example in case of gluten or wheat intolerance.

[20] Cf. http://nikeplus.nike.com/

[21] Cf. http://www.runtastic.com/

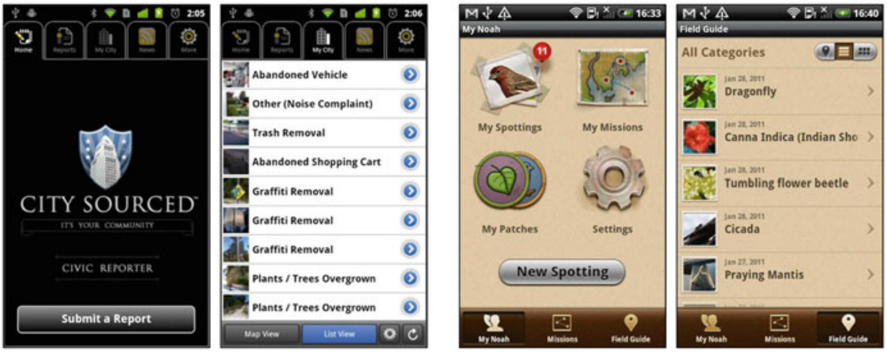


Fig. 10 Citizen participation apps: CitySourced and Project Noah (Source: respective websites)

Citizen Participation Apps

While the applications listed in the previous sections are intended to deliver some functionality or to provide some kind of support to the user, the wide spread of smart phones led also to a different typology of apps, centred on the user as active contributor.

This is the case of apps fostering citizens' participation. A good example is *CitySourced*,[22] a real time civic engagement platform; the mobile app provides a free and intuitive means to empower residents and to stimulate them to identify civic issues (like public safety, quality of life, environmental issues); the app then collects the contributions and reports them to city hall for resolution. The citizen thus becomes a "civic reporter", improves his awareness of the surrounding environment and gets recognition for his assistance.

In the realm of Citizen Science (Irwin 1995)—i.e. the involvement of volunteers to collect or process data as part of a scientific or research experiment—a mobile app like *Project Noah* (Ansari 2013)[23] gives nature lovers the possibility to explore and document wildlife, thus helping scientists with their ongoing research. Users can contribute with their "spottings" and "field missions"; in turn, the app provides location-based field guides to better explore the user's surroundings (Fig. 10).

Finally, in the latest years, an ever increasing adoption of crowdsourcing has happened to distribute work and tasks to available workers all around the world. By leveraging the physical position of those workers, crowdsourcing apps like *Roamler*[24] bring the tasks directly "where" they should be solved by a mobile workforce. Roamler lists a number of potential applications: out-of-stock monitor, category scan, price check, POS check, etc., thus directly targeting business customers; the same platform, however, could be leveraged for other types of missions, like micro-volunteering and social campaigns (Fig. 11).

---

[22] Cf. http://www.citysourced.com/

[23] Cf. http://www.projectnoah.org/

[24] Cf. http://www.roamler.com/

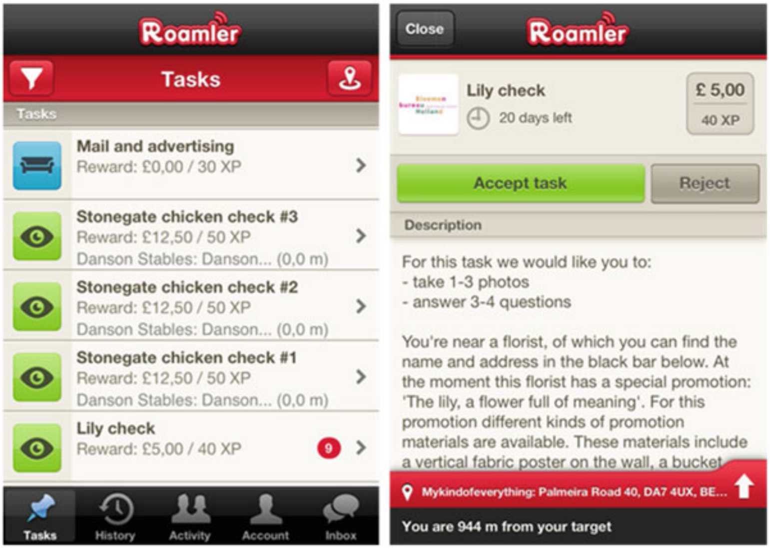


Fig. 11 Citizen participation apps: Roamler (Source: respective website)

Games with a Purpose Apps

The last category of location-based mobile apps is constituted by pure Human Computation (Law and Ahn 2011) examples and, specifically, by Games with a Purpose (von Ahn 2006) designed around a geo-spatial context.

CityExplorer (Matyas et al. 2008)[25] collects images, geographic positions and descriptions of points of interest in cities. The game, which takes inspiration from the Carcassone board game,[26] allows players to conquer POIs by posting their geographic coordinates, photos or tags; the provided information is then checked by other players: if discovered to be correct, the information's author gains points. The purpose of this GWAP is therefore to collect and verify POI data (Fig. 12).

The main purpose of *Eyespy* (Bell et al. 2009) is to identify the most visible and significant POIs in a city. This information can be useful to support navigation or to create tourist maps. In the game, players take photographs that are shared with other players, who then have to find where those pictures were taken. Points are scored by players for both confirming other players' images, but also for producing popular or highly recognized photos. In this way, players are concerned with submitting pictures that are likely to be confirmed by other players, thus diminishing cheating. Moreover, by selecting the photos with the highest hit number, it is possible to highlight interesting paths (Fig. 13).

[25] Cf. http://www.kinf.wiai.uni-bamberg.de/cityexplorer/

[26] Cf. http://en.wikipedia.org/wiki/Carcassonne_(board_game)

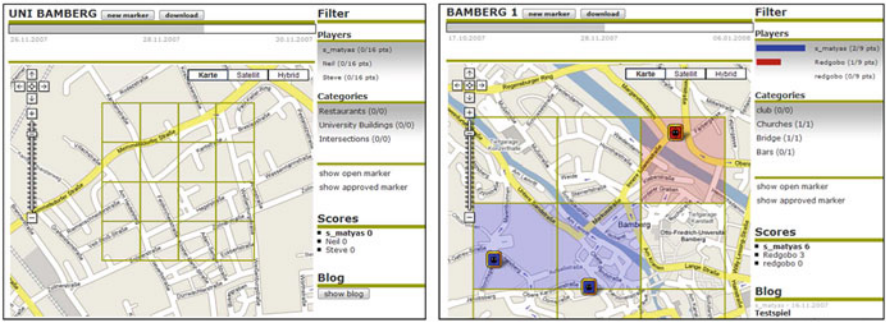


Fig. 12 Games with a purpose apps: CityExplorer (Source: Matyas et al. (2008) and website)

Fig. 13 Games with a purpose apps: Eyespy (Source: Bell et al. (2009))

Similar and more recent applications of this kind—mobile, location-based, GWAPs—were also developed by this chapter's author group.

Selecting and ranking the most representative photos from social media and linking them to the appropriate urban point of interest is the purpose of *UrbanMatch* (Celino et al. 2012). More specifically, the purpose of UrbanMatch is to derive meaningful links between a datasets containing the points of interest (POIs) in a urban environment and a dataset with the images depicting those POIs and retrieved from Web social media; among all photos taken in the proximity of a POI, UrbanMatch is designed for linking the most representative ones to that POI. UrbanMatch is presented to its users as a photo-pairing game; if they associate two photos, this hint is taken as a sign that the two images refer to the same POI; distracting options are given to reduce cheating. The pairs from all players are compared and aggregated to derive meaningful information (Fig. 14).

Fig. 14 Games with a purpose apps: UrbanMatch

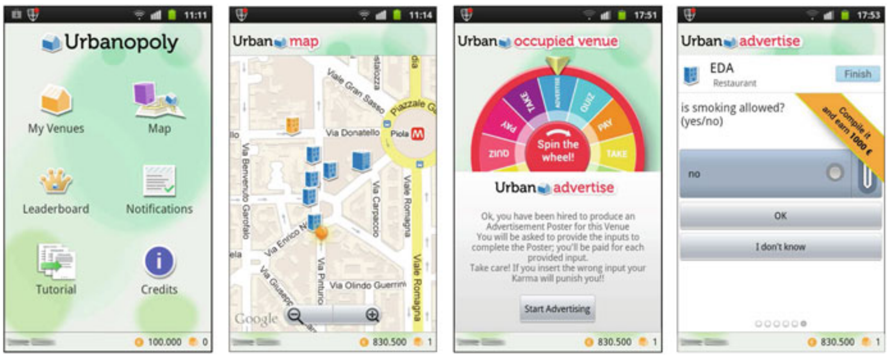


Fig. 15 Games with a purpose apps: Urbanopoly

A more complex GWAP inspired by the Monopoly board game[27] is *Urbanopoly* (Celino et al. 2012). Taking as input open geo-spatial data from the community project OpenStreetMap, Urbanopoly challenges its players to play mini-games in the form of questions, quizzes or quests in order to conquer venues and become a rich "landlord". The different mini-games are the expedient to insert different challenges within the app: some missions are data collection tasks, some other actions require the player to solve data validation tasks. An aggregation algorithm (Celino 2013) combines players' actions to consolidate up-to-date and reliable information. The gameplay and the competition with friends, on the other hand, provide the long-term incentive for players (Fig. 15).

[27] Cf. http://en.wikipedia.org/wiki/Monopoly_(game)

Table 1 Comparative analysis of location-based apps along different dimensions

App category	Location	Audience	Time	Workflow
Social networking	Medium-short distance	Multiple	Short term	synch
Games	Medium-short distance	Multiple	Long term	synch or asynch
Commerce/marketing	Specific POIs	Single/multiple	Medium-long term	asynch
Sharing	Path/specific POIs	Single/multiple	Long term	asynch
Health/well-being	Path/wide area	Single	Long term	asynch
Citizen participation	Wide area	Multiple	Short term	asynch
GWAPs	Specific POIs	Multiple	Medium term	asynch (short term)

Comparative Analysis of Location-Based Applications

To analyse the apps presented before, we propose four dimensional axes: location, audience, time and workflow.

The *location* axis is the "playing area", i.e. the geo-spatial boundary, the real world layer within which the application can function. For example, an app could require the user to be in the proximity of a single POI, or to move along a path, or more generically to be in a specific area. This axis influences the app design in several aspects, from the mechanics to the typology of tasks and actions required to the user (e.g., if the user is required to shoot a photo, he must be very close to the subject to be represented).

The *audience* axis reflects the number of users/workers needed to complete a task, thus splitting the apps in two groups: multi-user and single-user apps. The two categories have different requirements and mechanics. For example multi-user apps could be required to cope with real-time interaction between users; thus, in turn, could introduce a hard-to-be-achieved requirement, such as having two or more users in the same location at the same time.

The *time* axis expresses the "duration", the expected time-span in which the user continues to play the app; in the specific case of games, this is an important aspect, since it divides casual games—short and repeated game levels that let the user stop playing at any time after completing a level—from hardcore games—a storyline and a long-term strategy are needed to win the game. The time axis thus influences the type and number of tasks a user may solve by using the app.

The last axis is the *workflow*, which accounts for the coordination and simultaneity of the app execution between different users. In GWAPs, the synchronism between players is important and sometimes mandatory, because the validation process is based on the agreement/disagreement between players; on the other hand, in location-based apps, the user synchronism could be difficult to achieve.

Table 1 recaps the analysis of location-based apps along the proposed dimensions. While the described examples may be only partially representative for the app categories, we believe that those characteristics are generally valid for those genres and contribute to the app success. It is apparent that different combinations of the

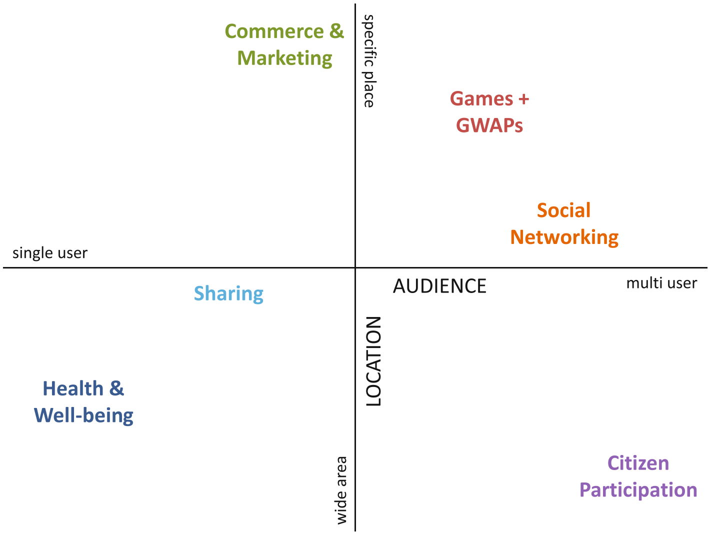


Fig. 16 Positioning of location-based apps along the location and audience axes

four "ingredients" result in different applications. Therefore, when designing a location-based app, it is important to correctly balance the different aspects, in order to achieve the desired goal.

Figure 16 illustrates the positioning of the analysed apps along the location and audience axes. Since all apps are location-based, the location axis is important for all of them, even if the specific location is more essential when the apps centre their functionalities (or their business) on precise POIs. While all applications assumes to have multiple users, the need for cooperation/competition between users varies widely among the analysed apps: the social aspects is stronger for networking, gaming and public participation apps, in which the interaction between users is an important trait of the app mechanics; on the other hand, personal and commerce-oriented apps can be used by single customers, even if sociality can introduce further motivation or enjoyment (e.g. competition on run performances for fitness apps).

Figure 17 illustrates the positioning of the analysed apps along the time and workflow axes. While all developers would like their apps to be addictive enough to keep the user running them in the long-term, the "fidelity" to the app is not a homogeneous requirement: participation or networking apps can survive to casual or non-returning users, while sharing or well-being apps need a long term-engagement. With regards to the workflow axis, most apps do not require simultaneous access from users; even GWAPs that use redundant workers to solve the task designed mechanics that postpone the cross-checking after the data collection; pure gaming apps show the highest variability in workflow approaches, while social synchronous use is mandatory for networking apps.

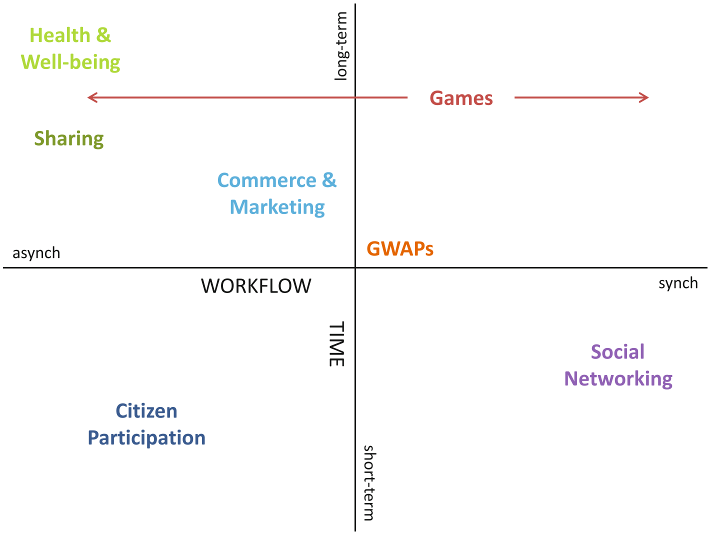


Fig. 17 Positioning of location-based apps along the time and workflow axes

A Reference Methodology for Location-Based Games with a Purpose

Like in every application field, there is no "magical formula" to design and develop a successful location-based GWAP; especially in the realm of mobile stores, the popularity of a single app heavily depends on different factors, especially because mobile apps are usually required to be either useful or fun. In this section, we would like to give some general guidelines and best practices to help prospective developers in designing location-based GWAPS that are both useful and fun (section "Guidelines and Best Practices to Design a Location-Based Games with a Purpose") and to provide a check-list of questions to guide the tailoring and improvement of that design (section "Check-List to Refine and Improve a Location-Based Games with a Purpose").

Guidelines and Best Practices to Design a Location-Based Games with a Purpose

The design of a location-based GWAP is similar to the design of any mobile app. Thus, in this section, rather than focusing on software design details, we would like to give some "rules of thumb" to help prospective developers reflect and think about their motivation and objectives. The following guidelines are given in no specific order, but a full understanding of all of them is recommended to be aware of all the facets of a location-based GWAP design.

Determine your location. So, you want to design a location-based app. But, what is your location? Define and characterize the area relevant for the app (e.g. a neighbourhood, a city, a region, the whole world); define and describe the relevant entities for your app (e.g. POIs, roads, people), their boundaries (e.g. indoor areas), and the requirements for proximity (i.e. how close should the player be to be able to play the game?).

Understand your audience. It is very important to clearly determine the players you intend to involve: who could use the app? What skill do they need to play the game? Distinguish along various axis, including age, language, education, mobility patterns, habits, etc.

Understand your context. The app will be run in a specific situation; not only the location is important, but also the environment and conditions in which the player is immersed. When will the app be used (e.g. morning vs. evening, free time vs. work, when seated vs. while running, for 2 min vs. for an entire hour, etc.)? Where will the game be played (e.g. indoor vs. outdoor, at office vs. on the bus, standing vs. seating, etc.)?

Motivate for long-term engagement. Most mobile games are intended for short rounds but also for several repeated matches. What reward scheme will you adopt to convince the first-time player to re-launch the app and play it again? Will the game be organized in levels of increasing difficulty? Will the game provide a long-term goal or theme that carries from one task to the next? Will there be points, badges, leaderboards, etc.? Are there deterministic as well as random wins and losses?

Learn from the masters. Studying and understanding the success factors of existing games and apps is important to detect the pros and cons of your design. Does the app reuse some known game mechanics that look familiar to the first-time payer? Is the game linked to other successful systems (e.g. social networks)? Can you exploit and leverage a pre-existing community (e.g. friendships) to provide further engagement?

Think outside the box. It is not obvious that a single app could let you fully reach your purpose. Can you combine the game—to target casual users—with other social applications and initiatives—to engage active participants—or pure crowdsourcing efforts?

Design for mobile devices. Never forget the specific characteristics of mobile devices, like the screen dimension. Make sure to adopt standard usability guidelines and follow simplicity as your main goal. Remember that Internet connection can be unstable (e.g. design for temporary local storage and delayed data transmission) and that positioning can be imprecise (e.g. do not rely on phone localization for short-distances and always allow for playing within a vicinity radius).

Check-List to Refine and Improve a Location-Based Games with a Purpose

Now you have designed your location-based GWAP and you would like to revise your assumptions and choices. The first way to do this is by testing, i.e. by involving early adopters and play the game for some time to understand merits and shortcomings; SWOT analysis (Strengths, Weaknesses, Opportunities and Threats (Hill and Westbrook 1997)) could be very useful at this stage, especially if you have early mock-ups and prototypes, which are not fully-functioning but give an idea of how the game is shaping up.

Besides this testing, in the following we provide a sort of check-list made of questions aimed to evaluate the current design and enable its tailoring and improvement. We divide the list into five categories that we believe are crucial for a successful location-based GWAP. Data and purpose. Which kind of data can be collected through the game (in terms of value, quality and quantity)? Are the game mechanisms well integrated with the data processing part?

Gameplay. Are the game rules easy to learn? Is the game theme/metaphor consistent with the purpose? Does the concept allow for future developments and additional extensions/possibilities?

Feasibility. Does the game require mobile device's sensors (e.g. GPS, camera, sound recorder, etc.)? Does the game require a continuous Internet connection? Can the game be played in closed spaces? What is the difficulty to prepare an initial set of data (bootstrap phase)?

Mobility. Does the game concept includes location-aware mechanics? Can it be played on a mobile device? Does it rely on the layer's physical presence in the environment?

Sociality. Does the game offer external rewards? Can the game build a community of users? Can it leverage an existing community? Does the game encourage collaboration among users (in the real/virtual world)?

Towards Citizen Computation

In this chapter, we introduced the concept of Location-based Games with a Purpose and we discussed their characteristics and their relevance. Besides the success of location-based mobile apps and the ever increasing popularity of Human Computation initiatives, we believe that this category of applications can be expected to become prominent in the research arena because of an emerging branch of Human Computation that we named *Citizen Computation*.

Citizen Computation sits at the crossroads of several research fields. It relies on *Human Computation* (Law and Ahn 2011) to provide a human-based solution to unresolved tasks; it understands from *Crowdsourcing* (Doan et al. 2011) the ability to involve a crowd of workers and learns from Citizen Science (Irwin 1995) how to recruit volunteers on the territory; finally it builds on *Social Computing* (Wang et al. 2007) to understand and leverage social ties and interactions. With specific reference to location-based GWAPs, we believe that Citizen Computation games represent a valuable development choice to achieve the objective to "change the world" (McGonigal 2011) in a playful and entertaining way.

Acknowledgements This work was partially supported by the PlanetData EU project (FP7-257641). The author would like to thank Marta Corubolo, Simone Contessa, Daniele Dell'Aglio, Emanuele Della Valle and Stefano Fumeo for the valuable contribution to the discussion.

References

Ansari Y (2013) Biodiversity monitoring and conservation. In: Michelucci P (ed) Human computation handbook. Springer, New York

Bell M, Reeves S, Brown B, Sherwood S, MacMillan D, Ferguson J, Chalmers M (2009) Eyespy: supporting navigation through play. Technology 123–132

Billinghurst M (2013) Augmented reality interfaces in human computation systems. In: Michelucci P (ed) Human computation handbook. Springer, New York

Celino I (2013) Human computation VGI provenance: semantic web-based representation and publishing. Trans Geosci Remote Sens

Celino I, Cerizza D, Contessa S, Corubolo M, DellAglio D, Valle ED, Fumeo S (2012) Urbanopoly–a social and location-based game with a purpose to crowdsource your urban data. In: 2012 international conference on social computing, Amsterdam. IEEE, pp 910–913

Celino I, Contessa S, Corubolo M, DellAglio D, Valle ED, Fumeo S, Krüger T (2012) Linking smart cities datasets with human computation–the case of urbanmatch. In: The semantic web–ISWC 2012, Boston, pp 34–49

Deterding S, Dixon D, Khaled R, Nacke L (2011) From game design elements to gamefulness: defining "Gamification". In: Proceedings of the 15th international MindTrek conference, Tampere. ACM, pp 9–15

Doan A, Ramakrishnan R, Halevy A (2011) Crowdsourcing systems on the world-wide web. Commun. ACM 54(4):86–96

Dou Z, Song R, Wen J.-R (2007) A large-scale evaluation and analysis of personalized search strategies. In: Proceedings of the 16th international conference on world wide web, Banff. ACM, pp 581–590

Entertainment Software Association(2012) Essential facts about the computer and video game industry. http://www.theesa.com/

Goodchild MF (2007) Citizens as sensors: the world of volunteered geography. GeoJournal 69(4):211–221

Hill T, Westbrook R (1997) SWOT analysis: it's time for a product recall. Long Range Plan 30(1):46–52

Hodson H (2012) Why Google's ingress game is a data gold mine. NewScientist.com

Irwin A (1995) Citizen science: a study of people, expertise and sustainable development. Routledge, London/New York

Kamvar M, Baluja S (2006) A large scale study of wireless search behavior: Google mobile search. In: Proceedings of the SIGCHI conference on human factors in computing systems, Paris. ACM, pp 701–709

Law E, Ahn LV (2011) Human computation, vol 5. In: Synthesis lectures on artificial intelligence and machine learning. Morgan and Claypool, San Rafael

Matyas S, Matyas C, Schlieder C, Kiefer P, Mitarai H, Kamata M (2008) Designing location-based mobile games with a purpose: collecting geospatial data with CityExplorer. In: Proceedings of the 2008 international conference on advances in computer entertainment technology, ACE'08, Atlanta. ACM, New York, pp 244–247

McGonigal J (2011) Reality is broken: why games make us better and how they can change the world. Penguin Press HC, Westminster

Nielsen Mobile Insights (2013) The mobile consumer: a global snapshot. Technical report, The Nielsen Company

von Ahn L (2006) Games with a purpose. IEEE Comput 39(6):92–94

Wang F-Y, Carley KM, Zeng D, Mao W (2007) Social computing: from social informatics to social intelligence. Intell Syst IEEE 22(2):79–83

Augmented Reality Interfaces in Human Computation Systems

Mark Billinghurst

Introduction

Augmented Reality (AR) is an interface technology that aims to seamlessly merge the digital and physical worlds. Although AR has been used to describe many different types of technology, Azuma provides the most widely accepted definition, saying that AR systems share three common characteristics (Azuma 1997):

1. They combine virtual imagery with a view of the real world
2. They support real time interactivity
3. The virtual imagery shown is registered in three dimensions

AR systems with these characteristic have been available since the 1960s when Sutherland developed a see through head mounted display (HMD) and used it to show simple computer graphics overlaid on the real world (Sutherland 1968). In the almost 50 years since researchers have explored how AR could be used in engineering, entertainment, medicine and a wide range of other application areas. Recently a growing number of commercial AR applications have begun to be delivered through smart phones, the web, on gaming consoles or other readily accessible technologies. Today hundreds of millions of people can have an AR experience with the technology that is in their pocket, home or office.

Around the same time that AR was becoming readily accessible, the field of Human Computation was beginning. As defined by Von Ahn's ground breaking 1995 dissertation (Von Ahn 2005), Human Computation (HC) is …*a paradigm for utilizing human processing power to solve problems that computers cannot yet solve.* In general, HC crowdsourcing systems use web technologies and distributed networking to allow remote humans to perform simple computational tasks.

M. Billinghurst (✉)
The HIT Lab NZ, University of Canterbury, Christchurch, New Zealand
e-mail: mark.billinghurst@hitlabnz.org

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_26,

In this chapter we describe how AR technology could be used as a front end for Human Computation systems. Although AR technology has been applied in many application domains there has been little research on how it can be combined with Human Computation. In the next section we review related work in AR and HC systems. Then we present early research prototypes that combine AR and HC to explore this potential. Finally we discuss promising directions for future research.

Related Work

The greatest opportunity for using Augmented Reality with Human Computation systems is in the area of mobile Augmented Reality. This is because mobile AR systems enable users to easily interact with their surrounding real environment, are currently the most widely deployed AR systems, and use the same network infrastructure as Human Computing applications. In this section we review how each of these fields developed independently before presenting research that shows how they can be combined together.

Augmented Reality

The first mobile AR applications were based on bulky backpack systems that combined custom made portable computers with a head mounted display, GPS and compass sensing and a variety of input devices. For example, the Touring Machine allowed users to walk around a university campus and see virtual tags appearing over the buildings explaining what departments were in them (Feiner et al. 1997).

In more recent years, mobile phones and handheld devices have become powerful enough to be able to provide AR experiences (Schmalstieg and Wagner 2007) and cumbersome backpack systems are no longer needed. Commercial applications such as Layar (http://www.layar.com/), Wikitude (http://www.wikitude.org/) and Junaio (http://www.junaio.org/), among others, can be used to see virtual representations of points of interest (POI) superimposed over the real world (see Fig. 1). In this case virtual content is shown over the phone's live video view of the real world, and the integrated GPS and compass sensors are used to locate the phone and what the user is looking at. Many mobile phone based AR applications have been developed, such as for tourist guides (El Choubassi et al. 2010), as a restaurant finder (http://www.yelp.com/), or providing directions to the subway (http://www.acrossair.com/apps_newyorknearestsubway.htm), among others.

Many of these systems use an AR browser approach where a thin client application is installed on the user's mobile phone and content is retrieved from remote servers depending on the user's location and the content type they are interested in. For example, if a user is interested in finding a restaurant to eat at they could subscribe to a restaurant information channel and request restaurant POI relative to

Fig. 1 Junaio AR interface for showing bus stop locations

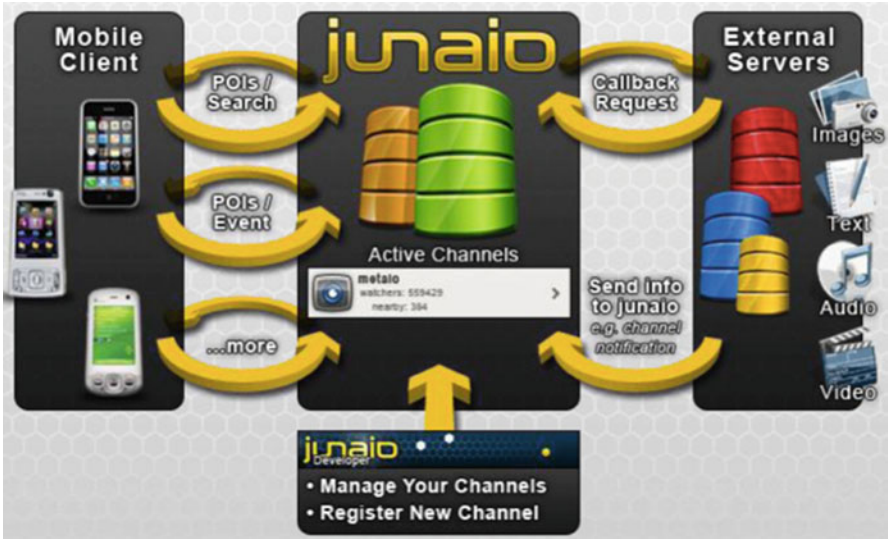


Fig. 2 The Junaio client/server architecture

their own location. This can be shown in an AR view on their mobile device. Junaio and other popular AR browsers use this approach to allow tens of millions of people to see AR content on their mobile devices. Figure 2 shows how Junaio routes requests for information content from the mobile AR client, through its own servers and to external content servers and back again.

Just like traditional web browsers, the client/server architecture of mobile AR browsers has a number of advantages, including: (1) having a single consistent interface for experiencing a wide variety of AR material, (2) needing to only install a small piece of browser software on the mobile device and then download local content of interest when wanted, and (3) being able to simply modify content at the server and then push it out to all subscribed users. However, as Langolatz points out, in many cases AR browsers display information created by professionals in an offline step that limits the dynamic and social aspects of content creation and the amount of content created (Langlotz et al. 2011b).

Content creation is an important problem for mobile AR. In 1999 Spohrer envisioned a system call the "WorldBoard", which was a combination of distributed online information systems and geo-referenced indexing (Spohrer 1999). Information could be freely created and published by users and was indexed by location, not abstract URL. Mobile AR is one technology that could allow users to post messages anywhere and to retrieve any information in any space.

One step towards Spohrer's vision is through adopting Web 2.0 technology. Initially the web was a mostly a source for information and used for one-way information retrieval. Only a few people created content, and most users accessed content without creating or modifying it. The web content itself was mostly static and did not allow users to interact with it or provide additional information. This evolved into Web 2.0 which is characterized by open communication, decentralization of authority, and freedom to share and re-use Web content (Barsky and Purdon 2006). Web 2.0 allowed such innovations as social networking, crowd sourced content, and human computation. This opened the way for services based on user participation, like Flickr, YouTube, and Facebook, among others.

In an approach called AR 2.0, Schmalstieg et al. (2011) describe how AR interfaces can be combined with Web 2.0 technology to create large scale AR experiences that combine input from web services, social networking, and user generated content. A location-based AR application uses data and services remotely stored and served by web mash-ups, visualized on the user's mobile device. The user can be offered data and services related to geospatial information corresponding to user location and geo-database web services. Content authoring can be performed using a computer or directly on the mobile device while on location. Taking advantage of open APIs and mash-ups, complex applications can be easily broken down into smaller components and leverage existing on-line services.

In recent years, developers have begun to develop initial AR 2.0 experiences. The first of these explore semi-automatic ways of developing AR browser experiences, and allowing users to generate their own content. For example, Twitter 360 (http://www.twitter-360.com/) is a mobile AR application that provides an AR interface to Twitter feeds and allows users to see geo-located tweets superimposed over the real world (See Fig. 3). In this way when the people they are following send tweets users can see both the content of the message and the location of the person sending the message. In a similar way Yelp (http://www.yelp.com/) and other mobile applications have an AR view where user generated content is automatically formatted in a way that can be overlaid on the live camera view.

Other applications such as Sekai Camera (http://sekaicamera.com/) explicitly allow users to add their own content from within an AR view (see Fig. 4). In this case users can take pictures, write a message, or record an audio clip and leave it at their current location as a public annotation for others to see. When another user visits the same location they will see the AR content left by others. Researchers

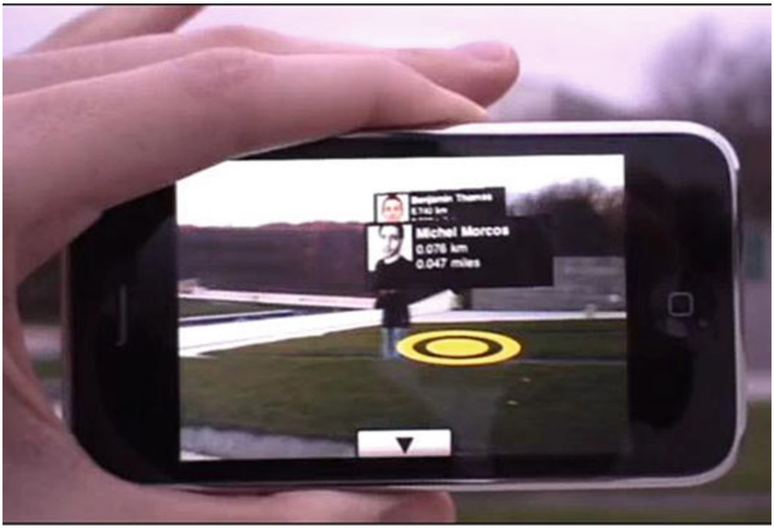


Fig. 3 Using Twitter 360 to see geo-located tweets

Fig. 4 Sekai camera interface showing user generated AR content

such as Langlotz et al. (2011a) and de-las-Heras-Quiros et al. (2010) are exploring other ways for mobile AR users to edit content in place on their mobile devices. For example, Libregeosocial (http://www.libregeosocial.org/) is an open source AR browser that supports content creation, allowing users to add virtual labels to objects in the real world, not just locations (de-las-Heras-Quiros et al. 2010). Langlotz's outdoor AR system uses GPS and panorama-based vision tracking combined with sensors for tracking the mobile phone orientation, and allows users to add virtual 2D and 3D objects to the AR scene (Langlotz et al. 2011a). Most recently, 13th Lab's Minecraft Reality allows people to view digital content created in the game Minecraft in an AR view on their mobile phones (http://minecraftreality.com/). This is one of the first examples of a mobile AR application that is able to view content created in a gaming environment.

As can be seen from this work, mobile AR has developed to the point where AR experiences can be readily deployed on mobile devices, but large-scale content creation remains a problem. Recent work integrating AR applications with Web 2.0 infrastructure has enabled a variety of automatic and user controlled methods for authoring AR scenes. However there is still a need for people who are able to support the content creation process and perform tasks such as image identification, data filtering, quality control and other actions. This is where Human Computation can be used. In the next section we provide a summary of research efforts in Human Computation and show how they have evolved to where they can contribute to AR experiences.

Human Computation

Following on from Von Ahn's early work (2005) there have been a wide range of Human Computation applications developed. Many of the first uses were for web-based applications that could be performed by almost anyone with a limited skill level. Typical of this is the ESP game that showed pairs of player's pictures and gave them points for arriving at the same keywords (von Ahn and Dabbish 2004). This was a very effective way to use Human Computing to semantically tag thousands of pictures. Subsequent research developed games for collecting geographic information about images (Arase et al. 2009), collecting common-sense facts (Von Ahn et al. 2006a), and locating objects in images (Von Ahn et al. 2006b). All of these efforts have validated Human Computation as an effective way of performing simple recognition and understanding tasks. The semantic games portal is a website with a link to dozens of similar web based semantic games (http://semanticgames.org/). These games are typically not designed for the mobile platform or fast response.

Some systems also rely on additional cultural knowledge and expertise for success. For example, Liu at. al. present a mobile application developed for travelers in Japan that allows them to have near real time translation of Japanese characters (Liu et al. 2010). The user takes a picture and then sends it to a number of native Japanese translators that perform a translation to English that is sent back to the requester. This approach has many possible applications, for example, a person may take a photo of a restaurant sign and ask what it means, or how to find their train on a train timetable. User studies found that if the translators understood clearly what the requester was asking they were able to give effective translations in a timely manner, and produced better results than automatic methods.

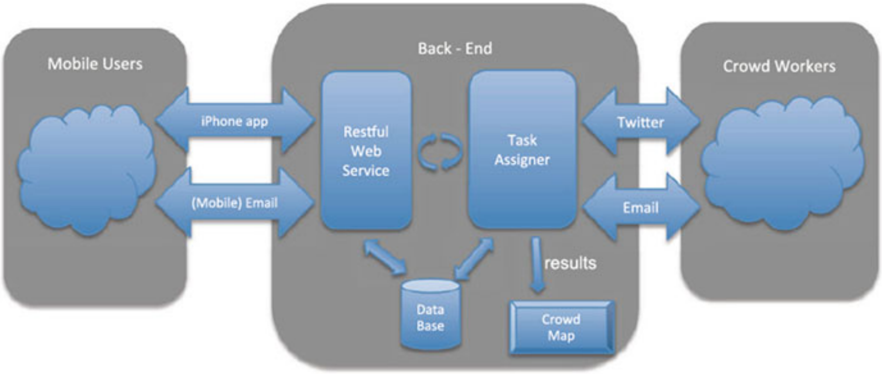


Fig. 5 UbiAsk architecture, typical of mobile HC applications

UbiAsk (Liu et al. 2011) is an extension of this approach that allows travelers in Japan to ask questions about any pictures taken, and not just request translation services. It adds game elements by allowing players to control virtual territories on a map of their city based on the number of tasks that they successfully complete. As part of this research a generalized architecture was developed that could be used for many mobile human computing applications (see Fig. 5). In this architecture, a mobile application is connected to backend services which have a task assigner that can use social network services (e.g. Twitter) to connect to external workers that perform the HC task.

These HC systems rely on people's skill and cultural knowledge to perform certain tasks. However, recently systems are being developed that use people's physical presence in an environment and are more location dependent. This can be particularly important when seeking to improve geospatial data quality. For example, UrbanMatch (Celino et al. 2012a) is a mobile location based game that uses player's familiarity with a city to link photos with points of interest in the city. Players are shown points of interest and known images from a trusted source (e.g. OpenStreetMap) and asked if photos from an untrusted source (e.g. Flickr) might also relate to the point of interest. In order to complete the task task, players must players much have local knowledge of the city where the points of interest are located. UrbanMatch uses a similar mobile client/remote server architecture as UbiAsk (see Fig. 6). User testing found that people were able to correctly identify over 99 % of the candidate pictures and 91 % of the players rated the game as easy to play.

A similar project to this is Urbanopoly (Celino et al. 2012b), which is a social, mobile and location-based game designed around the idea of the Monopoly board game (see Fig. 7). The goal of the research was to use people physically present in the environment together with location based technologies to improve the quality of street data collected by the OpenStreetMap community. Players play a series of mini games that allow them to earn money and buy venues. These games require them to check existing information about real businesses, or enter additional information such as what type of business they are, the food they serve if they are a

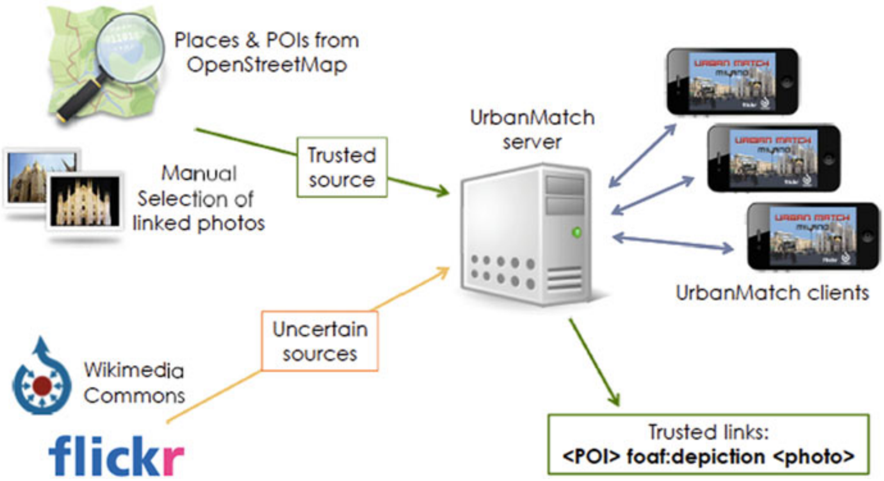


Fig. 6 UrbanMarch architecture

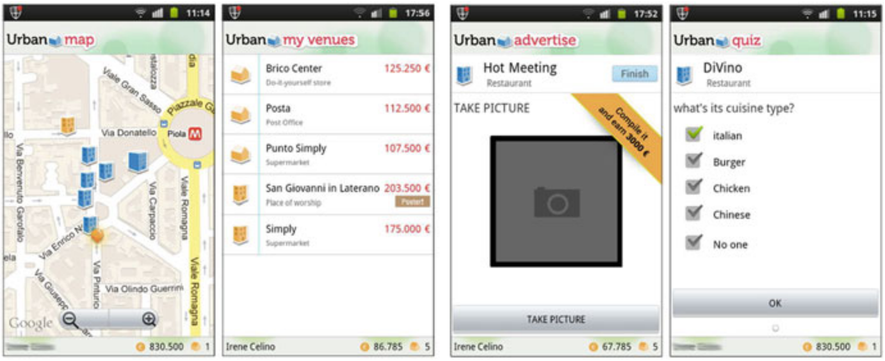


Fig. 7 The Urbanolopy interface showing the map and game elements

restaurant, etc. In this way the system can be used to update the OpenStreetMap database. The system is also tightly integrated with Facebook which enables players to share status and motivate each other to keep on playing.

As can be seen there has been an evolution in Human Computation research from systems that can be performed by almost anyone, to systems that require more skill and can only be completed by people at certain locations, or with local knowledge. This trend to support more location awareness in HC applications is a perfect complement to the growing interest in the social networking from the AR community. In the next section we give an overview of efforts to combine Augmented Reality and Human Computing.

Augmented Reality and Human Computation

As the previous sections show there is a lot of potential for AR systems to use HC to provide content, and to support processing in other ways. However there has been little research to date combining AR and HC systems. In this section we review the first research efforts in this area.

A simple example of how AR can be combined with HC is the social network based mobile AR work of Song et al. (2010). They have developed a mobile AR application that allows users to point their mobile phone camera at an object and then query what the object is. The results of the query are shown as an AR overlay on the mobile phone screen (see Fig. 8). This type of application has many possible uses such as a museum information system, or for tourists. However in the past similar systems have been limited by the size of the image database used for recognition. Their work uses a connection to social network services and human tagged pictures to overcome this limitation.

Song et al.'s system is made up of two parts; a mobile AR client and a mobile AR server, both connected to Twitter (see Fig. 9). Twitter allows users to provide links to images and text that provides information about these images. The server combines an image recognition module, a social network service (SNS) crawling module and a database of images. The SNS crawling module gathers images with their text tags from Twitter and enters them into an image database. A specific Twitter tag is used to identify image content submitted for the service. These images are then matched against image queries coming from the user using an image-matching algorithm, and the results provided back to them through the mobile AR client. This approach has three benefits; (1) users can participate in the mobile AR service directly, (2) the image database can be continuously extended with user submitted content, and (3) the image database is likely to have redundant images, which can be used by the server to improve recognition accuracy. The overall result is that Human Computation can be used to overcome the database limitations and improve image-matching accuracy in the mobile AR application.

Fig. 8 Using a mobile AR client to identify museum objects

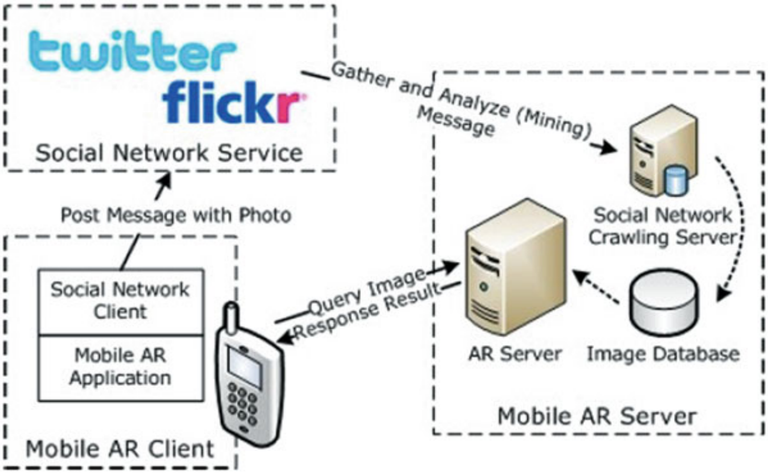


Fig. 9 Connecting a social network service to a mobile AR client

One of the best methods for exploring the use of AR and HC together is to put an AR interface onto an existing infrastructure for Human Computing. This is what the developers of PLASH have done in developing their FriendCompass application (Ho et al. 2010). PLASH (Platform for Location Aware Services with Human Computation) is an open software platform that connects location aware services with Human Computation. Is it comprised of Communication, Data, and Service Layers that connect together to support PLASH applications. The communications layer supports various wireless communication protocols and networking contexts, while the Data Layer is responsible for geo-location data representation, and the Service Layer provides services to support end user applications. Figure 10 shows the multi-layered architecture.

The PLASH platform provides all the infrastructure needs to develop a variety of Human Computing based applications, such as support for geo-queries, user authentication, and social network management. The authors show the flexibility of the platform by presenting several sample applications. These include a Tour Route Recommendation application that uses human computation and data mining techniques to provide personalized recommendations for tourists. This uses the current location and preferences of a user to match with trip history information collected from other users. Another example is a traffic application that allows users to share local information, such as road conditions, traffic jams, and accidents.

However the most relevant is FriendCompass that uses mobile AR to show shared friend location information and points of interest. Users are able to use their current location and orientation to find their friends within view. The server software uses the fundamental services provided by the Service Layer (e.g., authorization control services and location dependent service queries). Figure 11 shows the AR interface; the green icon on the screen is the user's friend. Using PLASH makes it easy for developers to build applications that combine AR and HC.

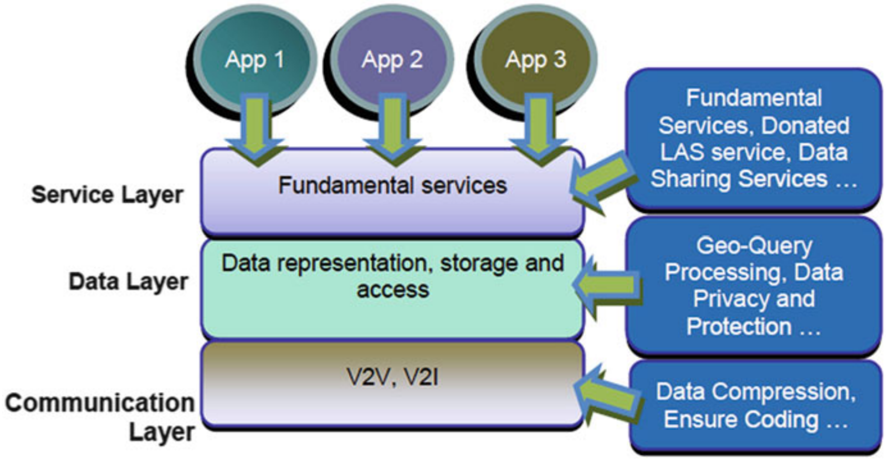


Fig. 10 PLASH multilayered architecture

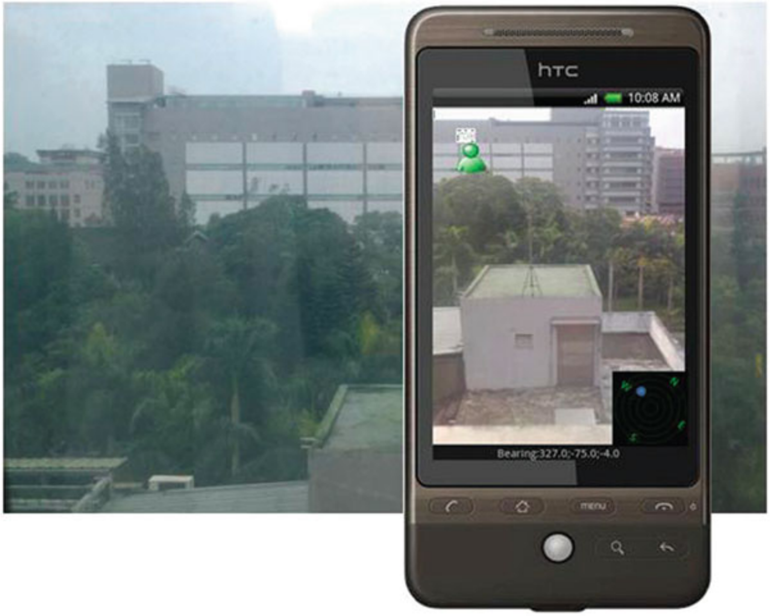


Fig. 11 FriendCompass, using HC to locate friends

Finally, a more complex example is provided by BOTTARI, an AR application for personalized and localized restaurant recommendations (Balduini et al. 2012). BOTTARI is unlike other restaurant recommendation services because it continuously analyses social media streams to understand how people feel about restaurants

Fig. 12 BOTTARI AR interface showing restaurant locations and ratings

in a certain area. It uses both a deductive and an inductive stream reasoner to analyze over 200 million Twitter messages about eating places in Insadong, Korea. The LarKC platform (Fensel et al. 2008) is used for the data processing and reasoning. In addition, BOTTARI uses manually created static descriptions of the same restaurants drawn from recommender sties like Yelp and PoiFriend, Korean restaurant websites, and a few Korean portals. Creating these static descriptions involved a considerable amount of manual work to finally produce a high-quality geo-referenced knowledge base in which each restaurant is described by 44 attributes.

The final result is presented in an AR Android application that shows users' restaurants and their recommendations close to their current location. Users can filter their restaurant search by using buttons that let them find places that are popular, interesting, emerging or personalized for them. Figure 12 is the AR view of the BOTTARI interface, showing virtual tags appearing over live camera views of the users world. The tags provide additional information such as the type of restaurant and the reputation. The interface can also show trending opinions about restaurants over time and more details about particular restaurants. User studies found that BOTTARI provided a much wider range of recommendations that what a tourist could obtain from tourist guides and Web 2.0 sites.

These systems have shown that Human Computation can be used to enhance the usefulness of AR systems and enable large scale-AR applications. The ability to use HC infrastructure to connect to people with local knowledge is particularly valuable. However there are more opportunities for on-going research. In the next section we present one particularly promising direction for future efforts.

Opportunities for Future Research

The AR/HC systems described to date have mostly used Human Computation to improve the quality of location based data in an AR application (e.g. BOTTARI) or used AR as an interface for showing the output of HC data sharing efforts (e.g. FriendCompass). They have been delivered on a handheld AR platform and don't necessarily rely on a real time connection between the user and the Human Computers. However, there are new developments in AR that could provide opportunities for AR and HC research.

One of these developments is the emergence of hands free wearable AR displays and computers that allow people to be always connected to remote collaborators from wherever they are and whatever tasks they are engaged in. Typical of these is the Google Glass system (see Fig. 13) that combines a very lightweight see-through monocular display with a camera, integrated computer, on board audio, and wireless connectivity.

Unlike handheld mobile AR systems, Google Glass provides a hands-free experience and also allows first person viewpoint sharing from the wearer's point of view. This could be used to enable new types of Human Computation applications that are focused on real time information sharing and task assistance. For example, if the user has a problem with their car they could connect to a real time Human Computer to guide them through repairing it. The remote human could both see the user's point of view as well as maintaining an audio connection with then and also using AR visual cues to overlay information on their field of view. Having this shared visual context will enable them to more easily understand the task the user is trying to achieve and provide assistance.

In the future head worn displays like Google Glass will evolve into contact lens based displays and more intimate technology (Parviz 2009). This could enable more organic social experiences where virtual versions of remote collaborators could appear to be sitting around the same conference table as the local person, and so collaborate as naturally as if they were all face-to-face (Billinghurst and Kato 2002). This research will go beyond what UbiTask and other earlier AR and HC systems

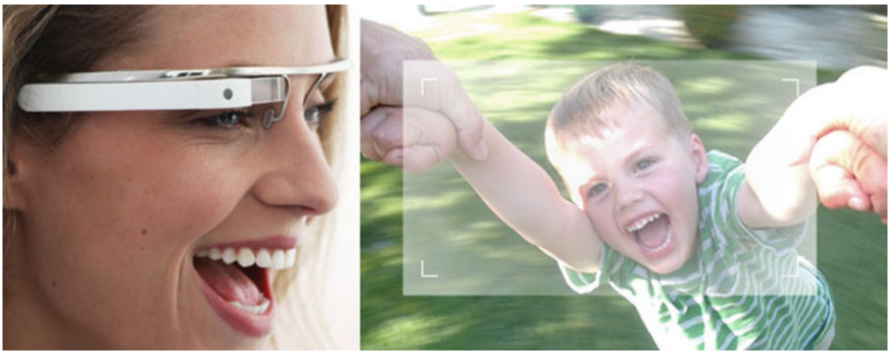


Fig. 13 Google glass system with first person video capture

have achieved by providing both more immediate response and also a greater understanding of the users context. It will allow users to connect globally with Human Computers that have just the right skill set to help them with their local needs.

Conclusions

In this chapter we have reviewed efforts in Augmented Reality and Human Computing and recent research that has combined them together. These early results have shown that Human Computing can be used to overcome some of the content and scale issues of mobile AR and enable useful location based applications. Combining an AR interface with a HC system is a natural way to present location based HC output. There are even more opportunities with the new generation of AR devices that are based around extremely lightweight head mounted displays with integrated cameras. In the future these systems will enable remote people to see what a user is doing and provide natural visual and audio cues to support their actions. This will enable an entirely new class of HC applications.

References

Acrossair subway application—http://www.acrossair.com/apps_newyorknearestsubway.htm

Arase Y, Xie X, Duan M, Hara T, Nishio S (2009) A game based approach to assign geographical relevance to web images. In: Proceedings of the 18th international conference on World wide web—WWW'09. ACM Press, New York

Azuma RT (1997) A survey of augmented reality. Presence: Teleoperators Virtual Environ 6(4):355–385

Balduini M, Celino I, Dell'Aglio D, Della Valle E, Huang Y, Lee T, Kim S, Tresp V (2012) BOTTARI: an augmented reality mobile application to deliver personalized and location-based recommendations by continuous analysis of social media streams. Web Semantics: Science, Services and Agents on the World Wide Web

Barsky E, Purdon M (2006) Introducing web 2.0: social networking and social bookmarking for health librarians. J Can Health Lib Assoc 27:65–67

Billinghurst M, Kato H (2002) Collaborative augmented reality. Commun ACM 45(7):64–70

Celino, I., Cerizza, D., Contessa, S., Corubolo, M., DellAglio, D., Valle, E. D., & Fumeo, S. (2012, September). Urbanopoly--A Social and Location-Based Game with a Purpose to Crowdsource Your Urban Data. In Privacy, Security, Risk and Trust (PASSAT), 2012 International Conference on and 2012 International Confernece on Social Computing (SocialCom) (pp. 910–913). Amsterdam, The Netherlands, September 3rd - 5th, 2012, IEEE Press

Celino I, Contessa S, Corubolo M, Dell'Aglio D, Della Valle E, Fumeo S, Kruger T (2012b) Linking smart cities datasets with human computation—The case of urbanMatch. Int Semantic Web Conf (2)

de-las-Heras-Quiros P, Roman-Lopez R, Calvo-Palomino R, Gato J, Gato J (2010) Mobile augmented reality browsers should allow labeling objects. A position paper for the augmented reality on the web W3C workshop. Available from http://www.w3.org/2010/06/w3car/mar_browsers_should_allow_labeling)objects.pdf

El Choubassi M, Nestares O, Wu Y, Kozintsev I, Haussecker H (2010) An augmented reality tourist guide on your mobile devices. Advances in multimedia modeling. Lecture notes in computer science, volume 5916/2010, Springer, pp 588–602

Feiner, S., MacIntyre, B., Höllerer, T., Webster, A. (1997). A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment. Personal Technologies, 1(4), 208–217

Fensel, D., van Harmelen, F., Andersson, B., Brennan, P., Cunningham, H., Della Valle, E., ... & Zhong, N. (2008, August). Towards LarKC: a platform for web-scale reasoning. In Semantic Computing, 2008 IEEE International Conference on (pp. 524–529).Santa Clara, CA, USA - August 4-7, 2008, IEEE Press

Ho YH, Wu YC, Chen MC (2010) PLASH: a platform for location aware services with human computation. Commun Mag, IEEE 48(12):44–51

Junaio software—http://www.junaio.org/

Langlotz, T., Mooslechner, S., Zollmann, S., Degendorfer, C., Reitmayr, G., Schmalstieg, D. (2012). Sketching up the world: in situ authoring for mobile augmented reality. Personal and ubiquitous computing, 16(6), 623–630

Langlotz T, Regenbrecht H, Schmalstieg D (2011) Large scale content creation for mobile augmented reality, University of Otago Postgraduate Day, 9. Available from http://infosci.otago.ac.nz/assets/ispg/IS-PostGradDay-2011/ISPG-Day-2011-proceedings.pdf#page=22

Layar software—http://www.layar.com/

Libregeosocial website—http://www.libregeosocial.org/

Liu, Y., Lehdonvirta, V., Kleppe, M., Alexandrova, T., Kimura, H., & Nakajima, T. (2010, December). A crowdsourcing based mobile image translation and knowledge sharing service. In Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia (p. 6). Limassol, Cyprus, December 1-3, 2010. ACM

Liu Y, Vili Lehdonvirta TAML, Nakajima T (2011) Engaging social medias-case mobile crowdsourcing. SoME'11

Minecraft Reality website: http://minecraftreality.com/

Parviz B (2009) For your eye only. Spectr IEEE 46(9):36–41

Schmalstieg, D., & Wagner, D. (2007, November). Experiences with handheld augmented reality. In Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on (pp. 3-18). 13-16 November 2007, Nara, Japan. 2007, IEEE

Schmalstieg D, Langlotz T, Billinghurst M (2011) Augmented reality 2.0. In: Brunnett G, Coquillart S, Welch G (eds) Virtual realities. Springer, Vienna, pp 13–37

Sekai Camera website—http://sekaicamera.com/

Semantic Games website—http://semanticgames.org/

Song, I., Kim, I. J., Hwang, J. I., Ahn, S. C., Kim, H. G., & Ko, H. (2010, December). Social network service based mobile AR. In Proceedings of the 9th ACM SIGGRAPH Conference on Virtual-Reality Continuum and its Applications in Industry (pp. 175–178). December 12th - 13th, Seoul, Korea, 2010. ACM

Spohrer J (1999) Information in places. IBM Syst J 38:602–628

Sutherland I (1968) A head-mounted three-dimensional display. In: Proceeding of the fall joint computer conference. AFIPS conference proceedings, vol 33, AFIPS, Arlington, pp 757–764

Twitter 360 website—http://www.twitter-360.com/

Von Ahn L (2005) Human computation, Ph.D. thesis, Carnegie Mellon University

von Ahn L, Dabbish L (2004) Labeling images with a computer game. In: Proceedings of the SIGCHI conference on human factors in computing systems (CHI '04), ACM, New York, pp 319–326

Von Ahn, L., Kedia, M., & Blum, M. (2006a). Verbosity: a game for collecting common-sense facts. In Proceedings of the SIGCHI conference on Human Factors in computing systems (pp. 75–78). Quebec, Canada. 22nd - 27th April, 2006.ACM

Von Ahn, L., Liu, R., & Blum, M. (2006bl). Peekaboom: a game for locating objects in images. In Proceedings of the SIGCHI conference on Human Factors in computing systems (pp. 55–64). Quebec, Canada. 22nd - 27th April, 2006. ACM

Wikitude software—http://www.wikitude.org/

Yelp restaurant finder—http://www.yelp.com/

Pervasive Human Computing

Joel Ross

Humans Computing Everywhere

Humans perform informal computations throughout their daily lives across a variety of localized situations: from the arithmetic of estimating the cost of a purchase at a grocery store, to the calculus of regulating vehicle speed to match surrounding traffic, to executing synchronous scheduling algorithms to make sure that someone picks up the kids from school on time. In this sense, human computation is already a pervasive phenomenon—a process that is performed by a vast number of people in a variety of contexts.

Most prominent human computation systems rely on this pervasiveness in order to enable human-driven problem solving and information processing on a large scale. Human computation is frequently crowdsourced through systems such as Amazon's Mechanical Turk (AMT 2013) in order to either harness the vast quantities of human processing required to make human computation more effective than machine systems, or to enable the benefits of collective intelligence and crowd wisdom (e.g., Lévy 2001; Surowiecki 2005) in solving computational problems. Indeed, the "remote person call" or "human-as-a-service" view of human computation (see Irani and Silberman 2013) relies on such computation to be available at all times: on its home page, AMT describes itself as offering "a global, on-demand, 24 ×7 workforce" (AMT 2013). Human computation systems require a near-constant connection between human computers and the mechanical systems that direct (Quinn and Bederson 2011) their computing.

J. Ross (✉)
University of Puget Sound, Tacoma, WA, USA
e-mail: jross@pugetsound.edu

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_27,

This requirement for constant access to computations performed by humans—a process that already occurs pervasively in a variety of locations—suggests that *pervasive computing* may offer a suitable interaction paradigm for supporting human computation-based systems. Pervasive computing[1] is a model of human-computer interaction (that is, interaction between a human and a computer) that involves moving away from traditional desktop interaction to focus on computing-in-context, embedding digital computer systems into the everyday physical world. Such computing systems may be passively embedded in the environment so that users are only peripherally aware of them (such as with ambient displays (Ishii and Ullmer 1997)), or may represent computing systems with which users actively engage. One of the most common examples of a shift away from the desktop can be found in the increasing ubiquity of mobile devices and smart phones specifically—the mobility and constant network access afforded by such devices allow them to be integrated into everyday interactions, so that their use becomes "pervasive" in everyday experience. Research in pervasive computing often focuses on the ideas of "computing everywhere" and "everything can be a computer." Indeed, emerging research and even consumer products that make use of mobile augmented reality (AR) systems and "wearable computing" (Mann 1997) continue to support embedding computers into people's everyday lives.

Pervasive computing thus offers an intriguing interaction paradigm for human computation. Just as pervasive technologies move digital computation away from the desktop machine into the everyday physical environments, pervasive *human* computation emphasizes moving the human computing into a variety of localized contexts. Indeed, pervasive computing as a form of interaction is highly interested in the context in which computation is used (e.g., Dourish 2004)—how computation can be embedded into the everyday lives of users. Such concerns remain valid even when the computation is performed by humans on the other end of a persistent network, rather than machines. Yet when considering pervasive human computation, we also need to perform a kind of inversion of this focus, since the human computers are the "users" of interest. Pervasive computing considers how computation may be used by humans in an everyday context; pervasive human computing introduces the question of how computation may be *performed* by humans in an everyday context.

In this chapter, I explore some of the uses of pervasive systems as platforms for performing human computation: porting current microtask-based interaction forms to mobile devices, and having humans act as computational controllers for mobile sensors. I discuss how these forms of human computation utilize or respond to the situatedness of the pervasive context in which they are performed. I follow this analysis with a reflection on some of the implications of considering human computation through the lens and goals of pervasive computing, particularly in terms of the visibility of the humans performing computation.

[1] Also known as *ubiquitous computing*, or "ubicomp" for short. Although "pervasive computing" and "ubiquitous computing" have been used to imply different emphases, in this article I will be using them interchangeably.

Mobile Human Computation

As mentioned above, mobile devices such as mobile phones are one of the most common platforms for moving computation into an everyday context and making it pervasive. Indeed, at a simple level, human computation can be made pervasive by porting existing systems and interaction patterns such as AMT for use on mobile devices. As an example, consider Harvard professor Jonathan Zittrain's vision of crowdsourced human computation combined with pervasive technologies:

> One can visualize in the near future a subway car packed with people, each far less attuned to the local environment and to each other than even with today's distractions of newspapers and iPods. Instead, they will stare into screens even for just a few minutes and earn as much money [via systems such as AMT] in that time as their respective skills and stations allow. (Zittrain 2008)

In this scenario, any extra minutes (extra mental "cycles," to use a mechanical metaphor) are devoted towards human computation rather than alternative activities such as media consumption.[2] While Zittrain problematizes this behavior (particularly contrasting for-pay activity with human contact or conversation), such mobile-based human computation need not be entirely profit driven. As a more positively framed alternative, those subway riders could be using their mobile devices to play *Foldit* (Khatib et al. 2011) instead of *Angry Birds*—performing socially beneficial human computation in a mobile context.

In this way, human computation can be made pervasive by making the context in which it is formed more pervasive, such as through mobile technologies. This strategy has been refined through a number of research projects (e.g., Eagle 2009; Gupta et al. 2012; Narula et al. 2011), enabling human computation particularly in the context of developing countries. A second common strategy for making human computation pervasive applies crowdsourcing techniques for data gathering to pervasive contexts, creating what Zittrain goes on to describe as "distributed human sensors" (Zittrain 2008). These systems have humans act as computer sensors and record information about their localized environment (e.g., Paulos et al. 2009; Tuite et al. 2011). I discuss these projects and methods in more detail in the following sections.

In both of these methods, humans perform computation pervasively in the contexts of their everyday lives—yet such methods may or may not fully utilize the pervasive context in which they occur. Pervasive computing gives computing *situatedness*: the computation occurs within a specific local and social situation, allowing that situation to serve as input to and shape the interaction with the computational system. In pervasive human computation, this situatedness may allow human computers to access localized and contextualized knowledge, actions, or behaviors, thereby influencing the computation they perform. In exploring pervasive human computation systems, it is important to consider the impacts and use of this situatedness: what makes pervasive human computation different from non-pervasive human computation?

[2] In his novel *Rainbow's End*, Vernor Vinge expands this vision to include cognitive labor performed through mobile, wearable AR systems.

Again, note that in this chapter I am interested in how human computers perform their computation pervasively, not in how human computation as a replacement for mechanical computation (computation performed by machines) may be used pervasively. There has been significant and admirable work in the latter context: for example, *VizWiz* (Bigham et al. 2010) uses human computation harnessed through AMT to perform pervasive image recognition to support blind people in interacting with their environments. Yet in such systems, the human computation is still performed non-pervasively—the humans doing the image recognition are likely still using the desktop model of interaction, working through AMT using a web browser. Such systems address problems in pervasive computing using human computation, rather than making the human computation itself pervasive, which is the topic of interest here.

Human Computation Tasks on the Go

One of the simplest and earliest ways to make human computation pervasive is to have human computers report the results of their computation through mobile devices. This enables people to perform human computation during their everyday life, in a variety of different contexts and environments. Such systems must be enabled by existing infrastructures for pervasive technologies (i.e., ubiquitous network connections,[3] energy for powering mobile devices, etc.)—pervasive human computing "piggy-backs" off of mechanical pervasive computing systems.

Yet despite these requirements, systems for enabling such pervasive human computation have primarily been explored in the context of developing regions. For example, *txtEagle* (Eagle 2009) built on the ubiquity of mobile devices and GSM reception in East Africa to deliver AMT-style human computation tasks to the mobile phones of workers in Kenya and Rwanda. These tasks—like those in AMT—were performed for pay, and offered as a way to supplement the low-income populations. Indeed, because of infrastructure in place for transferring mobile airtime (and the popularity of using airtime as a kind of currency), payments in either cash or airtime could easily be delivered to workers. The system's use is described with the following hypothetical scenario:

> *David, Maasai Herdsman, Kisumu, Kenya*. While David had been unable to complete formalized education, he, along with many of his Maasi peers, does own a mobile phone. David completes voice-tasks, helping Nokia train a speech recognition engine on his native Maasai dialect. When David wishes to complete a task, he 'flashes' the txteagle Asterisk box that calls him back, asking him to repeat specific key words and phrases. After 30 minutes of work, David has earned enough airtime to last him a week… (Eagle 2009)

Due to the limitations of available mobile phones (e.g., relying on numeric text entry), human computation tasks supported by *txtEagle* were primarily text- and

[3] Though even the computation of transmitting network data could be performed by humans, in what is informally called a "sneakernet".

audio-based: for example, human computers would perform transcription (of English words for those who were fluent) or translating text between their local languages to support software localization. Other systems have been developed to overcome these mechanical limitations. For example, *mClerk* (Gupta et al. 2012) uses proprietary protocols that predate MMS to send images to low-end phones in semi-urban India. This enables human computers in the region to perform optical character recognition (OCR) on scanned images.[4] Similar to *txtEagle*, *mClerk* pays human computers with mobile airline, administered manually through a "recharge shop."

Interestingly, in deploying the system, the researchers developing *mClerk* found that potential workers were skeptical of the system (perceiving it as a possible scam rather than a potential source of income). Yet once they overcame their skepticism, most users reported such human computation tasks were good for killing time. This study highlights some of the complications of developing mobile-based human computation systems: *computation activities need to be able to fit into existing activity structures*. For a human computation system to be operated pervasively, it needs to fill the same interaction gaps addressed by other mobile usage (see e.g., O'Hara et al. 2007)—for example, tasks that computers are able to complete in short bursts of time, or that can be performed while engaged in other activities. The micro-tasks common to systems such as AMT are usually suitable for such situations; nevertheless, such a restriction may influence the development of future pervasive computing systems.

The projects sampled here are all systems deployed within developing regions, raising the question of what factors may make such contexts amenable to pervasive human computing. I suggest that the main factor may be the "for pay" nature of crowdsourced human computation systems (such as AMT) that provide an interaction model for use of these systems. Although the economics of such systems are still being researched (see e.g., Horton and Chilton 2010; Silberman et al. 2010; Toomim 2011), in practice AMT-style tasks are performed for a relatively small wage.[5] As payment is the primary motivator in these markets, a low wage may restrict usage to those computers for whom the wage is still "worth the time": those in developing regions. Even non-pervasive human computation markets such as AMT see more work from lower-income regions such as India than higher-income countries such as the U.S. (Ross et al. 2010).

Thus designing pervasive human computation systems that are deployable in developed regions may require designs beyond "AMT on a cell phone", offering non-monetary motivations for performing computation. For example Heimerl et al. (2012) describe integrating human computation into a vending machine, using non-vital snacks as a reward instead of monetary payment. This design is exemplary of pervasive human computation, as the human computing is integrated into the

[4] *MobileWorks* (Narula et al. 2011) also supports human-performed OCR via mobile phones, but delivers images over a web application that requires a more powerful (and expensive) mobile phone.

[5] In 2009 (Ross et al. 2010) report workers from India make about USD 2.00/h on AMT, while in 2012 (Gupta et al. 2012) report the *mClerk* system payed around USD 2.84/h.

everyday environment. Other motivation structures may avoid extrinsic rewards all together, such as by "gamifying" human computation (e.g., von Ahn and Dabbish 2008; Carranza and Krause 2012). Such efforts can build on research into pervasive games (Montola et al. 2009) and games for harnessing collective intelligence (e.g., McGonigal 2008) to design interactions in which utilizable human computation pervades a game activity, which itself can pervade everyday life.

Whatever the motivation, while deploying AMT-style human tasks to mobile human computers does move the computation into a pervasive context, this form of interaction may not fully utilize the situatedness enabled by pervasive computing. Classical human computation tasks such as image identification rarely depend on or consider the context in which the computation is performed: indeed, identifying images on a mobile phone may even be made more difficult because of differing environmental lighting conditions! Systems such as *txtEagle* and *mClerk* do consider the social and cultural context of the computers to a small extent (e.g., when asking for translations between local languages), but these systems fail to consider the human computer's *specific* environment. Further research is needed into how the specific context in which human computation is performed may influence either the distribution or evaluation of AMT-style tasks in order to more effectively develop pervasive human computation systems.

In sum, the ubiquity of mobile devices offers a suitable platform for developing pervasive human computation systems—whether they simply provide a method for participating in existing crowdsourcing markets while on the go, or if they build on new forms of interaction for motivating contributions during short moments of free time. Yet motivating adoption of human computation platforms may require moving beyond the mobile device as a platform, embedding avenues for performing human computation in the artifacts that fill peoples' environments. Such embedding may help systems to better utilize the situatedness of the pervasive human computing, taking advantage of the computer's specific local and social context.

Human Sensing of Local Environments

While many existing human computation systems utilize the AMT-style "receive a task; complete a task; receive a reward" model of interaction, such systems do not fully utilize the mobile, pervasive nature of the interaction. Other forms of pervasive human computation work to expand the idea of what it means for humans to perform computation in order to take advantage of the localized context afforded by pervasive computing. These systems move beyond asking humans to act as just information processors, to asking them to emulate other aspects of mechanical computation.

The most prevalent of these other aspects is *sensing* the surrounding environment: in particular, having humans control and direct the use of embedded sensors. Also known as *participatory sensing*, this mode of interaction emphasizes crowdsourcing the use of sensors embedded in mobile devices, thereby enabling large groups of people to "gather, analyze and share local knowledge" (Burke et al. 2006).

Such interaction can be used to enable citizen science (e.g., Paulos et al. 2009), having humans direct the collection of pollution or noise data to better inform scientific research. Similarly, other systems such as *PhotoCity* (Tuite et al. 2011) have humans direct the use of an even more common type of sensor: the visual sensors that form the cameras found in most smart phones. In this system (framed as a game to motivate participation), humans use the cameras to intelligently provide photos that can be combined to successfully create a 3D reconstruction of a location. Thus rather than performing computation to *process* data, these human computers use their decision making skills to *produce* data that can then be processed.

As Reeves and Sherwood (2010) point out, the decision-making performed by humans in choosing how to direct the sensors is still a valid form of human computation. Such decisions "draw upon human agency and local practices" (Reeves and Sherwood 2010) to produce data more efficiently than may be produced by a fully automated sensor network (à la, Chong and Kumar 2003), much as human computation can be more efficient at the prototypical task of image identification. By putting humans in the loop in these pervasive sensing systems—turning them into pervasive human sensing systems—the computational efforts exerted by humans can outperform the computational efforts of the machines. Thus such human-directed sensing is a form of pervasive human computation: one that effectively utilizes the situated, localized nature of the computation being performed

Beyond simply directing mechanical sensors, pervasive human computation can even involve humans performing the sensing themselves. In this model of interaction, a system may query people for information that they can sense (e.g., "is there traffic?" "how's the weather?"), and then aggregate that data in order to produce computational models. To ease participation (and make such participation truly pervasive), the aggregating system can rely on reports that humans already produce, such as through social media. For example, people's reports of earthquakes on Twitter can be used to send alerts and notifications faster than traditional reporting systems (Sakaki et al. 2010), or provide situational awareness to support disaster response (Vieweg et al. 2010) *because* sensed data result from very specific contexts. These applications thus demonstrate how the situatedness of pervasive human computation can enable novel and effective systems.

This view that humans-as-sensors perform computation stretches the traditional understanding of what "human computation" entails (though Reeves and Sherwood (2010) note that even some tasks on the traditional human computation platform of AMT, such as writing product reviews, might not be considered "computation"). Zittrain's paper *Ubiquitous Human Computing* (Zittrain 2008) even suggests that systems that report biological vital signs from humans can be conceived as a form of human computation—computation that involves humans directly. Indeed, Zittrain suggests that such sensing could be used to support epidemiology–building on existing data mining systems such as Google Flu Trends (Google 2013). Quinn and Bederson (2011) argue that data mining systems are not human computation systems in themselves, but may they not be systems that involve or rely upon human computation?

In order to consider how human computation can best take advantage of contextual information available in pervasive systems, we may need to expand our understanding of what it means for a task to be computational. For example, social interactions are not normally considered to be computation, yet there may be identifiable "algorithms" which apply in these situations (such as the scheduling algorithm of how to plan one's day). If we want to make human computation pervasive, we may need to apply technomorphisms[6] to the wide range of actors and artifacts that exist within pervasive environments—using the lens of computer science and computation to look at traditionally non-computational systems. Such considerations can help us to take full advantage of the situated pervasive contexts in which pervasive human computation is performed.

Situating Pervasive Human Computation

Pervasive computing is computing that occurs in a variety of contexts: computation in the everyday world in which we live. Similarly, pervasive human computing moves human computers away from the desktop and "into the wild," allowing that computation to occur within a particular localized and social context—where and how the computation occurs matters! But how can we best utilize the contextualization afforded by pervasive human computing? How can performing human computation out in the world benefit existing forms of interaction (beyond simply increasing the availability of human workers), or otherwise enable the development of new systems? Future research is needed to further study the impacts of pervasive computing's situatedness on human computation, and how to best harness local contexts in human computation systems. Thus the significant open question is: *in what ways does the situatedness enabled by pervasive systems influence human computation?*

For one, research needs to explore how location influences computations performed: do humans tend to perform different types of computation (or perform computation in different ways) depending on their location? Are there problems that are dependent on localized computation but may be amenable to completion by human computers? Are there forms of human computation that could be immediately applied to problems in a local environment?

Second, research might consider how the presence of other nearby actors and artifacts—human or mechanical—can shape human computation performed pervasively. For example, research might consider the effectiveness of encouraging impromptu face-to-face collaborations, either between existing social groups or between co-located human computers. Other systems might use the pervasive presence of computers in order to help organize or control devices embedded in the environment. The question of how human computers may interact with their environments when computing in a pervasive context—how to best harness the potential benefits of this interaction—requires further study.

[6] A play on "anthropomorphism," referring to the attribution of technological characteristics to non-machines; see e.g., Vertesi (2008).

Finally, what are the influences of different cultural or social contexts? Cultural context is already a factor that needs to be considered when using existing human computation systems: translation tasks may require a certain fluency, or identification tasks may rely on knowledge of particular cultural touchstones. These issues may be further complicated when human computation occurs in a potentially more heterogeneous pervasive context. Similarly, the value or acceptability of systems may be influenced when presented within a social context that is not traditionally understood as computational—such as how the *mClerk* system was viewed as a potential scam (Gupta et al. 2012). The relationship between human computation, the connectivity and attention it requires, its framing of human labor, and other such factors need to be carefully considered in the development of pervasive human computation systems.

These are just some example questions that are ripe for future research; indeed, all these questions will need to be addressed in order to effectively utilize the situational context in which pervasive human computation is performed.

Invisible Human Computation

The research domain of pervasive computing is significantly based on the vision presented in Mark Weiser's foundational article, *The Computer for the twenty first Century* (Weiser 1995). In this paper, Weiser highlights the "seamlessness" of computer interaction enabled by pervasive computing—computers are so integrated with everyday artifacts and actions, that the computers "vanish into the background" and become invisible. The drive for computing technologies to become invisible, which has motivated large swaths of pervasive computing research, is clearly established from the article's first sentence: "The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until their are indistinguishable from it." Tolmie et al. (2002) refer to this idea as "unremarkable computing," suggesting that such seamless interaction results not from the design of the technology, but rather from how the technology is utilized in practice. Although there has been some criticism of invisible computing as a model of interaction (especially work on "seamful design" (Chalmers and Galani 2004); see also Bell and Dourish (2006)), it has remained the dominant vision of pervasive computing for decades.

But what happens when pervasive computing's idea of invisibility is applied to human computers? What happens when the humans that are doing the work "vanish into the background?" Such vanishing already occurs in non-pervasive human computation systems, such as how AMT obscures worker identities and renders them invisible by framing them as a form of infrastructure (Irani and Silberman 2013; Ross et al. 2010)—a part of the system's API. This obscuring leads to issues such as wage disparity (Silberman et al. 2010) in existing human computation systems—issues that likely would continue with pervasive human computation systems. Moreover, Weiser's vision of invisible computing suggests the idea of "scrap computers" (disposable computers, analogous to scrap paper); could making human computers invisible also cast them as disposable? We need to make sure that such obscuring does not become even more prominent when developing human computation systems for a pervasive context in which seamless interaction is the norm. While machines and technology can vanish into the background, we as developers and researchers have a moral obligation not to let our technomorphism of human computers cause the same to happen to them.

Notably, Weiser's goal in making computers invisible was to "make individuals more aware of the people on the other ends of their computer links" (Weiser 1995): users would be more cognizant of the others they are interacting with than the technology. Yet human computation systems—in addition to obscuring the computer (who happens to be a human)—often work to obscure the "user" of that human computation. Zittrain argues that obscuring the user (the requester or employer in for-pay systems) denies the human computers the moral choice about what they do or how their computational labor is used (Zittrain 2008). Indeed, legitimate human computation platforms such as AMT have been used for illicit purposes (such as allowing spammers to break CAPTCHAs), likely without the human computers being aware (Harris 2011). The problem of computation being decontextualized may be more significant in pervasive systems, particularly if the computation involves actions taken within a localized context—a human computer may be asked to act as a sensor and take a picture of a particular location without knowing the purpose of that surveillance.[7]

In these ways, considering human computation through the lens of pervasive computing highlights issues in how human computation systems often render the computer invisible, whether or not that computation is performed pervasively. In developing pervasive human computation systems, we should adopt a design stance that acknowledges—even emphasizes—the "seams" in the system. We should support awareness of the connections between the mobile human computers and the users of their computation, as well as limitations of the system that may be introduced by a particular localized context. Research should focus on revealing and harnessing the details of the human computation's context, and not let the computing fade invisibly into the background.

Conclusion

In this chapter, I have discussed the concept of pervasive human computation: a mode of interaction in which human computation is performed by people during their everyday lives in a variety of localized contexts. This form of human computation can range from current microtask-based interaction forms ported to mobile devices, to having humans control or act as mobile sensors to provide

[7] The dangers of crowdsourcing activity without context are effectively dramatized in Bruce Sterling's short story, *Maneki-Neko*.

human-gathered data to computational systems. Pervasive human computation has the potential to allow human computers to harness localized or contextualized information from their environment, thereby supporting a greater variety of systems and problem solving based on large-scale human-driven information processing.

When making human computing pervasive, the differentiating factor is the *context* in which the computation is performed: rather than sitting at a desk, human computers can be out in the world. It is this situatedness that makes pervasive computing significant—the computation occurs in a particular context. What is important is not that pervasive human computing occurs everywhere, but that it can occur *any*where—in a variety of specific locations and contexts. In developing systems, we need to be careful to not lose track of the particulars of the computation's context. Instead, we need to harness these specific contexts through systems that respect and make apparent the participating human actors (whether the computers or the users of the computation), in order to develop the most effective uses of pervasive human computation.

References

Amazon mechanical turk (2013). http://www.mturk.com/

Bell G, Dourish P (2006) Yesterday's tomorrows: notes on ubiquitous computing's dominant vision. Pers Ubiquitous Comput 11(2):133–143

Bigham JP, Jayant C, Ji H, Little G, Miller A, Miller RC, Miller R, Tatarowicz A, White B, White S, Yeh T (2010) VizWiz: nearly real-time answers to visual questions. In: Proceedings of the 23nd annual ACM symposium on user interface software and technology, UIST '10. ACM, New York, pp 333–342. http://doi.acm.org/10.1145/1866029.1866080

Burke J, Estrin D, Hansen M, Parker A, Ramanathan N, Reddy S, Srivastava MB (2006) Participatory sensing. In: ACM sensys world sensor web workshop, Boulder, CO

Carranza J, Krause M (2012) Evaluation of game designs for human computation. In: Eighth artificial intelligence and interactive digital entertainment conference, Stanford, Palo Alto, CA. http://www.aaai.org/ocs/index.php/AIIDE/AIIDE12/paper/viewFile/5535/5756.

Chalmers M, Galani A (2004) Seamful interweaving: heterogeneity in the theory and design of interactive systems. In: Proceedings of the 5th conference on designing interactive systems: processes, practices, methods, and techniques, DIS'04, Cambridge. ACM, New York, NY, pp 243–252. http://doi.acm.org/10.1145/1013115.1013149

Chong CY, Kumar S (2003) Sensor networks: evolution, opportunities, and challenges. Proc IEEE 91(8):1247–1256

Dourish P (2004) What we talk about when we talk about context. Pers Ubiquit Comput 8(1): 19–30. http://dx.doi.org/10.1007/s00779-003-0253-8

Eagle N (2009) Txteagle: mobile crowdsourcing. In: Internationalization, design and global development. Springer, Berlin/New York, pp 447–456

Google flu trends (2013). http://www.google.org/flutrends/

Gupta A, Thies W, Cutrell E, Balakrishnan R (2012) mClerk: enabling mobile crowdsourcing in developing regions. In: Proceedings of the SIGCHI conference on human factors in computing systems, CHI'12, Austin, TX. ACM, New York, pp 1843–1852. http://doi.acm.org/10.1145/2207676.2208320

Harris CG (2011) Dirty Deeds Done Dirt Cheap: A Darker Side to Crowdsourcing, Privacy, security, risk and trust (passat), IEEE third international conference on social computing, pp.1314–1317, 9–11 Oct. 2011. Boston, MA. IEEE. doi: 10.1109/PASSAT/SocialCom.2011.89

Heimerl K, Gawalt B, Chen K, Parikh T, Hartmann B (2012) CommunitySourcing: engaging local crowds to perform expert work via physical kiosks. In: Proceedings of the SIGCHI conference

on human factors in computing systems, CHI'12, Austin, TX. ACM, New York, pp 1539–1548. http://doi.acm.org/10.1145/2207676.2208619
Horton JJ, Chilton LB (2010) The labor economics of paid crowdsourcing. In: Proceedings of the 11th ACM conference on electronic commerce, EC'10, Cambridge, MA. ACM, New York, pp 209–218. http://doi.acm.org/10.1145/1807342.1807376
Irani LC, Silberman MS (2013) Turkopticon: interrupting worker invisibility in amazon mechanical turk. In: Proceedings of the SIGCHI conference on human factors in computing systems, CHI'13, Paris, France. ACM, New York, pp 611–620. http://doi.acm.org/10.1145/2470654.2470742
Ishii H, Ullmer B (1997) Tangible bits: towards seamless interfaces between people, bits and atoms. In: Proceedings of the SIGCHI conference on Human factors in computing systems, Atlanta, GA. ACM, New York, pp 234–241
Khatib F, Cooper S, Tyka MD, Xu K, Makedon I, Popović Z, Baker D, Players F (2011) Algorithm discovery by protein folding game players. Proc Natl Acad Sci 108(47):18949–18953. http://www.pnas.org/content/108/47/18949
Lévy P (2001) Collective intelligence. Read Digit Cult 4:253
Mann S (1997) Wearable computing: a first step toward personal imaging. Computer 30(2):25–32
McGonigal J (2008) Why I Love Bees: A Case Study in Collective Intelligence Gaming. The Ecology of Games: Connecting Youth, Games, and Learning. Edited by Katie Salen. The John D. and Catherine T. MacArthur Foundation Series on Digital Media and Learning. Cambridge, MA: The MIT Press, 2008. 199–228. doi:10.1162/dmal.9780262693646.199
Montola M, Stenros J, Waern A (2009) Pervasive games: theory and design. Morgan Kaufmann, Burlington
Narula P, Gutheim P, Rolnitzky D, Kulkarni A, Hartmann B (2011) MobileWorks: a mobile crowdsourcing platform for workers at the bottom of the pyramid. In: Association for the advancement of artificial intelligence. http://www.aaai.org/ocs/index.php/WS/AAAIW11/paper/viewPDFInterstitial/3962/4263
O'Hara K, Mitchell AS, Vorbau A (2007) Consuming video on mobile devices. In: Proceedings of the SIGCHI conference on human factors in computing systems, San Jose. ACM, pp 857–866. http://portal.acm.org/citation.cfm?id=1240754
Paulos E, Honicky RJ, Hooker B (2009) Citizen science: enabling participatory urbanism. In: Foth, M. (Ed.) 2009. Handbook of research on urban informatics: the practice and promise of the real-time city, Hershey, PA: Information Science Reference, IGI Global pp 414–436
Quinn AJ, Bederson BB (2011) Human computation: a survey and taxonomy of a growing field. In: Proceedings of the 2011 annual conference on human factors in computing systems, CHI'11, Vancouver, Canada. ACM, New York, pp 1403–1412. http://doi.acm.org/10.1145/1978942.1979148
Reeves S, Sherwood S (2010) Five design challenges for human computation. In: Proceedings of the 6th nordic conference on human-computer interaction: extending boundaries, NordiCHI'10, Reykjavik, Iceland. ACM, New York, pp 383–392. http://doi.acm.org/10.1145/1868914.1868959
Ross J, Irani L, Silberman MS, Zaldivar A, Tomlinson B (2010) Who are the crowdworkers?: shifting demographics in mechanical turk. In: Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems, Atlanta. ACM, pp 2863–2872. http://portal.acm.org/citation.cfm?doid=1753846.1753873
Sakaki T, Okazaki M, Matsuo Y (2010) Earthquake shakes twitter users: real-time event detection by social sensors. In: Proceedings of the 19th international conference on world wide web, WWW'10, Hong Kong. ACM, New York, pp 851–860. http://doi.acm.org/10.1145/1772690.1772777
Silberman MS, Ross J, Irani L, Tomlinson B (2010) Sellers' problems in human computation markets. In: Proceedings of the ACM SIGKDD workshop on human computation, HCOMP'10, New York. ACM, pp 18–21. http://doi.acm.org/10.1145/1837885.1837891
Surowiecki J (2005) The wisdom of crowds. Anchor, New York. http://portal.acm.org/citation.cfm?id=1095645
Tolmie P, Pycock J, Diggins T, MacLean A, Karsenty A (2002) Unremarkable computing. In: Proceedings of the SIGCHI conference on human factors in computing systems: changing our

world, changing ourselves, Austin, Minneapolois, MN. ACM, pp 399–406. http://portal.acm.org/citation.cfm?doid=503376.503448
Toomim M (2011) Economic utility of interaction in crowdsourcing. In: Workshop on crowdsourcing and human computation at CHI 2011, Vancouver. http://crowdresearch.org/chi2011-workshop/papers/toomim.pdf
Tuite K, Snavely N, Hsiao DY, Tabing N, Popovic Z (2011) PhotoCity: training experts at large-scale image acquisition through a competitive game. In: Proceedings of the SIGCHI conference on human factors in computing systems, CHI'11, Vancouver, Canada. ACM, New York, pp 1383–1392. http://doi.acm.org/10.1145/1978942.1979146
Vertesi J (2008) "Seeing like a rover": embodied experience on the mars exploration rover mission. In: CHI'08 extended abstracts on human factors in computing systems, Florence. ACM, pp 2523–2532
Vieweg S, Hughes AL, Starbird K, Palen L (2010) Microblogging during two natural hazards events: what twitter may contribute to situational awareness. In: Proceedings of the SIGCHI conference on human factors in computing systems, CHI'10, Atlanta, GA. ACM, New York, pp 1079–1088. http://doi.acm.org/10.1145/1753326.1753486
von Ahn L, Dabbish L (2008) Designing games with a purpose. Commun ACM 51(8):58–67
Weiser M (1995) The computer for the 21st century. Sci Am 272(3):78–89
Zittrain J (2008) Ubiquitous human computing. Phil Trans R Soc A 366(1881):3813–3821

Building Blocks for Collective Problem Solving

Kshanti A. Greene and Thomas A. Young

Introduction

Online communication media, such as social networks and news forums, have no lack of people providing their opinions on what the problems are with “big government,” “big business,” the economy, the environment, the climate, healthcare or<insert “breaking news” story of the day>. Some people even make suggestions on how these issues could be addressed. Occasionally people will even sign their names to a petition on an issue they feel strongly about. But do people ever feel satisfied that their ranting, raving or even thoughtful discussion does any good? What if people had a means to channel this energy towards actually solving the problems that they face? This chapter introduces building blocks that will enable groups to discuss and solve problems in a massively collaborative manner. Discussion will occur around the thoughts, observations, and ideas that people contribute. Linking these thoughts by association will form abstractions of the problem, and can also be used to form patterns for addressing different types of problems.

A *problem* is defined generally as a situation that may be improved by intervention. A *solution* is then defined as the approach or mechanism(s) for intervention. Problem-solving is the process of describing a problem and seeking out a solution. When multiple people are involved in this process, it is likely that the people will differ in their experience with the problem, as well as on their opinion of what needs to be improved (if anything), and on how to intervene to improve the situation. This introduces potential conflict, and makes collaborative problem solving more difficult.

Collective problem-solving is collaboration at a massive scale. Instead of a small group of collocated individuals working concurrently, we may have thousands of people working on the same problem or solution. This is likely to result in multiple interpretations of the problem and multiple solutions. The goal in this case is not to

K.A. Greene (✉) • T.A. Young
Social Logic Institute, Albuquerque, USA
e-mail: kshanti@gmail.com; greene@sociallogic.org

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_28, © Springer Science+Business Media New York 2013

define a single problem and solution, but to condense the many perspectives into a representative summary of the significant points of view. Essentially, we are mining the human contributions to understand the problem and seek possible solutions. This will help the decision-makers to evaluate which solutions will be most effective to achieve specific objectives, as defined by the collective.

There are many approaches to problem solving and computational tools to help model problems. A collaborative problem solving tool could do one of the following: select one of these existing mechanisms, provide access to many tools, or create its own approach. None of these options are ideal since the tool developer would need to decide between potentially alienating people who are reluctant to try a new approach, or developing a confusing "Swiss army knife" of multiple tools.

The approach described in this chapter provides a single tool that enables people to use their preferred problem solving approach, or to allow a new one to emerge. The tool provides building blocks (like Legos®;) that can be put together in infinite ways to compose a representation of the problem and possible solutions. We call these building blocks *problem solving pavers* (PSPs). The processes by which these living models are built are based on templates that can emulate an existing problem solving process. This chapter discusses these building blocks, and how groups of problem solvers (henceforth called *solvers*) can assemble them to emulate several existing approaches.

Building Blocks

In this chapter, a *problem graph* refers to a graphical model in which the nodes represent components of a problem or solution, and the edges represent relationships between those components (see Fig. 1 for an illustration). A problem graph is essentially a Markov random field (MRF) in which there are different types of nodes, and edges that represent dependencies between the nodes Kindermann and Snell (1980). An important characteristic of the MRF is that each node is *conditioned on* only its neighbors. This means that if a node receives the value of its neighbors, it does not need to know the value of other non-neighboring nodes to determine its own value. This is the *Markov property* that forms the basis for probabilistic graphical models. The Markov property has an important implication for collective problem solving. It means that the solvers can work locally, within their area of expertise, without knowledge of the whole problem space. As long as they can address the issues relevant to their area of interest (for instance factors that directly influence a particular decision), they can contribute their knowledge and experience.

A solver's building blocks (PSPs) are the nodes and edges that they will combine to describe aspects of a problem or solution. We specify the types of nodes and edges to make them more meaningful to a large population of solvers. These components are based on existing computational models (such as decision networks and semantic networks). In Markov random fields, the edges can be assigned a value, called a *conditional probability*. In English, this is the likelihood of one node *given* the value of another node. For instance, a node representing *rain* could be conditioned

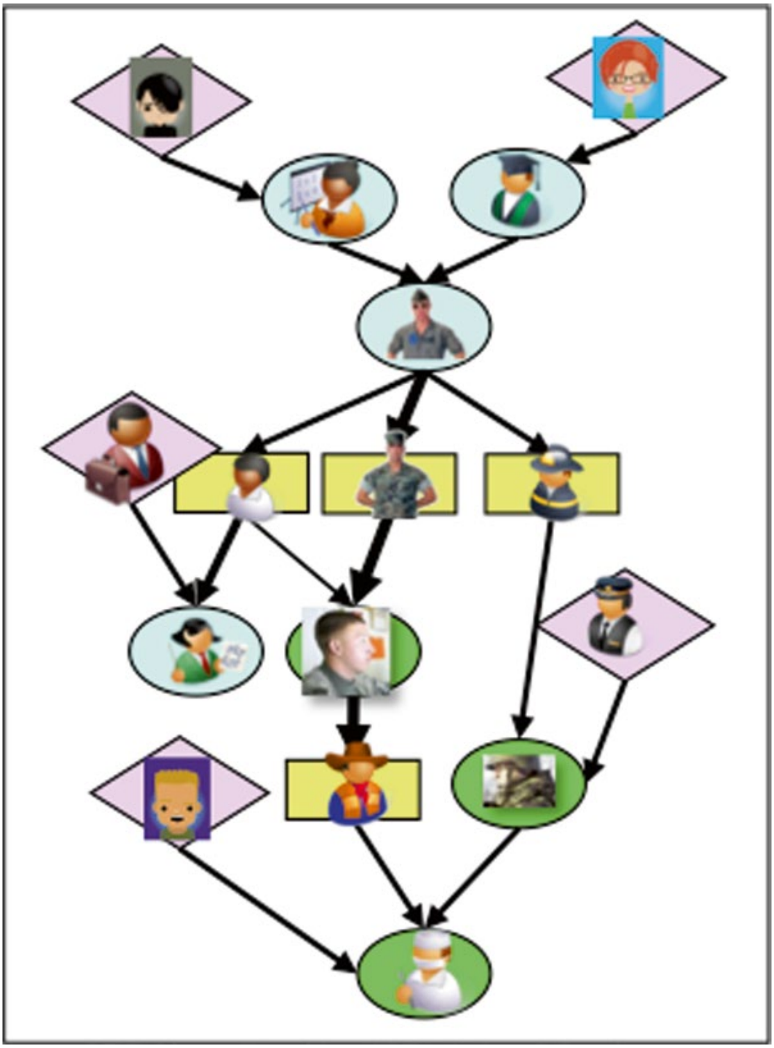


Fig. 1 Illustration of a graphical model in which different components are contributed by different solvers

on a node representing *clouds*, indicating that it is more likely to be raining if there are clouds than if there are not clouds. Mathematically, if there is an edge between two nodes X and X_i, $P(X|X_i)$ is the value of X given the value of X_i. However, X is really conditioned on all of its neighbors X_n, so to find X, we use $P(X|X_n)$.

In our approach, an individual solver only needs to indicate the existence of a relationship between two nodes. Similar to how some ant species form pheromone trails to a food site, repetitive trail marking increases the strength of the trail. The aggregate of these forms the value of the edge. If the edges represent conditional probability, and more people indicate the existence of a relationship between X and X_i than X and X_j, then $P(X|X_i) > P(X|X_j)$. In English, this means that X is more dependent on X_i than X_j.

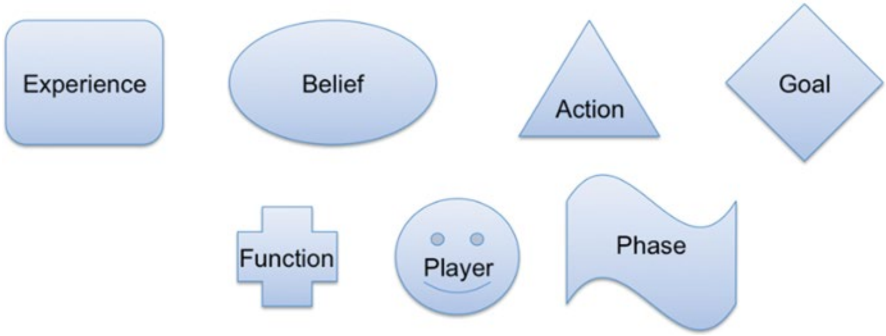


Fig. 2 Different types of PSPs used for describing problems and solutions

Seven types of nodes, shown in Fig. 2, are now described briefly. These nodes will have the same shapes in figures throughout the chapter.

- *Experience*: An observation, event, or documented piece of evidence.
- *Belief*: A hypothesis, opinion, state, or even experience whose value is not certain. A *belief* is equivalent to a "random variable" in graphical models.
- *Action*: An intervention that can change the situation. An *action* can also represent a decision point.
- *Goal*: A state, belief, or experience that is desired. A *goal* can also be expressed to represent a constraint, utility or risk (as in Bayesian decision networks Shachter (1986)).
- *Player*: Represents a person or entity that is involved in a problem or solution. A *player* is not meant to represent the solver, unless the solver is also an actor in the problem/solution space.
- *Function*: Evaluates a mathematical function using the values on its incoming edges as input. For instance, a *function* can be used to sum costs or find the most likely (highest probability) hypothesis.
- *Phase*: A phase is a meta-node that "encapsulates" other nodes. It represents a stage in the problem-solving process. A series of phase nodes form a *template* for problem solving. Other types of nodes that are created during a specific phase will be connected to that phase node.

Templates can be used to represent and initiate a problem-solving process. Each *phase* may be associated with some instructions for the solvers. Templates will typically be defined by an advanced solver.

There are five types of edges that can connect these node types, discussed next. Which type to use may be determined contextually based on the node type, or it can be specified by the solvers:

- *Conditional*: Indicates that the value of a node is dependent on or correlated with another. For instance, "rain" is correlated with "clouds". This edge type can also simply indicate a relationship (for instance between two *players*).
- *Hierarchical*: Represents a categorical relationship, such that a concept "belongs to" another concept (e.g. "ice cream" belongs to "cold" or "tasty").

- *Ordinal*: Indicates a temporal ordering (such as *actions* that need to be done in a certain order), or a qualitative ordering (such as the relative importance of a set of *goals*).
- *Comparable*: Indicates that two nodes have essentially the same meaning in the problem context (such as "connection" and "relationship").
- *Numeric*: Carries a numeric value with it, to be used as input to a *functional* node. Individual solvers can provide specific values for these edges (such as a cost estimate). These values can be aggregated in different ways (such as mean, median, etc.).

All edge types can be associated with some uncertainty. Consider the following example: "tasty" and "gross" are two different concepts, and 90 % of the population connects "ice cream" to "tasty" while the remaining 10 % connect it to "gross." Unlike a semantic network in which relationships represents "facts," we can infer that for this population, ice cream is *more likely* to be tasty than gross.

The problem-solving models created using these PSPs are "executable," meaning that the relative strength of each of a set of possible solutions can be found and presented to the solvers. These solutions may be the best that the contributing solvers can derive. However, in general no guarantee will be made that the solutions are optimal. Finding an optimal solution is not our goal. Optimality is not attainable in many situations, due to conflicting objectives, incomplete information, and complex interdependencies. Ultimately, our goal is to achieve some form of equilibrium, allowing people to take action and make choices that take their consequences into consideration. The first step is to enable a large population of people to be able to describe problems and solutions.

The problem solving pavers that we describe are components in a collective problem-solving system, called *ePluribus*, that is in development by Management Sciences Inc. for a DARPA Small Business Research Innovation (SBIR) grant. Solvers will use *ePluribus* through a web browser or mobile app.

Modeling Existing Problem-Solving Approaches

There are so many problem-solving techniques and models, that there is not enough space in this chapter to discuss how these building blocks can be applied to many of them. We have selected a few important approaches and computational models to illustrate the generality of this set of building blocks. In the following sections, we first show the template, which is the problem-solving process, and then we illustrate the process using the PSPs. The shapes of the nodes in these diagrams correspond to the shapes in Fig. 2. The colors of the nodes indicate in which phase the node is created.

The purpose of this article is to demonstrate the generality of our approach to many problem solving techniques. Each of the techniques in the next section was designed for a specific purpose and is familiar to a specific set of experts. However, we do not necessarily recommend "trying these at home." In section "Solution Paths: A More Intuitive Approach" we discuss a much simpler and intuitive

Table 1 Brief description and purpose of each problem-solving technique described in this chapter

Technique	Description	Purpose
Means Ends Analysis	Describe current and goal states and then apply operators to reduce distance to goal	Developed to imitate human reasoning in logic programs. Can be used to determine how to reach goal
Root Cause Analysis	Describe an undesired situation and then identify the causes for events or symptoms	Diagnostic mechanism used to uncover initial causes of a problem
Statistical Hypothesis Testing	Identify a default and alternative hypothesis and evaluate the likelihood of the default given observations	Used to evaluate whether new evidence supports replacing a dominant theory
Decision Matrix	List decision options and criteria. Find the highest scoring decision option based on how well each meats the criteria	Used to identify the best decision option given a number of criteria
Semantic Network	Describe concepts and the relationships between them	Used to describe knowledge so that it can be applied to solve a computational problem
Constraint Network	Identify situational variables and the constraints between them. Find values for variables that meet constraints	Used to constrain a solution space to those options that meet certain specific goals
Solution Paths	Describe current and desired situations and identify best solution paths from current to desired	General purpose approach that enables a group to identify potential solutions from diverse perspectives

approach that helps solvers discover *solution paths* from an undesired situation to a more desirable situation. Table 1 lists each technique described in this chapter and summarizes their approach and intended purpose.

Means-Ends Analysis

Means-ends analysis (MEA) was introduced by Newell and Simon in their ground-breaking work on the Artificial Intelligence system, General Problem Solver (GPS) (Newell and Simon 1963). The goal of that system was to enable a machine to use "human" reasoning. The authors applied MEA to logic problems, but the same process can be applied to more human-centered problems, in this case figuring out how to travel from a person's home to her aunt's house (based on[1]). The template is shown in Fig. 3. The problem is illustrated using our PSPs in Fig. 4 and discussed after the figure.

- *Identify actions*: At the top of Fig. 4 are actions (triangles) that should be taken when the state in the oval is true (where *D* is distance from *Aunt's*).

[1] Example of Means-Ends Analysis, Rutgers University. http://www-rci.rutgers.edu/ cfs/472_html/Planning/GPS_472.html

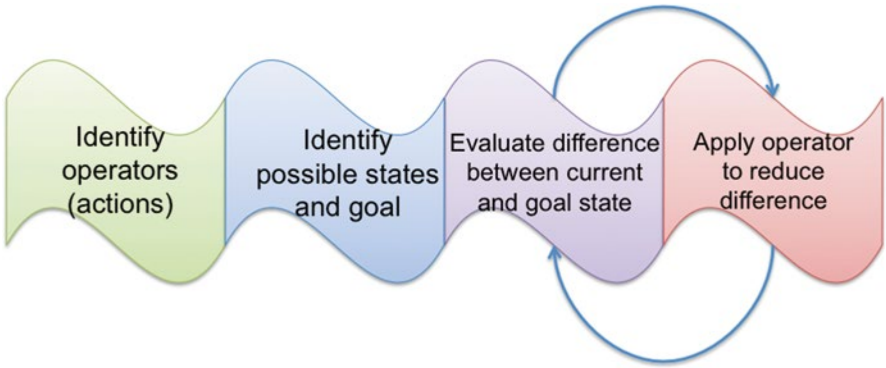


Fig. 3 Template of the means-ends analysis process. The *arrows* indicate phases that can be repeated until completion

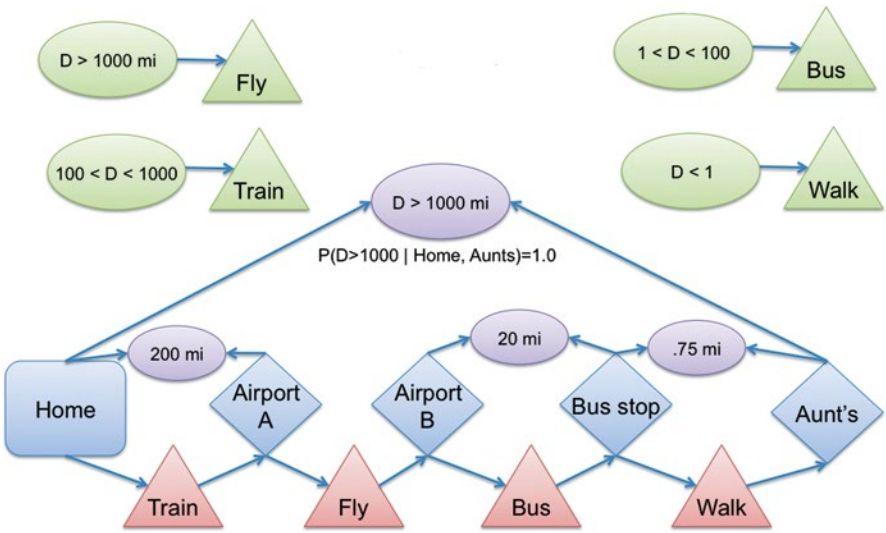


Fig. 4 The means-ends analysis process using PSPs

The diamonds represent goals or sub-goals, which in this case are stops between *Home* and *Aunt's*.

- *Identify possible states and goals*: Once the current state (*Home*) and goal (*Aunt's*) have been identified, the solvers may want to consider subgoals or states that are between the current state and final goal. These are represented by diamonds in Fig. 4.
- *Evaluate difference between current and goal states*: The solver first observes that the distance between *Home* and *Aunt's* is greater than 1,000 miles.
- *Apply operator (action) to reduce difference*: The distance causes the action *Fly* to be applied (which connects two subgoals *Airport A* and *Airport B*).

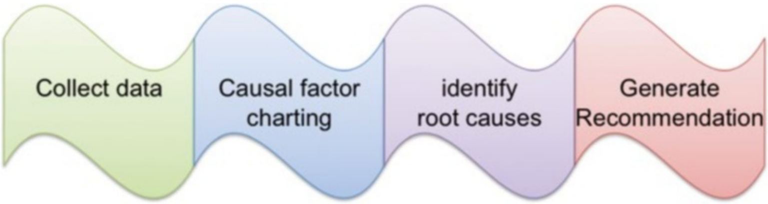


Fig. 5 Template of the root cause analysis process

The solver still needs to figure out how to get to and from the airport, resulting in recursively applying the actions *Train*, *Bus*, and *Walk*. Eventually the diagram shows that the person can indeed reach her Aunt's house using the available actions.

When multiple solvers contribute in means-ends analysis we may see alternative actions, subgoals and hypotheses about the state at a given time. One objective in this process would be to have consensus at least on the ultimate goal. Multiple paths to this goal may emerge due to variances in preferences.

Root Cause Analysis

Root cause analysis (RCA) is a diagnostic mechanism often used in engineering fields to uncover the initial causes of a problem, instead of simply looking at the symptoms. The prolific Japanese inventor and founder of Toyota Motors, Sakichi Toyoda, has been credited with its introduction (Fatima 2011). RCA first collects data from a situation including symptoms and the events that led up to the situation. Analysis proceeds by asking why an unexpected or undesired event occurred. This process may be recursive. The initial (root) causes are then identified and recommendations to prevent these from occurring again are proposed. This process is shown in the template in Fig. 5.The following example uses PSPs to model a RCA example in Rooney andHeuvel (July 2004), in which the authors identify the root causes of a kitchenfire.

- *Collect Data*: The process begins in Fig. 6 at the bottom. The green rounded rectangles represent the chain of events (experiences) leading to a fire and causing it to destroy the kitchen.
- *Causal factor charting*: The causes of the events (in blue and purple ovals) are then explored, in particular what caused the fire to start and why the fire extinguisher did not work.
- *Identify root causes*: The nodes in purple are then identified as the "root" causes, which are events that can typically be avoided by doing something in a more appropriate manner. These may be indirect causes of the undesired events, but are key to their manifestation.
- *Generate recommendation*: Finally, recommended actions to inhibit the root causes in future situations, are presented (in red triangles).

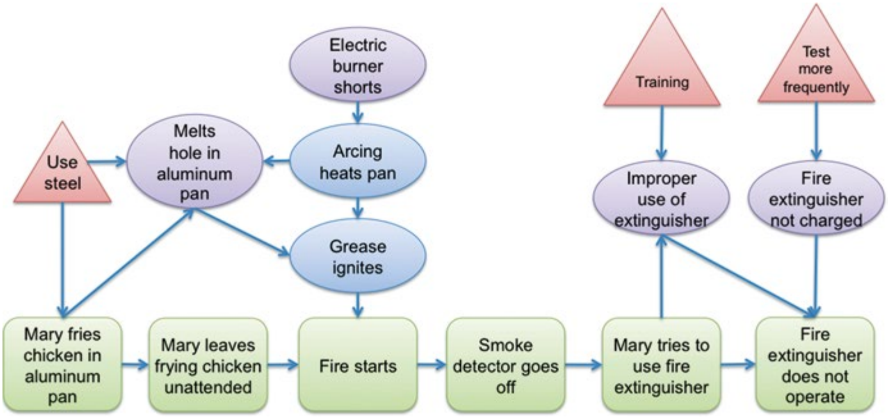


Fig. 6 The root cause analysis process using PSPs

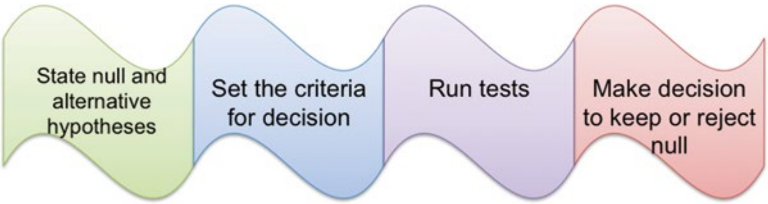


Fig. 7 Template of the statistical hypothesis testing process (From Gravetter and Wallnau (2012))

Multiple solvers contributing in the root cause analysis process may result in a more diverse set of possible causes and potentially conflicting opinions about the causes. Conflicting opinions are not necessarily damaging, unless there is disagreement on whether a recommended action is actually a root cause. In this case, more feedback and experimentation may be needed to resolve these discrepancies.

Statistical Hypothesis Testing

Statistical hypothesis testing (SHT) is commonly used by scientists and analysts when considering whether new evidence warrants displacing a dominant theory. SHT uses sampling to compute the likelihood that experimental results occur given the null hypothesis (Gravetter and Wallnau 2012). If the likelihood of the null hypothesis being true given the observations is less than some predefined *level of significance* (los) then the null hypothesis can be rejected in favor of the alternative. The template for the statistical hypothesis test is shown in Fig. 7.

The diagram in Fig. 8 illustrates a hypothesis testing process using PSPs, in which the experiment tests a null hypothesis that "children watch on average

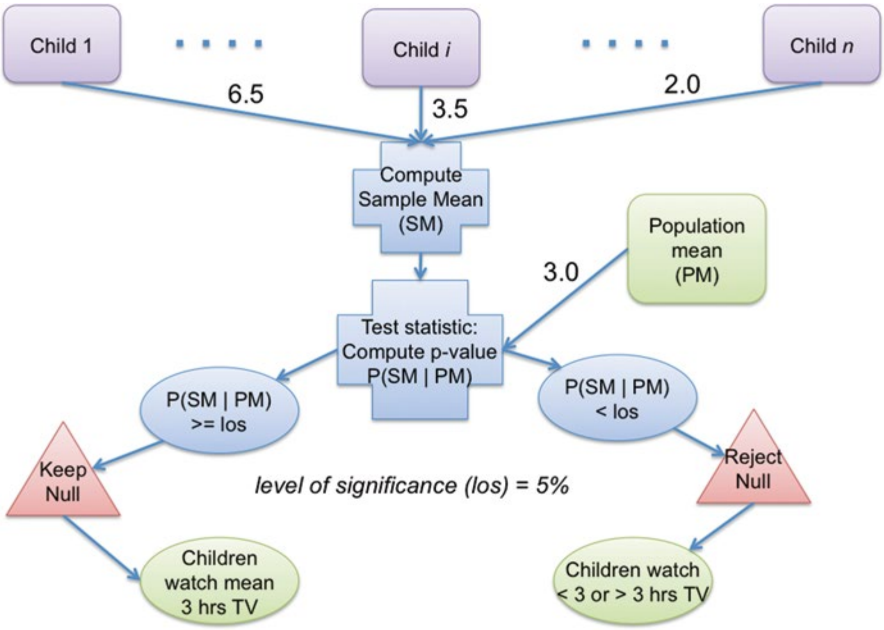


Fig. 8 The statistical hypothesis testing process illustrated using PSPs

3 hours of television (TV) per day." This problem and its representation are now described.

- *State the null and alternative hypotheses*: The green ovals at the bottom of the diagram show the null hypothesis (on the left) and the alternative, that the children watch more or less than 3 h of TV (on the right).
- *Set the criteria for decision*: The *population mean* (PM) of 3 h is the current null hypothesis. The *sample mean* (SM) will be computed as the average number of hours watched given the observed sample. In this case the null hypothesis will be rejected when the likelihood of observing the sample mean given the null hypothesis is less than the level of significance (set at 5 %). The top blue cross-shaped functional node uses the times from the *n* children as input and computes the sample mean. The bottom functional node takes in the sample mean and the population mean and computes the *p-value*, also known as the *test statistic*. This is the probability of seeing the sample mean given the population mean ($P(SM|PM)$).
- *Run tests*: The tests would ask or measure the number of hours that each of the *n* children watches TV per week. These numerical values then feed into the top functional node to compute the sample mean.
- *Make decision to keep or reject null hypothesis*: If the probability of the sample mean given the population mean is equal to or greater than the los, then the decision to retain the null hypothesis (left branch of diagram) will be fired. If $P(SM|PM) < los$, then the right branch will be taken and the null hypothesis will be rejected.

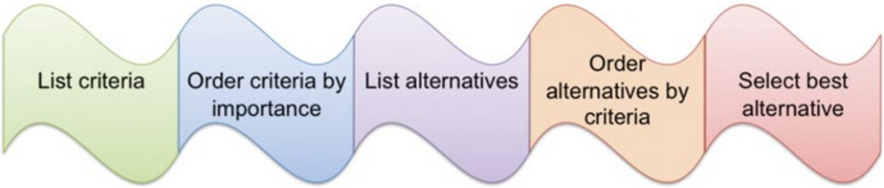


Fig. 9 Template of the decision matrix formation process

Multiple solvers contributing to a statistical hypothesis test could result in multiple alternative hypotheses and potential disagreement over how to compute the test statistic. Multiple alternative hypotheses can be addressed using an approach called simultaneous statistical inference Miller (1981). As for disagreement in other areas, the need for consensus would likely be correlated with the importance of the test. If the test has high significance, then the group building the experiment should settle on the details before conducting. In this case, multiple perspectives could be important to ensure that all issues are considered.

Decision Matrix

Decision matrices are used to select an option from a set of possible options. The options are evaluated by how well they meet a set of criteria. As in the process shown in Fig. 9, the criteria for a decision are identified and evaluated for to their importance. Then the alternatives are listed and evaluated by how well each meets a certain criteria. The final step is to compute the score for each alternative as a weighted sum of its scores for each criteria.

The decision matrix process using PSPs is illustrated in Fig. 10 and proceeds as follows:

- *List criteria*: The green diamond goal nodes, C_1… C_K, represent the decision criteria.
- *Order criteria by importance*: Criteria are ordered by importance using the ordinal edges. The relative importance is shown by the thickness of the edges leading to the *Best* node (thicker means more important).
- *List alternatives*: The purple triangle action nodes, A_1… A_n, represent the decision alternatives.
- *Order alternatives by criteria*: The thickness of the orange lines from each action to each criteria indicates each action’s ability to meet criteria. Ordering the actions can be done in one of at least of two ways. First, the alternatives can be ordered according to how well they meet each criteria using the ordinal edges. Second, the solvers can simply mark the alternatives that meet a criteria by creating or verifying the edge between the action and goal nodes.
- *Select best alternative*: The best alternatives will be those that have the heaviest edge weights to either the most important criteria, or spread evenly to all the criteria.

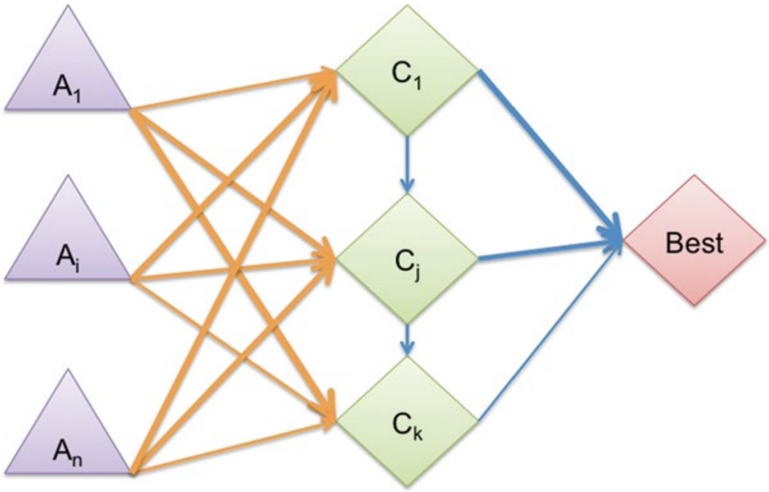


Fig. 10 A decision matrix represented using PSPs submitted by one hypothetical solver

One could use functional nodes to compute the score of each alternative, or observe visually which alternatives seem best.

When multiple solvers contribute to the decision matrix, there could be significant disagreement about the relative importance of the criteria and how well each decision option meets each criteria. There are at least two ways to evaluate the preference order of criteria. A first approximation would be to combine (in this case sum) the ranks provided by each solver. For instance, suppose three criteria, $\{C_1, C_2, C_3\}$, are ranked by three individuals as follows (from most to least important):

$$R_1 = \{C_2, C_1, C_3\}$$
$$R_2 = \{C_1, C_2, C_3\}$$
$$R_3 = \{C_2, C_3, C_1\}$$

When each criteria is assigned a weight based on its rank ordered, from most to least important, with the weight 3 given to the most important, 2 to the second most important and 1 to the least important, the scores for each criteria are as follows:

$$C_1 = 2 + 3 + 1 = 6 \text{ (ranked second by } R_1\text{, rst by } R_2 \text{ and third by } R_1\text{)}$$
$$C_2 = 3 + 2 + 3 = 8 \text{ (ranked rst by } R_1\text{, second by } R_2 \text{ and rst by } R_1\text{)}$$
$$C_3 = 1 + 1 + 2 = 4 \text{ (ranked third by } R_1\text{, third by } R_2 \text{ and second by } R_1\text{)}$$

In this case, C_2 is the most important criteria according to the group of solvers because it has the highest overall weight. A second approach addresses the impossibility theorem, introduced by Kenneth Arrow, that identifies possible contradictions when aggregating multiple preference orderings (Arrow 1950). In this approach, we

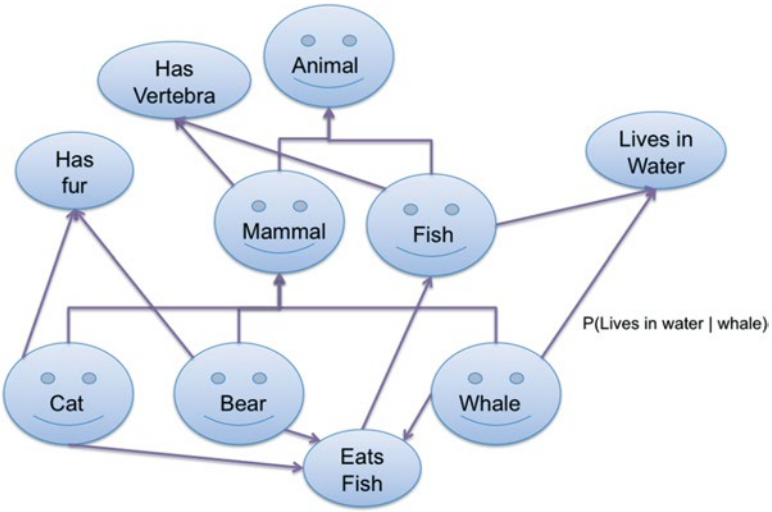


Fig. 11 A semantic relationship about a set of animals. The bent edges are hierarchical relationships and the straight edges are conditional

would begin by ordering the possible permutations of the criteria by the number of people that support each ordering. This approach maintains the identity of very divergent groups, and allows competing approaches to be represented and considered. The foundations of this approach are discussed in Greene et al (2010).

Semantic Network

Semantic networks are used to represent knowledge, and can be an important part of a problem solving process. The nodes represent the concepts, and edges between them are typically annotated with a semantic relationship (Sowa 1992). For instance, a directed edge between a predatory animal and its prey might be marked with the relationship "eats". The process for defining a semantic network is fairly simple, and alternates between *Identifying concepts* and *Identifying relationships*. The building blocks that we have described do not specifically allow the edges to be annotated with anything other than a numeric value. Instead, all concepts should be contained in a node, including semantic relationships.

To simulate an annotated edge, the original concept is connected to a node containing the semantic relationship. Figure 11 shows relationships between some animals and their behavior. In the original network,[2] the *whale* and *fish* nodes had edges pointing to a *water* node. The edges were marked with the label *lives in*.

[2] Based on a semantic network at http://en.wikipedia.org/wiki/Semantic_network, accessed 3/10/13.

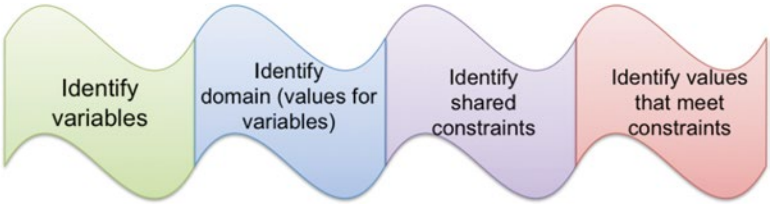


Fig. 12 Template for the constraint network process

In the network in Fig. 11 the animal nodes are instead connected to a node *Lives in water*. Note that in this case, the entire verb phrase should contained in one node. While we could have a node representing *water*, a node *lives in* has no meaning on its own, and therefore cannot really be assigned a probabilistic value.

Multiple solvers contributing to a semantic network could result in disagreement about the semantic relationships between concepts. In this case, the relationships will become probabilistic. For instance, not all bears have access to fish. Solvers could add other semantic relationships, such as *Eats berries* and *Eats elk*. The frequency of edge validations would result in different likelihoods for $P(Eats\_fish|Bear)$, $P(Eats\_berries|Bear)$, and $P(Eats\_elk|Bear)$. Alternatively, a more comprehensive bear classification could be described.

Constraint Network

The last problem-solving example is based on constraint satisfaction problems, in which pre-specified constraints must be met for a solution to be valid. We use a constraint network as a foundation, in which an edge represents a constraint between variable nodes (Dechter 1992). Figure 12 shows the process for creating a constraint network.

The constraint satisfaction process with PSPs is now described using a crossword puzzle example from Dechter (1992). Figure 13 shows a crossword puzzle on the left and the words that can be used to fill it on the right. The allowed words are arranged to meet the constraints in the diamonds at the bottom of the figure.

- *Identify Variables*: The green ovals in Fig. 14 represent the crosswords in Fig. 13. For example, *W1* represents the horizontal, five letter word that starts in the top left of the crossword puzzle. These variables are connected to constraint nodes (green diamonds) that indicate how many letters the word must have.
- *Identify domain (values for variables)*: The allowed values for the variables $W1 \ldots W5$ are shown in Fig. 13.
- *Identify shared constraints*: In the crossword example, a constraint between variables means that their crossword representations share a letter. For example, $W1$ must share a letter with $W2$ and $W3$. These constraint relationships are shown in Fig. 14 using the purple diamonds. The diamonds contain two numbers that represent word indices in each word that must contain the same letter.

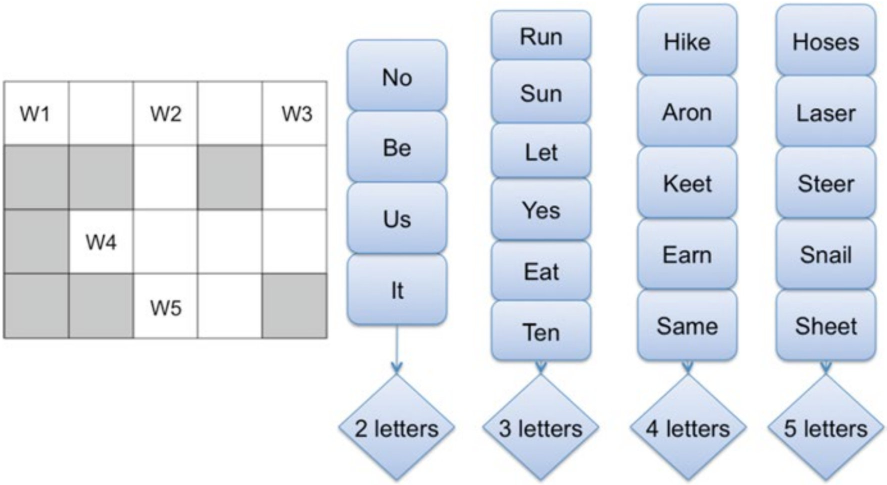


Fig. 13 A crossword with five words sharing constraints (on the *left*) and possible values for constraint variables (on the *right*)

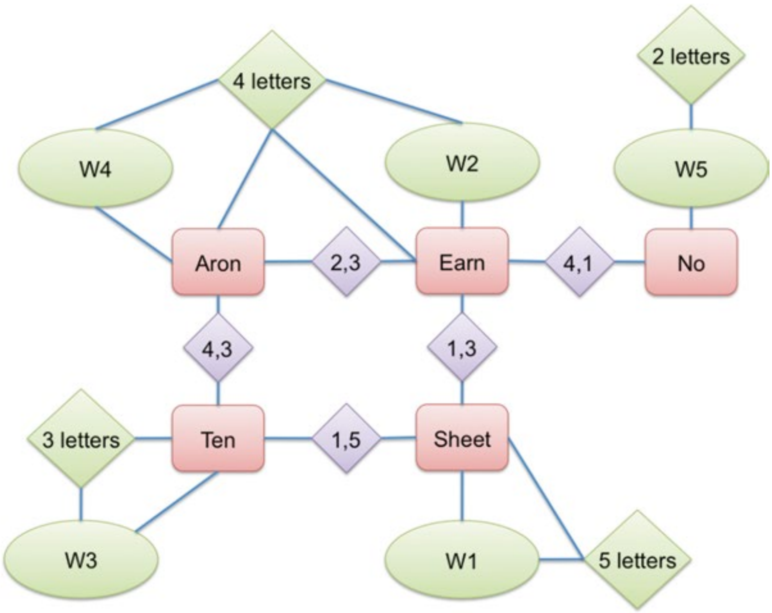


Fig. 14 A constraint satisfaction solution described using PSPs

For example, the diamond containing "1,5" on the path between *W*3 and *W*1 means that the first letter of *W*3 must be the same as the fifth letter of *W*1. In this diagram, the first letter in the purple diamonds is associated with the variable on the top or left, and the second is associated with the variable on the bottom or right.

- *Identify values that meet constraints*: The last step is to find words (from Fig. 13) that meet the constraints in purple (in Fig. 14). If a constraint is met between two words then they will be connected with a path. This is represented by pairs of red rectangles that both connect to the same purple constraint node. A solution to the constraint satisfaction problem is found when all constraints are met. The diagram shows a solution in which all shared constraints are connected with edges.

Multiple objectives submitted by different solvers could result in shared constraints. For instance, suppose one person wants to see a Foreign movie and her partner would like to see a movie starring Ingrid Bergman. Meeting both constraints means that the pair see a Foreign movie starting Ingrid Bergman. It is possible that multiple constraints are mutually exclusive, meaning that they cannot both be addressed at the same time. In these cases, the network cannot be solved until one of the constraints is removed or modified. This collaborative process could reveal other constraints that the community can agree on.

Solution Paths: A More Intuitive Approach

Each of the approaches described in the previous section is good for a specific purpose, and many of them provided a foundation for our problem solving building blocks. However, none of the approaches can be used in *all* situations. We introduce a simple approach that may contain elements from the existing approaches, but can be applied in multiple situations. This approach, called *solution paths* is most like the means-ends analysis described in section "Means-Ends Analysis". We observed that problem-solving always involves one or both of the following stages: (1) attempting to understand a given situation and (2) attempting to improve the situation through intervention. To accomplish the second stage, it helps to know what "improve" means to those that would like the problem solved. Figure 15 illustrates the template for the *solution paths* approach, which helps solvers identify paths from an undesired situation to a desired situation. This process could also be used to identify paths that might cause a desired situation to devolve into undesired.

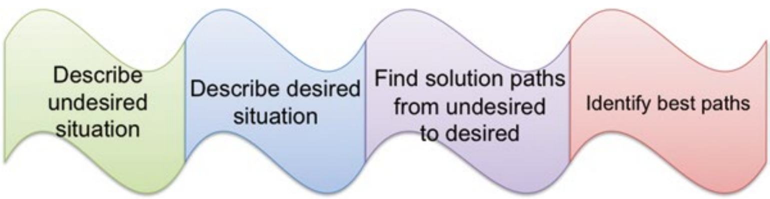


Fig. 15 Template of the *solution paths* approach developed by the authors

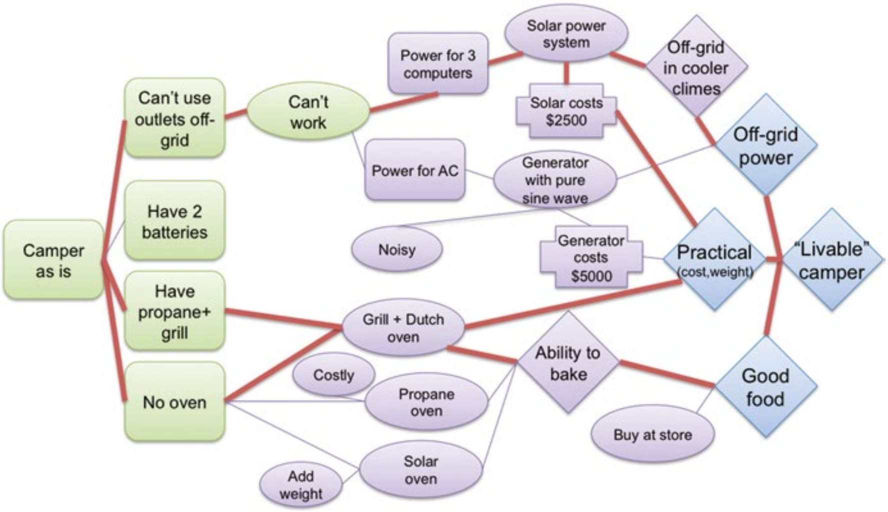


Fig. 16 A solution paths approach described using PSPs

Figure 16 illustrates the problem solving process undertaken by the authors (a married couple) who wish to be able to live in their camper for extended periods of time. In order to do so, they need to address the inability to use the power outlets while not connected to "the grid." The process is described next.

- *Describe undesired situation (left side):* The solvers first describe the camper as it is, including what it lacks and what it has. The main problem is that they cannot work for long periods in the camper when off-grid because they cannot plug in their computers. They also have no oven.
- *Describe desired situation (right side):* The ultimate goal is a "livable" camper, in which they can have off-grid power and good food. Of course the solutions should be practical, meaning not too costly and they should not add a lot of weight to the camper.
- *Find solution paths between undesired and desired:* Once the current situation and goals have been identified, the solvers then can start the main work of finding possible solution paths. This includes evaluating the specifics of what is needed to enable the goals to be met, and providing potential solutions that address these needs (such as a generator and solar power system). In addition, they must consider the secondary issue of wishing to have good food. Some of the solutions have undesired side effects, such as the noisiness of the generator. The solvers also added function nodes to compute the costs of each solution (details not shown).
- *Identify best paths:* A solution path should connect the left side to the right side. In this case there are a number of sub-goals, and the best solutions will address all of these. This means that there may be branches in the paths, but that all

branches should be taken to complete a solution. The *best* paths are those with the greatest support. In this case, this means that both solvers agreed that the path was good. These paths are indicated by a heavier, red line.

The reader might observe that this problem-solving process bares similarities to the decision matrix. In this case, the solvers are evaluating different options based on multiple criteria. An important distinction is that the solution paths approach allows the solvers to address multiple sub-problems. In this case, they are attempting to find a solution that meets two diverging needs- power and cooking. Another distinction is that this approach allows the solvers to specifically identify the issues that prevent or enable paths to be created (for instance, the need to power three computers). Finally, there is really no need to have all problems, goals and potential solutions enumerated prior to beginning the process. In fact, the authors did not identify the "dutch oven" possibility until well into the process.

When multiple solvers are involved in the process, or the problem is complex, multiple solution paths may emerge. The goal would be to identify the most viable paths that achieve a more desired situation. To get everyone on the same page initially, the group should at least agree on one statement to describe the undesired situation and one statement to describe the goal (or desired situation). There may be multiple competing or parallel subgoals, but at least with a common goal there is one shared element to unite the collective.

Conclusions

The set of problem solving pavers discussed in this chapter was selected to enable human problem-solvers to describe and explore problem and solution spaces. The goal is to allow many individuals to work distributedly on the same problem. By decomposing the problem into interrelated components, the problem can be addressed at multiple levels of abstraction and solvers can focus on the parts of the problem they are familiar with.

The examples in this chapter were contrived in order to make a case that the building blocks are generalizable enough to be applied to many different problem-solving approaches. A new problem-solving template was introduced that enables solvers to combine elements from different approaches to lay solution paths from an undesired state to a desired state. The best solutions, according to the community of solvers, will be the paths with the strongest support. Future work will address interfaces for collective problem solving and demonstrate these building blocks used in real problem-solving situations.

Acknowledgements Without the thoughtful contributions from Thomas Young, Pietro Michelucci, George Luger, Joe Kniss, Melanie Moses, and Steven Garcia, this work would not have been possible.

References

Arrow KJ (1950) A difficulty in the concept of social welfare. J Polit Econ 58:328–346. http://econpapers.repec.org/RePEc:ucp:jpolec:v:58:y:1950:p:328

Dechter R (1992) Constraint networks. In: Encyclopedia of articial intelligence, 2nd edn. Wiley, pp. 276–285

Fatima A (2011) How has the root cause analysis evolved since inception? Bright Hub PM. Http://www.brighthubpm.com/risk-management/123244-how-has-the-root-cause-analysis-evolved-since-inception/

Gravetter F, Wallnau L (2012) Statistics for the behavioral sciences, 9th edn. Cengage Learning, Andover

Greene K, Kniss J, Luger G (2010) Representing diversity in communities of bayesian decision-makers. In: Proceedings of the IEEE international conference on social computing, social intelligence and networking symposium, Minneapolis

Kindermann R, Snell JL (1980) Markov random fields and their applications. American Mathematical Society, Providence

Miller R (1981) Simultaneous statistical inference. Springer, New York

Newell A, Simon H (1963) GPS: a program that simulates human thought. In: Feigenbaum EA, Feldman J (eds) Computers and thought. McGraw-Hill, New York

Rooney J, Heuvel LV (2004) Root cause analysis for beginners. Qual Prog 37(7):45

Shachter RD (1986) Evaluating influence diagrams. Oper Res 34(6):871–882. doi:http://dx.doi.org/10.1287/opre.34.6.871

Sowa J (1992) Semantic networks. In: Shapiro SC (ed) Encyclopedia of artificial intelligence. Wiley, New York

Adaptive Agents in Combinatorial Prediction Markets

Anamaria Berea

Introduction

Research has demonstrated that opinion pools and prediction markets outperform single expert opinion when forecasting the outcome of complicated aggregated events. Decomposition-Based Information Elicitation and Aggregation (DAGGRE) is a program sponsored by Intelligence Advanced Research Projects Activity (IARPA) and executed at George Mason University (GMU). DAGGRE is finding methods to improve over the unweighted average or plain-vanilla prediction markets (Hanson 2007).

Prediction markets are fundamentally markets for information. They use a web based platform where any user (forecaster) can log in and make his or her own bets with respect to the forecasting problems. Prediction markets are a useful and efficient way for aggregating human judgment and the wisdom of the crowds (Lyon and Pacuit 2013).

For example, Fig. 1 shows such a forecasting problem on the DAGGRE prediction market. The forecaster adjusts the probability based on her own judgment. When the question closes (the event happens or the deadline of the question expires), she is rewarded or punished for how close her change in forecast was to the actual outcome of the event.

Prediction markets are therefore a mathematical set comprised of forecasting problems, probability values, users (or forecasters) and incentives (points or money in certain cases).

A. Berea (✉)
C4I Center of Excellence, Volgenau School of Engineering, George Mason University, Fairfax, VA, USA
e-mail: aberea@c4i.gmu.edu

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_29,

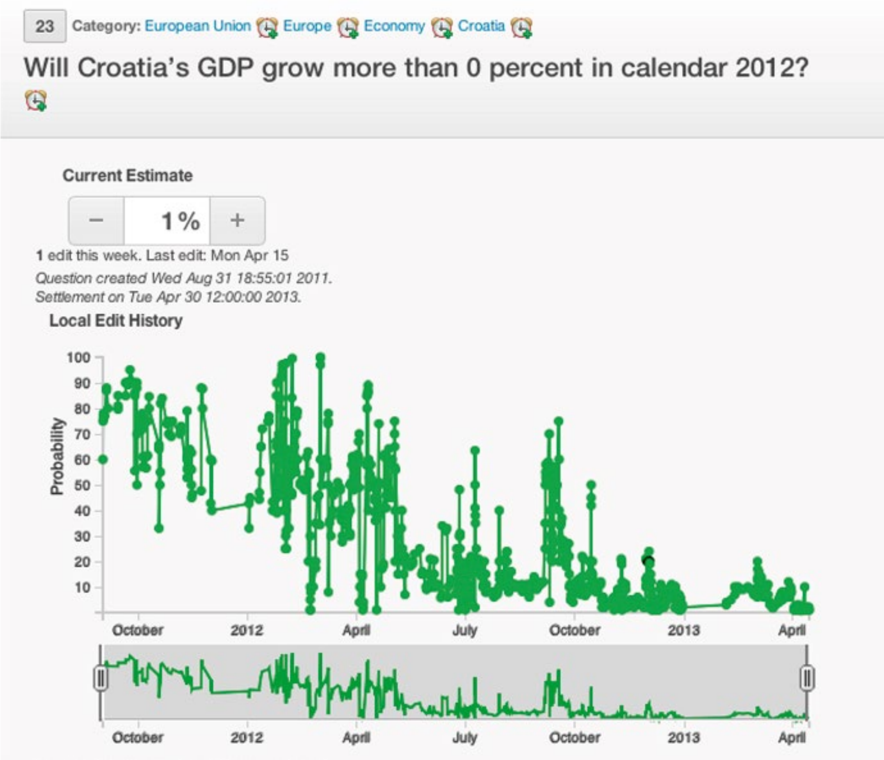


Fig. 1 An example of forecasting problem from the DAGGRE prediction market

DAGGRE started as a plain prediction market and, after roughly 1 year, it became the first generalized combinatorial market in the world. The combinatorial prediction market allows for the events that are being forecasted to be linked among each other, so that the probability of the outcome of one event may become the assumption for another event (Hanson 2003). In this way, the users are using the information the market itself provides and not only external sources. Participants can make trades on combinations of events, i.e.: "Will there be an uprising in Region A" assuming the truth or false value for "Will there be an uprising in Region B".

The combinatorial prediction market is fundamentally a Bayes Net (Pennock and Xia 2011), where the nodes are represented by each question and the links between them represent dependencies. Bayesian Networks are belief networks expressed as probabilistic graphical models (Ben-Gal 2007). The nodes are the variables and the edges are the conditional dependencies between the nodes. The probability values of the nodes in a Bayes Net are conditional on the probabilities of the other nodes. Bayesian Networks are used in many fields and particularly in decision support systems, because they are a very helpful tool for causal reasoning and inference.

Besides a few thousand human users, DAGGRE also employs automated traders (autotraders), using simple algorithms to determine a desirable estimate for trading, without human intervention (or, while in testing, with limited human intervention). Autotraders are designed to examine a set of questions that appear on the DAGGRE prediction market website, and make trades on selected world events. These world events can be any type of macro level social phenomena, such as elections, riots, international agreements, epidemics, a.s.o. (Cameron 1963).

These autotraders are adapting to the market forecasts, in the sense that they read the estimates given by the market and only afterwards places the bid. They adapt to the human trades (Holland and Miller 1991).

Methodology

Most multi-agent systems and simulations involve algorithms that are designed to "interact" with each other. In the case of DAGGRE, the algorithms are interacting with humans and are behaving adaptively to the aggregate human judgment, or the "wisdom of the crowds" (Surowiecki 2005).

The experiments described below shows comparatively the forecasting accuracy of a Bayesian Network model, a flat prediction market, a combinatorial prediction market and an auto trader (which methodology performed better and under which conditions). We started by forecasting a hypothesis ("Grexit" in this case, described in the following section) independently on the flat prediction market and the Bayes Net model. After we compared the performance between the two methodologies and the DAGGRE flat prediction market was switched to a combinatorial prediction market, we decided to combine the two methodologies into an auto trader and to asses the forecasting boosting power in accuracy of both methodologies combined.

The auto trader is fundamentally an algorithm we designed based on the original Bayes Net model that was trading on the combinatorial prediction market.

The BAYES NEt algorithm uses a knowledge-based approach (Matsumoto et al. 2011; Sun et al. 2012). Experts specify a Bayes net model and the conditional probabilities for each node (the weights of the dependencies between nodes). In the Bayes Net model, each node represents a question about a geopolitical event that is also up for bidding on the DAGGRE prediction market. The model aims to forecast the probability of an event to happen (also named "target" or "hypothesis" question), conditional on the other events described by the Bayes Net (named "supporting" questions). The algorithm then reads the current market estimates for the set of supporting questions as 'soft evidence' to infer the likelihood of the hypothesis question, and it only trades on the hypothesis question. The soft evidence is the partial evidence that comes from the real world which is likely to have an impact on the outcome. For example, more information from the opinion polls or from mass media before an election will trigger an adjustment and a revision of the probability of the event to happen.

In order to assess the performance and forecasting accuracy, either on the market or for the Autotraders presented in this paper, we calculate the Brier score. The Brier score (Brier 1950) is a measurement of the accuracy of probabilistic predictions.

$$BS = \sum (forecast - outcome)^2 \,/\, no.forecasts \tag{1}$$

The *outcome* can only be 0 (False) or 1(True).

As a distance metric, lower is the better. The Brier score ranges from 0…2, and is the sum of the squared differences between the forecast and the outcome averaged over the number of forecasts. For example, on a binary (Yes/No) question, simply guessing 50 % all the time yields a score of 0.5. The closer to 0, the better the forecasting accuracy.

All Autotraders receive points in order to make trades. As do human traders, they gain points if they invest "well" and lose points in they disinvest. In other words, they are rewarded and punished equally as a human trader would be if they made the same trading decisions.

The Autotraders are communicating with the market daily, by reading the last values of the estimates and by placing their bids.

Grexit: The Flat Prediction Market

One of the world events that we have been forecasting on DAGGRE both in the flat and in the combinatorial prediction market was the exit of Greece from the European Union (not only the Eurozone, but from the Union itself), named by mass-media as Grexit.

The premise for this event was that Greece has defaulted on all its loans, yet it also received a huge cut on its private debt. We started with a singular question in the flat market, such as "Will Greece remain in the European Union by June 1, 2012?" and we developed a Bayes Net model offline (see Fig. 2).

The forecasting problem of Greece exiting the European Union ("Grexit") is the hypothesis and therefore the core node in the Bayen Net model we developed offline.

The decision nodes in the Bayes Net reflect the only two possible official ways for Greece to exit the EU:

1. Will Greece be ejected from the EU by June 1 2012? Outcome: Y/N
2. Will Greece withdraw from the EU by June 1 2012? Outcome: Y/N

These decision nodes were added as separate questions in the market. The model relies on the assumption that there are only two decision makers that would have an impact on the hypothesis: either the EU decides for Greece to be ejected or Greece decides to withdraw. For either decision maker, the actual event would not happen unless the vote in the EU is unanimous. Nevertheless, the model assumes that such a decision would mostly be influenced by the vote of Germany.

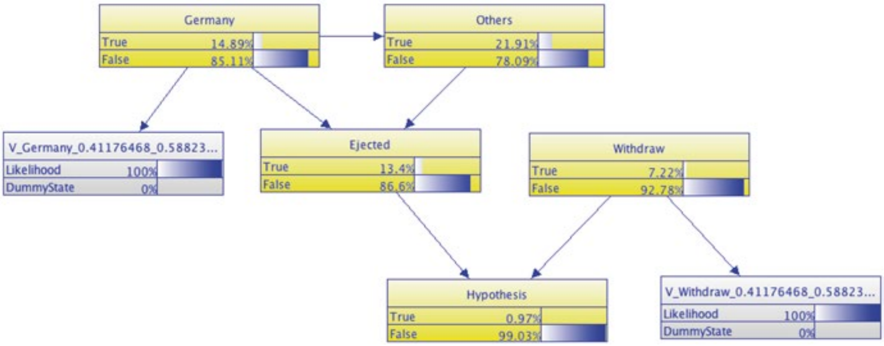


Fig. 2 The offline Bayes Net model of Grexit. "Germany" represents the node for Germany voting power to eject Greece and "Others" represents the other European Union members (other than Greece itself) that also have voting power. The true/false nodes are the forecasting questions on the market; the likelihood/dummystate nodes are the likelihood updates for the respective forecasting questions

In this offline Bayes Net, the probability of the core node is given only by soft observations. This means that the core forecasting problem is being revised based only on the indirect prediction market information that comes from the supporting nodes (the values the prediction market gives to "Germany" and "Withdrawal"). In this case, they are represented by the blue nodes (see Fig. 2). The blue nodes are the ones that continuously update the probabilities of the supporting (evidence) nodes in the model.

The core node is the final outcome of the model (the hypothesis), that we needed to assess based on the daily updating of evidential nodes (blue). The evidential nodes are the inputs in the model. The evidential nodes show the status of the prediction market for nodes "Germany" and "Withdraw" respectively. The probabilities in the evidential nodes are updated by introducing the likelihood values. For example, if we introduce a likelihood of 60 % true (40 % false) in the "Germany" node and propagate through the network, the core probabilities get updated. Propagation is performed using Jeffreys' Rule (Jeffreys 1946).

Jeffrey's rule revises the probability of a function based on another function; in this way, it conditions the probability of an event on another event and updates the belief about an event ("Grexit") based on the beliefs of other events (Germany's voting influence and Greece's own view of the EU membership).

For the core node, we compared this observation given by the offline model with the estimate from the market:

1. We update the core with the likelihood and propagate through the network (blue nodes).
2. We compare the probabilities of the parent nodes with the market.
3. We analyze the differences from the updated BN probabilities and the market in forecasting accuracy.

Figure 3 shows that the offline Bayes Net model performed better than the market. This is an expected result, given that the human users did not link the three

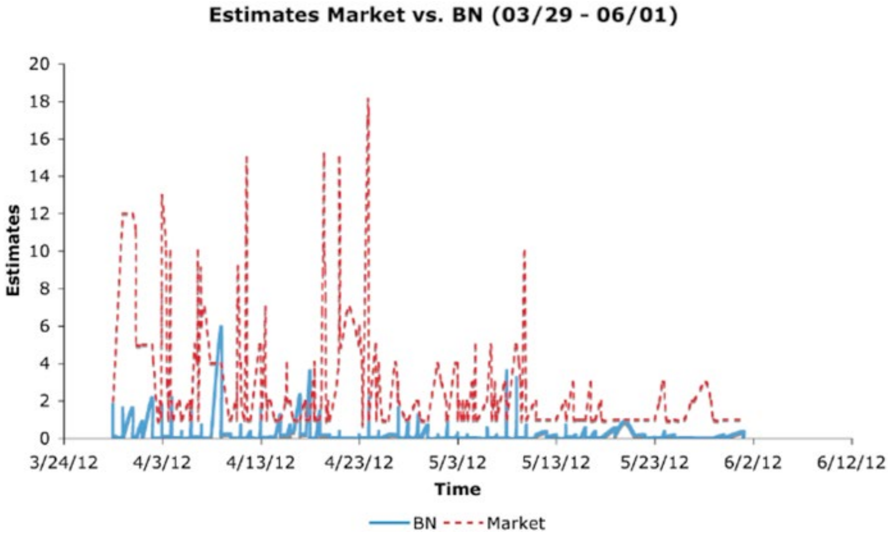


Fig. 3 The offline Bayes Net model versus the flat prediction market on Grexit. The time series of the estimates (probabilities) shows that the probability of Greece exiting the European Union is closer to the outcome (0 = False) for the offline model than for the market. The raw estimates (probabilities) that are given by the forecasters range from 0 to100

forecasting problems causally in the market and estimated on each of them independently. This result also implies that linking the events gives better and more refined forecasting.

Figure 4 shows on top how the Brier score of the market performed while the question was live, closing with a value of 0.1.

The Brier score of the offline Bayes net model was 0.00012, which shows an improvement of three orders of magnitude relative to the flat prediction market.

"Grexit": The Combinatorial Prediction Market

After the launch of the combinatorial prediction market, we also re-launched the same three questions with respect to Grexit, only that now the users were able to link them in the market as well and make their estimates given the probabilities from the parent nodes. One important thing to note though is that, although this capability was available, the users would not always use the combinatorial feature to place their bets. In other words, if they wished so, the human users were able to use the links between the events implied by the Bayes Net model that we only used offline in the flat prediction market; they were able to use the supporting questions as assumptions for the hypothesis.

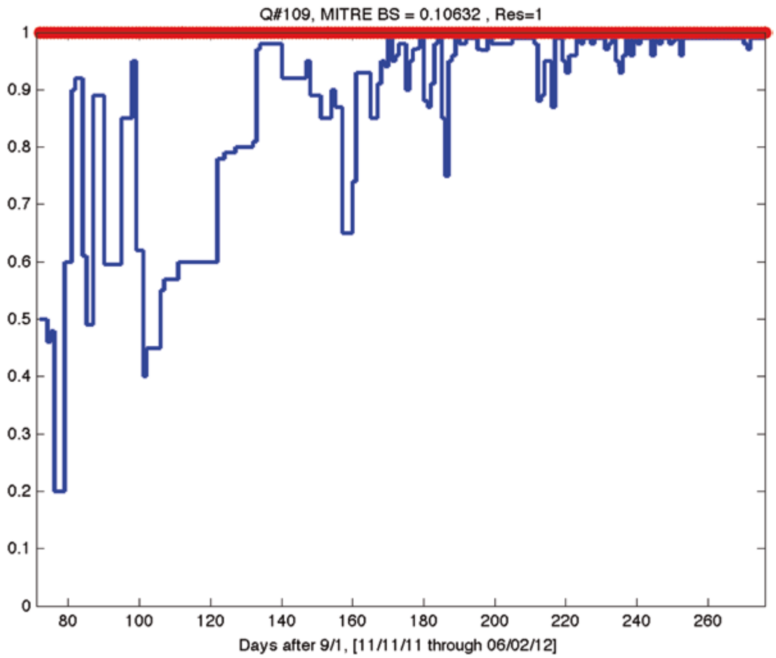


Fig. 4 The Brier score of Grexit on the flat prediction market. The *bolded line* represents the final outcome (1 = True) and the *plotted line* represents the daily Brier score aggregated for the entire market

The re-launched questions were also ordinal—we allowed for a flexible resolution deadline, i.e.:

Will Greece exit the EU:

(a) Between August 7, 2012 and October 1, 2012.
(b) Between October 1, 2012 and January 1, 2013.
(c) Between January 1, 2013 and April 1, 2012.
(d) Not before April 1, 2013.

For robustness, the same time intervals were used for the causal nodes of "Germany" and "Withdraw".

All questions closed with the outcome (d) Event will not happen before April 1, 2012.

Figure 5 shows the aggregated Brier score of Grexit given by the combinatorial market.

The overall Brier score of Grexit on the combinatorial prediction market is 0.019 and the Brier scores for the supporting questions were 0.01 for "Withdrawal" and 0.008 for "Germany". In this case, the human users performed *better* than in the flat prediction market, with an improvement of one order of magnitude in the Brier score.

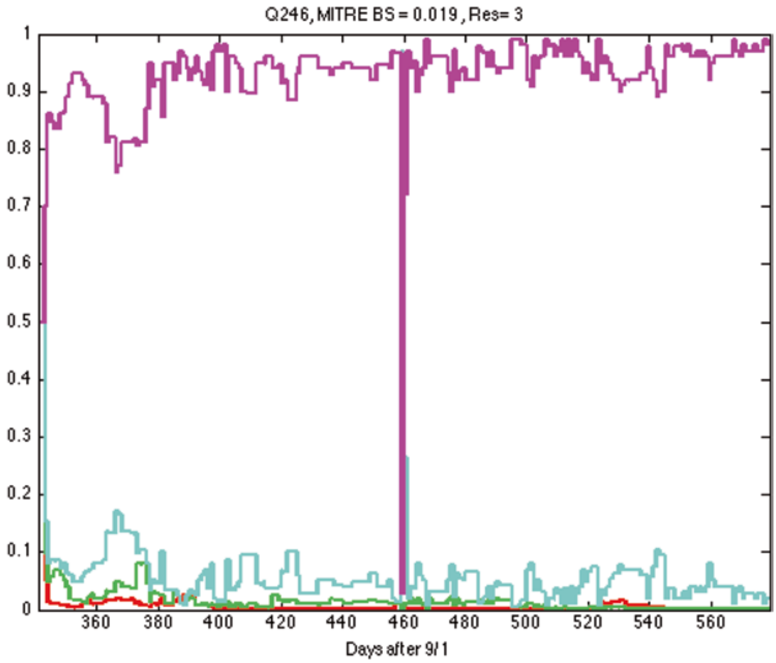


Fig. 5 The Brier score of Grexit on the combinatorial prediction market. Each line represents the probability/estimates time series for each of the four options (time intervals). The line at the top of the plot corresponds to option *(d) Not before April 1, 2013;* and the large point changes in values show that the forecasters (human or auto trader) can be wrong in assessing the outcome. This is mitigated by the final Brier score, which is an averaging measure

"Grexit": Adaptive Agents in the Combinatorial Prediction Market

In the third stage of the experiment, alongside human users, we also introduced the Bayes Net Autotrader, an algorithm that would trade every day on the core question of Grexit, taking into account the estimates given by the combinatorial market on the parent nodes ("Germany" and "Withdraw"). The autotrader performance is close to the one of the market, as the updating is performed every day.

Figure 6 shows the Brier score for the autotrader. Similarly to the market, the autotrader's forecasting accuracy is better than the one given by the offline Bayes Net in the flat prediction market, but only marginally better than the forecasting accuracy of the combinatorial market.

The overall Brier score of the autotrader was 0.0033, an improvement of one order of magnitude relative to the combinatorial market and of two orders of magnitude relative to the flat prediction market.

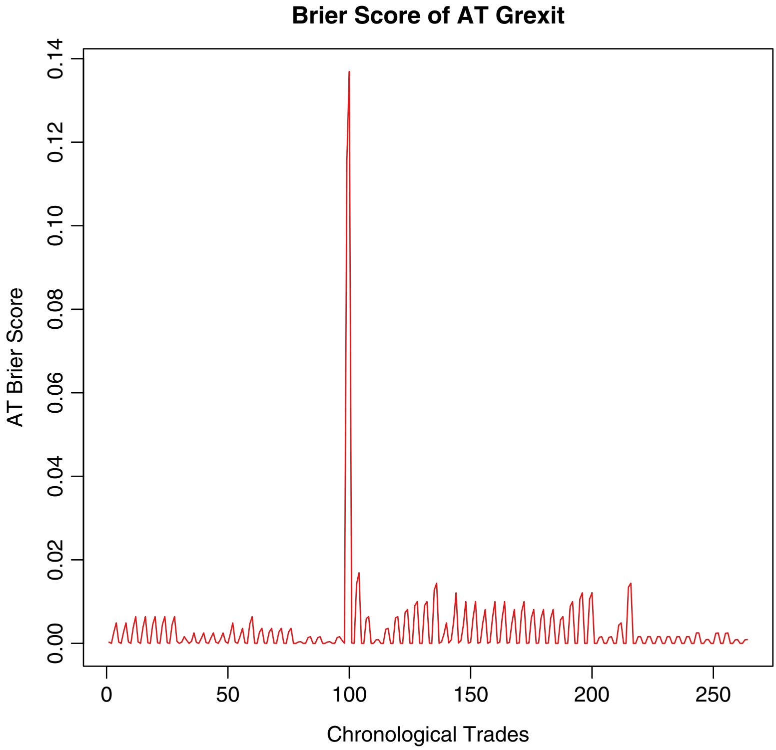


Fig. 6 The Brier score of Bayes Net Grexit Autotrader on the combinatorial prediction market

The results given by this series of experiments show that there can be designed an auto trader for the prediction markets to improve forecasting accuracy of methods such as crowdsourcing and the wisdom of the crowds. The auto trader is based on probabilistic models designed by the experts, but it is *adaptive* to the information received from crowdsourcing and the prediction market.

Conclusions

The combinatorial prediction markets allow for a better forecasting environment, both for human information crowdsourcing and for adaptive autotrading. The ongoing experiments with Bayes Net decompositions and adaptive agents in real time show that combinatorial prediction markets are a very powerful tool for improving forecasting accuracy in real time.

Acknowledgements The author is very grateful for the software and coding support provided by Rob Alexander and Dan Maxwell from KaDSci and Shou Matsumoto and Chris Karvetski from George Mason University. Supported by the Intelligence Advanced Research Projects Activity (IARPA) via Department of Interior National Business Center contract number D11PC20062. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DoI/NBC, or the U.S. Government.

References

Ben-Gal I (2007) "Bayesian networks" (PDF). In: Ruggeri F, Kennett RS, Faltin FW (eds) Encyclopedia of statistics in quality and reliability. Wiley, Chichester

Brier GW (1950) Verification of forecasts expressed in terms of probability. Mon Weather Rev 75:1–3

Cameron WB (1963) Informal sociology: a casual introduction to sociological thinking. Random House, New York

Hanson R (2003) Combinatorial information market design. Inf Syst Front 5(1):107–119

Hanson R (2007) Logarithmic market scoring rules for modular combinatorial information aggregation. J Predict Mark 1(1):3–15

Holland JH, Miller JH (1991) Artificial adaptive agents in economic theory. American Economic Review Papers and Proceedings, 81:365–370

Jeffreys H (1946) An invariant form for the prior probability in estimation problems. Proc R Soc Lond Ser A Math Phys Sci 186(1007):453–461

Lyon A, Pacuit E (2013) The Wisdom of crowds: methods of human judgement aggregation. In: Michelucci P (ed) Handbook of human computation. Springer

Matsumoto S, Carvalho RN, Ladeira M, da Costa PCG, Santos LL, Silva D, Onishi M, Machado E, Cai K (2011) UnBBayes: a java framework for probabilistic models in AI. In: Java in academia and research. iConcept Press.

Pennock D, Xia L (2011) Price updating in combinatorial prediction markets with Bayesian networks. In Proceedings of the Twenty-Seventh Conference Annual Conference on Uncertainty in Artificial Intelligence (UAI-11), pages 581–588, Corvallis, Oregon. AUAI Press

Sun W, Hanson R, Laskey KB, Twardy CR (2012) Probability and asset updating using bayesian networks for combinatorial prediction markets. In: Proceedings of the 28th conference on uncertainty in artificial intelligence (UAI-2012), Catalina

Surowiecki J (2005) The wisdom of crowds. Anchor Books, New York

Risks and Rewards of Crowdsourcing Marketplaces

Jesse Chandler, Gabriele Paolacci, and Pam Mueller

The present chapter focuses on the risks and rewards of using online marketplaces to enable crowdsourced human computation. We discuss the strengths and limitations of these marketplaces, with a particular emphasis on the quality of crowdsourced data collected from Amazon Mechanical Turk. Data quality is by far the most important consideration when designing computational tasks, and it can be influenced by many factors. We emphasize Mechanical Turk because it is currently one of the most popular and accessible crowdsourcing platforms and offers low barriers of entry to researchers interested in exploring the uses of crowdsourcing. In addition to describing the strengths and limitations of this platform, we provide general considerations and specific recommendations for measuring and improving data quality that are applicable across crowdsourcing markets.

Crowdsourcing is the distribution of tasks to a large group of individuals via a flexible open call, in which individuals work at their own pace until the task is completed (for a more detailed definition see Estellés-Arolas and González-Ladrón-de-Guevara 2012). Crowd membership is fluid, with low barriers to entry and no minimum commitment. Individuals with heterogeneous skills, motivation, and other resources contribute to tasks in parallel. Crowdsourcing leverages the unique knowledge of individual crowd members, the sheer volume of their collective time and abilities, or both to solve problems that are difficult to solve using computers, or smaller and more structured groups.

J. Chandler (✉)
University of Michigan/PRIME Research, Ann Arbor, USA
e-mail: jjchandl@umich.edu

G. Paolacci
Erasmus University Rotterdam, Rotterdam, Netherlands

P. Mueller
Princeton University, Princeton, USA

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_30, © Springer Science+Business Media New York 2013

The unique strengths of groups are generally used to solve one of two basic kinds of problems. Some problems have no obvious a priori solution, but correct answers seem obvious in hindsight (e.g. insight problems; Dominowski and Dallob 1995) or can be verified. In these cases, crowds can generate responses from which the "best" response can be selected according to some criteria. The volume and diversity of workers with different perspectives, strategies and knowledge can lead to quick, unorthodox, and successful solutions. The Internet has furthered this approach to problem solving by creating virtual meeting places where people can post problems for others to solve. For example, Innocentive (Allio 2004) is a website that has helped companies find solutions to technical challenges like preventing oxygen from passing through rubber, or adding fluoride powder to toothpaste without dispersing it into the air. Often solutions to these specialized, technical problems are provided by amateurs, hobbyists, or experts in apparently unrelated fields (Lakhani 2008).

Tasks that require resources beyond those available to a single individual or work group are also well-suited to crowdsourcing. The compilation of the Oxford English Dictionary is one early example of this approach. A unique feature of this dictionary is that it includes not only definitions, but also published examples of word use. Examples were collected on slips of paper by a large body of volunteers and then aggregated by editors (Winchester 2004). Advances in machine computation have made it easier to manage projects of this scale. For example, The Open Science Collaboration coordinates the real time collaborative efforts of scientists and citizen-scientists to systematically code, replicate and communicate social scientific findings using freely available web-software (Open Science Collaboration 2013).

A subset of time-intensive tasks are tasks that are easy for people to solve, but difficult for machines to solve. These assignments are particularly amenable to crowdsourcing. In many cases, a crowd's responses can be automatically aggregated, eliminating the need to comprehensively review responses. The volume of workers performing each task can allow ideosyncratic perspectives, strategies and knowledge to be homogenized through aggregation, leaving consistent performance across a task even though each individual completed only a small portion of it. Consequently, advancing machine computation has increased the applications of crowdsourcing through the development of human-machine hybrid systems that tackle ambitious projects such as describing the contents of images in near real time (e.g., VizWiz; Bigam et al. 2010), classifying millions galaxies (Galaxy Zoo; Lintott et al. 2008), or determining the shapes that proteins fold into (Foldit; Cooper et al. 2010). Each of these projects emerged as a result of the uneven ability of machine computation to handle the various necessary task elements.

While some platforms for marshaling crowds have been developed to solve specific large problems, "crowdsourcing marketplaces" have also emerged to match workers and requesters with more modest needs. The most prominent example is Mechanical Turk (MTurk), a crowdsourcing website launched by Amazon in 2005 to assist with the maintenance of its own websites (e.g. identifying duplicate products; Pontin 2007). Corporations and individuals alike use crowds recruited from MTurk to conduct human computation operations. Twitter, for instance, relies on MTurk workers to categorize search queries to make them more meaningful to other users. Machine computing can easily identify a spike in the popularity of a query (e.g., “Big Bird” in Fall 2012), but not its semantic properties. Trending queries are passed on to MTurk workers, who can easily determine that this is a result of political events (Mitt Romney’s comments in the US Presidential Debate) rather than *Sesame Street*.

Scientists have also been quick to harness crowd computing for academic research, relying on crowds to complete a variety of time-consuming tasks including generating corpora of stimuli for machine learning experiments (Lane et al. 2010; Lau et al. 2009); rating and classifying words according to meaning (e.g., Li et al. 2008); transcribing speech (Gruenstein et al. 2009; Marge et al. 2010); proofreading text for errors (Tetreault et al. 2010); verifying citations (Molla and Santiago-Martinez 2011) and coding observational data (e.g. Hsieh et al. 2010). Others are experimenting with building more complex workflows, where workers collaborate on complex multi-stage projects, or in which workers are treated as agents with a plurality of diverse responses, rather than a means of measuring the average beliefs of a population (Nickerson et al. 2011; Yu and Nickerson 2011)

Strengths of Crowdsourcing Marketplaces

Transaction Cost Effectiveness. The major advantage of marketplaces is that they make crowdsourcing accessible to requesters with limited financial and technical resources. The fixed costs of crowdsourcing (servers, record keeping, technical support, etc.) can be shared by many requesters and the technical challenges can be handled by dedicated specialists. Other less tangible efficiencies are also realized through sharing a common platform. Workers only need to be recruited into the market once, reducing marketing costs. Moreover, they only need to learn how to use a single standardized interface and can share their experiences with others, making it easier for them to find, understand, and successfully complete work (Ipeirotis and Horton 2011).

Crowd Accessibility. Crowds require a certain critical mass to function. Potential workers are unlikely to invest time visiting websites unless they have a reasonable chance of finding work (a special case of a two-sided market, see Rochet and Triole 2003). Some crowdsourcing projects, like digitizing every book in the world, or identifying all the stars in the sky, are large enough to warrant their own dedicated framework (e.g., reCaptcha; von Ahn et al. 2008). However, the majority of human computation problems are quick to complete, intermittent, or frequently change in content or required knowledge. A common market ensures a steady enough supply of tasks to help maintain a persistent crowd, even while individual requesters recruit and dismiss workers on demand. MTurk was able to achieve this scale initially by serving as a labor market for Amazon’s own in-house human computation needs.

Efficient Matching and Task Completion. Microtask sites pay workers according to the tasks they complete, rather than an hourly wage. Piece rates ensure that workers are paid according to their productivity, and even assuming minimal variation in worker ability and task demands, workers should be able to sort themselves into assignments they do best (Becker and Murphy 1992). Piece rates also benefit requesters. Since each worker proceeds at their own pace, receiving new work only when old work is completed, the completion time for a project will be driven by the average pace at which tasks are completed, as opposed to traditional methods of dividing labor that are often constrained by the pace of the slowest worker (Davis 1965).

Low Market Prices. Aside from a minimal payment to the web service (MTurk charges 10 % of worker payments to cover overhead and financial transaction fees), the only cost faced by requesters to crowdsource their tasks is worker compensation. Horton and Chilton (2010) estimated the median reservation wage of MTurk workers to be less than $2 per hour, i.e., less than 20 % of the wage of the average general secretary in the United States (Bureau of Labor Statistics 2010). Current rates are likely higher, but even a rate of $6 per hour is sufficient for a task to be posted to one of the various forums where workers share well-paying HITS (e.g. http://www.reddit.com/r/HITsWorthTurkingFor, www.turkernation.com).

There are a number of reasons that workers within certain crowds accept wages which traditional workers would not: they can select tasks that are relatively interesting or meaningful (Kaufmann et al. 2011), they can work from any location, and they can use time that has little other economic value (e.g., completing work between or even in parallel with other tasks). MTurk also allows requesters to recruit workers from regions or countries with lower costs of living and lower minimum wages. However, we should also note that the US workers are often comprised of people with limited traditional sources of income (Shapiro et al. 2013) and that researchers may want to consider the ethical implications of the wages they offer workers when making payment decisions (for discussions see Horton 2011; Kittur et al. 2013; Silberman et al. 2010).

Trust and Reputation Transparency. Exchanging goods or labor requires a certain amount of trust. In offline communities, reputational information is spread informally through a community. Online, requesters and workers must interact anonymously with each other, making them vulnerable to fraud or exploitation. The division of work into smaller tasks paid as piecework prevents the need to engage in long-term commitments between workers and recruiters. Workers can try working with a requester once with minimal risk and increase their commitment if the first transaction proceeds smoothly.

Centralizing work within an online marketplace makes it possible to share information about potential exchange partners so participants can identify and avoid or sanction untrustworthy partners, even when they are effectively anonymous (Resnick et al. 2000). MTurk, for example, tracks the proportion of tasks that workers successfully complete, and requesters can use this information as a recruitment criterion. Particularly unscrupulous workers can be blocked by individual requesters, and multiple blocks can result in workers being banned from the marketplace. Similarly, workers maintain ratings of requesters (e.g. www.turkopticon.com) that can guide other workers' decisions about who they work for.

Data Quality. The low cost of labor, combined with the conventional wisdom that "you get what you pay for," can lead to skepticism about the true value of work performed by crowds of strangers working for below minimum wage. Empirical examinations have found that data quality is not something that can be solved through wages: poorly paid crowds produce data of the nearly the same quality as well paid crowds (albeit slowly; Rogstadius et al. 2011; Mason and Watts 2009), community volunteers (Goodman et al. 2012), or undergraduate students (Paolacci et al. 2010, for a general discussion see Gneezy et al. 2011). There are forces that ensure quality even when payment is low: many tasks that are difficult for machines are trivially easy for people to do, and for more difficult tasks, reputational concerns may dissuade workers from submitting poor quality work. Further, since most crowdsourcing tasks recruit workers using an open call, high wages attract more workers of all skill levels to the task equally. Instead, features of task design, instruction clarity and worker selection may play a greater role in determining work quality in crowds.

Recruitment Flexibility. Crowdsourcing marketplaces allow requesters to specify that workers possess certain attributes in order to complete a task. Worker recruitment on MTurk can be restricted to residents of a specific country, or to workers who have completed more than a certain number of tasks with a specified rate of accuracy. Moreover, as discussed below, with minimal coding knowledge requesters can create and assign ad hoc "qualifications" to workers based on nearly any measurable attribute that grant specific workers access to tasks. Thus, smaller bespoke crowds can be constructed out of the workforce to complete highly specialized tasks.

Crowds are easy to program. For those with little experience programming machines, a major advantage of crowds is that they are comparatively easy to instruct. People are experienced at communicating with each other, and actively work to make sense of their environment. People also interpret the pragmatic meaning of a request in far more detail than a literal reading would suggest, drawing upon contextual details and assumptions based on their own experience as communicators (e.g., that all relevant information is provided, and all provided information is relevant; for a discussion see Grice 1989). As a result, crowds are tolerant to errors and ambiguity, and can easily go beyond the information provided to complete a task as the requester intended. In contrast, even when completing a task as simple as rating the positivity of words, a machine requires numerous variables to be defined including the universe of words to be rated, the context in which they might be used and the purpose the requester will use them to ensure an appropriate range and distribution of responses.

Limitations of Crowdsourcing Marketplaces

Although crowdsourcing marketplaces offer a number of compelling opportunities, there are also some potential challenges that may interfere with the accuracy of human computation. Speed and cost are inversely related to each other, and both are constrained by marketplace features beyond the control of individual requesters. Data quality may vary by marketplace, but also varies highly across tasks and workers and is thus under the direct control of requesters. We review several issues that pertain specifically to data quality.

Lack of motivation. While workers are to some extent intrinsically motivated to participate in crowdsourcing tasks (e.g., von Ahn 2006), motivation is fickle and workers are inclined to avoid the most difficult elements of a task (Mason and Watts 2009, Study 2). In this sense they can be regarded as "satisficers" who are likely to do only the minimal amount required to ensure payment (Simon 1972). For example, if workers are asked to search for information on the Internet and are paid a reward even if they indicate that the requested information is not available, they may be inclined to report that the information does not exist without a thorough search.

Cognitive limitations. Workers are people, and consequently suffer from a long but predictable set of cognitive and perceptual biases. This has led behavioral experimentalists within diverse disciplines to use workers as a subject pool for research (Goodman et al. 2012; Paolacci et al. 2010; Rand 2012). However, for the same reason, human computation researchers need to acknowledge that crowdsourced workers are not infallible computational agents, but rather are boundedly rational individuals that selectively allocate limited and depletable cognitive resources (for a general overview see Kahneman 2011). While these biases lead to perceptions beliefs and decisions that are "good enough" under most circumstances, they also produce systematic errors. These features may make crowdsourcing less suitable for some tasks where the requester seeks objectively correct answers through the aggregation of worker responses because aggregation cannot remove systematic bias.

Instruction ambiguity. The same cognitive abilities that make it possible for people to "program" a crowd with minimal instructions can pose problems for requesters because these processes will draw upon all information—both intentionally and unintentionally communicated—to understand a task. There are numerous examples of how design features such as response formats, question order and the affiliation of a communication partner guide inferences about the interviewer's intent and thus influence the responses provided (e.g., Bao et al. 2011; for a review see Schwarz 1999). Unfortunately, these features may be selected or communicated arbitrarily by requesters, without considering the effects they can have on worker's responses.

Workers may also make inferences about what a requester wants by drawing on their prior experiences with other requesters. For example, Goodman and colleagues (2012) conducted a decision making study in which they asked workers to guess the number of countries in Africa (adapted from Tversky and Kahneman 1974). Although the authors did not explicitly ask workers to look this information up, an unusually large proportion of them gave answers that matched information available on the Internet. One explanation for this is that it is normative for MTurk workers to provide factually correct information, which led workers to believe that the requesters desired a factually correct answer rather than a subjective impression. Although little research has directly investigated this issue on Mechanical Turk, the importance of tacit norms in other workplaces has been extensively documented (Wenger 1998).

Worker (non-)naivety. Workers may complete the same task several times or share information about tasks with each other. Prior knowledge about the contents or objectives of a task may benefit some crowdsourcing tasks. However, it is possible for workers to have *too much* information. At the most basic level, if the requester is interested in measuring the average rating of a target to smooth out the idiosyncratic beliefs of workers, it is obviously preferable to ensure that several different individuals rate it, rather than the same individual several times. Indeed, all "wisdom of crowds" tasks (Lyon and Pacuit this volume) that aggregate worker responses *require* that judgments are made independently; when worker responses are not independent, errors will be correlated with each other and cannot be canceled out through aggregation (e.g., Anderson and Holt 1997; Hullman et al. 2011). Independence across different tasks may also matter in more complex workflows. For example, if workers are required to complete several related tasks in stages, such as transcribing text and then rating other workers' transcriptions for accuracy, requesters would want to avoid situations in which the same worker translates and evaluates the accuracy of their own translation.

The sheer size and anonymity of crowds makes it easy to underestimate the likelihood of duplicate workers. After all, with thousands of tasks and thousands of workers, what is the probability that the same worker would end up processing the same information twice? Two factors make this more likely than it might otherwise seem. First, workers tend to follow favorite requesters by subscribing to websites that alert them whenever favored requesters make work available for completion (e.g., www.turkalert.com). Second, workers complete varying numbers of tasks, with most of the work completed by a small group of extremely prolific workers. For example, we found that in a sample of 16,000 completed task submissions, the most prolific 1 % of workers was responsible for completing 10 % of the work, and the most prolific 10 % were responsible for providing 41 % of the observations (Chandler et al. in press, see also Berinsky et al. 2012; Grady and Lease 2010).

While Amazon by default prevents workers from completing the same task twice as a part of a single batch of tasks, additional measures (such as the use of Qualifications or third party software; Chandler et al. in press; Goldin and Darlow 2013; Pe'er et al. 2012) must be used to ensure that workers across different tasks are kept unique.

Workers may also share information with each other about the nature of a task, or collude in the responses they provide (Kazai and Milic-Frayling 2009). Workers gather in forums (e.g., http://www.reddit.com/r/HITsWorthTurkingFor, mturkforum.com) to share information and opinions about tasks (e.g., particularly interesting and lucrative HITs), which could potentially lead them to have foreknowledge of certain task details. Thus, tasks that rely heavily on initial impressions of a target of judgment, or tasks that screen out workers based on specific responses, should be designed with care to minimize worker foreknowledge.

Worker Honesty. Some tasks may require that people post information that is not directly verifiable or that has no factually correct response. For example, a requester may want to solicit opinions about a particular image or idea, or may want to know a worker's geographical location to assess their knowledge about local businesses.

In general, workers provide factually accurate information (Shapiro et al. 2013) but deception can increase substantially if workers benefit from lying (Suri et al. 2011). In particular, on MTurk, large numbers of non-US workers claim to be US residents in order to receive cash payments (perhaps because workers in most other countries are paid with Amazon credit rather than cash).

Ensuring Data Quality in Crowdsourcing Marketplaces

Data quality is determined by numerous factors, some of which are under the control of requesters. Obtaining quality data is most straightforward for tasks that can be divided into many smaller components. This makes it easier for workers to select elements of the task that they enjoy or are good at while minimizing the learning curve. Further, smaller tasks are often completed more efficiently because minimally motivated workers can still provide useful data (Mason and Watts 2009). Additional steps can be added to ensure quality control. For example, Mechanical Turk workers can successfully proofread and condense complex text, when a task is broken into smaller subtasks of finding problems, fixing problems and verifying proposed fixes (Bernstein et al. 2010).

For complex tasks, it may also be necessary to test worker ability before hand, and restrict access to workers who possess the necessary skills, or to consider other online labor markets (e.g. oDesk) that match requesters with more specialized workers. Regardless of the software platform requesters use to recruit workers, they should also consider what software is best suited to the collection of work. Even sites like MTurk that allow tasks to be created using their own website also allow tasks to be created on a separate webpage or software program that is linked to or embedded within the web interface (Mason and Suri 2012). Thus, requesters should not feel constrained by the platform used to distribute the work.

Task Design. There are many potential uses of crowdsourcing websites, and there is no one-size-fits-all solution to task design. In general, the approach requesters take when designing a task is more important than the specific design choices they make. Tasks should always be pilot tested, first by the requester and then by a small pool of workers, before being fully distributed to workers. Crowd interest is greatest when a HIT is first posted (Chilton et al. 2010), and minor mistakes can quickly become expensive. MTurk provides a "requester sandbox" in which the technical details of tasks can be tested by a requester. For pilot testing on workers, requesters should provide both the task of interest, and questions about the task of interest, to identify potential improvements in design (Collins et al. 2004). They should also have a clear benchmark against which the quality of work can be evaluated.

Although comparatively little research has been done on task design itself (for exceptions see Grady and Lease 2010; Khanna et al. 2010), there is a large literature on survey design that is relevant to requesters, which may be useful when considering data quality issues identified in pilot testing. Surveys are similar to crowdsourcing tasks in that instructions are communicated to workers rather than jointly discussed, and responses are collected through similar standardized methods. Consequently, it may be useful to requesters to consult a general overview of web survey construction when designing tasks (e.g. Couper 2008) in addition to more general resources on web design (e.g., Krug 2009).

Screening Workers. As discussed earlier, MTurk allows requesters to select workers for inclusion in tasks based on whether or not they possess specific attributes. In general, workers with more experience and a higher reputation should be less likely to provide poor quality work. There is also evidence of differences in the quality of work provided by workers from different geographical locations, perhaps reflecting language difficulties or differences in education (Khanna et al. 2010; Kazai et al. 2012). Alternatively, or additionally, requesters can create their own qualifications to screen workers according to more specific criteria such as their competence on particular tasks (e.g., Chua et al. 2009; Zhou et al. 2011; for details on how to implement these procedures in Mechanical Turk see Chandler et al. in press.

Preventing Satisficing. Since many workers are motivated by money to complete tasks as efficiently as possible, satisficing (providing minimally adequate responses; Krosnick 2006) is a major concern. Instructions or task elements can be presented sequentially with delays between each new piece of information to slow workers down (Kapelner and Chandler 2010). Satisficing can be further reduced by introducing features that require workers to think about the "correct" response rather than simply providing their first impressions. One study asked workers and experts to evaluate the quality of Wikipedia pages. Worker ratings and expert ratings were uncorrelated, except when workers were also required to include answers to objectively verifiable questions (Kittur et al. 2008). Similarly, other researchers found that accuracy improved when workers were asked to predict how other workers would respond to a question rather than simply offer their own opinion ("Bayesian truth serum"; Shaw et al. 2011; for a discussion see Prelec 2004).

Worker motivation can also be increased. Crowds perform better on meaningful tasks (Chandler and Kapelner 2013, see also Reed et al. this volume). Another alternative is to simply pay workers to pay attention. MTurk allows requesters to award bonuses to workers above and beyond the initial rate paid for completing work. Thus, requesters can structure a task to make it monetarily rewarding for workers to pay attention. To illustrate, in a pair of virtually identical studies (conducted by the third author of the present chapter), MTurk workers were paid either a total sum for participating ($1) or a smaller initial sum ($.30) with the remainder ($.70) paid as a bonus for successfully recalling details about the experimental manipulation. Although both sets of workers had the same potential earnings, those paid a smaller sum plus a performance bonus were more likely to correctly answer the factual multiple choice questions (98.2 %) than participants who were paid a lump sum (87.0 %), $\chi^2(1, N=494)=23.03$, $p<.001$ (see also Shaw et al. 2011). Interestingly, the success of bonuses in promoting attention seems to be independent of the bonus amount (Chandler and Horton 2011).

Identifying Poor Quality Workers. There are a number of strategies that can be used to identify poor quality workers. Responses by workers who frequently disagree with their peers can be excluded (Elson and McKeown 2010; Sheerman-Chase et al. 2011). Alternatively, "gold-standard" questions with factually correct answers, or "catch-trials" with obviously correct responses can be included along with the task of interest to measure worker ability and attentiveness (e.g., Sayeed et al. 2011). Tasks submitted along with incorrect responses to these questions can be excluded from analysis under the assumption that other components of the task are likely to also be incorrect. Additionally, or alternatively, all of the responses provided by workers who fail a predetermined number of such checks can be excluded.

Multiple choice questions are frequently used to measure data quality because they are easily scored. The assumption is that workers who do not take the task seriously, or who do not understand the instructions, will likely respond at random, and are thus likely to select incorrect responses. In general, the sensitivity of gold-standard multiple choice questions to detect quality responses increases asymptotically: All else being equal, a single, four-item multiple choice question will only identify the 75 % of random responders who select one of the three incorrect answers, while two four-item multiple choice questions will identify the 96 % of random responders who select an incorrect answer on either or both questions. The actual ability of multiple choice questions to detect random responding is also dependent on the quality of the response alternatives (cf., Case and Swanson 2001).

Measuring Data Quality. Data quality is often quantifiable and measurable. Reliability of categorical or continuous ratings can be evaluated based on its agreement with ground-truth, expert ratings or worker consensus. The critical question is whether agreement is sufficiently better than chance, although the level of agreement necessary is highly task dependent. Crowdsourced data is unusual in that not all workers complete all elements of a task. Reliability of data with this property can be measured using Krippendorff's alpha (Krippendorff 2004, for SAS and SPSS macros see Hayes and Krippendorff 2007). High reliability scores between workers is a function of both task difficulty and the number of raters and is a necessary precondition for valid responses. If reliability is low, it could suggest poorly communicated instructions or a plurality of acceptable answers. Reliability can be increased by refining worker instructions and increasing the number of workers who perform each task.

Cleaning and aggregating responses. Responses by different workers can also be combined. In general, aggregating the ratings of many independent judgments, even through averaging or a simple majority, will increase their accuracy, as idiosyncratic errors cancel each other out (Galton 1907). More complex methods of aggregating responses can improve data quality yet further. Some approaches use quantitative methods to improve quality, trimming responses that are likely to be outliers (Jung and Lease 2011) or estimating worker quality and then weighting their responses on specific tasks accordingly (Hosseini et al. 2012; Tang and Lease 2011). Other approaches use workers themselves to review and combine responses in an interactive, iterative process (Nickerson et al. 2011).

As a final note, aggregation does not increase the likelihood of a correct solution unless each judgment is independent. If a majority of answers are identical but agreement is not independent—either because workers have discussed their responses beforehand or because care was not taken to avoid duplicate respondents (see limitations section)—then the value of the majority's opinion may be suspect. Likewise, aggregation will not provide a correct solution for problems in which workers are systematically wrong, either because they lack the necessary information to reach a correct conclusion or because cognitive biases lead workers to draw incorrect conclusions.

Conclusions

Crowdsourcing marketplaces present an opportunity for researchers who require human computation services, especially for tasks that are small, require a variety of different skills or interests, or are intermittent in their availability. They offer a persistent workforce that is available on demand for an affordable price. However, data provided by workers is not inevitably high quality: tasks must be designed to maximize the likelihood and ease with which workers can provide useful responses.

While specific design considerations largely depend on the researcher's goals, task design can be improved iteratively through pilot testing, and a number of principles exist that can improve the quality of data collected on crowdsourcing marketplaces. In particular, crowd members are heterogeneous and requesters can take advantage of this by preselecting workers who are most capable of performing specific tasks. Further, tasks can be optimized so that workers can understand them and feel motivated to complete them correctly. Finally, despite varying rates of participation by individual workers, quality can be measured, and to a certain extent improved, through aggregating responses. In this sense, the output of the crowd can be greater than the sum of its parts.

Online marketplaces have developed rapidly in the past few years. While it is notoriously difficult to predict what will happen in the future (e.g., Tetlock 2005), there are a few developments that seem particularly plausible. Network effects give Mechanical Turk a large competitive moat against alternative platforms, but individuals are working to counter some of its limitations within its current framework. Requesters are beginning to use it as merely a gateway to request labor, and are directing workers to complete tasks on other software platforms that allow dynamic and real-time collaborative tasks.

Perhaps more crucially, workers and requesters alike are developing the means to increase market transparency. While Amazon has implemented minimal channels for transmitting information directly between requesters and workers, and indirectly between various requesters, much of the increased transparency discussed in this chapter is a result of requesters and workers finding their own means of communicating with each other outside of Amazon's platform.

However, information exchange is still relatively limited. There is no public register of market participants, and workers can only be recruited using a narrow range of metadata. Additionally, requesters are unable to access information about general market conditions or task completion rates that would allow them to optimize tasks and compensation rates, or to directly match tasks with workers of varying levels of skill and motivation. Often, requesters must build their own panel of workers (which takes time) based on information that was privately collected, or shared in informal, insecure ways. Perhaps worse, workers have no access to requesters' profiles, making the relationship between Requesters and Workers inherently asymmetrical. Some workers rely on independent websites that allow workers to rate and subscribe to requesters. However, in general workers are unable to determine which tasks pay fairly and which qualifications are worth the unpaid effort necessary to complete them. For requesters, completions times thus depend heavily on whether their tasks are credentialed in an external forum (Chandler et al. in press). More generally, poor quality requesters run the risk of creating something close to a "market of lemons" in which the highest quality workers refuse to participate because of these issues (Akerlof 1970; for a discussion see Horton 2010). All of these issues hinder the effectiveness of MTurk as a labor market, and we anticipate that workers and requesters will continue to increase information exchange and transparency.

Another interesting question is what tasks online labor markets will be used for in the future. As machine perception and language processing improve, it is likely that demand for human and human-machine hybrid computational solutions will no longer be needed for these tasks. Just as steam drills replaced railroad workers, and office productivity software has replaced middle class white collar employees, so too will software replace crowds, for some tasks. It remains to be seen whether crowdsourcing, especially microtask labor markets, are merely a solution to temporary deficiencies in the advance of machine computing, or if, as has occurred in other labor markets, new tasks will continue to emerge as a technology advances. For instance, as workflow management platforms become more automated, iterative tasks may become possible. As research about task decomposition develops, there will be opportunities to use microtask markets for problems that require increasingly complex and creative solutions. As more data becomes digitized and interconnected, there will be more opportunity to search for interrelations between increasingly disparate topics. Finally, a larger sociological question that remains to be answered and is how these changes within crowdsourcing marketplaces may impact other labor markets (see Felstiner this volume) and society at large (see Nardi this volume).

References

Akerlof GA (1970) The market for" lemons": quality uncertainty and the market mechanism. Q J Econ 84:488–500

Allio RJ (2004) CEO interview: the InnoCentive model of open innovation. Strategy Leadersh 32(4):4–9

Anderson LR, Holt CA (1997) Information cascades in the laboratory. Am Econ Rev 87: 847–862

Bao J, Sakamoto Y, Nickerson JV (2011) Evaluating design solutions using crowds. In: Proceedings of the 17th Americas conference on information systems, Detroit, MI, USA

Becker GS, Murphy KM (1992) The division of labor, coordination costs, and knowledge. Q J Econ 107(4):1137–1160

Berinsky AJ, Huber GA, Lenz GS (2012) Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. Polit Anal 20:351–368. doi:10.1093/pan/mpr057

Bernstein MS, Little G, Miller RC, Hartmann B, Ackerman MS, Karger DR, Crowell D, Panovich K (2010) Soylent: a word processor with a crowd inside. In: Proceeding UIST 2010, ACM Press, pp 313–322

Bigham JP, Jayant C, Ji H, Little G, Miller A, Miller RC, … Yeh T (2010) VizWiz: nearly real-time answers to visual questions. In: Proceedings of the 23nd annual ACM symposium on user interface software and technology, ACM, New York, pp 333–342

Case SM, Swanson DB (2001) Constructing written test questions for the basic and clinical sciences, 3rd edn. National Board of Medical Examiners, Philadelphia

Chandler D, Horton J (2011) Labor allocation in paid crowdsourcing: experimental evidence on positioning, nudges and prices. In: Workshops at the twenty-fifth AAAI conference on artificial intelligence. AAAI Press, Menlo Park, California

Chandler D, Kapelner A (2013) Breaking monotony with meaning: motivation in crowdsourcing markets. J Econ Behav Organ 90:123–133

Chandler J, Mueller P, Paolacci G (in press) Methodological concerns and advanced uses of Amazon mechanical Turk in psychological research. Manuscript submitted for publication

Chilton LB, Horton JJ, Miller RC, Azenkot S (2010) Task search in a human computation market. In: Proceedings of the ACM SIGKDD workshop on human computation, ACM, New York, pp 1–9

Chua CC, Milosavljevic M, Curran JR (2009) A sentiment detection engine for internet stock message boards. In Pizzato LA, Schwitter R (eds) Proceedings of the Australasian language technology association workshop 2009, Sydney, pp 89–93

Collins A, Joseph D, Bielaczyc K (2004) Design research: theoretical and methodological issues. J Learn Sci 13(1):15–42

Cooper S, Khatib F, Treuille A, Barbero J, Lee J, Beenen M, Leaver-Fay A, Baker D, Popović Z (2010) Predicting protein structures with a multiplayer online game. Nature 466(7307): 756–760

Couper M (2008) Designing effective web surveys. Cambridge University Press, New York

Davis LE (1965) Pacing effects on manned assembly lines. Int J Prod Res 4(3):171–184

Dominowski RL, Dallob PI (1995) Insight and problem solving. In: Sternberg RJ, Davidson JE (eds) The nature of insight. MIT Press, Cambridge, pp 33–62

Elson DK, McKeown KR (2010) Automatic attribution of quoted speech in literary narrative. In: Proceedings of the twenty-fourth AAAI conference on artificial intelligence. The AAAI Press, Menlo Park, pp 1013–1019

Estellés-Arolas E, González-Ladrón-de-Guevara F (2012) Towards an integrated crowdsourcing definition. J Info Sci 38(2):189–200

Galton F 1907 Vox populi. Nature 75:450–451

Gneezy U, Meier S, Rey-Biel P (2011) When and why incentives (don't) work to modify behavior. J Econ Perspect 25:191–209

Goldin G, Darlow A (2013) TurkGate (Version 0.4.0) [Software]. Available from http://gideongoldin.github.com/TurkGate/)

Goodman JK, Cryder CE, Cheema A (2012) Data collection in a flat world: the strengths and weaknesses of mechanical Turk samples. J Behav Decis Making 26:213–224

Grady C, Lease M (2010) Crowdsourcing document relevance assessment with mechanical Turk. In: Proceedings of the NAACL HLT 2010 workshop on creating speech and language data with Amazon's mechanical Turk. Association for Computational Linguistics, pp 172–179

Grice HP (1989) Studies in the way of words. Harvard University Press, Cambridge

Gruenstein A, McGraw I, Sutherland A (2009) A self-transcribing speech corpus: collecting continuous speech with an online educational game. In: Proceedings of the speech and language technology in education (SLaTE) workshop. Warwickshire

Hayes AF, Krippendorff K (2007) Answering the call for a standard reliability measure for coding data. Commun Methods Meas 1:77–89. doi:10.1080/19312450709336664

Horton JJ (2010) Online labor markets. Springer Berlin Heidelberg, pp 515–522

Horton JJ (2011) The condition of the Turking class: are online employers fair and honest? Econ Lett 111(1):10–12

Horton JJ, Chilton LB (2010) The labor economics of paid crowdsourcing. In: Proceedings of the 11th ACM conference on electronic commerce, ACM, pp 209–218

Hosseini M, Cox I, Milić-Frayling N, Kazai G, Vinay V (2012) On aggregating labels from multiple crowd workers to infer relevance of documents. Adv Inf Retr 182–194

Hsieh G, Kraut RE, Hudson SE (2010) Why pay?: exploring how financial incentives are used for question & answer. In: Proceedings of the 28th international conference on human factors in computing systems, pp 305–314. doi: 10.1145/1753326.1753373

Hullman J, Adar E, Shah P (2011) The impact of social information on visual judgments. In: Proceedings of the 2011 annual conference on human factors in computing systems, ACM, New York, pp 1461–1470

Ipeirotis P (2010) Demographics of mechanical Turk. CeDER-10–01 working paper, New York University

Ipeirotis PG, Horton JJ (2011) The need for standardization in crowdsourcing. CHI

Jung HJ, Lease M (2011) Improving consensus accuracy via Z-score and weighted voting. In: Proceedings of the 3rd Human Computation Workshop (HCOMP) at AAAI Press, Menlo Park, California

Kahneman D (2011) Thinking, fast and slow. Farrar, Straus and Giroux, New York

Kapelner A, Chandler D (2010) Preventing satisficing in online surveys: a 'kapcha' to ensure higher quality data. In: The world's first conference on the future of distributed work, San Francisco, CA (CrowdConf2010)

Kaufmann N, Schulze T, Veit D (2011) More than fun and money. worker motivation in crowdsourcing–a study on mechanical turk. In: Proceedings of the seventeenth Americas conference on information systems, Detroit

Kazai G, Milic-Frayling N (2009) On the evaluation of the quality of relevance assessments collected through crowdsourcing. In: SIGIR 2009 workshop on the future of IR evaluation, Boston, MA, p 21

Kazai G, Kamps J, Milic-Frayling N (2012) The face of quality in crowdsourcing relevance labels: Demographics, personality and labeling accuracy. In: Proceedings of the 21st ACM international conference on Information and knowledge, ACM, New York, pp 2583–2586

Khanna S, Ratan A, Davis J, Thies W (2010) Evaluating and improving the usability of mechanical turk for low-income workers in India. In: Proceedings of the first ACM symposium on computing for development, ACM, New York, p 12

Kittur A, Chi EH, Suh, B (2008) Crowdsourcing user studies with mechanical turk. In Proceedings of the SIGCHI conference on human factors in computing systems, ACM, New York, pp 453–456

Kittur A, Nickerson J, Bernstein M, Gerber E, Shaw A, Zimmerman J, … Horton J (2013) The future of crowd work. In: Sixteenth ACM conference on Computer Supported Coooperative Work (CSCW 2013), Forthcoming

Krippendorff K (2004) Reliability in content analysis. Hum Commun Res, 30(3):411–433

Krosnick JA (2006) Response strategies for coping with the cognitive demands of attitude measures in surveys. Appl Cogn Psychol 5(3):213–236

Krug S (2009) Don't make me think: a common sense approach to web usability. New Riders, Berkeley, CA

Lakhani KR (2008) InnoCentive. com (A). Harvard Business School Case, 608–170

Lane I, Weibel A, Eck M, Rottmann K (2010) Tools for collecting speech corpora via Mechanical-Turk. In: Proceedings of the NAACL HLT 2010 workshop on creating speech and language data with Amazon's mechanical turk, Association for Computational Linguistics, Stroudsbug, PA, pp 184–187

Lau T, Drews C, Nichols J (2009) Interpreting written how-to instructions. In: Kitano H (ed) Proceedings of the 21st international joint conference on artificial intelligence, Morgan Kaufmann, San Francisco, pp 1433–1438

Li B, Liu Y, Agichtein E (2008) *CoCQA*: co-training over questions and answers with an application to predicting question subjectivity orientation. In: Proceedings of the 2008 conference on empirical methods in natural language processing. Association for Computational Linguistics, Stroudsburg. doi: 10.3115/1613715.1613836, pp 937–946

Lintott CJ, Schawinski K, Slosar A, Land K, Bamford S, Thomas D, … Vandenberg J (2008) Galaxy zoo: morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey. Mon Not R Astron Soc 389(3):1179–1189

Marge M, Banerjee S, Rudnicky AI (2010) Using the Amazon mechanical turk for transcription of spoken language. In: Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE international conference on (5270–5273), Institute of Electronics and Electrical Engineers, Washington, DC. doi:10.1109/ICASSP.2010.5494979

Mason W, Suri S (2012) Conducting behavioral research on Amazon's Mechanical Turk. Behav Res Methods 44(1):1–23

Mason W, Watts DJ (2009) Financial incentives and the performance of crowds. In: Proceedings of the ACM SIGKDD workshop on human computation, ACM, New York, pp 77–85

Molla D, Santiago-Martinez ME (2011) Development of a corpus for evidence based medicine summarisation. In: Proceedings of Australasian language technology association workshop, Australasian Language Technology Association, Melbourne, pp 86–94

Nelson L, Held C, Pirolli P, Hong L, Schiano D, Chi EH (2009) With a little help from my friends: examining the impact of social annotations in sensemaking tasks. In: Proceedings of the 27th international conference on human factors in computing systems, ACM, New York, pp 1795–1798. doi:10.1145/1518701.1518977

Nickerson JV, Sakamoto Y, Yu L (2011) Structures for creativity: the crowdsourcing of design. In: CHI workshop on crowdsourcing and human computation, pp 1–4

Open Science Collaboration (2013) The reproducibility project: a model of large-scale collaboration for empirical research on reproducibility. In: Stodden V, Leisch F, Peng R (eds) Implementing reproducible computational research (A Volume in The R Series). Taylor and Francis, New York

Paolacci G, Chandler J, Ipeirotis P (2010) Running experiments on Amazon mechanical turk. Judgm Decis Making 5:411–419

Pe'er E, Paolacci G, Chandler J, Mueller P (2012) Screening participants from previous studies on Amazon mechanical turk and qualtrics. Available at SSRN 2100631

Prelec D (2004) A bayesian truth serum for subjective data. Science 306(5695):462–466

Pontin J (2007) Artificial intelligence, with help from the humans. New York Times*, 25*. Retrieved from http://www.nytimes.com/2007/03/25/business/yourmoney/25Stream.html

Rand DG (2012) The promise of mechanical turk: how online labor markets can help theorists run behavioral experiments. J Theor Biol 299:172–179

Resnick P, Kuwabara K, Zeckhauser R, Friedman E (2000) Reputation systems. Commun ACM 43(12):45–48

Rochet JC, Tirole J (2003) Platform competition in two-sided markets. J Eur Econ Assoc 1(4):990–1029

Rogstadius J, Kostakos V, Kittur A, Smus B, Laredo J, Vukovic M (2011) An assessment of intrinsic and extrinsic motivation on task performance in crowdsourcing markets. In ICWSM

Sayeed AB, Rusk B, Petrov M, Nguyen HC, Meyer TJ, Weinberg A (2011) Crowdsourcing syntactic relatedness judgements for opinion mining in the study of information technology adoption. In: Proceedings of the 5th ACL-HLT workshop on language technology for cultural heritage, social sciences, and humanities, Association for Computational Linguistics, Stroudsburg, pp 69–77

Schwarz N (1999) Self-reports: how the questions shape the answers. Am psychol 54(2):93

Shapiro DN, Chandler J, Mueller PA (2013) Using mechanical turk to study clinical populations. Clin Psychol Sci 1:213–220
Shaw AD, Horton JJ, Chen DL (2011) Designing incentives for inexpert human raters. In: Proceedings of the ACM 2011 conference on computer supported cooperative work, ACM, New York, pp 275–284
Sheerman-Chase T, Ong EJ, Bowden R (2011) Cultural factors in the regression of non-verbal communication perception. In: 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops), Barcelona, Spain, pp 1242–1249
Silberman M, Irani L, Ross J (2010) Ethics and tactics of professional crowdwork. XRDS Crossroads ACM Mag Stud 17(2):39–43
Simon HA (1972) Theories of bounded rationality. Decis Organ 1:161–176
Suri S, Goldstein DG, Mason WA (2011) Honesty in an online labor market. In: von Ahn L, Ipeirotis PG (eds) Papers from the 2011 AAAI workshop. AAAI Press, Menlo Park
Tang W, Lease M (2011) Semi-supervised consensus labeling for crowdsourcing. In: Proceedings of the ACM SIGIR workshop on crowdsourcing for information retrieval, ACM, New York
Tetlock P (2005) Expert political judgment: how good is it? How can we know? Princeton University Press, Princeton
Tetreault JR, Filatova E, Chodorow M (2010) Rethinking grammatical error annotation and evaluation with the Amazon mechanical turk. In: Proceedings of the NAACL HLT 2010 fifth workshop on innovative use of NLP for building educational applications, Association for Computational Linguistics, pp 45–48
Tversky A, Kahneman D (1974) Judgment under uncertainty: heuristics and biases. Science 211(January):453–458
von Ahn L (2006) Games with a purpose. Computer 39(6):92–94
Von Ahn L, Maurer B, McMillen C, Abraham D, Blum M (2008) Recaptcha: human-based character recognition via web security measures. Science 321(5895):1465–1468
Wenger E (1998) Communities of practice: learning, meaning, and identity. Cambridge University Press, Cambridge
Winchester S (2004) The meaning of everything: The story of the Oxford English Dictionary. Oxford University Press
Yu L, Nickerson JV (2011) Cooks or cobblers?: crowd creativity through combination. In: Proceedings of the 2011 annual conference on human factors in computing systems, ACM, New York, pp 1393–1402
Zhou DX, Resnick P, Mei Q (2011) Classifying the political leaning of news articles and users from user votes. In: Proceedings of the fifth international AAAI conference on weblogs and social media. The AAAI Press, Menlo Park, pp 417–424

Designing Systems with Homo Ludens in the Loop

Markus Krause

Introduction

Digital games are a feasible option to provide a valuable benefit for getting into the loop[1] and taking part in a human computation (*HC*). Games can be used in a range of scenarios from acquiring common sense knowledge to finding problem solving strategies. How deeply a task can or should be merged into a game concept differs. Some tasks are more suitable as they have aspects that resemble game mechanics or games. However many tasks that do not seem to fit well can be shaped to be the basis for interesting game mechanics. Many projects already demonstrated how games can empower HC. *ESP* (Von Ahn and Dabbish 2004) the first HC game turns the quite boring task of labeling images into a successful game. It produced 1.3 million labels with around 13,000 players in a 4 month period. The game pairs two players over the internet. The game shows both players the same picture and lets them enter keywords that describe the content of that image. If both players agree on a keyword, they both score and the next picture is shown. One might argue that such a game is boring as it is very repetitive. Although that might be true even exceptionally successful games such as Farmville do use very simple and repetitive elements in their gameplay.

However, HC games can also be versatile. An example is *OnToGalaxy* (Krause et al. 2010) that integrates HC tasks such as ontology population into an action game. The player takes on the role of a space ship commander investigating interstellar space to rescue earth. Meanwhile, the player solves tasks that support artificial systems processing natural language. The game attracted around 500 players in the first 10 h of its release. But HC games can do more. They can support

[1] Moni Naor used the term humans in the loop in a position paper explaining the basic concepts of human computation (Naor 1996).

M. Krause (✉)
Leibniz University, Appelstrasse 9A, 30167 Hannover, Germany
e-mail: public.markus.krause@gmail.com

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_31,

computational systems in dealing with some of the most complex problems. *FoldIt* (Bonetta 2009) is a game that presents simplified three-dimensional protein chains to players, and provides a score according to the predicted quality of the folding done by the player. The human support significantly enhances the results and may aid in finding a cure for cancer or HIV. A game of similar complexity is *Phylo* (Kawrykow and Roumanis 2011). This game again solves a biological problem called multiple sequence alignment. From such alignments, it is possible to trace the source of certain genetic diseases. On the other hand playful or game-like elements introduce challenges and issues that cannot be ignored. As the basic idea of integrating tasks into games is to conceal their work character it introduces a certain distraction. Furthermore designing digital games is a field of research on its own. This chapter will provide a guideline to identify tasks that can benefit from ludic (game like or playful) elements. This guideline will investigate how to reshape a task to fit into a digital game. The chapter will address the challenge to ensure data quality and illustrate common pitfalls of systems with homo ludens[2] in the loop. To put all these in a meaningful and digestible context the chapter will propose a design process and illustrate all process phases along a small real world example.

Homo Ludens in the Loop

Most HC projects share a common structure. An entity called *requester* has a certain *task* to complete. The task is split into packages of a certain workload suitable for processing called *requests*. The implemented system distributes these requests to *contributors*. Contributors respond to these requests and the system aggregates these *responses* into *answers*. Sometimes more than a single response is needed to derive an answer. All answers accumulated constitute the *solution* for the initial task that is then read for further processing. More detailed models that describe HC systems can be found in this book as well as in various publications. Figure 1 illustrates this general model of a HC project as it is used in this chapter.

This chapter is dealing with the parts of a HC system that handles request distribution and aggregation of responses. Many HC projects use crowdsourcing as a method to acquire responses to their requests. At this point it is necessary to distinguish between crowdsourcing and HC and how both terms are used throughout this chapter. HC is a paradigm in which human mental abilities solve tasks not effectively or efficiently solvable by machines alone. Crowdsourcing is a method to outsource tasks to a group of contributors mostly over the Internet. Both terms have strong family resemblances and are often used synonymously but will be distinguished in this chapter.

[2] The term *homo ludens* (*Man the Player* or *Playing Man*) dates back to Dutch cultural theorist Johan Huizinga. He highlights the importance of play in the human quest for meaning in his book *Homo Ludens: A Study of the Play-Element in Culture* (Huizinga 1944).

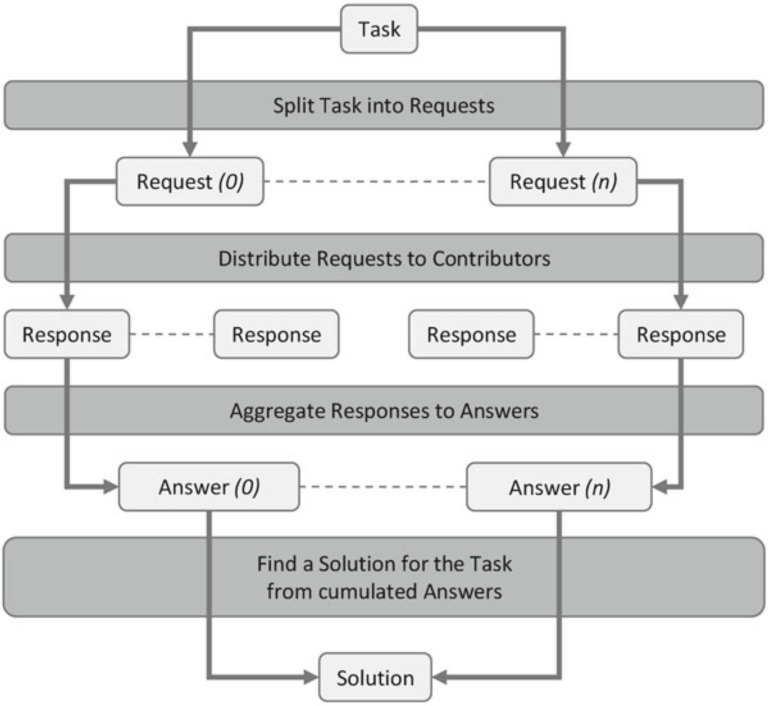


Fig. 1 General model of a problem solving human computation system

The primary perspective on HC systems is the perspective of the requester. The requester has a certain problem she wants to solve. The problem is formulated as a task expected to be solvable by a (human computation) system. This task has certain requirements such as the number of requests to be solved, the necessary answer quality to achieve, and so on. These requirements are formulated and define the systems implementation. The system is presented to contributors via a distribution channel and constitutes an overture towards the contributors. This overture is the basis on which a contributor decides to participate or not. Figure 2 shows this concept.

In paid HC this overture is independent from the system. The overture can be changed by raising the offered payment. Changing the overture of ludic systems is different as a change of the overture always affects the system and thereby may conflict with certain requirements. Another important aspect is that game elements do introduce sources of distraction. The pivot point when designing ludic systems for HC is therefore to balance requirements and overture.

This chapter will distinguish two general forms of systems with ludic elements. The first form of systems adds game elements to a task: these systems are designed with a focus on the requirements and do not alter or adapt them. The second form of systems merges a task into a digital game. Systems of this form are designed with a

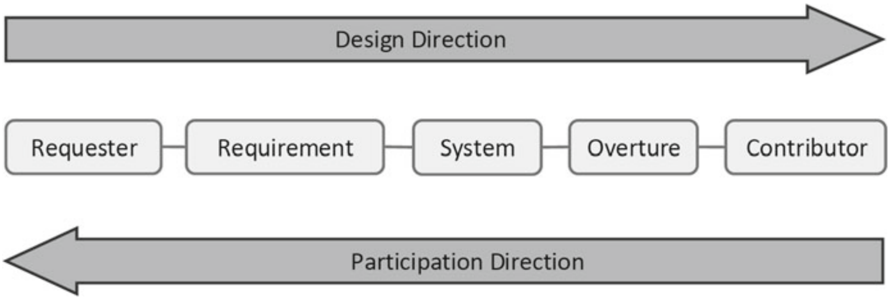


Fig. 2 General model of two perspectives on a HC system

focus on the overture and with the perspective of contributors in mind. These two forms define a range rather than two distinguished categories. Which form is suitable depends on the goals and task of a project.

Design Concept

Whether a HC system can benefit from ludic elements is hard to decide in the early stages of a project. However, many of the initial conceptual decisions can be shared between both forms of HC systems. Identifying tasks or subtasks is one common design aspect. Systems with humans in the loop also aggregate valuable data by observing interactions between humans and machines. Therefore a project needs a mapping between these interactions and the task. Finally both forms need methods to ensure data quality. When designing a system, involvement with these aspects should result in a set of requirements. These requirements in the case of a ludic system are the basis for the mechanics of the game.

Identification

Before precisely defining an HC task it is advisable to take a general perspective on tasks that are suitable for HC. By definition these tasks are either not effectively or efficiently solvable by a computational system alone. Yet many tasks that fall into this category are also complex for humans. Therefore it might be unintuitive which task is a good candidate for HC. There are however some categories of tasks that are in general suitable for HC.

A largely unsolved challenge for computational systems is human level perception of aesthetics, like judging the quality of motion, a sound, or an image. Humans are very good at interpreting various perceptions. Aesthetics in this regard means perception by means of the senses and judgment means the interpretation of these impressions. Different approaches like the systems of Talton (Talton et al. 2009) and Dawkins (Richard 1987) explore this field. They use human aesthetic judgment to create natural looking lightning of virtual environments or to model objects in two and three-dimensional space. Tasks of aesthetic judgment are especially challenging to evaluate by a computational system as they are most often subjective.

Combinatorial optimization tasks are also a common problem in computer science. Different approaches show that human mental abilities can outperform current computational systems. Humans are able to solve some of these problems in an intuitive manner and thereby overcome issues like local minimum/maximum traps (Corney et al. 2010). In contrast to an algorithm, which is based on the logical reasoning of its designer, intuition is the ability to gain insight into something; to form an opinion, or to find an ad-hoc solution; without a conscious reasoning process. HC systems such as *FoldIt* (Cooper et al. 2010), and *Phylo* (Kawrykow and Roumanis 2011) illustrate that.

From the perspective of a computational system the human world is full of ambiguities. Humans can shed light on some of these ambiguities. Examples for contextual reasoning and common sense are tasks such as resource annotation—like image or audio annotation—and natural language understanding. HC is applicable for various context related tasks. Prominent examples are image labeling (Von Ahn and Dabbish 2004), audio annotation (Barrington et al. 2009), as well as natural language understanding (Chamberlain et al. 2008).

Finally humans can easily act as agents in their physical environment. The ability of a computational system to manipulate the physical world is usually limited. Humans can easily interact with their physical environment. Examples utilizing human interaction with the physical environment are given by Matyas (Matyas et al. 2008) and Tuite (Tuite et al. 2010).

These general task patterns can help to identify potential candidate tasks for HC. The following list shows some requirements that are defined by the task and how these requirements can affect the design of the system and its ludic elements.

Number of Requests: Despite being relatively simple, labeling tasks require numerous requests, as seen in the image labeling *ESP* game (Von Ahn and Dabbish 2004). These systems most often parallelize their process as described by Little et al. (2010). When thousands or millions of tasks have to be processed a key issue is simplicity and clarity of the interaction design. This means that a contributor intuitively understands the task at hand. A strategy to simplify tasks can be to let contributors select answers from predefined sets, instead of formulating free responses. Examples for such a simplification are given by Dasdan et al. (2009) as well as Krause and Aras (2009). In some cases a computational system can automatically select candidate answers and verify them by asking a contributor. If such a preselection is not possible the task can be split into two sub-tasks. The first HC task then provides candidate answers and the second task ranks them (Aras et al. 2010).

The more requests a project needs to find a solution the more expensive it will be. This holds true for paid crowdsourcing as well as for systems solely relying on their ludic nature to attract contributors. Tasks with a vast amount of requests need a decent long-term motivation for contributors to be solved with a game. Tasks with a very small number of requests may not justify the overhead of a game unless the individual requests are very complex.

Complex Tasks: HC systems can also be used to solve tasks that necessitate strong commitment from their contributors, such as *Phylo* (Kawrykow and Roumanis 2011) or *FoldIt* (Bonetta 2009). These tasks are more challenging. Contributors need more concentration and are required to have a certain experience to solve them. These systems typically have fewer requests. In cases where huge amounts of complex tasks have to be processed a direct use of HC can be expensive. Initial training data for artificial systems can however be acquired with HC. Such systems can then handle these tasks more accurately than before. Various approaches in this direction were presented by (Brew and Greene 2010; Lease 2011; Quinn et al. 2010). Such tasks can benefit from ludic elements. Training of an artificial intelligence system is an interesting game mechanic itself. Games such as Black and White (2001) were based on training an artificial intelligence.

Time Frame: The time frame in which the project needs responses is also a key factor. Digital games have a certain lifecycle and are in many ways a consumable product. Games need a certain development time and this time is always longer than anticipated. Games once release need some time to attract player their count will afterwards reach a peak then degrade. On gaming portals this time frame is only some weeks or even days. Designing games with a real long term motivation is complex and the outcome is mostly unpredictable. When time is a crucial issue crowdsourcing the task might be more suitable.

Observing Player Behavior

HC systems generate useful data by observing human interactions with computational systems. Well thought out interaction design and a sound survey strategy can help to reduce error rates or unwanted behavior. The following list gives some hints on important requirements.

Doing a Task in Different Ways: Using different workflows for the same task can reduce error rates as described by (Lin et al. 2012). The paper explores how dynamic switching between workflows and therefore different interaction designs can improve data quality. Especially in ludic environments different possible strategies to solve a task can be an option to enhance user experience.

Secondary Variables: It is possible to use task independent data to detect unwanted behavior. Language evaluation for instance can take advantage of a language independent feature vector that contains values about user behavior to predict whether a user's input is reliable (Kilian et al. 2012).

Evaluating Data

By definition, a HC task is hard to solve by computational systems. As a result an answer given by a contributor is hard to evaluate with a computer. This is a special challenge for systems with homo ludens. As with normal games, digital games require players to carry out certain actions following sets of rules. These actions are mirrored by a meaningful change in the game world which can be called progress. In a HC game this progress involves giving contributors feedback about their responses. Play sessions of most HC games are relatively short. Due to the casual nature of these games the return rate of contributors is also low. Therefore HC games do need a strategy that allows giving instant feedback. Some possible methods are explained below.

Human based Evaluation: Various methods of user-centered evaluation strategies have been presented. A common approach is to pair contributors and only accept those answers they both can agree on. Instances for this approach are various (Von Ahn et al. 2006; Bernstein et al. 2009). Standard methods for human based evaluation are *Input-* and *Output-Agreement*. *Output-Agreement* games are a generalization of the *ESP* game. Two strangers are randomly chosen. In each round, both are given the same input and must produce outputs based on the input. Game instructions indicate that players should try to produce the same output as their partners. Players cannot see one another's outputs or communicate with one another. Both players must produce the same output. They do not have to produce this output at the same time but must produce it while the input is displayed onscreen (Law and Von Ahn 2009). In *Input-Agreement* games two players are shown either the same object or different objects and each is asked to type a description of their given object. Based on these descriptions, the players must decide whether they have been given the same object (Law and Von Ahn 2009). There are various other possible methods involving human judgment e.g. sequential evaluation. In such a scenario contributors responses are the requests for another human computation task. The task in the second HC cycle then evaluates the responses of the first. All methods involving human evaluation however limit the game design space as they delay the feedback to a contributor's response unless multiple players are online at the same time. Solutions to the timing problem are recording play sessions and let contributor play against these recordings.

Trust Metric: Another way of a human centered approach is calculating certain trust values for each contributor. These values are calculated based on the user responses to requests with known optimal responses which are interspersed to test the users' reliability. Examples can be found in different publications (Aras et al. 2010; Ipeirotis et al. 2010; Krause et al. 2010).

Algorithmic Evaluation: In some cases a computer is able to calculate the quality of a given response to a certain degree. For example: in language related tasks it is possible to check for known words and correct grammar as presented by Aras et al. (2010). In other cases, represented by *FoldIt* (Cooper et al. 2010) or *Phylo* (Kawrykow and Roumanis 2011), the quality of a response can be calculated precisely. Yet, due to combinatorial explosion it cannot be calculated in advance.

In some situations it might be possible to evaluate a given response but not in a suitable time frame. In such situations it can be a reasonable solution to use a less accurate algorithm for immediate feedback. A more precise version can afterwards be used for a final evaluation.

Motivation

When humans are part of a computational process motivation becomes an important topic. Not only because the HC system has to foster contributor engagement. It is also necessary to have the right type of motivation. Boldly tricking contributors with psychological manipulation to do work for free as Tom Sawyer did is not an acceptable method for scientific endeavors. The aim of this chapter is as said before to give hints about how well designed tasks can provide contributors with valuable experiences in exchange for their mental effort. Before introducing some common game mechanics for systems with homo ludens in the loop it is inevitable to present a short discussion of common pitfalls.

Mixing intrinsic and extrinsic Motivation: Mixing incentives is in general not an issue. Adding ludic elements to tasks on paid crowdlabor markets even enhances contributor performance as results from Toomim et al. indicate. Adding vast amount of incentive elements which are perceived as an external reward can have negative effects. Kohn (1999) describes in his work the negative effect of so called extrinsic motivation. Examples for such extrinsic motivations are achievements that are not related to the game or raffles. A more in-depth explanation of extrinsic and intrinsic motivation is given by Ryan and Deci (2000). Contributor will merely use the game as a tool rather than playing it. In such a case the player will most probably stop playing the game or need another source of motivation.

Overjustification: A connected common pitfall is overjustification. An excessive use of incentive elements can have a negative effect. For instance, if a child is constantly motivated by its parents to draw, often times this leads to a loss of interest over time. This effect of "over justification" was described by Lepper (Lepper et al. 1973) and can also occur while playing games (Kohn 1999). When player perceive badges or scores to be not an inherent result of their achievements these values will turn play into work. Games are played because of their mechanics if the incentives are outside of the inherent mechanics of the game they become pointless. Player will not play a game because of a score value or some badges. They play because of the path to their score and the meaning behind it.

Private Crowd: Building a private group of players from scratch is a tedious endeavor. There are various online gaming platforms where millions of players do crowd. Utilizing these platforms along other distribution channels such as *Facebook* is most often the optimal way. Game distribution platforms such as *Valves Steam*[3]

[3] http://store.steampowered.com/

are also interesting but have a much higher expectation on the quality of a game than casual online game portals. The same holds true for crowdsourcing in general finding a private crowd for an HC task is not advisable. An exception is when the system needs input from a small group of experts only.

Negative Motivation: Loss aversion is an efficient mechanism. The idea is to provide contributors with goods or benefits right from the start. The contributor can however lose these benefits if she behaves in a certain undesired way. Although efficient this mechanism is easy to misuse and can leave the player unsatisfied.

Engagement and Distraction: Adding ludic elements also adds sources for distraction. Games such as OnToGalaxy (Krause et al. 2010) have a valuable gaming experience but also distract player from focusing on requests. This is not a strong issue when more intuitive responses are welcome but can cause inaccurate answers in some cases. To spot such pitfalls early on play testing first prototypes is essential.

Mechanics

Game design in its entirety is an independent scientific discipline. A good starting point to investigate this field is (Crawford 1984; Salen and Zimmerman 2004). This chapter however aims at giving easy access to the field of HC with digital games. Therefore it will layout some mechanics that can provide a starting point for a project that wants to use a game or ludic elements. Mechanics can be understood as the formal rules of a game. These rules define possible actions players can take, the winning conditions, how rules are enforcement, etc.

Achievements: A reason games are intriguing is that actions within the game are reflected directly within the game world. Achievements are one example of this meaningful change in the game world. In real life good quality work is often expected but often not explicitly rewarded or honored. Badges are for instance a method to emphasize good work or exceptional deeds. Good options to give achievements are: reaching an accuracy rate, finishing an assessment, or solving an exceptionally hard task. Trophies are another form of achievement; for instance the first person that solves a request claims this request as a trophy. This is primarily interesting for systems that have requests that require a strong commitment and are complex to solve.

Bonuses: A bonus is a reward that works especially well for paid crowdsourcing. If a contributor does a large number of tasks with high accuracy a bonus is a good option to gratify her effort. However the system needs a well working method to ensure quality as otherwise bonuses can be an incentive for cheating. With inaccurate quality management it is also possible that good contributors can be denied a deserved bonus.

Collaboration: Especially for complex tasks collaboration is useful. Many tasks can benefit from sequential or even real-time collaboration. When more than one skill is needed to solve a task or subtask, making these connections visible can be a great enhancement. To integrate collaboration advisory systems are a good option.

Providing the opportunity for contributors to ask others for help is one way to allow collaboration. Helping other contributors can also be an incentive for more experienced players. This can even be combined with achievements were one contributor is granted the option to act as a teacher. Additionally social aspects yet alone can be a major motivational aspect for players.

Progress: Progress can be both a mechanism as well as a general concept. One reason that games are alluring is that they provide constant feedback on ones actions. The idea behind this concept is that every action of a player results in a meaningful and perceptible change in the game world. Simple examples are progress bars that illustrate improvement. Another example could be a project that aims at training a computational system for instance a search algorithm. This algorithm uses user generated ontology. While the contributor adds new relations to the ontology she can observe how search results change according to her input.

Meaning: It is very motivating to see how one's own actions have an impact. Projects that can benefit from this mechanism are projects that aim at solving a concrete problem. Examples are *FoldIt* and *Phylo* were contributor solve requests to serve a greater good.

Alignment: One argument for ludic elements as an incentive for HC is that games provide intrinsic motivation. This is only partially true. Game elements give a higher probability that contributors use the system for the sake of interacting with it thus "playing the game." This does not mean that contributors are intrinsically interested in the requirements or the task itself. To achieve true intrinsic interest a project has to align the requirements of the task with the desire of the contributor. An example is given by Krause et al. (2012). The approach uses a web based quiz for a human subject survey. The incentive is similar to those in quizzes or polls that provide an analysis of the contributor based on given responses. However intrinsic motivation can also introduce bias. Consider the following example in which contributors have to classify images as being appropriate for children. The intrinsic motivation of a very conservative parent might be an undesired influence on the outcome of the task.

Example

To illustrate the content of this chapter we have designed a human computation game. The game is called *Empathy* and aims at finding the most common labels for images. The game is very similar to *ESP* (Von Ahn and Dabbish 2004) and other *GWAP*[4] games. We chose this design as the *GWAP* games are a well known design and still relative complex to build. The game idea is inspired by the successful Family Feud game show. In this show two groups compete against each other.

[4] www.gwap.com

The goal of the game is to find the most popular responses to a survey question that was posted to a group of 100 individuals. Game shows are a good source of inspiration. They tend to have simple yet engaging mechanics and are designed to reach a broad group of people. Furthermore they often have tasks included that are already close to HC tasks. Designing a human computation game can follow the standard game design process. Some books that deal with game design in general are *Rules of Play* (Salen and Zimmerman 2004) and *The Art of Computer Game Design* (Crawford 1984). Most HC games are games with simple rules that do not require a lot of commitment. Such games are recently called casual games and there design is well explained in: *Casual Game Design: Designing Play for the Gamer in ALL of Us* (Trefry 2010). The standard game design process consists of three main phases. In the *pre-production* phase a brainstorming can give initial ideas. After the brainstorming a certain set of promising concepts should be tested with a paper prototype or other similar methods. As a rule of thumb around five prototypes are most often enough to find a working concept. More concepts can be hard to test in a reasonable amount of time. In the *production phase* the concept is implemented and tested with people that are not directly involved in implementing the game. Even games with a production timeline of some days or weeks should be tested frequently at least every time a feature is added. Games are defined by user experience therefore a designer should try to get as much of impressions on her game as possible before it is released. The final phase is the *post-production*. The reception of a game by its audience is never predictable. Therefore, after the game is released it is necessary to check player responses. This will also foster the design process of the next game. Besides handling game design issues it is also necessary to check gathered responses regularly as player will try to trick the system. Detecting such behavior early on can help to enhance evaluation methods in time.

For *Empathy* the concept of *Family Feud* was modified in various ways as some of the mechanics were not implementable in the desired time frame of not more than 2 days. As we expected the value of the collected data to be roughly equivalent to two working days. Therefore *Empathy* does not support interaction between groups of players. The game however implements the basic idea of finding the most frequent answer. The player sees an image and has to find a label that she thinks other players have chosen to tag the image. Figure 3 shows a screenshot of the concept art of the game. The game design itself is simple and easy to implement. The game provides different mechanics as described above. One mechanic it uses is *Achievements*. Players can earn badges for doing a certain amount of requests, finding words that are not already in the database, not using swear words, etc. The game also allows the player to experience *Progress*. When the player responded with an answer that is also the most frequent answer to the shown image the game will add this image to a list. This list is shown to the player at the right side of the game screen. Although, collaboration would have been an intuitive mechanic for the game idea it was not used, due to time constraints. The task of the game is relatively simple and each request just needs some seconds to be completed.

To estimate the quality of responses the system uses a variety of methods. It allows only one answer per player per request. This way no single contributor can

Fig. 3 Concept art for the *Empathy* game

pollute the database by doing the same image multiple times. The system uses *wordnet* (Miller 1995) to detect swear, harassment, and slang words. The system detects overly frequent usage of words for instance when a user types one word over and over again.

The term frequency is compared to the term frequency in the database. Furthermore the system detects semantic similarity between words again based on *wordnet*. This allows scoring responses from a contributor more flexible than with a comparison on string level. Also integrated is a spellchecker that identifies misspelled words and also allows detecting random strings. The game includes these methods to give the contributor feedback on their actions. This constant feedback is a key element for fraud detection and prevention. However the system does not force the player into a certain behavior, it just ignores their answers and reports back that a certain word seems to be improper. The system also calculates a trust value for every response based on the measured statistics and the response history of the contributor.

To distribute the game it was released on two online gaming platforms namely *kongregate*[5] and *newgrounds*.[6] Both platforms allow for publishing flash and html5 games through their website. It is possible to reach an even broader audience through social networks such as *facebook*.

[5] www.kongregate.com

[6] www.newgrounds.com

Fig. 4 Screenshot of the game *Empathy*

Experiment

To demonstrate the power of the presented idea we conducted a small experiment to compare the costs and data quality of both approaches. We implemented two different versions of the game idea explained above. Both games were submitted to the same two online gaming platforms within 3 weeks. The design time for both games was roughly 2 days not including the time to implement the server side technology and fraud detection system. Both components are available online.[7] The first game (*Empathy*) does not use any fraud detection mechanism. Figure 4 shows an image of the game.

The second version (*GuessIt*) uses the fraud detection methods as explained earlier. A screenshot is depicted in Fig. 5. Additionally the task was submitted to an online crowdsourcing platform in two different versions. The first version (*CF1*) was submitted only to contributors from the US as proposed by the crowdsourcing provider. The second version (*CF2*) was published to all available contributors. Both versions used a set of images with a set of acceptable answers to identify untruthful behavior. These acceptable answers are hand coded labels for an image. These data is called *gold standard data*.

[7] https://code.google.com/p/gamelab/

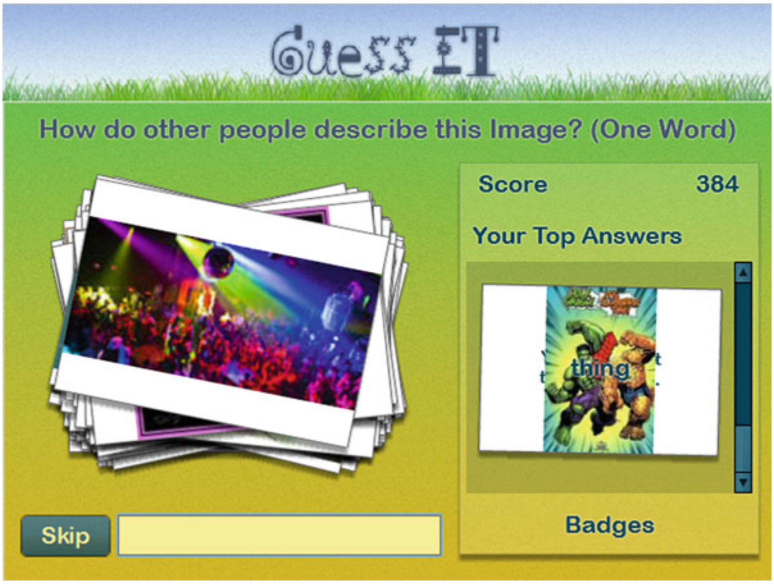


Fig. 5 Screenshot of the *GuessIt* game

Results

Both games were published to two gaming platforms: *kongregate* and *newgrounds*. These platforms allow player to rank games (higher values are better). The *Empathy* game was ranked with 2.42 out of 5 on *kongregate* and 2.72 out of 5 on *newgrounds*. It gathered a total of ~13,000 responses for ~3,600 images from 272 contributors in the first 4 days of its release and was played for 156 h. The game was developed in 2 days including the server side implementation. The average response quality of *Empathy* is low. The responses include swear words, spelling errors, and other undesired artifacts. One contributor submitted the same swear word 1,338 times. To analyze the response quality we hand coded 500 responses collected with *Empathy* to either be acceptable or not. Only 69.6 % of the collected responses are acceptable.

GuessIt was ranked 2.53 on *kongregate* and 2.73 on *newgrounds*. These scores seem to be relatively low however the mean ratings for other games at the same time on *kongregate* (2.47) and *newgrounds* (2.86) are comparable. *GuessIt* gathered ~14,000 responses for the same ~3,600 images from 529 contributors in the first 3 days of its release. The game was derived from *Empathy* but provided feedback for the contributor about the quality of individual responses. The development time was roughly one and a half day. Not including the development time for the used server side fraud detection methods. The average response quality was much higher. 98.8 % of the responses collected are acceptable. To calculate the average response quality we hand coded a random sample of 500 responses from the responses collected with *GuessIt*. The responses still contained swear words and spelling errors

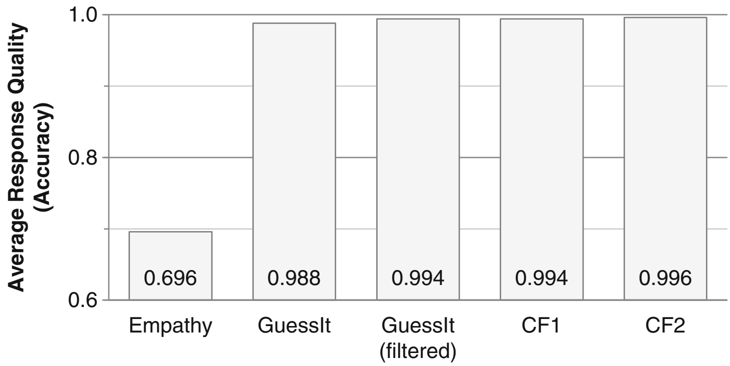


Fig. 6 Comparison of average response qualities for different experiments. *CF1*and *CF2* are results from paid crowdsourcing. *GuessIt* filtered represents the quality considering responses with high trust values only

but only 6 out of the 500 are unacceptable! Only three responses with improper content have a positive trust value and therefore are included into the final database. This gives *GuessIt* a final accuracy of 99.4 %.

The reference test *CF1* contained ~500 images from the initial ~3,600. For each image four responses were requested. The experiment was set up via Crowdflower and published to Mechanical Turk. Only contributors from the US were allowed to participate. The experiment collected 3,209 responses (2,260 trusted) from 86 contributors in a period of 8 days. More information on the completion time of HC tasks on Mechanical Turk can be found in Wang et al. (2011). The price for this task was $37.02. The accuracy was very high and no intentionally wrong answers were found. There were only some spelling errors. Without prior experience it took half a day to design the task and prepare the gold data.

The reference test *CF2* contained the same ~500 images from the initial ~3,600. For each image four responses were requested. The experiment was set up via Crowdflower and published to Mechanical Turk. Contributors were not restricted to come from the US. The task was cloned from *CF1* and took only half an hour to set up. The job collected 3,240 responses (2001 trusted) from 74 contributors. The average response quality was again very high. The price for the task was also $37.68. Figure 6 shows a comparison of average response qualities from all experiments.

Conclusion

A recurrent challenge for HC systems is motivation. Contributors support human computation projects for many reasons. When developing human computation systems a key challenge is to offer a valuable reward for contributors. This chapter illustrated how to design HC systems with ludic elements to make their use an inherently pleasurable experience. The chapter described concepts, methods, and pitfalls of systems with homo ludens in the loop. It demonstrated that a system designed along the described principles is able to:

- Attract players on online gaming platforms
- Produce data of high quality similar to paid crowdsourcing
- Does not need intensive game design skills
- Can be used more than once with only minor modifications

Using digital games to motivate contributors is reasonable. Games allow for great flexibility in design and use of evaluation methods not possible with paid crowdsourcing.

References

Von Ahn L, Dabbish L (2004) Labeling images with a computer game. In: CHI ’04 proceedings of the 22nd international conference on human factors in computing systems. ACM Press, Vienna, 319–326

Von Ahn L, Kedia M, Blum M (2006) Verbosity: a game for collecting common-sense facts. In: CHI ’06 proceedings of the 24th international conference on human factors in computing systems, ACM Press, Montréal, p 78

Aras H, Krause M, Haller A, Malaka R (2010) Webpardy: harvesting QA by HC. In: HComp’10 proceedings of the ACM SIGKDD workshop on human computation, ACM Press, Washington, DC, p 4.

Barrington L, O’Malley D, Turnbull D, Lanckriet G (2009) User-centered design of a social game to tag music. In: HComp’09 proceedings of the ACM SIGKDD workshop on human computation, ACM Press, Paris, pp 7–10

Bernstein M, Tan D, Smith G, Czerwinski M, Horvitz E (2009) Collabio: a game for annotating people within social networks. In: UIST ’09: proceedings of the 22nd annual ACM symposium on user interface software and technology, ACM Press, Victoria, pp 97–100

Bonetta L (2009) New citizens for the life sciences. Cell 138(6):1043–1045

Brew A, Greene D (2010) The interaction between supervised learning and crowdsourcing. In NIPS’10 Computational Social Science and the Wisdom of Crowds (pp. 1–4). Vancouver, British Columbia, Canada.

Chamberlain J, Poesio M, Kruschwitz U (2008) Phrase detectives: a web-based collaborative annotation game. In: I-Semantics’08 proceedings of the international conference on semantic systems, ACM Press, Graz

Cooper S, Khatib F, Treuille A, Barbero J, Lee J, Beenen M, Leaver-Fay A, Baker D, Popović Z, Players F (2010) Predicting protein structures with a multiplayer online game. Nature 466(7307):756–760

Corney JR, Kowalska I, Jagadeesan AP, Lyn A, Medellin H, Regli W (2010) CrowdSourcing Human Problem Solving Strategy. In CrowdConf’10 Proceedings of the 1st International Conference on Crowdsourcing. San Francisco, CA, USA

Crawford C (1984) The art of computer game design. McGraw-Hill/Osborne Media, Berkeley

Dasdan A, Drome C, Kolay S (2009) Thumbs-up: a game for playing to rank search results. In: WWW ‘09 proceedings of the 18th international conference on world wide web, ACM Press, New York, pp 1071–1072

Huizinga J (1944) Homo ludens: a study of the play-element in culture, 1st edn. Taylor & Francis, London

Ipeirotis PG, Provost F, Wang J (2010) Quality management on Amazon mechanical Turk. In: HComp’10 proceedings of the ACM SIGKDD workshop on human computation, ACM Press, Washington, DC, pp 0–3

Kawrykow A, Roumanis G (2011) Phylo—a human computing framework for comparative genomics

Kilian N, Krause M, Runge N, Smeddinck J (2012) Predicting Crowd-based Translation Quality with Language-independent Feature Vectors. In HComp'12 Proceedings of the AAAI Workshop on Human Computation. Toronto, ON, Canada: AAAI Press

Kohn A (1999) Punished by Rewards: The Trouble with Gold Stars, Incentive Plans, A's, Praise, and Other Bribes (2nd ed., p. 448). Boston, MA, USA: Mariner Books

Krause M, Aras H (2009) Playful tagging: folksonomy generation using online games. In: WWW '09 proceedings of the 18th international conference on world wide web, ACM Press, Madrid, pp 1207–1208

Krause M, Takhtamysheva A, Wittstock M, Malaka R (2010) Frontiers of a paradigm—exploring human computation with digital games. In: HComp'10 proceedings of the ACM SIGKDD

Krause M, Smeddinck J, Takhtamysheva A, Markov V, Runge N (2012) Playful Surveys: Easing Challenges of Human Subject Research with Online Crowds Challenges of Human Subject Research with. In HComp'12 Proceedings of the AAAI Workshop on Human Computation. Toronto, ON, Canada

Law E, von Ahn L (2009) Input-agreement: a new mechanism for collecting data using human computation games. In: CHI '09 proceedings of the 27th international conference on human factors in computing systems. ACM Press, Boston, pp 1197–1206

Lease M (2011) On quality control and machine learning in crowdsourcing. In: HComp'11 proceedings of the AAAI workshop on human computation, pp 97–102

Lepper MR, Greene D, Nisbett RE (1973) Undermining children's intrinsic interest with extrinsic reward: a test of the 'Overjustification' hypothesis. J Personal Soc Psychol 28:129–137

Lin CH, Mausam Weld DS (2012) Crowdsourcing: Dynamically Switching between Synergistic Workflows. In HComp'12 Proceedings of the AAAI Workshop on Human Computation. Toronto, ON, Canada: AAAI Press

Lionhead Studios (2001). Black and White. Electronic Arts

Little G, Chilton LB, Goldman M, Miller RC (2010) Exploring iterative and parallel human computation processes. In: Proceedings of the 28th of the international conference extended abstracts on human factors in computing systems—CHI EA'10, 4309, ACM Press, New York, doi:10.1145/1753846.1754145

Matyas S, Matyas C, Schlieder C (2008) Designing a location-based game for the collection of geospatial data: the cityExplorer case study. In: ACE '08: proceedings of the international conference on advances in computer entertainment technology. ACM Press, Yokohama, pp 244–247

Miller GA (1995) WordNet: a lexical database for English. Commun ACM 38(11):39–41

Naor M (1996) Verification of a human in the loop or identification via the Turing test. Rehovot

Quinn AJ, Bederson BB, Yeh T, Lin J (2010) Crowdflow: integrating machine learning with mechanical Turk for speed-cost-quality flexibility

Richard D (1987) The blind watchmaker: why the evidence of evolution reveals a universe without design. W.W. Norton & Company, New York

Ryan R, Deci E (2000) Intrinsic and extrinsic motivations: classic definitions and new directions. Contemp Educ Psychol 25(1):54–67. doi:10.1006/ceps.1999.1020, http://www.ncbi.nlm.nih.gov/pubmed/10620381

Salen K, Zimmerman E (2004) Rules of play. MIT Press, Cambridge

Talton JO, Gibson D, Yang L, Hanrahan P, Koltun V (2009) Exploratory modeling with collaborative design spaces. Proc ACM SIGGRAPH Asia 28(5):1–10

Trefry G (2010) Casual game design: designing play for the gamer in ALL of Us. Morgan Kaufmann Publishers

Tuite K, Snavely N, Hsiao D-Y, Smith AM, Popović Z (2010) Reconstructing the world in 3D: bringing games with a purpose outdoors. In: FDG'10 proceedings of the fifth international conference on the foundations of digital games. ACM Press, Monterey, pp 232–239

Wang J, Faridani S, Ipeirotis P (2011) Estimating the completion time of crowdsourced tasks using survival analysis models. In WSDM 2011 Workshop on Crowdsourcing for Search and Data Mining (CSDM 2011). Hong Kong, China: ACM Press

Human-Computer Interaction Issues in Human Computation

Stuart Reeves

What Is Human-Computer Interaction and How Does It Relate to Human Computation?

Human-computer interaction explores the construction of novel interactive systems (hardware and software), the evaluation and study of interactive systems in use, and the construction of theoretical understandings of those evaluations. Of course, this is a narrow view of HCI and does not fully account for its relationship to other disciplines which have a role within or relationship to it, such as art and design, or software engineering. For the purposes of this chapter, however, it will suffice.

Human computation has become an application domain for HCI, often in the context of crowdsourcing systems. Given that interactions between human and machine form the foundations of human computation, the fit between the two is natural. Most prominent examples of this relationship stem from early work by von Ahn and Dabbish (2004), which synthesised von Ahn's cryptography research with Dabbish's work on collaborative systems and computer-supported cooperative work (CSCW). Through this, von Ahn and Dabbish produced influential work that brought human computation concerns onto the HCI agenda, particularly through various demonstrations of interactive 'games with a purpose' (von Ahn and Dabbish 2008). The canonical example within HCI is the 'ESP Game' (von Ahn and Dabbish 2004), in which paired players attempt to match descriptive tags for images, resulting in the rapid collection of human-constructed annotations for large numbers of images as a 'byproduct' of human-computer interactions. This work has developed

S. Reeves (✉)
School of Computer Science, University of Nottingham, Nottingham, UK
e-mail: stuart@tropic.org.uk

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_32,

into a large literature concerned with developing and evaluating (new and existing) interactivity in human computation systems, ranging from web or desktop-based 'games with a purpose' (such as the ESP Game), to citizen science applications (e.g., Galaxy Zoo[1]), to mobile systems (e.g., Bell et al. (2009)).

Contributions to HCI

The key contributions of this body of work to HCI is the development of a novel interaction technique, i.e., solving hard computational tasks with human action, as well as an exploration of various associated HCI issues that inform design. The notion of 'interactive human computation systems' as an interaction technique inverts the more usual and familiar HCI relationship between humans and machines, in which machines are the computational actors. In the inverted technique, *humans are seen as computational nodes* or components within a human-computer assembly, as opposed to a more common HCI perspective which seeks to understand how *machine computational resources* come to feature and be employed within human-human and human-computer interactivity. Thus, within human computation literatures, human activity has been seen as potentially offering vast resources of computation for solving hard computational tasks. Interactivity is a further key part: the knowledge that in theory, interactive computation provides a greater computational power than non-interactive algorithmic systems (Wegner 1997) supports this notion. The exploration of this unusual configuration of human and machine has necessarily resulted in its particular associated HCI issues being explored. Generally within the human computation literature published at HCI venues this has tended to focus on how to design interactive human computation systems which are correct (i.e., "producing the correct answer in the presence of noise") and efficient (Law and von Ahn 2011) in terms of 'quality control', or managing issues of 'cheating' or 'gaming the system' through input and output agreement systems (Law and von Ahn 2009) (e.g., as in the ESP Game's matching of pairs of players (von Ahn and Dabbish 2004)). This literature has also addressed issues of motivating human 'components' of human computation systems (e.g., Bell et al. (2009); Reeves and Sherwood (2010)), as well as how human and machine contributions can be organised, both in terms of workflows and aggregation strategies (Reeves and Sherwood 2010; Quinn and Bederson 2011).

HCI Challenges to Human Computation

However, as Quinn and Bederson argue, "human computation has a tendency to represent workers as faceless computational resources", for instance, not considering "issues related to ethics and labor standards" (Quinn and Bederson 2011).

[1] http://www.galaxyzoo.org

This is in many ways atavistic towards HCI's strong promotion of user-centred design and focus on user experience as a foundational element of system design. Prior work (e.g., Reeves and Sherwood (2010)) has also critiqued the notion of conceptualising humans as "processing nodes for problems that computers cannot yet solve" (von Ahn and Dabbish 2008) or as a "remote server rackspace" of "distributed human brainpower".

The tenor of this argument reflects wider historical forces in HCI; the human computation approach within HCI largely reflects what could be seen as a 'traditional' HCI approach, bound by normative forms of design (e.g., there is no sense of participatory design processes where users co-design the system) and evaluation (e.g., formal evaluation techniques such as usability testing). Thus to a great extent, human computation as it features within HCI has remained generally unaffected by the significance of the 'turn to the social' that impacted HCI during the 90s (Button and Dourish 1996; Bannon 1991; Grudin 1990b) and helped bring about a revolution in how we conceptualise 'the user'. Briefly put, this 'turn' in HCI involved a move beyond individualist cognitive formulations of the 'user' to social conception of the 'user' and greater consideration of the importance of coordination and collaboration amongst groups of users.

The rest of this chapter focuses on the implications of this shift for human computation and how it might conceive of these 'processing nodes'. In order to address this we must now turn to consider the 'cognitive turn' as well as the 'social turn' in HCI (section "Cognitive and Social 'Turns' in HCI: Conceptualising 'the User'") before examining its implications for human computation itself (section "Directions for Developing HCI in Human Computation").

Cognitive and Social 'Turns' in HCI: Conceptualising 'the User'

HCI initially emerged from a convergence between computer science and psychological, cognitive and social psychological models of interaction (Dourish 2006). Of the cognitive and social 'turns' in HCI, we find their traces most prominently within the evaluative traditions and practices of HCI. There is a significant body of literature within HCI concerned with developing the methods and perspectives with which to conduct evaluation of computer interfaces in use. This ranges widely in purpose, from evaluations concerned with an individual's task efficiency and interface usability (e.g., see 'GOMS' as mentioned below) for work/productivity applications to ethnographic evaluations of user experience of artistic performances (e.g., Reeves (2011)). The range of these various methods and perspectives has increased with the growing spread of digital technologies and their attendant interfaces into ever more aspects of our everyday lives.

Early approaches at the start of the 1980s for evaluating the usability of computer interfaces were derived largely from human factors and cognitive psychology, which considered both the perceptual qualities of interface elements (e.g., ergonomics) as well as the cognitive processes which users engage in when interacting with machines. This perspective informed a range of evaluative practices concerned with examining task performance and its relationship to user interface design. Cognitive conceptions such as the human processor model (Card et al. 1983) have provided the basis of task decomposition techniques, such as GOMS (Goals, Operators, Methods, Selection) and its variants (see John (1995)), that offer predictive task performance indicators for expert users engaged in limited tasks, such as data entry and so on.

Key to these early approaches to conceptualising 'the user' in HCI is in constituting the human element as an individual, delimited by the descriptions of cognitive psychology, with individual capabilities described in terms of motor, sensor, memory and computational processing capacities. Models of cognitive faculties explain the possibility of human action in the world by theorising inner conceptual/mental representations constructed by the human to represent systems that exist 'out there' in the world. The human consults these internal representations as a resource when interacting with the world. In this cognitive approach as articulated in HCI, 'the user' has particular goals and subgoals which, so arranged, provide a plan of action to bring about an overall goal such as 'write a letter' or 'send a text message' (e.g., see Card et al. (1983)). Through this planning view, the cognitive approach seeks to model and therefore predict the 'human factors' in interactions between human and machine, the reasoning being that through this, designers may themselves be able to systematically explore a design space to find optimal solutions to interface design problems.

However, the advent of a different, socially-oriented approach, drawing particularly on the social sciences, did, towards the end of the 1980s and the start of the 1990s, begin to challenge this dominant individualist cognitive perspective within HCI in a number of ways (Bannon 1992). Not only was the model of 'the user' transformed, but so was understanding of the role of the technological artefact. As Grudin argues, this shift developed into a more holistic view of interaction, exploding the typical HCI definitions of the interface to situate both the technology and 'the user' into complex socio-technical constellations, and instead reveal a role for HCI in the design of this itself *as* the interface (Grudin 1990a). Beyond this, recognition of this growing importance of understanding the social features of interaction were equally found in developments of psychological approaches, such as the emergence of Distributed Cognition theory (Hutchins 1995) and its application in an HCI context (Rogers 1994).

Underlying Perspectival Shifts in HCI: Phenomenology and Workplace Studies

A key influence in this perspectival shift occurring in HCI was the instrumental effect of a range of workplace ethnographies which unpacked the character of coordination and collaboration with, around and through interactive technologies. Put simply, the individual was no longer a relevant unit of analysis. Instead, as Heath et al. describe, the term ‘collaboration’ provided a useful “gloss to capture a complex configuration of momentary arrangements through which two or more individuals, sequentially or simultaneously participate in particular tasks or activities” (Heath et al. 1995). Many of these ethnographies of workplace technology were driven by an underlying orientation towards sociological phenomenology such as symbolic interactionism, conversation analysis and ethnomethodology (e.g., see Szymanski and Whalen (2011)). In contrast with cognitivist accounts, which derive from a Cartesian perspective of mind-body dualism, the phenomenological perspective gives primacy to ‘subjective’ experience; in phenomenological sociology this matter is transformed into investigation of ‘intersubjectivity’, that is, developing theoretical and empirical understandings of how seemingly incommensurable ‘subjective’ individual experiences are *negotiated* such that individuals may engage in concerted social actions (such as, in Heath et al. (1995), organising market trades in a dealing room).

Drawing upon this philosophical background, these allied approaches have thus formed part[2] of the reorientation in HCI towards considering the phenomenology of interaction, i.e., the nature of ‘user experience’ (rather than, say, the ‘information processing’ capacities of the user). In terms of their contribution to understanding how interactive technologies are experienced in the ‘lifeworlds’ of users, this body of (mostly) ethnographic work has unpacked the ways in which interaction with and around technologies is a fundamentally *socially organised* phenomenon. That is, they detail just how meaning is actively produced, achieved, maintained and repaired by participants in those interactions. This stands in contrast with a traditional cognitive view that would ascribe meaning in terms of input/output to/of an individual’s cognitive workings. For instance, Heath et al. (1995) explore how careful verbal and bodily (e.g., gestural) conduct is employed to sensitively produce moments of collaboration in order to make coordinated decisions regarding bidding for stocks within the trading room. In this way the meaning of a given trade does not reside in an individual’s mental representation, but is *produced* through a social orientation to ongoing collaborative action.

Directions for Developing HCI in Human Computation

Now that the broad outline of the cognitive and social turns in HCI have been discussed, this section explores in more detail how cognitivist ideas have found a natural home in some conceptualisations of human computation systems. As part of this, the following also unpacks what the implications of HCI’s ‘social turn’ might be for human computation.

[2] Obviously there are other influences on HCI which have shaped the ascendancy of ‘user experience’ as a core concern, however these are beyond the scope of this chapter.

Cognitive Alignments Between HCI and Human Computation

There is a strong similarity between a cognitive conception of the human in HCI and the standard human computation role of humans. Firstly we begin with the term itself, i.e., 'human computation': it is here that the computation model being applied to human action is initially ascribed. In this sense the term itself configures the field with certain assumptions about the nature of this human action, i.e., that it is readily characterisable in terms of computation. Secondly, the explicit ways in which human elements are described in the literature confirms this view. For instance, humans have been characterised as serving as information processors, or computational nodes (e.g., as before, von Ahn and Dabbish (2008)), as well as definitional forms of "human computation algorithms" being derived from Donald Knuth's computational ones (Law and von Ahn 2011). Building upon this perspective, there has been a focus within human computation on game theoretic accounts of human agency, such as in the design of questions or in order to incentivise/motivate users (e.g., (Jain and Parkes 2009; Law and von Ahn 2011, p. 61)). Traditional game theory models rely upon a computational view of human agency (e.g., that human agency involves rational calculation of outcomes), and a transcendent understanding of rational action. This is as opposed to a situated, local view of rational action in which order and meaning is locally produced (Suchman 1987).

It is no coincidence that the cognitivist approach emerged across a range of disciplines (e.g., psychology, linguistics, computer science, neuroscience, etc.) in parallel with the development of digital computing during the 1950s: the computer was seen as providing a suitable and appropriate metaphor for developing understandings of the human. The broad appeal of this metaphor cannot be underestimated; computational metaphors have been a driving force in the development of theoretical models across a range of disciplines including biology, linguistics, anthropology, physics and art (Cantwell Smith 2010). Yet metaphors can sometimes prove problematic, in that they may distort the nature of phenomena as well as directing focus away from their nature in favour of the simplifications afforded by the metaphor.

What HCI's Social Turn Means for Human Computation

With the 'social turn' in HCI, critiques of this cognitive, computational metaphor view have flourished. A key text here is Suchman's influential work that argues against a cognitivist, plan-based model of human action, instead transforming the rational plan into a resource which may be drawn upon in the situated, moment-by-moment mundane actions of humans who are ongoingly achieving the construction of social order (Suchman 1987). While in Suchman's case the analysis was of experts using photocopiers to perform basic tasks, for human computation, this radically changes how we conceive of the 'node' in human computation systems. Instead of 'information processors' manipulating data orchestrated by digital computer management we must see humans in these systems as accomplishing social order: developing intersubjective or shared understandings in and through organising their physical and verbal actions moment-by-moment, designing and crafting those interactions so as to be intelligible and meaningful to others, and engaging in ad-hoc but coherent and concerted social actions with one another. In this view the cognitive notions of goals and plans are the construction of the analyst imposed onto innate social order rather than an underlying, transcendent theory of human action. Similarly, the utility of analogies between algorithms and "human computation algorithms", considering time complexity, efficiency and correctness (Law and von Ahn 2011), can potentially obscure considerable design differences (Reeves and Sherwood 2010).

Some work in human computation has begun dismantling the conception of humans as processing 'nodes' in computation networks by studying the situated ways in which meaning is produced through interactions between users engaging in human computation tasks. For instance, in 'Eyespy', we developed a mobile human computation game in order to produce sets of photos which were useable for navigation tasks as a byproduct of that play (Bell et al. 2009). Like the ESP Game, we relied upon human competencies in order to construct a high quality data set (in our case, of 'good' navigational images, as opposed to 'good' textual tags for images). Players of the game gained points for creating photo tags of landmarks which other players subsequently attempted to locate and visually confirm (based on GPS proximity), in turn gaining points themselves. Successful players oriented their in-game actions towards designing photographs that leveraged local knowledge (of 'good' landmarks), 'findability' and how recognisable they were. These human competencies relied upon commonsense knowledge, i.e., 'what anyone knows' about a given geographic area, and what would constitute a 'good photo' for other players (see von Ahn et al. (2006) for an attempt to collect a generalised set of such commonsense knowledge). It is precisely this notion of how players oriented towards each other, produced their content in ways that were crafted as appropriate to the framing of the game and the prospective recipients of their photos. In short, players' actions are not algorithmic but *interpretive* and *socially organised* within the human computation system, contradicting characterisations of computational nodes or cognitive information processors in which interpretation and social action is part of internal mental processes rather than an accomplished negotiation between humans.

Human Computation System Design

This view of human computation developed as part of studies within HCI has three key messages for designers to consider when constructing the next interactive human computation system (Reeves and Sherwood 2010). They impact two (of five) foundational questions of human computation suggested by Law and von Ahn (2011): firstly, how to guarantee solutions are accurate, efficient and economical; and secondly, how to motivate human components in their participation, expertises and interests.

1. The broadest point is the importance of user experience. HCI's lessons, via a focus on user experience and its 'turn to the social' mean that *the human perspective should be considered a foundational issue* to inform the design and construction of interactive human computation systems. Echoing Quinn and Bederson (2011), once again, this means human issues need to be considered as the initial step rather than something to evaluate post-hoc (e.g., "issues related to ethics and labor standards", also see Irani and Silberman (2013)).
2. *Meaning is situated and locally produced*. This does not mean that human computation systems cannot produce generalised results or reusable products, however it does mean that such things are not readily analogous with machine-based algorithms or necessarily aligned with cognitive descriptions of human agency. Instead, when we consider how (for instance) image tags are designed in the ESP Game, we should view this as the coordinated production of meaning between players rather than input and output transactions.
3. How the human computation system is approached and experienced by its human participants *fundamentally frames their interaction with it*. Therefore the products of those actions cannot be separated from the social and situated circumstances in which it was produced (see above). This matter of framing is a key design feature, for instance, the way a task delegated to users is introduced and the relationship that is configured between them and designers shapes the way in which that task is carried out (Brown et al. 2009). In other words, human computation tasks do not get performed in isolation. Instead, the seemingly secondary features of interface and task design (e.g., tutorials, what type of task it is communicated as, such as for money (e.g., Mechanical Turk) or scientific progress (e.g., Galaxy Zoo)) can radically change how the human components of computation systems act.

References

von Ahn L, Dabbish L (2004) Labeling images with a computer game. In: Proceedings CHI '04, pp 319–326

von Ahn L, Dabbish L (2008) General techniques for designing games with a purpose. Commun ACM 2008:58–67

von Ahn L, Kedia M, Blum M (2006) Verbosity: a game for collecting common-sense facts. In: Proceedings CHI '06, pp 75–78

Bannon L (1991) From human factors to human actors: the role of psychology and human–computer interaction studies in systems design. In: Greenbaum J, Kyng M (eds) Design at work: cooperative design of computer systems. Lawrence Erlbaum Associates, Hillsdale, pp 25–44

Bannon LJ (1992) Perspectives on CSCW: from HCI and CMC to CSCW. In: Proceedings of international conference on human-computer interaction (EW-HCI), pp 148–158

Bell M, Reeves S, Brown B, Sherwood S, MacMillan D, Chalmers M, Ferguson J (2009) Eyespy: supporting navigation through play. In: Proceedings of SIGCHI conference on human factors in computing systems (CHI), ACM Press, New York, pp 123–132

Button G, Dourish P (1996) Technomethodology: paradoxes and possibilities. In: Proceedings of the SIGCHI conference on Human factors in computing systems, pp 19–26

Cantwell Smith B (2010) Age of significance (Introduction). http://www.ageofsignificance.org/aos/en/aos-v1c0.html. Accessed 15 Apr 2013
Card S, Moran T, Newell A (1983) The psychology of human-computer interaction, CRC Press
Dourish P (2006) Implications for design. In: Proceedings of ACM conference on human factors in computing systems (CHI), ACM, New York, pp 541–550
Grudin J (1990a) Interface. In: Proceedings of the 1990 ACM conference on computer supported cooperative work (CSCW), ACM, New York, pp 269–278
Grudin J (1990b) The computer reaches out: the historical continuity of interface design. In: Chew JC, Whiteside J (eds) Proceedings of the SIGCHI conference on human factors in computing systems (CHI '90), ACM, New York, pp 261–268
Heath C, Jirotka M, Luff P, Hindmarsh J (1995) Unpacking collaboration: the interactional organisation of trading in a city dealing room. Comput Support Coop Work 3:147–165
Hutchins E (1995) Cognition in the wild. MIT Press, Cambridge
Irani L, Silberman MS (2013) Turkopticon: interrupting worker invisibility in Amazon mechanical Turk. In: Proceedings of CHI 2013, ACM, New York, 28 Apr–2 May 2013
Jain S, Parkes DC (2009) The role of game theory in human computation systems. In: KDD workshop on human computation, pp 58–61
John BE (1995) Why GOMS? Interactions 2(4):80–89
Law E, von Ahn L (2009) Input-agreement: a new mechanism for collecting data using human computation games. In: Proceedings of the 27th international conference on human factors in computing systems, (CHI) ACM, New York, pp 1197–1206
Law E, von Ahn L (2011) Human computation. Morgan & Claypool Synthesis Lectures on Artificial Intelligence and Machine Learning
Quinn AJ, Bederson B (2011) Human computation: a survey and taxonomy of a growing field. In: Proceedings of the SIGCHI conference on human factors in computing systems (CHI '11), ACM, New York, pp 1403–1412
Reeves S (2011) Designing interfaces in public settings: understanding the role of the spectator in human-computer interaction. Springer
Reeves S, Sherwood S (2010) Five design challenges for human computation. In: NordiCHI '10: Proceedings of the 6th Nordic conference on human-computer interaction, ACM, New York, pp 383–392
Rogers Y, Ellis J (1994) Distributed cognition: an alternative framework for analysing and explaining collaborative working. J Inf Technol 9(2):119–128
Suchman L (1987) Plans and situated actions: the problem of human-machine communication. Cambridge University Press, New York
Syzmanski MH, Whalen J (eds) (2011) Making work visible: ethnographically grounded case studies of work practice. Cambridge University Press
Wegner P (1997) Why interaction is more powerful than algorithms. Commun ACM 40(5):80–91

Collective Action and Human Computation

From Crowd-Workers to Social Collectives

Jasminko Novak

Introduction

The notion of human computation describes a class of approaches to distributed problem solving that integrate human contributions with computational techniques within computational systems. They leverage human cognitive capabilities to solve problems that are easy for human users but difficult for purely computational techniques (von Ahn 2006; Quinn and Bederson 2011). Typical examples of such problems involve the digitization, semantic analysis and classification of multimedia content for search and information retrieval (von Ahn and Dabbish, 2004; von Ahn et al. 2008; Chen et al. 2009). While this remains a major class of human computation applications, recent work has addressed a number of different areas: from common-sense facts collection (Lieberman et al. 2007), to text editing and composition (Bernstein et al. 2010; Kittur et al. 2011), language translation (Hu et al. 2011) or environmental monitoring (Dutta et al. 2009).[1]

Though such applications pursue a number of variations in implementing the idea of human computation (e.g. in specific task designs, participants recruitment or solution aggregation) they tend to share two underlying assumptions: they adhere to the question "How can human users help solve semantically complex problems?" and they adopt the perspective of individual users in designing the crowd-tasks, incentive structures or process workflows.

The individual user perspective means that the problem-solving process is conceived as a structured, system-initated workflow in which a large number of individual users contribute solutions to atomic primitives (micro-tasks) that may represent elements of a larger task. The solutions of individual users (e.g. pair-wise

[1] See (Quinn and Bederson 2011; Fraternalli et al. 2012) for an extensive overview.

J. Novak (✉)
University of Applied Sciences Stralsund/European Institute for Participatory Media,
Wilhelmstr. 67, 10117 Berlin, Germany
e-mail: j.novak@eipcm.org

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_33, © Springer Science+Business Media New York 2013

comparison of identity of two images, tagging of a piece of content or transcription of a word) are considered independently of each other and processed by statistical means to produce aggregate solutions (e.g. majority voting). In such a process, the users are considered as a part of a computational process operating without any awareness of each other's work process, individual results or the broader task context. While such a conceptual model has proven well-suited for specific classes of tasks, it basically mimics the tayloristic approach to work modeling and organization that considers the synergies of work contributions of many individuals in a rather mechanistic manner (Nagar 2011). Accordingly, the majority of existing approaches provides little or no support for direct communication and collaboration between the users and models the user participation as private exchanges between the task-owner and the task-solver.

In contrast, experiences from the large body of knowledge on collaborative problem solving and collaborative knowledge production point to the importance of group interaction and communication in different kinds of collaborative social formations (e.g. online communities, social networks) (Preece 2000; Brown and Duguid 2000; Kittur and Kraut 2008; Woolley et al. 2010). They highlight the role of voluntary, open group participation where individual contributions form a collective good freely accessible and benefiting all users (Kittur and Kraut 2008; Wasko and Teigland 2002). They also show how the very nature of the context of open collaboration can also enable new modes of collective production, leading to otherwise unattainable solutions (though posing a number of challenges for effective coordination and operationalization).

This begs the question of how such more open, participatory models of *collective action* can inform the development of new kinds of human computation systems and approaches: Can we conceptualize specific classes of human computation as instances of different forms of social collaboration? How can we design human computation systems where the involvement of a large number of human users as providers, aggregators or "processors" of information leads to outcomes that benefit the entire collective rather than only individual contributors and task owners? How can the theory of collective action inform the design of such collaborative approaches to human computation? Addressing this little investigated perspective on human computation is the goal of this chapter.

Reframing Human Computation

Human Computation and Self-Interested Individuals

Existing models of human computation tend to share several important characteristics: (1) they focus on exploiting the capabilities of human users to solving semantically complex problems within a *computational approach*, (2) the problem solving paradigm is based on the decomposition of complex tasks into atomic units (microtasks) that are trivial to solve for human users regardless of their background and level of expertise,[2] (3) their realization is based on voluntary participation of relatively large numbers of users over the Internet (crowdsourcing) and (4) they treat users largely as a mass of isolated individuals rather than as a distributed collective. While such models have proven to work very well for a number of application classes (see overview in Quinn and Bederson 2011) they also pose a number of limitations, some of which are starting to be questioned (Nagar 2011). The focus on decomposing complex problems in micro-tasks requiring little cognitive effort (e.g. image tagging, pairwise comparisons, transcripts of images or text snippets), obscures the challenges of tackling more complex classes of tasks that cannot be easily broken down to trivial primitives: *intrinsically creative* processes such as authorship cannot be effectively broken down to linear models of simpler tasks such as crowdsourced summarizing proposed in (Kittur et al. 2011) and neither can ill-structured, *wicked problems* (Rittel and Webber 1973; Star 1989) which are most apt to benefit from collective contributions.

Similarly, the conception of users as a mass of isolated, self-interested individuals obscures the large body of knowledge about the potential and importance of group communication and interaction for collaborative solving of complex problems and knowledge production (e.g. Wikipedia, online communities). Finally, the limitations of the individual perspective are also reflected in corresponding incentive designs that are largely extrinsic. They tend to focus on the individual user (payment per solved micro-task, player incentives in games-with-a-purpose) and influence both process and tasks designs at their very core (e.g. no designs for direct communication between the users, game like structures that hardly mask the underlying tasks as a reason of their existence). The focus on the individual perspective has also been reflected in models of ownership of the problem and of the results produced: typically the problem owners (e.g. businesses, public institutions, research organizations), own both the problems and the solutions while the benefits to the crowd *workers* are contained to monetary renumeration (micro-payments) or game satisfaction (games with a purpose). Mediators such as crowdsourcing platforms or service providers (e.g. Amazon Mechanical Turk, CrowdFlower, Microtask.com) provide the systems and tools for implementing the process workflows, managing user participation and performing the solution evaluation and aggregation.

Collaborative Approaches to Human Computation

At the same time, a number of explorations into applying the human computation *metaphor* to more complex classes of tasks and collaborative processes have started to recognize the importance of group interaction and collaboration in collective

[2] Though possibilities for solving more complex tasks have been investigated, these approaches still follow the decomposition principle (Kittur et al. 2011; Dean and Ghemawat 2008) and accommodate some means of qualification tests for selecting workers with appropriate skill levels.

intelligizing (Nagar 2011). Examples include collaborative information gathering and structuring in decision-making (Greene et al. 2010) and information management (Tungare et al. 2010), social business process management in organizations (Brambilla et al. 2011) or the cooperative management of environmental resources (Fraternali et al. 2012). While retaining the idea of combining human and machine intelligence, such approaches are starting to re-interpret the conceptual model of human computation more freely, relaxing a strict distinction between human computation and crowdsourcing (proposed in (Quinn and Bederson 2011)) and opening up to broader models of collaboration and collective intelligence. Crowdsourcing can thereby be seen as a specific mechanism of task distribution and participant recruitment through an open call to a large group of individuals (Howe 2006). Furthermore, instead of using centralized platforms for publishing tasks and acquiring participants, increasing attention is being given to using existing online social networks that already gather huge numbers of participants into webs of intricate social relationships (Bozzon et al. 2012).

This work is starting to build bridges between human computation, crowdsourcing and more general classes of techniques and applications for collaborative problem solving, information sharing and collaborative knowledge production. In doing so it acknowledges that some of the most successful approaches to harnessing collective intelligence involve models in which groups of participants engage in collaborative exchanges producing online information goods that are available for use and benefits to the entire collective—and not only to individual task owners or platform providers. Classical examples of such *online collective action* (Bimber et al. 2005) include information sharing and collaborative knowledge production through simple mailing lists, online communities and networks of practice (Wasko and Faraj 2005), structured question and answer systems (Ackerman and McDonald 1996; Bian et al. 2008) or community-based expertise location (Ackerman et al. 2002). More recent examples also involve the applications of crowdsourcing to solving pressing social problems through collective information gathering, idea generation and argumentative deliberation for e.g. addressing climate change (Klein and Iandoli 2008), participatory city management (Novak and Preuße 2011, http://www.thecity2.org/, http://opencities.net) and local community development (http://www.mindmixer.com/, http://www.nexthamburg.de/). Yet other involve human users as information providers in different forms of human-in-the-loop systems e.g. for aiding crisis disaster and management (Okolloh 2009; Goolsby 2010; Meier 2013), or crowdsourcing data collection and annotation for collective knowledge production in citizen science (Kanefsky et al. 2001; Bonney et al. 2009; Luther et al. 2009).

Though exhibiting a number of specific differences, all these examples share a basic underlying goal and conceptual premise: the involvement of a large number of diverse human users as providers, aggregators or "processors" of information and knowledge should lead to outcomes that benefit the entire collective rather than only individual contributors or commissioners of work assignments. However, there is still little understanding of how we can effectively conceptualize and design such human computation systems that support collaborative engagement of participants in the production of *public goods* benefiting entire social collectives. How can we identify a conceptual framework and specific design patterns that can help us in designing for collective action as a specific class of human computation? To address this question, the next sections consider related findings from the *theories of collective action* and show how they can inform the conceptualization and design of systems for collaborative human computation.

Online Collective Action

Collective action theories describe the behaviour of participants in collaborative problem solving and other forms of voluntary cooperation towards the creation of public goods (Marwell and Oliver 1993; Fulk et al. 1996). Public goods are shared resources from which all individual members of a collective can benefit even if they have not contributed to creating the good (e.g. a public park or public television) (Kollock 1998). Their creation, provision and/or maintenance require voluntary contributions of at least some individuals from the collective.

Public goods exhibit several special characteristics: (1) they are non-excludable, meaning that no individual can be excluded from the use of the public good, regardless of whether s/he contributed to its creation or not (Head 1962 referenced in Wasko and Teigland 2004), (2) the cost of providing such goods does not change with the number of people benefiting from it (Olson 1965) and (3) they are non-rival, meaning that their consumption by one person doesn't reduce the amount available to others (Hardin 1982 in Bimber et al. 2005). This makes it easy for individuals to free-ride on the efforts of others, since they can enjoy the benefits of using the public good without having to contribute to its creation. But practice shows that collective action occurs in spite of this assumed free-rider problem.

Accordingly, theories of collective action have focused on understanding why, how and under what circumstances self-interested individuals still engage into voluntary collective action and contribute efforts to the production and/or maintenance of public goods. This makes such theories of great relevance to providing a theoretical foundation for conceptualizing the major issues, problems and opportunities for collaborative models of human computation. While originally having considered material goods (Olson 1965), the theories of public goods and collective action have also been applied to voluntary cooperation in online settings, such as knowledge exchanges in online communities and social networks (Wasko and Teigland 2002; Wasko and Faraj 2005), online activism and collaborative creation of shared knowledge bases (Fulk et al. 1996; Bimber et al. 2005). This work highlights a number of factors influencing the dynamics of collective action involving self-interested individuals, which are related to similar problems in human computation. The next section reviews the major findings of relevance in this context.

Collective Action in Online Knowledge Exchanges

Online collective action occurs frequently through the exchange of messages between unfamiliar individuals that use web-based tools (forums, wikis, discussion boards, social network services) to share problems, information, knowledge and experiences about their professional practice or shared interests. While different forms of such social collectives are distinguished with different terms online communities, communities of practice, networks of practice (Brown and Duguid 2000) they all share the basic underlying mechanisms of the creation of a collective good. Similar mechanisms apply to the cooperative construction of shared knowledge bases, where individual members provide different pieces of information or knowledge into a shared repository (e.g. Wikipedia, collaborative tagging). And such a model also underlies other examples where knowledge is only a vehicle to creating new artifacts, products or services such as open source software or open product design and implementation (e.g. quirky.com).

Thus, the online exchange of knowledge in community networks can be considered as an archetypical model for understanding the mechanisms of online collective action. The posting of and responding to messages between the members of the network (e.g. online community, network of practice) represents a form of collective action that creates a publicly available body of knowledge accessible for all members of the network (Wasko and Teigland 2002). Since messages are publicly visible to all interested members the resulting body of knowledge is non-excludable i.e. can be accessed by all members regardless of their contributions. The cost of creating a message does not change with the number of members that may benefit from its content and the use of knowledge contained in a message by an individual member does not diminish its value for others.

Such conditions lead to the well-known paradox of individual vs. collective benefits: while there is a natural tendency to free-ride i.e. capture the benefits of using the public good (reading helpful responses) without own contributions, if everybody did so the public good would not be created at all. This is a core challenge for harnessing collective action and is also found in human computation: both approaches employing crowds composed of isolated individuals as well as approaches exploring more collaborative settings need to address this question.

Predictors and Mechanisms of Collective Action

A number of different explanations have been proposed to account for this paradox and explain why individuals contribute and engage in collective action in spite of the possibilities for free-riding (see (Wasko and Teigland 2004) for a detailed overview). One line of explanations proposes that the *heterogeneity of the resources and interests of the participants* increases the chances for the collective good to be realized, since there may be individuals with more resources (money, time, expertise, energy or influence (Oliver et al. 1985)) and higher interest in the public good who will create and maintain it for the benefit of the collective. This line of reasoning is in line with considerations of the role of diversity in increasing the effectiveness of collective intelligence in decision-making (Bonabeau 2009). Thereby, individual interest in contributing may be motivated by a number of different factors, such as the level of professional expertise and organizational tenure (which were found to increase likeliness of providing useful advice in online networks (Constant et al. 1996 in Wasko and Teigland 2002)). Similarly, group size has been found a good predictor of collective action, explained with the assumption that larger groups have automatically more potential resources for action (Scott and El-Assal 1969 in Wasko and Teigland 2002).

At the same time, theoretical contributions have argued that sustaining collective action in larger groups is in fact more difficult since individual contributions are more likely not to be noticed or be perceived as unnecessary, resulting in free-riding (Hardin 1982). This discrepancy may be explained in situations when the cost of the contributions for providing the good changes with the number of users: if the costs rise with the number of users, larger groups will be less likely to exhibit collective action than small ones; otherwise free-riders are not a burden for others and thus won't influence the success of collective action (Macy 1990 in Wasko and Teigland 2002).

Other work suggests that dense networks consisting of *direct social ties* between all participants may facilitate collective action, since a large amount of direct ties furthers the propensity to cooperation. However, other forms of patterns of social interactions may also suffice or be the determining force sustaining collective action. *Reciprocal gift exchanges* may sustain cooperation by building on the expectation that given help will be reciprocated in the future (Kollock 1999). As may *generalized exchanges* where one's contribution is not returned directly by its recipient but by a third party (Ekeh 1974) (e.g. one's questions are not answered by members to whom one had replied himself but by others). Such generalized exchanges based on indirect reciprocity or interest-driven contributions have been empirically verified as a determining factor for specific networks of practice (Wasko and Teigland 2002; Wasko and Faraj 2005).

Contributions may also be asymmetrically skewed with a small subset of the collective providing the most contributions ensuring the production of the collective good—a well-known phenomena from online communities and collective intelligence applications (Preece and Shneiderman 2009). In such cases, rather than merely identifying the large portion of members as free-riders, the notion of *critical mass* (Oliver et al. 1985) suggests that there may exist well-functioning classes of collective action that are dependent on a large enough number of individuals who provide the most contributions and significantly deviate from the average. This is also in line with empirical observations of the asymmetrical contributions in online information and knowledge exchanges where small groups of highly active members account for most contributions (see Preece and Shneiderman 2009 for an overview).

The presence of such critical mass can be verified empirically by examining the extent to which social ties are centralized i.e. concentrated to a small number of individuals (Wasko and Teigland 2002). This suggests that social network analysis may be used to identify networked collectives that may be particularly suitable for developing specific classes of approaches (e.g. engaging only a subset of members characterized by high degrees of community contributions) or for identifying key participants for acquiring critical mass of respondents (e.g. by using out-degree or centrality measures). Finally, affective factors such as trust and social capital (Putnam 2000; Nahapiet and Ghoshal 1998) can positively influence collective action. The assumption of these approaches is that the development of such strong, positive social relationships between members of the collective and with respect to the collective as a whole can suppress free-riding and incentivize active contributions. Similarly, social norms as standards of acceptable collective behaviour have been identified as effective means for preventing free-riding and furthering collective action in repeated interactions (Putnam 2000) much like explicit sanctions for non-compliance.

"Me-Centric" Participation and "We-Centric" Collaboration

Another perspective on the dynamics of online collective action that can inform the development of models integrating collective action into human computation is the dichotomy between "me-centric" participation (Koch 2008) and "we-centric" collaboration discussed in (Novak 2009). In collaboration research, user participation in communities and networks of practice has been largely conceptualized from the "we-perspective" of communities. Active participation and member contributions are motivated by intrinsic, often group-oriented or altruistic motives such as community citizenship, enjoyment of social interaction, reciprocity and reputation (Tedjamulia et al. 2005). The defining characteristic of communities is their self-organization and autonomous constitution of social norms, acceptable behaviors and uses of community resources (Preece 2000). The collective knowledge built up through member participation is considered a collective property of the community and a "public good" of its own, freely available for use and consumption (at least within community confines) (Kollock 1999; Preece 2000; Wasko and Faraj 2000).

Such a conceptualization of online collective action in which group interaction is necessary to create and maintain the public good is partly aligned with the theories of collective action. As noted in the previous section, the collective action theories model users as self-interested individuals who *in spite* of this provide contributions to the collective good. In fact, for the majority of participants in online communities, the benefits of such goods reside largely in personal usefulness of contributions created by others. The greatest proportion of community users are passive participants who consume community information or services (e.g. finding answers to their needs in a discussion forum). The so-called "lurkers" typically amount to 80–90 % of community members (Tedjamulia et al. 2005) and are attracted by extrinsic motivation: the prospect of easily accessible, credible information, highly relevant to their needs (Nonnecke and Preece 2001).

When aiming to integrate collective action into human computation, the discrepancy between intrinsic and altruistic "we-based" motivation of user engagement and the "me-centric" perspective of extrinsically motivated participation (e.g. monetary rewarded micro-tasks) induced by external task owners may pose problems with respect to acceptable modes of use and exploitation of the community's "public good". This may damage the fundamental motivational mechanisms needed for the collective good to emerge in the first place. For example, the community literature makes a strong point about the challenge of establishing company-sponsored communities, initiated and actively facilitated by commercial actors with the goal of exploiting the community activity for organizational or commercial purposes (Hagel and Armstrong 1997; Tedjamulia et al. 2005). While companies address this by offering extrinsic benefits accrued from the commercial exploitation (e.g. profit sharing, gifts, reputation (Tedjamulia et al. 2005)), empirical research in motivation theory suggests that such extrinsic incentives tend to undermine intrinsic motivation (the "crowding out" effect (Frey and Jegen 2002)). On the other hand, insights from motivation research itself show that extrinsic incentives may also reinforce intrinsic motivation (Frey and Jegen 2002). This suggests that the integration of community-based collaboration into traditional models of human computation cannot rely exclusively on the motivational mechanisms of the "we-centric" group perspective. Moreover, existing models of crowdsourcing and human computation are already largely based on "me-centric" individual participation not requiring group collaboration. An integration of "me-centric" participation (Koch 2008) with "we-centric" collaboration may thus allow us to identify models and design guidelines for supporting the integration of collective action in human computation.

The Principal-Agent Problem in Human-Computation

As suggested in (Novak 2009) the change of focus from "we-centric" group perspective with an implied sense of togetherness and pursuit of shared purpose, to individually motivated "me-centric" perspective can cause cooperation problems, in particular in customer-producer relationships. Such relationships are at the heart of current approaches to human computation, where task-owners exploit the work of crowd-participants for their own purpose. This relates human computation to the well-known principal-agent theory describing transactions between self-interested parties with differing goals in uncertainty conditions, characterized by information asymmetry and opportunistic behavior (see overview in Pavlou et al. 2007).

The uncertainty resides thereby in two main sources: the incongruence in goals and the inability of customers to monitor the producers' behavior. This leads to two main problems: (1) *adverse selection* where a customer (principal) selects a producer (agent) with inappropriate quality due to the inability to assess the producer's true characteristics (e.g. caused by possible misrepresentation by the producer) and

(2) *moral hazard* where the producer (agent) pursues goals not in the interest of the customer (principal), such as reducing the required effort and quality of delivered results to increase his profits (Pavlou et al. 2007).

A principal-agent relationship can be directly observed in traditional "individualistic" human computation, but more strikingly, it can also be applied to collective action. In the former case, the principals are the task-owners who entrust the task solution to the (mass of) individual crowd members representing the agents. The value proposition of task owners for motivating crowd participation is the promise of monetary rewards (micro-task platforms) or an entertaining experience (games with a purpose). Obtaining this requires active user contribution in the problem solving process. Though based on voluntary engagement and not sanctioned with contractual relationships, such models suggest a principal-agent relationship normally found in more formalized arrangements (Novak 2009). The task-owners or work commissioners (principals) delegate the responsibility for solving elements of the task to crowd participants (agents) who act on their behalf. Task-owners want to get high quality solutions for as little money as possible, whereas crowd workers want to invest as little time as possible in the solutions and receive as much money as possible. Since they have different interests and goals as well as different levels of information about each other (e.g. regarding the true capability of crowd workers to solve a given task or the appropriate amount of monetary compensation for solving a task), they engage in transactions between self-interested parties under uncertainty conditions.

The problem of adverse selection (due to hidden information on behalf of the workers) now corresponds to the problem of worker selection (assessing the suitability of the workers for a given task). The problem of moral hazard (due to hidden action by the workers) corresponds to the problem of free riding and unsatisfactory performance induced by extrinsic incentives, i.e. the workers not delivering promised performance by trying to game the system (e.g. entering random solutions to minimize time and effort for task solution). Hidden information (no appropriate access to information on true worker qualifications) and hidden action (difficulty in identifying and sanctioning actual worker actions) are typically mitigated through signaling, screening, monitoring and self-selection (Pavlou et al. 2007).

In signaling, the agent (worker) explicitly communicates his characteristics to the principal (task owner) in trust inducing ways (e.g. quality guarantee certificates). In screening, the principal engages actively in obtaining additional information about agent characteristics, e.g. through performance tests (qualification tasks) or assessment information from third-parties (e.g. reputation or levels of contribution in social communities). Monitoring can be applied by e.g. tracing the changes in the quality of task solutions of a crowd-worker over time. Self-selection can be implemented by enticing crowd-workers' self-assessment of qualification for specific tasks or by enticing their own selection of tasks based on what they truly believe to be able to solve effectively (e.g. by digressive remuneration based on observed quality of workers' solutions for specific task classes).

Moral hazard associated with hidden action can also be mitigated through signals and incentive systems, bonding, behavior and performance monitoring (e.g. in training tasks in worker selection). Reporting systems can reduce the information asymmetry and allow the principal to better assess agent behaviour. Examples include process tracking (e.g. visualization dashboards providing snapshots of the worker desktop such as in odesk.com) or reputation systems where past principals evaluate the agent's performance (e.g. eBay). While some of these mechanisms can and are being implemented in individualistic crowds, others require the incorporation in social collectives. The use of explicit and implicit reputation information for screening (worker selection) and prevention of moral hazard (free riding, manipulation, spamming) requires integration with collective reputation systems and social networks from which such information can be obtained.

The discussion in the previous sections has shown the need of extending the individualistic model of human computation with more open, participatory models supporting group interaction and communication. The analysis of the theories of collective action and the principal agent framework has pointed out how we can relate them to problems of collaboration in human computation. The next section introduces a conceptual design framework demonstrating how the described concepts and mechanisms can be applied to designing concrete systems integrating collective action in human computation.

A Conceptual Design Framework for Collaborative Human Computation

In individualistic approaches to human computation the crowd participants are treated as a mass of isolated individuals. The computational system mediates only a basic form of their implicit interaction by implementing some mechanism of aggregation of individual task solutions (e.g. majority voting). This section introduces a model of collaborative human computation that enhances the human computation process with more cooperative human-to-human interaction. A structurally similar approach has been successfully applied in to the domain of interactive value creation: there, the involvement of user knowledge in the design of new products mediated through single-user interaction with a computer system (e.g. user-innovation toolkits) has been extended with a more cooperative process of human-to-human interaction (Novak 2009). We build on this work in developing a conceptual model for collaborative human computation exploiting different modalities of collective action.

Human Computation and Expert-Based Crowdsourcing

A special kind of human computation applications that is inherently related to collective action is expert-based crowdsourcing. This class of applications addresses a situation where the crowd-task cannot be solved without domain knowledge and/or the problem itself cannot be effectively broken down into atomic micro-tasks (e.g. due to the nature of the problem or the nature of the dataset defining the tasks). An illustrative example of such a problem context is high-level semantic structuring and contextualization of historical multimedia collections depicting people, places and events of historical relevance, such as the one used by historians and researchers of European integration.[3] As reported in (Harloff 2012; Dionisio et al. 2013), in order to support high-level semantic queries and entity-based knowledge discovery (e.g. "Who is this person?", "What venue is depicted in this photo ?", "Who were the critical actors in the negotiation of the Maastricht Treaty?") a simple extension of automatic methods with general purpose crowds doesn't lead to satisfactory results. This is due both to the heterogeneity of the collection (e.g. photos of persons from different time periods, different scene settings, different content quality) (Dionisio et al. 2013) as well as to only partial availability of metadata and the importance of domain knowledge.

Due to the high heterogeneity of historical photo collections depicting both well-known and unfamiliar individuals in different historical contexts, resolving the identity (and the semantic context) of persons depicted in such images turns out to be a daunting task—both for computational methods and the domain experts. For example, realizing an effective person identification using a face recognizer requires a large enough set of already identified reference portraits—both for automatic and for crowd-based comparisons of unknown faces with the reference portraits. But the creation of such a reference set turns out to be a challenge even for the domain experts: establishing the identity of persons in a historical photo collection amounts to the research task itself for which the experts need the system support in the first place (Harloff 2012; Dionisio et al. 2013). Moreover, given only partial availability of appropriate meta-data, the inference of further contextual information required by the users (e.g. the venue depicted in the image, historical and social context) makes this a particularly hard problem for current computational approaches. Accordingly, in such cases the classical application of human computation involving general-purpose crowds to verify and extend the results of automatic identification is not feasible. Rather, what is needed is a model for involving the expert-based crowd that can provide the required domain knowledge in the task solution process (e.g. identification of unknown persons in the photos, related places, events and social contexts). But this is also a dynamic research process whose goals and needs evolve within daily practice of the users (Dionisio et al. 2013) resulting in dynamically evolving task types and definitions.

Designing a human computation system that involves a crowd of experts into solving such tasks poses a number of challenges regarding the task design, incentive mechanisms and performance requirements. Experts (such as in the aforementioned context) cannot be motivated by symbolic monetary remunerations typically offered in micro-task crowdsourcing schemes. They typically work under time constraints and pursue goal-oriented work. Thus it is also unlikely to imagine a pure "gaming experience" motivation within games-with-a-purpose scenarios. Posing trivial

[3] http://www.cvce.eu

micro-tasks (such as pairwise comparisons of an unresolved person face with an identified reference portrait) is also bound to fail in engaging them into solving the tasks. Moreover, such atomic micro-tasks require no expert knowledge by definition. In order to profit from the domain knowledge, the tasks need to be posed at a higher level of complexity, such that they are both cognitively engaging and can effectively take advantage of the experts' domain knowledge. The performance of the crowd-system i.e. the speed of receiving solutions to their problems also plays a major role for the users—while general purpose crowds tend to be applied in an offline mode to pre-structure the results, the above scenario frequently requires near real-time solutions.

One possible way to address this is by conceptualizing the expert crowd not as a mass of unrelated individuals solving meaningless tasks but as a networked collective in which individual members exchange information in order to solve shared problems. Individual tasks and solutions (e.g. determining the identity of an unknown person in a photograph, the depicted venue or the type of social relations between the depicted persons) now correspond to the question-and-answer messages exchanged between individual participants. This relates the problem of embedding expert-based crowdsourcing into human computation to the problem of online collective action.

Applying the lens of collective action requires that the contributions of individual experts in solving the tasks (answers to the questions of others) be made visible and available to others, thus creating a public good. The use of this knowledge by some participants (e.g. accessing the identities or types of social relations of unfamiliar persons in images that have been identified through questions-answer exchanges of other participants) does not diminish its benefits for others. The role of the human participants becomes the provision of domain knowledge and reasoning to solve semantic tasks more complex than atomic micro-tasks (e.g. answering a question "Who is this person?" or "What was the role of this person at this event?"). The role of the system becomes the provision of effective mechanisms for routing messages to appropriate participants, aggregating the solutions and providing communication and collaboration channels that stimulate appropriate motivational mechanisms.

To identify possibilities for enticing experts to engage into such exchanges we can now refer to the mechanisms from collective action theory, such as reciprocal or generalized exchange. We also know that the achievement of the collective good (sufficient extent of generated new knowledge to make it a useful resource for the experts) depends on the achievement of critical mass of experts engaging in the question-answer exchanges. While this requires the recruitment of a sufficient number of participants, theories of online collective action suggest that this need not necessarily be large numbers, as long as the user group exhibits sufficient heterogeneity of interests and resources (Sect. "Online Collective Action").

Finally, the suppression of negative effects of free-riding and the emergence of trust and social capital (establishment of positive social ties) as critical success factors requires effective means for containing the problems of adverse selection (hidden information) and moral hazard (hidden action) from the agency theory (Sect. "The Principal-Agent Problem in Human-Computation"). The role of principal in

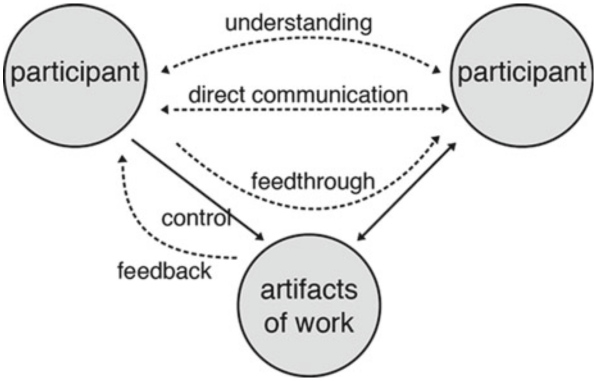


Fig. 1 People-artifact framework of collaboration (Dix et al. 1993)

this case can be assigned to the entire social collective (or the institutional provider of the content archive), whereby the individual participants can be understood as agents, having been entrusted with the production of the collective good. This relates the agency theory to issues such as discouraging unfair use of gained knowledge by the researchers (moral hazard, e.g. by not crediting its originators in appropriate ways) or strategic manipulation of patterns of relationships to support their individual points-of-view (hidden action). Both of which may adversely impact the willingness of experts to contribute their knowledge in solving the tasks and hinder the creation of the collective good.

Embedding Collective Action in Human Computation

To turn these theoretical insights into a conceptual framework informing the design of concrete human computation systems we build on the approach proposed in (Novak 2009). As a basic theoretical model we adopt the well-known people-artifact framework of collaboration (Dix et al. 1993) addressing functional relationships between actors in a cooperative process and tools to support it (Fig. 1). Its focus on the role of shared artifacts and the information flows between cooperating actors allows us to integrate two orthogonal dimensions of *collaborative* human computation: (1) the cooperation between human and computational intelligence and (2) the cooperation between human users.

Since we are addressing complex tasks that are ill-structured by definition and that require contextual knowledge to be interpreted and solved, the shared artifact should allow all parties to express and relate their local worlds of knowledge to each other: the problem space of the human users, the solutions proposed by the computational intelligence and the solutions proposed by the human users. Such settings can benefit from the design and use of a special kind of shared artifacts, the so-called *boundary objects*: artifacts that connect different perspectives of heterogeneous actors on a given problem or a domain of knowledge, without requiring the establishment of one shared perspective (Star 1989). The simultaneous availability of different perspectives allows actors with different goals, interests or worlds of knowledge to interpret the information contained in the shared object in different ways, appropriate for their specific needs. This allows them to exchange and develop new knowledge without giving up their own interests and points of view (Boland and Tenkasi 1995 in Novak 2007). In our case this involves both participants with different backgrounds and research tasks as well as the computational system as a participating actor in a cooperative process (human-machine cooperation) (Novak 2009).

This relates the problem of knowledge integration between human users and with the computational system to the theory of "perspective-making" and "perspective taking" (Boland and Tenkasi 1995). According to this theory, to effectively support knowledge integration in heterogeneous settings shared artifacts must enable and relate several processes to each other: they should allow diverse actors to express one's own perspective in one's own terms (perspective making), to develop an understanding of the perspective of other actors (perspective taking) and to internalize these insights by expressing them anew in their own terms (Boland and Tenkasi 1995). On one hand, this requires that the problem space be described in a way that expresses the needs of human users in terms of their knowledge context as well as in terms of the criteria of the computational system, allowing the mapping of the problems into the space of possible solutions. In our example, this could mean displaying possible identities of unknown persons in the images with user explanations of their suggestions alongside with the system-generated face bounding boxes, the related reference portraits and confidence values of the automatic face detection.

On the other hand, the different task solution possibilities generated by the computer system should also be visualized in the shared artifact (solution space) allowing active exploration by the users in a way that can be used to communicate with each other and with the computational engine (e.g. for explicit or implicit relevance feedback). In our example, this could take the form of displaying social graphs connecting persons co-occurring in images and other inferred relations (events, venues, types of social relations) together with the original images and system-generated (or user-based) confidence values leading to such results.

The corresponding model formalizing this in a theoretically grounded conceptual design framework is depicted in Fig. 2. It shows how collaborative human computation can be supported through interactive boundary objects that mediate shared understanding and integration of the knowledge of human users (explicit/implicit) in a collaborative solution process with the computational system. This model integrates the "people-artifact framework" (Dix et al. 1993) of collaboration with requirements of knowledge integration through boundary objects and the principal-agent perspective, to realize a process integrating human-computer and human-human collaboration.

The central principle is the creation of an open environment in which the shared artifact visualizes the perspectives of different human users and of the computational system and relates them to each other. This includes the visibility and shared manipulation of all information resources normally available only to individual actors (e.g. the task-owner, the isolated crowd-worker, the computational system) to *all* parties involved in a collaborative human computation system.

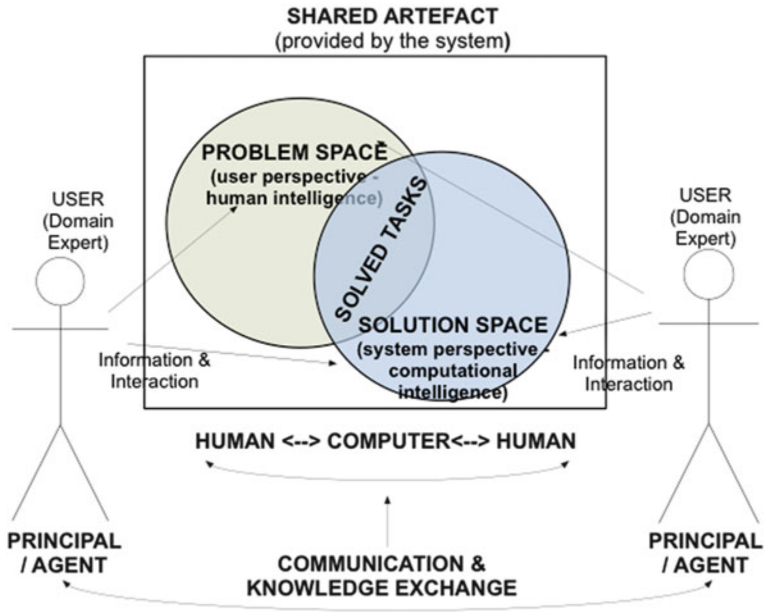


Fig. 2 Conceptual model of collaborative human computation (This model is an adapted version of the collaboration model in interactive value creation proposed in (Novak 2009) now transposed to human computation)

The transfer of expert knowledge about the problem domain occurs through direct user involvement in task selection and task design (e.g. free formulation of questions to other users or selection from predefined templates), their referencing to parts of the problem space (e.g. marking up a person to be identified in an image, or a venue to be resolved) and direct exploration of possible solutions (provided by the system or other users). From the user interaction with different parts of the problem or solution space the system can infer an “understanding” of user knowledge (e.g. detecting faces in images that the automatic detection failed to detect; Dionisio et al. 2013) and use this to generate new or better solutions (e.g. improving confidence values of automatic face recognition by removing false results in conflict with results of human experts).

Displaying the problem space and the space of possible solutions to all users allows different forms of explicit and implicit collaboration to take place. A user can mark the relevant portion of the problem space (e.g. a person face in an image) and submit a task to be solved by other users (e.g. a question on the identity of the marked person, his/her historical context or social relations to other persons in the image). Prior to that, the user can also verify the available results of system-generated solutions, or the user may be able to infer the solution for this specific problem through exploration of solutions from other users. Answers to user tasks from other participants are a form of explicit collaboration. They can be aggregated by computational means in the system (e.g. majority voting) but they can also be displayed in ways that allow users to reach individual decisions or resolve conflicting results through direct communication. Making solutions created by individual users through their own exploration of the problem and solution space also a part of the shared artifact (e.g. annotating a person face with a corresponding name, entering a historical or social relation linking two persons) makes them discoverable by others when they encounter the need to solve the same task. This is a form of implicit collaboration between human users, but as already mentioned it also enables implicit collaboration between human users and the system (e.g. relevance feedback).

Displaying the available solutions of all tasks is a fundamental condition for the creation of the collective good: the explicit (response messages) and implicit solutions (self-annotations) of the participants to individual problems are equally available to all others and not only to the task or solution "owners". Making the individual contributions clearly attributable to its source (e.g. to the system or an individual user) is a method for signaling the trustworthiness of the results and its sources, hence minimizing the problems of hidden information and hidden action (e.g. other users using the results without attributing proper credits to the solution owners or manipulating results to support their personal strategic purpose). In addition, such joint interaction with shared artifacts in an equitable setting can facilitate the creation of social ties and trust, creating effects which can also alleviate moral hazard and free-riding (cf. social capital).

The shared visualization of the problem and solution space may also exploit the heterogeneity of user resources and interests as a means of reaching critical mass required for sustaining collective action. Individual experts can freely choose which tasks to tackle and thus address problems closest to their individual interests. Other users may discover their relevance at a later time while addressing the same tasks or as part of their explorative problem-solving process; this in turn may stimulate their own contributions to possible solutions. The users can also engage in different kinds of contributions requiring different effort, based on available time and resources (e.g. explicitly answering a question, formulating a task to be solved, creating possible solutions through annotations or posting a comment on resolving a conflict between different solution alternatives). This may facilitate the bootstrapping process without requiring large numbers of active contributors and support the expansion of the pool of potential users by being able to recruit participants with diverging interests and resources.

Prototypical Application to Designing a Crowd-Based System

The described conceptual design framework has been prototypically instantiated by applying it to the design of an interactive system supporting the work of researchers in digital humanities (Schreibman et al. 2004). This system specifically aims at supporting the work of historians and political scientists accessing multimedia collections of historical materials, with a high proportion of photographs of historical events. A concrete test application example is the collection of the Centre Virtuel de la Connaissance sur l'Europe[4]) related to the historical development of European integration. These images contain a mix of identified prominent political personalities and unidentified persons from a broader environment of the process of European integration. The application (termed History of Europe)[5] should help the researchers to identify the context of a given photograph, such as identifying the participants, the event and the venue represented in the photograph. This is very difficult to achieve by purely automatic approaches since the metadata is only partially available, possibly inconsistent and the photo collection is very heterogeneous. Identifying depicted participants based on face recognition turns out to be a daunting task due to highly variable quality of the content, a limited availability of reference portraits as well as the various viewpoints, scene compositions and the very different periods of time in which the photos were taken (Harloff, 2012). For similar reasons, the inference of contextual information needed by the researchers to understand the historical and social context of individual persons (events, venues, kinds of social ties to other persons) represents a very hard problem for pure computational approaches.

Accordingly, a human computation approach is adopted that combines computational analysis and human intelligence. Machine algorithms are used to scan photographs in the collection searching for persons (face detection) and identifying their names through comparison to reference portraits (face recognition). The identified faces corresponding to the same person are clustered and a social graph is built connecting persons with social links based on their co-occurrence in the photos (Dionisio et al. 2013). Such a social graph serves as a tool for contextualizing a historical event and its participants and can be used by the researchers to identify the most probable event associated with a given photo, its attendees and related material and historical context. This can further be extended with content from additional sources and the researchers working hypotheses (e.g. identity of a person, historical event in question, venue of the event) can be validated through crowdsourcing.

The technical implementation of this application is undertaken based on the CUbRIK[6] system architecture (Dionisio et al. 2013). This architecture supports different modalities of the integration of computational techniques with crowdsourcing (e.g. task injection with explicit, implicit and passive crowdsorcing), the creation of task-specific worker pools and the distribution of tasks across different social media channels (e.g. Twitter, Facebook). To overcome the deficiencies encountered with general-purpose crowdworkers in previous experiments (Harloff 2012; Dionisio et al. 2013), the conceptual model presented in the previous section has been applied to the design of an interactive system involving an expert-based crowd into the production of a collective good. Involving domain experts promises a much higher performance and quality of solutions, since experts can rely on their domain knowledge to identify persons and other contextual information. The design of a proof-of-concept for the interactive user application is depicted in Figs. 3 and 4. It implements the proposed conceptual design framework of collaborative human computation in the following way.

[4] http://www.cvce.eu/

[5] Developed within the EU project CUbRIK: http://www.cubrikproject.eu/

[6] http://www.cubrikproject.eu/

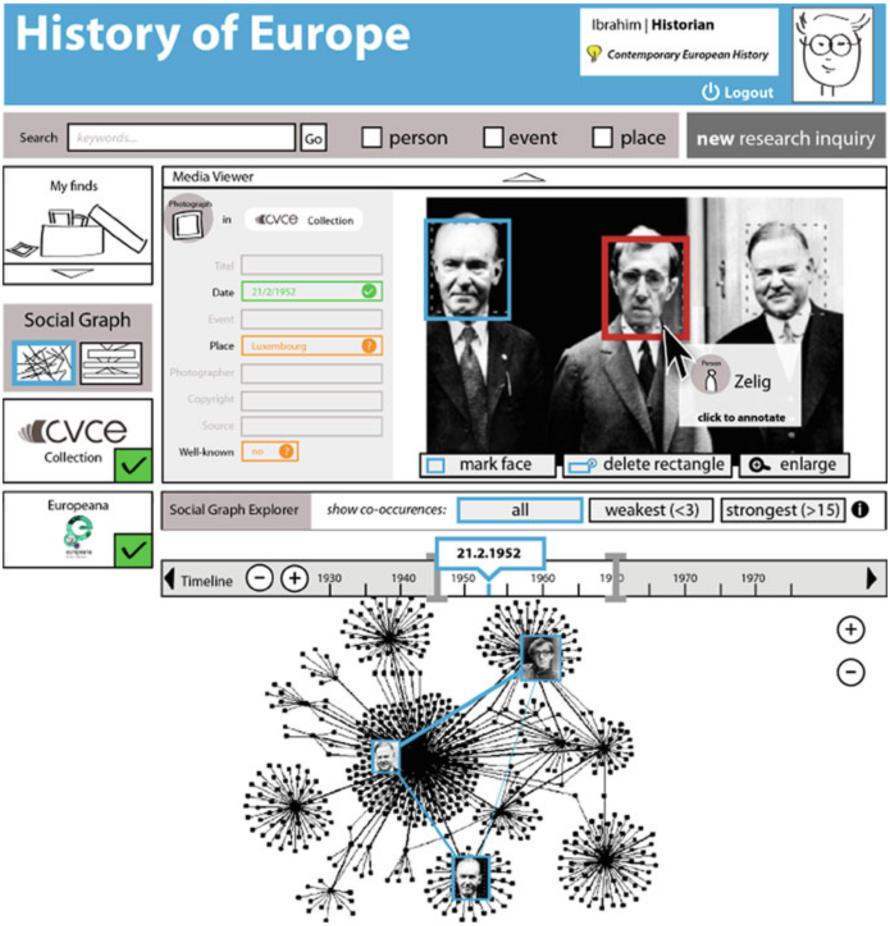


Fig. 3 History of Europe (HoE): proof-of-concept design of a collaborative human computation application incorporating collective action

The user interface displays simultaneously the individual photographs of groups of people that contain persons to be identified (problem space) and the social graph depicting faces of individual persons and their co-occurrence connections (solution space). In using this tool as part of their everyday workflow, the researchers can mark unfamiliar persons and place an explicit request for identification to other experts (ask colleagues) relayed through existing social networks or retrieve newest results of the automatic identification (ask computer). This is compatible with existing work practices of the target users (researchers in the digital humanities) who are already using existing social media (esp. Twitter), to distribute image-related queries such as “Who is this person?” among colleagues. This explicit crowdsourcing is based on community ties, i.e. researchers relying on their colleagues to provide answers to them in the future, though in form of generalized rather than reciprocal

Fig. 4 History of Europe (HoE): explicit collaboration through expert-based crowdsourcing

exchange patterns (the answers are not necessarily provided by the same users "owing" the favour).

By exploring the social graph visualization the users can also infer possible identities based on the relationships of the unknown person to already identified persons in the related images and associated information (events, venue, social relations etc.). Accordingly, they can insert their inferred result about the identity and context of a specific person as an annotation into the system (marking the corresponding face with a box and inserting its name, event, social relations to other persons etc.). In doing so, they are presented with the so-far generated results both from the automatic recognition and from other users (Fig. 5).

Alternatively, they can also vote on existing solution suggestions from other users. The annotations and votes are also visible to other users, including explanations for suggestions that were made and an indication of each user's level of expertise. Explanations, expert level indication of users and author-based majority voting

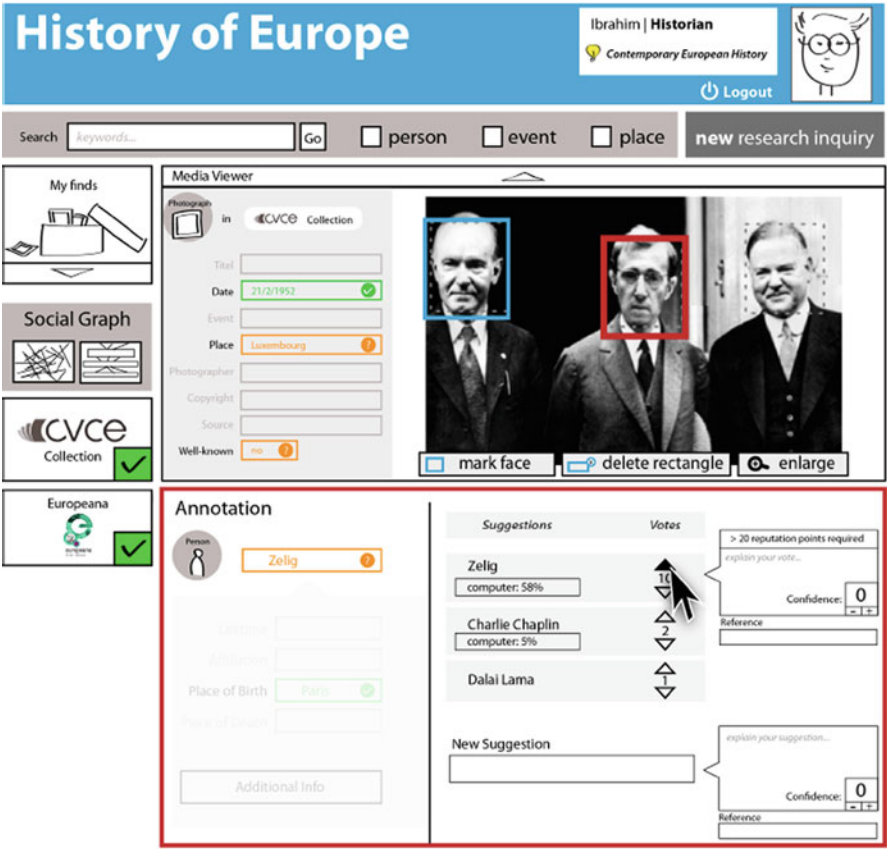


Fig. 5 History of Europe (HoE): implicit collaboration through collaborative annotation

allow for results that are more trustworthy and intelligible than those of the general-purpose crowds aggregated through statistical majority voting algorithms.

In this way, the aggregation of the results of individual crowd solutions is performed through collaborative reasoning of human experts. On the other hand, such a visualization, that makes the results of personal annotations available to all other users is a form of implicit crowdsourcing—the "answers" being provided without the need to initiate explicit task requests. Finally, the visualization of all results to all users regardless of their contribution results in a public good created through collective action (non-excludability). Thereby, the cost of providing a contribution (creating an annotation, answering a task) and the benefits of its consumption remain the same regardless of the number of users subsequently using the results (non-rivalry). The combination of different modalities of use and participation allows different kinds of users to take part—according to their different levels of expertise, interests and available resources. In this way, the system builds on the heterogeneity of interests and resources of the users as a means of overcoming the bootstrapping problem and of reaching the critical mass for sustaining collective action (Sect. "Reframing Human Computation").

Conclusions

This contribution has considered how human computation can be conceptualized as a specific method within a broader context of collaborative systems in order to enable the design of new models of human computation applications. Current individualistic models of human computation work very well for specific classes of problems—that can be broken down into trivial primitives solvable with little cognitive effort and without specific expertise by a crowd of isolated, self-interested individuals. As such, they exhibit different limitations which prevent their application to more complex problem classes requiring creative reasoning or specific domain knowledge. By showing how human computation relates to the large body of knowledge on social collaboration in information sharing and knowledge production this chapter has shown how current human computation can be extended with more cooperative models supporting human-to-human group interaction and communication.

The analysis of the theories of collective action and of the principal agent framework has pointed out how they can inform the conceptualization and design of systems for collaborative human computation. The proposed conceptual design framework shows how the discussed concepts and mechanisms can be applied to designing concrete systems integrating collective action in human computation. It provides a theoretically grounded starting point for collaborative human computation applications that engage groups of participants in solving tasks and creating knowledge as a *public good* benefiting entire social collectives rather than only individual contributors and task owners. The practical validity of the proposed conceptual model has been demonstrated by applying it to the design of a crowd-based system for solving semantically difficult tasks requiring expert knowledge. This illustrates its applicability to informing the design of concrete applications integrating human computation with collective action. Obviously, such a prototypical application can provide only a limited validation of the suitability and practical usefulness of the proposed model for the design of concrete applications. As this is one of the first attempts to conceptualize collaborative human computation in a theoretically grounded but practically applicable model, we hope it can entice further explorations in human computation for new forms of collective action.

Acknowledgements The research leading to the results described in this chapter has been partly funded within the European Union's Seventh Framework Programme FP7/2007–2013 under grant agreement n° 287704 (the CUbRIK project). Different ideas and results presented in this chapter have benefited from stimulating exchanges and cooperations with the project partners. My special thanks for their collaboration in the design and implementation of the proof-of-concept for the History of Europe application go to Piero Fraternalli and Marco Tagliasacchi from Politecnico di Milano, Martha Larson from TU Delft, Lars Wieneke from CVCE and Erik Harloff and Isabel Micheel from the European Institute for Participatory Media.

References

Ackerman MS, McDonald DW (1996) Answer garden 2: merging organizational memory with collective help. In: Proceedings of the ACM conference on computer-supported cooperative work (CSCW'96). ACM, New York, NY, USA, pp 97–105

Ackerman MS, Pipek V, Wulf V (2002) Sharing expertise: beyond knowledge management. MIT Press, Cambridge

Bernstein MS, Little G, Miller RC et al (2010) Soylent: a word processor with a crowd inside. In: Proceedings of the 23rd annual ACM symposium on User interface software and technology 2010. ACM, New York, NY, USA

Bian J, Liu Y, Agichtein E, Zha H (2008) Finding the right facts in the crowd: factoid question answering over social media. In: Proceedings of the 17th international conference on world wide web (WWW '08), ACM, New York, pp 467–476

Bimber B, Flanagin AJ, Stohl C (2005) Reconceptualizing collective action in the contemporary media environment. Commun Theory 15:365–388. doi:10.1111/j.1468-2885.2005.tb00340.x

Boland JR, Tenkasi RV (1995) Perspective making and perspective taking in communities of knowing. Organ Sci 6(4)

Bonabeau E (2009) Decisions 2.0: the power of collective intelligence. MIT Sloan Manage Rev 50(2):45–52

Bonney R, Cooper CB, Dickinson JL, Kelling S, Phillips T, Rosenberg K, Shirk J (2009) Citizen science: a developing tool for expanding science knowledge and scientific literacy. Bioscience 59:977–984

Bozzon A, Brambilla M, Ceri S (2012) Answering search queries with crowdSearcher. In: Proceedings of the 21st international conference on World Wide Web (WWW '12), ACM, New York, pp 1009–1018

Brambilla M, Fraternali P, Vaca C (2011) BPMN and design patterns for engineering social BPM solutions. In: Proceedings of the 4th international workshop on BPM and social software (BPMS 2011). Clermont-Ferrand, France

Brown JS, Duguid P (2000) The social life of information. Harvard Business School Press, Boston

Chen K, Wu C, Chang Y, Lei C (2009) A crowdsourceable QoE evaluation framework for multimedia content. In: Proceedings MM 2009. ACM, New York, NY, USA, pp 491–500

Constant D, Sproull L, Kiesler S (1996) The kindness of strangers: the usefulness of electronic weak ties for technical advice. Organ Sci 7(2):119–135

Dean J, Ghemawat S (2008) Map reduce: simplified data processing on large clusters. Commun ACM 51(1):107–114

Dionisio M, Fraternali P, Harloff E, Martinenghi D, Micheel I, Novak J, Pasini C, Tagliasacchi M, Zagorac S (2013) Building social graphs from images through expert-based crowdsourcing. In: Proceedings of SoHuman 2013–2nd international workshop on social media for crowdsourcing and human computation at ACM web science 2013 (to appear)

Dix AJ, Finley J, Abowd GD, Beale R (1993) Human-computer interaction. Prentice Hall, New York

Dutta P, Aoki PM, Kumar N, Mainwaring A, Myers C, Willett W, Woodruff A (2009) Common sense: participatory urban sensing using a network of handheld air quality monitors. In: Proceedings of the 7th ACM conference on embedded networked sensor systems 2009. ACM, New York, NY, USA, pp 349–350

Ekeh PP (1974) Social exchange theory: the two traditions. Harvard University Press, Cambridge

Fraternali P, Castelletti A, Soncini-Sessa R, Ruiz CV, Rizzoli AE (2012) Putting humans in the loop: social computing for water resources management. Environ Model Softw 37:68–77

Frey BS, Jegen R (2002) Motivation crowding theory. J Econ Surv 15(5) Blackwell

Fulk J, Flanagin AJ, Kalman ME, Monge PR, Ryan T (1996) Connective and communal public goods in interactive communication systems. Commun Theory 6(1):60–87

Goolsby R (2010) Social media as crisis platform: the future of community maps/crisis maps. ACM Trans Intell Syst Technol 1(1) Article 7

Greene K, Kniss J, Luger G (2010) Representing diversity in communities of Bayesian decision-makers. In: Proceedings of the 2ndIEEE international conference on social computing. Washington, DC, USA

Hagel J, Armstrong A (1997) Net gain: expanding markets through virtual communities. McGraw-Hill

Hardin R (1982) Collective action. Johns Hopkins University Press, Baltimore. Boston: Harvard Business School Press, 1997

Harloff E (2012) Who is this person? Konzeption und prototypische Evaluierung einer Crowdsourcing-Anwendung für Multimedia-Suche (Concept and prototypical evaluation of a crowdsourcing application for multi-media search). Bachelor Thesis, University of Applied Sciences Stralsund, Department of Business Studies, Sept 2012

Head JG (1962) Public goods and public policy. Public Financ 17(3):197–219

Howe J (2006) Crowdsourcing: a definition. http://crowdsourcing.typepad.com/cs/2006/06/crowdsourcing_a.html. Accessed 15 Apr 2013

Hu C, Bederson BB, Resnik P (2011) Monotrans2: a new human computation system to support monolingual translation. In: Proceedings CHI 2011. ACM, New York, NY, USA

Kanefsky B, Barlow NG, Gulick VC (2001) Can distributed volunteers accomplish massive data analysis tasks? In: Proceedings of the lunar and planetary science conference XXXII 2001. Houston, Texas

Kittur A, Kraut RE (2008) Harnessing the wisdom of crowds in wikipedia: quality through coordination. In: Proceedings of the 2008 ACM conference on computer supported cooperative work, ACM, New York, NY, USA, pp 37–46

Kittur A, Smus B, Khamkar S, Kraut RE (2011) Crowdforge: crowdsourcing complex work. In: Proceedings of the 24th annual ACM symposium on user interface software and technology 2011, ACM, New York, NY, USA, pp 43–52

Klein M, Iandoli L (2008) Supporting collaborative deliberation using a large-scale argumentation system: the MIT collaboratorium. In: Proceedings of directions and implications of advanced computing; conference on online deliberation (DIAC-2008*/OD2008)*. Berkeley, CA, USA

Koch M (2008) CSCW and enterprise 2.0—towards an integrated perspective. In: Proceedings of the 21th Bled conference 2008, Bled

Kollock P (1998) Social dilemmas: the anatomy of cooperation. Annu Rev Sociol 24:183–214

Kollock P (1999) The economies of online cooperation: gifts, and public goods in cyberspace. In: Smith MA, Kollock P (eds) Communities in cyberspace. Routledge, London, pp 220–239

Lieberman H, Smith D, Teeters A (2007) Common consensus: a web-based game for collecting commonsense goals. In: Proceedings IUI 2007, Honolulu, Hawaii, USA

Luther K, Counts S, Stecher KB, Hoff A, Johns P (2009) Pathfinder: an online collaboration environment for citizen scientists. In: Proceedings of the SIGCHI conference on human factors in computing systems, CHI '09, ACM, New York, pp 239–248

Macy M (1990) Learning theory and the logic of critical mass. Am Sociol Rev 55:809–826

Marwell G, Oliver P (1993) The critical mass in collective action: a micro-social theory. Cambridge University Press, New York

Meier P (2013) Human computation for disaster response. In: Michelucci P (ed) Handbook of human computation, Springer (to appear)

Nagar Y (2011) Beyond the human-computation metaphor. In: Privacy, security, risk and trust (passat) 2011 IEEE third international conference on and 2011 IEEE 3rd international conference on social computing (socialcom), 9–11 Oct 2011, pp 800–805

Nahapiet J, Ghoshal S (1998) Social capital, intellectual capital, and the organizational advantage. Acad Manage Rev 23(2):242–266

Nonnecke B, Preece J (2001) Why lurkers lurk. In: Proceedings AMCIS '01, Boston

Novak J (2007) Helping knowledge cross boundaries: using knowledge visualization to support cross-community sense-making. In: Proceedings of HICSS-40, Hawaii international conference on system sciences, Hawaii, Jan 2007

Novak J (2009) Mine, yours…Ours? Designing for principal-agent collaboration in interactive value creation. In: Proceedings of Wirtschaftsinformatik 2009, Vienna, Feb 2009

Novak J, Preuße S (2011) Designing visual systems for social data analysis in open government applications. In: Proceedings workshop on data-centric interactions on the Web at INTERACT 2011, International Conference on human-computer interaction, WS-CEUR, Lissabon, vol 817. Dec 2011

Okolloh O (2009) Ushahidi or 'testimony': web 2.0 tools for crowdsourcing crisis information. Particip Learn Action 59(1):65–70

Oliver PE, Marwell G, Teixeira R (1985) A theory of critical mass i: group heterogeneity, interdependence and the production of collective goods. Am J Sociol 91:522–556

Olson M (1965) The logic of collective action. Harvard University Press, Cambridge

Pavlou P, Huigang L, Yajiong X (2007) Understanding and mitigating uncertainty in online exchange relationships: a principal—agent perspective. MIS Q 31(1) pp 105–136

Preece J (2000) Online communities: designing usability and supporting sociability. Wiley, New York, NY, USA

Preece J, Shneiderman B (2009) The reader-to-leader framework: motivating technology-mediated social participation. AIS Trans Hum-Comput Interact 1(1):13–32

Putnam RD (2000) Bowling alone. Simon & Schuster, New York

Quinn AJ, Bederson BB (2011) Human computation: a survey and taxonomy of a growing field. In: CHI 2011, 7–12 May 2011, ACM, New York, NY, USA

Rittel HWJ, Webber MM (1973) Dilemmas in a general theory of planning. Policy Sci 4:155–169, Elsevier Scientific Publishing Company, Amsterdam

Schreibman S, Siemens R, Unsworth J (2004) "The Digital Humanities and Humanities Computing: An Introduction" in Schreibman, Susan, Siemens, Ray and John Unsworth, A Companion to Digital Humanities, Malden, MA: Blackwell, pp. xxiii–xxvii

Scott JW, El-Assal M (1969) Multiversity, university size, university quality and student protest: an empirical study. Am Sociol Rev 35:627–649

Star SL (1989) The structure of ill-structured solutions: boundary objects and heterogeneous distributed problem solving. In: Huhn M, Gasser L (eds) Readings in distributed artificial intelligence. Morgan Kaufman, San Francisco, CA

Tedjamulia SJ, Dean DL, Olsen DR, Albrecht CC (2005) Motivating content contributions to online communities: toward a more comprehensive theory. In: Proceedings HICSS'05 2005, IEEE Computer Society, Washington, DC, USA

Tungare M, Hanrahan B, Quintana-Castillo R, Stewart M, Pérez-Quiñones M (2010) Collaborative human computation as a means of information management. In: Proceedings of the 2nd international workshop on collaborative information seeking at CSCW 2010, Savannah, Georgia, USA

von Ahn L (2006) Games with a purpose. Computer 39(6):92–96, IEEE Computer Society

von Ahn L, Dabbish L (2004) Labeling images with a computer game. In: Proceedings CHI 2004, Vienna, Austria

von Ahn L, Maurer B, McMillen C, Abraham D, Blum M (2008) ReCAPTCHA: human-based character recognition via web security measures. Science 321(5895):1465–1468

Wasko MM, Faraj S (2000) It is what one does: why people participate and help others in electronic communities of practice. J Strateg Inf Syst 9:155–173

Wasko MM, Faraj S (2005) Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. MIS Q 29(1):35–57

Wasko MM, Teigland R (2002) The provision of online public goods: examining social structure in a network of practice. In: ICIS 2002 proceedings. Paper 15, Barcelona, Spain

Wasko MM, Teigland R (2004) Public goods or virtual commons? Applying theories of public goods, social dilemmas, and collective action to electronic networks of practice. J Inf Technol Theory Appl (JITTA) 6(1):25–41

Woolley AW, Chabris CF, Pentland A, Hashmi N, Malone TW (2010) Evidence for a collective intelligence factor in the performance of human groups. Science 330(6004):686–688

Cultural Evolution as Distributed Computation

Liane Gabora

Introduction

The origin of life brought about unprecedented change to our planet; new forms emerged creating niches that paved the way for more complex forms, completely transforming the lands, skies, and oceans. But if biological evolution is effective at bringing about adaptive change, human cultural evolution is arguably even more effective, and faster. Cultural change doesn't take generations; it works at the speed of thought, capitalizing on the strategic, intuitive creative abilities of the human mind.

This chapter outlines current and potential future steps toward the development of a human computation program inspired by the speed and effectiveness of how culture evolves. The overarching goal of the kind of research program outlined in this chapter is to develop a scientific framework for cultural evolution by abstracting its algorithmic structure, use this algorithmic structure to develop human-machine hybrid structures with previously unforeseen computational power, and apply them to solve real problems. The proposed approach can be thought of as a "repeatable method" or "design pattern" for fostering cultural emergence, defined by specific computational methods for modeling interactions at the conceptual level, the individual level, and the social level, and applying them to generate accumulative adaptive, open-ended cultural novelty.

L. Gabora (✉)
Department of Psychology, University of British Columbia, Okanagan Campus,
Kelowna, BC V1V 1V7, Canada
e-mail: liane.gabora@ubc.ca

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_34,

Two Approaches to a Scientific Framework for Culture

Cultural evolution entails the generation and transmission of novel behavior and artifacts within a social group, both vertically from one generation to another, and horizontally amongst members of a generation. Like biological evolution, it relies on mechanisms for both introducing variation and preserving fit variants. Elements of culture adapt, diversify, and become more complex over time, and exhibit phenomena observed in biological evolution, such as niches, drift, epistasis, and punctuated equilibrium (Bentley et al. 2004; Durham 1991; Gabora 1995). However, we lack a precise understanding of how culture evolves.

We begin by summarizing two approaches that have been taken to developing a formal understanding of the process by which culture evolves: Darwinian approaches, and Communal Exchange approaches.

Darwinian Approaches

Dawkins' (1976) proposal that culture evolves through reiterated variation and selection inspired formal Darwinian models of cultural evolution (Boyd and Richerson 1985, 2005; Gabora, 1996; O'Brien and Lyman 2000; Cavalli-Sforza and Feldman 1981; Henrich and Boyd 1998, 2002). It also inspired some archaeologists to apply methods designed for documenting the evolution of biological organisms to chart the historical evolution of artifacts (e.g., O'Brien and Lyman 2000; Shennan 2008). Aside from the questionable assumptions underlying this approach (Atran 2001; Fracchia and Lewontin 1999; Gabora 2004, 2006a; Gabora 2011; Skoyles 2008; Temkin and Eldredge 2007), it aims to model how cultural variants spread, not how they come into existence, strategically building on and opening up new niches for one another.

Holland (1975) elucidated the algorithmic structure of natural selection, resulting in the *genetic algorithm* (GA), and subsequently genetic programming (GP) (Koza 1993), optimization tools with diverse applications to everything from scheduling tasks (Hou et al. 1994) to pipeline design (Goldberg and Kuo 1987) to music and art (Bentley and Corne 2002; DiPaola and Gabora 2009). The term *cultural algorithm* has referred to a GA that includes a 'belief space' used to prune the search space (Reynolds 1994), not an algorithm inspired by how culture itself evolves. GAs are effective for multi-constraint problems with complex fitness landscapes, but would not do well on problems that require reformulating or *restructuring* the problem from another perspective. GAs are *breadth-first* (generate *many* solutions *randomly*, and some by chance may be effective), whereas cultural evolution, which relies on cognitive processes such as learning, is *depth-first* (generate *few* solutions making use of strategic analysis or spontaneous associations, either intentional or unintentional).

Communal Exchange

Mounting evidence suggests that a non-Darwinian framework is appropriate for, not just cultural evolution, but the earliest stages of organic life itself (Gabora 2006b; Kauffman 1993; Vetsigian et al. 2006; Williams and Frausto da Silva 2003), and aspects of modern microbial life (Woese 2002). There is widespread support for the hypothesis that the earliest protocells were self-organized autocatalytic networks that evolved (albeit haphazardly) through a non-Darwinian process involving horizontal transfer of innovation protocols, referred to as *communal exchange* (Vetsigian et al. 2006). Communal exchange differs substantially from natural selection. Acquired change is retained, and information is transmitted communally, not by way of a self-assembly code from parent to offspring. Formal methods for modeling reaction networks can be used to investigate the feasibility of the emergence of the kind of self-sustaining structure that could evolve through communal exchange.

It has been suggested that the basic unit of cultural evolution is, not an autocatalytic network per se, but an associative network that is (like an autocatalytic network) *autopoietic,* i.e., the whole emerges through interactions amongst the parts (Gabora 1998, 2001, 2004; 2008a, 2013; Gabora and Aerts, 2009). A communal exchange based computational model of cultural evolution has been developed (Gabora 1995, 2008b, c). EVOC (for EVOlution of Culture) consists of neural network based agents that invent new actions and imitate actions performed by neighbors. The assemblage of ideas changes over time not because some replicate at the expense of others, as in natural selection, but because they transform through inventive and social processes. Agents can make generalizations concerning what kinds of actions are fittest, and use this acquired knowledge to modify ideas for actions between transmission events. EVOC exhibits typical evolutionary patterns, e.g., cumulative increase in fitness and complexity of cultural elements over time, and an increase in diversity as the space of possibilities is explored, followed by a decrease as agents find and converge on the fittest possibilities. EVOC has been used to model how the mean fitness and diversity of cultural elements is affected by factors such as leadership, population size and density, borders that affect transmission between populations, and the proportion and distribution of creators (who acquire new ideas primarily by inventing them) versus imitators (who acquire new ideas primarily by copying their neighbors) (Gabora 1995, 2008a, b; Gabora and Firouzi 2012; Gabora and Leijnen 2009; Leijnen and Gabora 2010).

A communal exchange inspired method for organizing artifacts into historical lineages has also been developed. *Worldview Evolution*, or WE for short, uses both superficial (e.g., 'beveled edge') and abstract (e.g., 'object is thrown') attributes, as well as analogical transfer (e.g., of 'handle' from knife to cup) and complementarity (e.g., bow and arrow) (Gabora et al. 2011; Veloz, Temkin & Gabora 2012). It represents objects not in terms of a convenient list of discrete measurable attributes, but in terms of how they are actually conceptualized, as a network of interrelated properties, using a *perspective* parameter that can be weighted differently according to their relative importance. Preliminary analyses show that the conceptual network approach can recover previously unacknowledged patterns of historical relationship that are more congruent with geographical distribution and temporal data than is obtained with an alternative cladistic approach that is based on the assumption that cultural evolution, like biological evolution, is Darwinian.

These two computational models, EVOC and WE, show that a communal exchange approach to cultural evolution is computationally tractable. However such models will not begin to approach the open-ended ingenuity and complexity of human cultural evolution until they incorporate certain features of the cognitive process by which cultural novelty is generated.

The Generation of Cultural Novelty

We said that cultural evolution is a depth-first evolution strategy. A depth-first evolution strategy entails processes that adaptively bias the generation of novelty. A number of key, interrelated processes have been identified that, in addition to learning, accomplish this in cultural evolution. We now look briefly at some of these processes, as well as efforts to model them.

Recursive Recall and Restorative Restructuring

Recursive recall (RR) is the capacity for one thought to trigger another, enabling progressive modification of an idea. Donald's (1991) hypothesis that cultural evolution was made possible by onset of the capacity for RR has been tested using EVOC (Gabora and Saberi 2011; Gabora and DiPaola 2012; Gabora, Chia and Firouzi 2013). A comparison was made of runs in which agents were limited to single-step actions to runs in which they could recursively operate on ideas, and chain them together, resulting in more complex actions. While RR and no-RR runs both converged on optimal actions, without RR this set was static, but with RR it was in constant flux as ever-fitter actions were found. In RR runs there was no ceiling on mean fitness of actions, and RR enhanced the benefits of learning.

Although these findings support Donald's hypothesis, the novel actions generated with RR were predictable. They did not open up new cultural niches in the sense that, for example, the invention of cars created niches for the invention of things like seatbelts and stoplights. EVOC in its current form could not solve *insight problems*, which require restructuring the solution space (Boden 1990; Kaplan and Simon 1990; Ohlsson 1992). Restructuring can be viewed as a form of RR that entails looking at the problem from a new perspective, making use of the mind's *self-organizing, restorative* capacity.

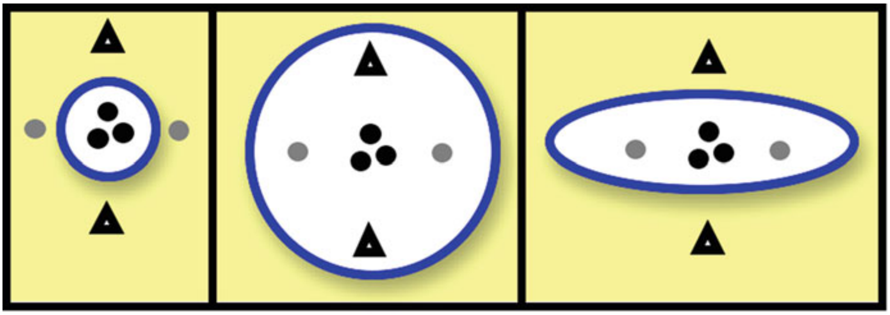


Fig. 1 *Convergent thought* (*left*) activates key properties only, represented by *black dots*. *Divergent thought* (*centre*) activates not just key properties but also peripheral (less salient) properties, represented by both *grey dots* and *black triangles*. The *grey dots* represent peripheral properties that are relevant to the current context (goal or situation); the *black triangles* represent peripheral properties that are irrelevant to the current context. *Associative thought* (*right*) activates key properties and context-relevant peripheral properties

Contextual Focus (CF) and Divergent Versus Associative Thought

It has been proposed that restorative restructuring is aided by *contextual focus* (CF): the capacity to spontaneously and temporarily shift to a more divergent mode of thought (Gabora 2003). Divergent thought entails an increase in activation of the associates of a given item (Runco 2010). Thus for example, given the item TABLE, in a convergent mode of thought you might call to mind accessible associates such as CHAIR, but in a divergent mode of thought you might also call to mind more unusual associates such as PICNIC or MULTIPLICATION TABLE. CF has been implemented in EVOC (the computational model of cultural evolution). Low fitness of ideas induces a temporary shift to a more divergent processing mode by increasing the 'reactivity', α, which determines the degree to which a newly invented idea can differ from the idea on which it was based.

Current research on the architecture of memory suggests that creative thought is actually not divergent but associative, as illustrated in Fig. 1 (Gabora 2010; Gabora and Ranjan 2013). While divergent thought refers to an increase in activation of *all* associates, associative thought increases only activation of those relevant to the context. Because memory is distributed and content-addressable, associations are forged by way of shared structure, in associative thought items come together that, though perhaps seemingly different, *share properties or relations,* and are thus more likely than chance to be *relevant* to one another, perhaps in a previously unnoticed but useful way.

A processing mode that is not just divergent but associative could be simulated in a model such as EVOC capitalizing on the ability to learn generalizations (e.g., symmetrical movements tend to be fit) to constrain changes in α. It would also be interesting to investigate the topological and dynamical properties of fitness landscapes for which divergent versus associative forms of CF is effective. CF is expected to be most beneficial for fitness landscapes that are rugged and subject to infrequent, abrupt change, with associative CF outperforming divergent CF.

Concept Interaction

Since creative processes such as restructuring involve putting concepts together in new contexts, a model of cultural evolution should be built upon a solid theory of concepts and *how they interact*. However, people use conjunctions and disjunctions of concepts in ways that violate the rules of classical logic; i.e., concepts interact in ways that are non-compositional (Osherson and Smith 1981; Hampton 1987; Aerts 2009; Aerts et al. 2009a). This is true both with respect to properties (e.g., although people do not rate 'talks' as a characteristic property of PET or BIRD, they rate it as characteristic of PET BIRD), and exemplar typicalities (e.g., although people do not rate 'guppy' as a typical PET, nor a typical FISH, they rate it as a highly typical PET FISH). Because of this, concepts have been resistant to mathematical description.

This non-compositionality can be modeled using a generalization of the formalisms of quantum mechanics (QM) (Aerts and Gabora 2005; Aerts, Broekaert, and Gabora 2011; Gabora and Aerts 2002a, b; Kitto et al. 2011). The reason for using the quantum formalism is that it allows us to describe the chameleon-like way in which concepts interact, spontaneously shifting their meanings depending on what other concepts are nearby or activated. The following formal exposition, though not essential for grasping the underlying concepts, is provided for the mathematically inclined reader. In QM, the state $|\psi\rangle$ of an entity is written as a linear superposition of a set of basis states $\{|\phi_i\rangle\}$ of a complex Hilbert space H. Hence $|\psi\rangle = \Sigma_i c_i|\phi_i\rangle$ where each complex number coefficient c_i of the linear superposition represents the contribution of each component state $|\phi_i\rangle$ to the state $|\psi\rangle$. The square of the absolute value of each coefficient equals the weight of its component basis state with respect to the global state. The choice of basis states is determined by the observable to be measured. The basis states corresponding to this observable are called *eigenstates*. Upon measurement, the state of the entity *collapses* to one of the eigenstates. In the quantum inspired State COntext Property (SCOP) theory of concepts, the basis states represent states (instances or exemplars) of a concept, and the measurement is the context that causes a particular state to be evoked. SCOP is consistent with experimental data on concept combination (Aerts 2009; Aerts et al. 2009a, 2012; Aerts et al. in press; Hampton 1987), and with findings that a compound's constituents are not just conjointly activated but bound together in a context-specific manner that takes relational structure into account (Gagné and Spalding 2009). The model is being expanding to incorporate larger conceptual structures (Gabora and Aerts 2009), and different modes of thought (Veloz et al. 2011). This theoretical work is complemented by empirical studies aimed at establishing that (i) some concept combinations involve interference and entangled states, and (ii) creative products are external evidence of an internal self-organization process aimed at resolving dissonance and restoring equilibrium through the recursive actualization of potentiality (Gabora 2010; Gabora et al. 2012; Gabora and Saab 2011; Henderson and Gabora 2013; Riley and Gabora 2012).

Harnessing the Computational Power of Cultural Evolution

We have looked at some of the key milestones that have been crossed in the development of a scientific framework for how culture evolves. These milestones include a crude but functional computational model of cultural evolution, research into the cognitive mechanisms underlying the generation of cultural novelty, and preliminary efforts to computationally model these mechanisms. The rest of this chapter presents new, untested, yet-to-be-implemented ideas for how to go about harnessing the speed and power of cultural evolution in the development of a human computation research program.

Computational Model of Restorative Restructuring

A first step is to develop a model of problem restructuring using a "reaction network" inspired model that has as its basic unit, not catalytic molecules, but interacting concepts. There are various methods for going about this, for example using Concat, or Holographic Reduced Representations to computationally model the *convolution* or 'twisting together' of mental representation (Aerts et al. 2009b; Dantzing, Raffone and Hommel 2011 Eliasmith and Thagard 2001; Thagard and Stewart 2011). Another promising route is to use a quantum-inspired theory of concepts such as SCOP that incorporates the notion of context-driven actualization of potential (Aerts and Gabora 2005a, b; Gabora and Aerts 2002a, b). A concept is defined in terms of (1) its set of states or exemplars Σ, each of which consists of a set L of relevant properties, (2) set M of contexts in which it may be relevant, (3) a function ν that describes the applicability or *weight* of a certain property for a specific state and context, and (4) a function μ that describes the transition probability from one state to another under the influence of a particular context.

The procedure is best explained using an example, such as the idea of using a tire to make a swing, i.e., the invention of a tire swing (from Gabora Scott, and Kauffman 2013). The concept TIRE consists of the set Σ of states of TIRE, and in the context 'winter', TIRE might collapse to SNOW TIRE. Suppose that the network's initial conception of TIRE, represented by vector $|p\rangle$ of length equal to 1, is a superposition of only two possibilities (Fig. 2). The possibility that the tire has sufficient tread to be *useful* is denote by unit vector $|u\rangle$. The possibility that it should be discarded as *waste* is denoted by unit vector, $|w\rangle$. Their relationship is given by the equation $|p\rangle = a_0|u\rangle + a_1|w\rangle$, where a_0 and a_1 are the amplitudes of $|u\rangle$ and $|w\rangle$ respectively. If a tire us useful only for transportation, denoted $|t\rangle$ then, $|u\rangle = |t\rangle$. States are represented by unit vectors and all vectors of a decomposition such as $|u\rangle$ and $|w\rangle$ have unit length, are mutually orthogonal and generate the whole vector space, thus $|a_0|^2 + |a_1|^2 = 1$.

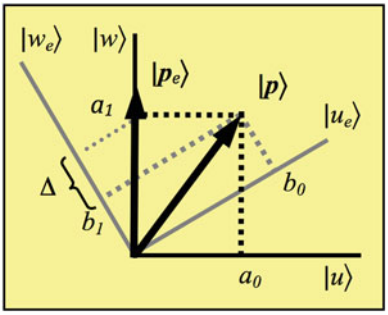


Fig. 2 Graphical depiction of a vector $|p\rangle$ representing particular state of TIRE, specifically, a state in which the tread is worn away. In the default context, the state of tire is more likely to collapse to the projection vector $|w\rangle$ which represents wasteful than to its orthogonal projection vector $|u\rangle$ which represents useful. This can be seen by the fact that subspace a_0 is smaller than subspace a_1. Under the influence of the context playground equipment, the opposite is the case, as shown by the fact that b_0 is larger than b_1. Also shown is the projection vector after renormalization

The conception of TIRE changes when activation of the set L of properties of TIRE, e.g. 'weather resistant', spreads to other concepts in the network for which these properties are relevant. Contexts such as playground equipment that share properties with TIRE become candidate members of the set M of relevant contexts for TIRE. The context playground equipment, denoted *e,* consists of the concepts SWING, denoted $|s_e\rangle$, and SLIDE, denoted $|l_e\rangle$. The restructured conception of TIRE in the context of playground equipment, denoted $|p_e\rangle$, is given by $b_0|u_e\rangle + b_1|w_e\rangle$, where $u_e\rangle = b_2|t_e\rangle + b_3|t_e s_e\rangle + b_4|t_e l_e\rangle$, and where $|t_e s_e\rangle$ stands for the possibility that a tire functions as a swing, and $|t_e l_e\rangle$ stands for the possibility that a tire functions as a slide. The amplitude of $|w_e\rangle$, $|b_1|$, is less than $|a_1|$, the amplitude of $|w\rangle$. This is because $|b_0| > |a_0|$, since $|b_0|$ consists of the possibility of a tire being used not just as a tire, but as a swing or slide. Because certain strongly weighted properties of SLIDE, such as 'long' and 'flat', are not properties of TIRE, $|b_4|$ is small. That is not the case for SWING, so $|b_3|$ is large. Therefore, in the context playground equipment, the concept TIRE has a high probability of collapsing to TIRE SWING, an entangled state of the concepts TIRE and SWING. Entanglement introduces interference of a quantum nature, and hence the amplitudes are complex numbers (Aerts 2009). If this collapse takes place, TIRE SWING is thereafter a new state of both concepts TIRE and SWING.

This example shows that a formal approach to concept interactions that is consistent with human data (Aerts 2009; Aerts et al. 2009a, 2012; Aerts et al. in press; Hampton 1987) can model the restructuring of information (e.g., TIRE) under a new context (e.g., playground equipment). Note how in the quantum representation, probability is treated as arising not from a lack of information per se, but from the limitations of any particular context (even a 'default' context).

The limitations of this approach are as interesting as its strengths. It is not possible to list, or even develop an algorithm that will list all possible uses or contexts for any item such as a tire or screwdriver (Longo et al. 2012). This is what has been referred to as the *frame problem*. As a consequence, human input is particularly welcome at this juncture to define the relevant contexts, e.g., the possible uses of a tire. Studies would be run using data collected from real humans to determine the extent to which the model matches typicality ratings and generation frequencies of exemplars of concepts in particular contexts by human participants, as per (Veloz et al. 2011). SCOP models of individual concepts can be embedded into an associative "reaction network" to study the associative structure of interrelated sets of concepts a whole, and the conditions under which it restores equilibrium in response to the introduction of new states of concepts that results from placing them in new contexts.

Using this SCOP-based cognitive "reaction network" it would be possible to test the hypothesis that contextual focus (the ability to shift between different modes of thought depending on the context) increases cognitive efficiency. If the amplitude associated with $|w\rangle$ for any concept becomes high—such as for TIRE if the weight of the property 'tread' is low—this signals that the potentiality to re-conceptualize the concept is high. This causes a shift to a more associative mode by increasing α, causing activation of other concepts that share properties with this concept, as described previously.

Enhanced Computational Model of Cultural Evolution

Let us now examine how a model of restorative restructuring such as the SCOP-based one we just looked at could be used to develop a cognitively sophisticated computational model of cultural evolution. We will refer to this 'new and improved' model as EVOC2.

So that the EVOC2 agents have something to make artifacts from, their world would contain resource bases from which objects are extracted and wastes are generated. Extracted objects can be joined (lego-style) to construct other objects. Agents have mental representations of resources and objects made from resources. Objects derived from the same resource are modeled in their conceptual networks as states of a concept. Newly extracted or constructed objects have a fitness that defines how useful or wasteful they are *with respect to the other objects an agent has encountered*. Thus existing objects provide contexts that affect the utility of new objects, and an agent's knowledge of existing objects defines its *perspective*.

The artificial culture can now evolve as follows:

Invent. Agents invent as in EVOC, except that they invent not actions but objects, using resources in adjacent cells. Extracting an object from a resource creates waste objects.

Detect and Actualize Potential for Adaptive Change. If a waste object p is accumulating adjacent to A1, A1 recursively modifies p by considering it from A1's perspective. This continues until p is in a new less wasteful state p_{A1*} which is an eigenstate with respect to A1's perspective. This process may modify not just p, but A1's perspective. Perspectives change in response to the ideas and objects an agent interacts with; thus a perspective can encompass more than one context.

Contextual focus. The previous step may involve temporarily assuming a more associative processing mode in response to the magnitude of potential for adaptive change.

Transmission. Modified object, p_{A1*}, becomes input to the associative networks of adjacent agents.

Context-dependent Restructuring. If p_{A1*} is wasteful (has potential to change) with respect to the perspective of another agent, A2, then A2 recursively modifies p_{A1*} until it is an eigenstate with respect to A2's perspective, at which point it is referred to as p_{A1*A2*}. Since A1's perspective is reflected in p_{A1*}, assimilation of p_{A1*} modifies A2's perspective in a way that reflects exposure to (though not necessarily incorporation of or agreement with) A1's perspective. This continues until p settles on stable or cyclic attractor, or we terminate after a set number of iterations (since a chaotic attractor or limit cycle may be hard to distinguish from a non-stable transient).

Evaluate. The user assesses the usefulness of the culturally evolved objects for the agents, as well as object diversity, and wastefulness.

EVOC2 will be deemed a success if it not only evolves cultural novelty that is cumulative, adaptive, and open-ended (as in EVOC with RR), but also (a) *restructures* conceptions of objects by viewing them from different perspectives (new contexts), (b) generates inventions that open up niches for other inventions, and (c) exhibits contextual focus, i.e., shifts to an associative mode to restructure and shifts back to fine-tune. It is hypothesized that these features will increase the complexity of economic webs of objects and recycled wastes.

Elucidating the Algorithmic Structure of Biological Versus Cultural Evolution

The design features that made EVOC2 specific to the problem of waste recycling can eventually be replaced by general-purpose counterparts, resulting in a *cultural algorithm* (CAL[1]). It will be interesting to **c**ompare the performance of a CAL with a GA on standard problems (e.g., the Rosenbrock function) as well as on insight tasks such as real-world waste recycling webs that require restructuring. Waste recycling is a particularly appropriate application because it explicitly requires considering how the same item offers a different set of constraints and affordances when considered with respect to a different goal, a different demographic, or a different aesthetic sensibility (one person's trash is another person's treasure). In general the CAL is expected to outperform the GA on problems that involve not just multiple *constraints* but multiple *perspectives,* e.g., economic and environmental.

[1] Cultural algorithm is abbreviated CAL because CA customarily refers to cellular automaton.

A long-term objective is to develop an integrated framework for evolutionary processes that encompasses natural selection, cultural evolution, and communal exchange theories of early life. Another long-term objective is to advance knowledge of how systems evolve. Early efforts toward a general cross-disciplinary framework for evolution Processes were modeled as *context-dependent actualization of potential*: an entity has potential to change various ways, and how it *does* change depends on the contexts it interacts with (Gabora and Aerts 2005, 2008). These efforts focused on distinguishing processes according to the degree of non-determinism they entail, and the extent to which they are sensitive to, internalize, and depend upon a particular context. With the sorts of tools outlined here, it will be possible to compare the effectiveness of communal exchange, Darwinian, and mixed strategies in different environments (simple versus complex, static versus fluctuating, and so forth. This will result in a more precise understanding of the similarities and differences between biological and cultural evolution, and help us recognize other evolutionary processes that we may discover as science penetrates ever deeper into the mysteries of our universe.

Summary and Conclusions

Culture evolves with breathtaking speed and efficiency. We are crossing the threshold to an exciting frontier: a scientific understanding of the process by which cultural change occurs, as well as the means to capitalize on this understanding. The cultural evolution inspired human computation program of research described in this chapter is ambitious and interdisciplinary, but it builds solidly on previous accomplishments.

We examined evidence that culture evolves through a non-Darwinian communal exchange process, and discussed a plan for modeling the autopoietic structures that evolve through biological and cultural processes—i.e., metabolic reaction networks and associative networks. This will make it possible to undertake a comparative investigation of the dynamics of communally exchanging groups of these two kinds of networks. This research is necessary to achieve a unification of the social and behavioral sciences comparable to Darwin's unification of the life sciences.

Efforts are underway toward the development of a computational model of cultural evolution that incorporates the kind of sophisticated cognitive machinery by which cultural novelty evolves. These include the combining of concepts to give rise to new concepts sometimes with emergent properties, and the capacity to shift between different modes of thought depending on the situation. An important step is to embed formal models of concepts in a modified "reaction network" architecture, in order to computationally model how clusters of interrelated concepts modify one another to achieve a more stable lower energy state, through a process we referred to as *context-driven restorative restructuring*. Efforts are also underway toward the development of a computer program for identifying patterns of historical relationship amongst sets of artifacts. Human input is used to define *contexts*—perspectives or situations that define which features or attributes are potentially relevant. One long-term objective of this kind of research program is to develop a cultural algorithm: an optimization and problem-solving tool inspired by cultural evolution. This will allow us to investigate how strategies for recursively re-processing and restructuring information, or shifting between different processing modes, affect the capacity to evolve cumulative, adaptive, open-ended novelty.

The ideas presented in this chapter are speculative, ambitious, and innovative both conceptually and methodologically, but they have far-reaching implications and potentially diverse applications. The human computation program proposed here could promote a scientific understanding of the current accelerated pace of cultural change and its transformative effects on humans and our planet. It may foster cultural developments that are healthy and productive in the long term as well as the short term, and help us find solutions to complex crises we now face.

Acknowledgements This research was conducted with the assistance of grants from the National Science and Engineering Research Council of Canada, and the Fund for Scientific Research of Flanders, Belgium.

References

Aerts D (2009) Quantum structure in cognition. J Math Psychol 53:314–348

Aerts, D., Aerts, S., & Gabora, L. (2009). Experimental evidence for quantum structure in cognition. In: P. Bruza, W. Lawless, K. van Rijsbergen, & D. Sofge (Eds.) Lecture Notes in Computer Science: Quantum Interaction (pp. 59–79). Berlin: Springer

Aerts D, Gabora L, Sozzo S (in press) How concepts combine: a quantum theoretic model. Topics in Cognitive Science

Aerts D, Gabora L (2005a) A state-context-property model of concepts and their combinations I: the structure of the sets of contexts and properties. Kybernetes 34(1&2):151–175

Aerts D, Gabora L (2005b) A state-context-property model of concepts and their combinations II: a Hilbert space representation. Kybernetes 34(1&2):176–205

Aerts D, Aerts S, Gabora L (2009a) Experimental evidence for quantum structure in cognition. In: Bruza P, Lawless W, van Rijsbergen K, Sofge D (eds) Proceedings of the third international conference on quantum interaction. German Research Center for Artificial Intelligence, Saarbruken, pp 59–70

Aerts D, Czachor M, De Moor B (2009b) Geometric analogue of holographic reduced representation. J Math Psychol 53:389–398

Aerts D, Broekaert J Gabora L, Veloz T (2012) The guppy effect as interference. In: Proceedings of the sixth international symposium on quantum interaction, Paris, 27–29 June

Atran S (2001) The trouble with memes: inference versus imitation in cultural creation. Hum Nat 12:351–381

Bentley PD, Corne D (eds) (2002) Creative evolutionary systems. Morgan Kaufmann, San Francisco

Bentley RA, Hahn MW, Shennan SJ (2004) Random drift and cultural change. Proc R Soc Br Biol Sci 271:1143–1450

Boden MA (1990/2004) The creative mind: Myths and mechanisms, 2nd edn. Routledge, London

Boyd R, Richerson P (1985) Culture and the evolutionary process. University Chicago Press, Chicago

Boyd R, Richerson P (2005) The origin and evolution of cultures. Oxford University Press, Oxford

Cavalli-Sforza LL, Feldman MW (1981) Cultural transmission and evolution: a quantitative approach. Princeton University Press, Princeton

Dantzing SV, Raffone A, Hommel B (2011) Acquiring contextualized concepts: a connectionist approach. Cogn Sci 25:1162–1189

Dawkins R (1976) The selfish gene. Oxford University Press, Oxford

DiPaola S, Gabora L (2009) Incorporating characteristics of human creativity into an evolutionary art algorithm. Genet Program Evolvable Mach 10(2):97–110

Dittrich P, Speroni di Fenizio P (2008) Chemical organization theory. Bull Math Biol 69:1199–1231

Dittrich P, Winter L (2007) Chemical organizations in a toy model of the political system. Adv Complex Syst 1(4):609–627

Dittrich P, Ziegler J, Banzhaf W (2001) Artificial chemistries—a review. Artif Life 7(3):225–275

Donald M (1991) Origins of the modern mind. Harvard University Press, Cambridge

Durham W (1991) Coevolution: genes, culture, and human diversity. Stanford University Press, Stanford

Eliasmith C, Thagard P (2001) Integrating structure and meaning: a distributed model of analogical mapping. Cogn Sci 25:245–286

Fracchia J, Lewontin RC (1999) Does culture evolve? Hist Theory 38:52–78

Gabora L (1995) Meme and variations: a computer model of cultural evolution. In: Nadel L, Stein D (eds) 1993 lectures in complex systems. Addison-Wesley, Boston, pp 471–486

Gabora L (1996) A day in the life of a meme. Philosophica 57:901–938

Gabora L (1998) Autocatalytic closure in a cognitive system: a tentative scenario for the origin of culture. Psycoloquy 9(67) [adap-org/9901002]

Gabora, L. (2000). Conceptual closure: Weaving memories into an interconnected worldview. In (G. Van de Vijver & J. Chandler , Eds.) Closure: Emergent Organizations and their Dynamics. Annals of the New York Academy of Sciences, 901, 42–53

Gabora, L. (2001). Cognitive mechanisms underlying the origin and evolution of culture. Doctoral Dissertation, Free University of Brussels

Gabora L (2003) Contextual focus: a cognitive explanation for the cultural transition of the middle/upper Paleolithic. In: Alterman R, Hirsch D (eds) Proceedings of the 25th annual meeting of the cognitive science society. Lawrence Erlbaum, Boston, pp 432–437

Gabora, L. (2004). Ideas are not replicators but minds are. Biology & Philosophy, 19(1), 127–143

Gabora L (2006a) The fate of evolutionary archaeology: survival or extinction? World Archaeol 38(4):690–696

Gabora L (2006b) Self-other organization: why early life did not evolve through natural selection. J Theor Biol 241(3):443–250

Gabora L (2008a) The cultural evolution of socially situated cognition. Cogn Syst Res 9(1):104–113

Gabora, L. (2008b). EVOC: A computer model of the evolution of culture. In V. Sloutsky, B. Love & K. McRae (Eds.), 30th Annual Meeting of the Cognitive Science Society. July 23–26, Washington DC (pp. 1466–1471). North Salt Lake, UT: Sheridan Publishing

Gabora, L. (2008c). Modeling cultural dynamics. Proceedings of the Association for the Advancement of Artificial Intelligence (AAAI) Fall Symposium 1: Adaptive Agents in a Cultural Context,. Nov 7-9, The Westin Arlington Gateway, Arlington VA, (pp. 18–25). Menlo Park, CA: AAAI Press

Gabora, L. (2010). Recognizability of creative style within and across domains: Preliminary studies. Proceedings of the Annual Meeting of the Cognitive Science Society (pp. 2350–2355). August 11–14, Portland, OR

Gabora L (2011) Five clarifications about cultural evolution. J Cogn Cult 11:61–83

Gabora, L. (2013). An evolutionary framework for culture: Selectionism versus communal exchange. Physics of Life Reviews, 10(2), 117–145

Gabora L, Aerts D (2002) Contextualizing concepts. In: Proceedings of the 15th international FLAIRS conference (special track ‘Categorization and concept representation: models and implications’, Pensacola Florida, American Association for Artificial Intelligence, pp 148–152, 14–17 May

Gabora L, Aerts D (2002b) Contextualizing concepts using a mathematical generalization of the quantum formalism. J Exp Theor Artif Intell 14(4):327–358

Gabora L, Aerts D (2005) Evolution as context-driven actualization of potential: toward an interdisciplinary theory of change of state. Interdiscip Sci Rev 30(1):69–88

Gabora, L., & Aerts, D. (2008). A cross-disciplinary framework for the description of contextually mediated change. In (I. Licata & A. Sakaji, Eds.) Physics of Emergence and Organization, (pp. 109–134). Singapore: World Scientific.

Gabora, L., & Aerts, D. (2009). A model of the emergence and evolution of integrated worldviews. Journal of Mathematical Psychology, 53, 434–451

Gabora L, DiPaola S (2012) How did humans become so creative? In: Proceedings of the international conference on computational creativity, Dublin, Ireland, pp 203–210, May 31–June 1

Gabora, L., & Firouzi, H. (2012). Society functions best with an intermediate level of creativity. Proceedings of the 34th Annual Meeting of the Cognitive Science Society (pp. 1578–1583). Held August 1–4, Sapporo Japan. Houston TX: Cognitive Science Society

Gabora, L. & Leijnen, S. (2009). How creative should creators be to optimize the evolution of ideas? A computational model. Electronic Proceedings in Theoretical Computer Science, 9, 108–119

Gabora L, Ranjan A (2013) How insight emerges in distributed, content-addressable memory. In: Bristol A, Vartanian O, Kaufman J (eds) The neuroscience of creativity. MIT Press, New York

Gabora L, Saberi M (2011) How did human creativity arise? An agent-based model of the origin of cumulative open-ended cultural evolution. In: Proceedings of the ACM conference on cognition and creativity, Atlanta, 3–6 November 2011

Gabora L, Leijnen S, Veloz T, Lipo C (2011) A non-phylogenetic conceptual network architecture for organizing classes of material artifacts into cultural lineages. In: Proceedings of the Annual Meeting Cognition Science Society, Boston, 20–23 July 2011

Gabora L, O'Connor B, Ranjan A (2012) The recognizability of individual creative styles within and across domains. Psychol Aesthet Creativity Arts 6(4):351–360

Gabora, L., & Saab, A. (2011). Creative interference and states of potentiality in analogy problem solving. Proceedings of the 33rd Annual Meeting of the Cognitive Science Society (pp. 3506–3511). July 20–23, Boston MA

Gabora, L., Scott, E., & Kauffman, S. (2013). A quantum model of exaptation: Incorporating potentiality into biological theory. Progress in Biophysics & Molecular Biology, 113(1), 108–116

Gagné CL, Spalding TL (2009) Constituent integration during the processing of compound words: does it involve the use of relational structures? J Mem Lang 60:20–35

Goldberg DE, Kuo CH (1987) Genetic algorithms in pipeline optimization. J Comput Civ Eng ASCE 1(2):128–141

Hampton J (1987) Inheritance of attributes in natural concept conjunctions. Mem Cogn 15:55–71

Henderson, M. & Gabora, L. (2013). The recognizability of authenticity. Proceedings of the 35th Annual Meeting of the Cognitive Science Society (pp. 2524–2529). Held July 31–Aug. 3, Berlin. Houston TX: Cognitive Science Society.

Henrich J, Boyd R (1998) The evolution of conformist transmission and the emergence of between-group differences. Evol Hum Behav 19:215–242

Henrich J, Boyd R (2002) On modeling cognition and culture: why replicators are not necessary for cultural evolution. J Cogn Cult 2:87–112

Holland J (1975) Adaptation in natural and artificial systems. MIT Press, Cambridge

Hou ESH, Ansari N, Ren H (1994) A genetic algorithm for multiprocessor scheduling. IEEE Trans Parallel Distrib Syst 5(2):113–120

Kaplan CA, Simon HA (1990) In search of insight. Cogn Psychol 22:374–419

Kauffman S (1993) Origins of order. Oxford University Press, New York

Kitto K, Ramm B, Sitbon L, Bruza PD (2011) Quantum theory beyond the physical: information in context. Axiomathes 12(2):331–345

Koza J (1993) Genetic programming. MIT Press, London

Leijnen S, Gabora L (2010) An agent-based simulation of the effectiveness of creative leadership. In: Proceedings of Annual Meeting Cognitive Science Society. Portland, pp 955–960, 11–14 August 2010

Longo G, Montevil M, Kaufman S (2012) No entailing laws, but enablement in the evolution of the biosphere. In: Proceedings of the fourteenth international conference on genetic and evolutionary computation, pp 1379–1392

O'Brien MJ, Lyman RL (2000) Applying evolutionary archaeology: a systematic approach. Kluwer, Norwell

Ohlsson S (1992) Information-processing explanations of insight and related phenomena. In: Keane MT, Gilhooly KJ (eds) Advances in the psychology of thinking, vol 1. Harvester Wheatsheaf, New York, pp 1–44

Osherson D, Smith E (1981) On the adequacy of prototype theory as a theory of concepts. Cognition 9:35–58

Reynolds RG (1994) An introduction to cultural algorithms. In: Proceedings of the 3rd annual conference of evolutionary programming, World Scientific, River Edge, pp 131–139

Riley, S. & Gabora, L. (2012). Evidence that threatening situations enhance creativity. Proceedings of the 34th Annual Meeting of the Cognitive Science Society (pp. 2234–2239). Held August 1–4, Sapporo Japan. Houston TX: Cognitive Science Society

Runco M (2010) Divergent thinking, creativity, and ideation. In: Kaufman J, Sternberg R (eds) The Cambridge handbook of creativity. Cambridge University Press, Cambridge, pp 414–446

Shennan S (2008) Evolution in archaeology. Annu Rev Anthropol 37:75–91

Skoyles JR (2008) Natural selection does not explain cultural rates of change. Proc Natl Acad Sci 105(22):E27–E27

Tëmkin I, Eldredge N (2007) Phylogenetics and material cultural evolution. Curr Anthropol 48:146–153

Thagard P, Stewart TC (2011) The AHA! experience: creativity through emergent binding in neural networks. Cogn Sci 35:1–33

Veloz T, Gabora L, Eyjolfson M, Aerts D (2011) A model of the shifting relationship between concepts and contexts in different modes of thought. In: Proceedings of the fifth international symposium on quantum interaction, Aberdeen, 27 June 2011

Veloz T, Tëmkin I, Gabora L (2012) A conceptual network-based approach to inferring cultural phylogenies. In: Proceedings of the annual meeting of the cognitive science society, Sapporo, 2012

Vetsigian K, Woese C, Goldenfeld N (2006) Collective evolution and the genetic code. Proc Natl Acad Sci 103:10696–10701

Williams RJP, Frausto da Silva JJR (2003) Evolution was chemically constrained. J Theor Biol 220:323–343

Woese CR (2002) On the evolution of cells. Proc Natl Acad Sci 99:8742–8747

Collective Search as Human Computation

Winter Mason

Introduction

As the world has grown and become more connected, with the majority of the world living in population-dense environments, the days of the lone inventor, renaissance man, and independent craftsman are waning. Increasingly, navigating the complex problems that face complex societies require collective efforts, sometimes at a massive scale.

One way complex problems are solved collectively is through competition, working separately and closely guarding information about steps towards a solution. Although this may appear on the surface to be antithetical to the idea of collective problem solving, because *many* individuals are working towards the same goal, it is a collective effort, and in fact it has been discovered that competitions are an effective way of producing high-quality solutions to problems (e.g., Innocentive.[1]) Another way groups solve problems is for the group to work collaboratively, which typically involves advanced planning and frequent communication between members. Most people think of this form when considering collective problem solving: the small team of individuals, sitting around a table brainstorming and developing solutions. In fact, this form has been extensively studied in a variety of literatures, including psychology and management science (Marquart 1955; Osborn 1957; Stasser and Titus 1985). However, many forms of collective problem solving fall somewhere in between these two extremes. In some situations there are multiple individuals seeking to discover the best solution to a problem, but they are freely sharing information

[1] http://www.innocentive.com/

W. Mason (✉)
Stevens Institute of Technology, Hoboken, NJ 07030, USA
e-mail: m@winteram.com

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_35,

because it is to the collective benefit if anyone discovers the solution. For example, this is true for non-profits trying to find the best solution to disaster management and recovery. Collective search is a particular type of collective problem solving that often takes this intermediate form, with individuals sharing information about their search process (where they are and what they have discovered) in order to maximize the effectiveness of the collective.

When thinking about search, the relationship to problem solving may not be immediately obvious. In fact, it is likely that the first thing that enters your mind when considering "search" is either the perennial quest to find one's keys or a text entry field in your internet browser. However, *search* can be construed more broadly than this, as the act of sampling the parameter space of an unknown objective function that maps the input parameters to values. While this may not be how one normally thinks about looking for lost keys, one can consider points in physical space to be the input parameters (x,y,z) with an objective function that is zero everywhere except the location where one's keys happen to exist. For web search—from the perspective of the user—the input parameters would be the query terms, and the objective function would then map the features of the returned documents to a "relevance" value that indicates how useful the document is to the user that issued the query. Even subjective problems can be framed this way—only the value returned by the unknown objective function is different for each person: two people searching the fridge for the best midnight snack might come away with two completely different items but be equally happy with the result.

Perhaps the clearest connection between search and problem solving can be seen in Newton's method for calculating a square root. If one wanted to know the solution to $x = \sqrt{2704}$, it is possible to calculate it exactly using digit-by-digit calculation. However, it is also possible to search for the solution: one begins with some random initial guess x_0, and then using the formula $x_{i+1} = \frac{f(x_i)}{f'(x_i)}$ this initial guess is iteratively updated to get progressively closer to the true square root of the number—in this case, 52. Here, the input parameter is the guess x_i and the objective function could be, for instance, to find the minimum of $|x_i^2 - 2{,}704|$ (which is minimized when the true root has been found). The formula used in Newton's method is the algorithm by which one explores for the solution, the decision about what point in the space to sample next.

The essence of computation is "the process of mapping some input representation to some output representation using an explicit, finite set of instructions" (Law and von Ahn 2011). *Human computation* is simply computation executed by one or more humans. Therefore, the relationship between collective search and human computation is a direct one—when humans are searching, they are doing human computation. This is true for individuals, but the algorithms for search become more complex and in many ways more interesting when search is executed by a collective.

In this chapter I consider the factors that affect collective search, and review different approaches to understanding collective search, both how it is done normally and how it can be done best. I conclude by highlighting important open questions and sketching out directions for future research.

Factors Affecting Collective Search

When it comes to studying collective search, there are many things that affect how search is conducted, and the probability of success and efficiency with which the solution is found (or the value is improved). These factors include the form of the objective function that maps parameters to values, the means by which an individual can sample the problem space, the nature and structure of communication between individuals, and the motivations/incentives for the individuals.

Defining Problems

An objective function is made up of three key parts: (1) the input parameters, (2) the form of the function, and (3) the output value that is being maximized (or minimized). Characterizing the input parameters of a search task superficially appears trivial, as they are the variables that can be manipulated in order to obtain a value from the unknown objective function. However, in most real problems the dimensionality is unknown or arbitrary. In other words, in many cases one does not know which variables relate to the output variable of interest: this is related to the well-known problem of feature selection (Langley 1994). In some search tasks the input variables are easily identifiable; when looking for one's keys, the only relevant variables are the spatial dimensions. In others it is much more difficult; in the search for a cure for a disease, the possible inputs are nearly without limit and identifying the relevant inputs is itself a formidable challenge. Additionally some so-called "wicked" problems are nearly impossible to decompose into separable input dimensions, something discussed towards the end of the chapter.

Once the inputs have been identified, each has a particular domain. The correct dosage for a particular medication in the treatment for a disease is an input parameter that has a domain in the positive real numbers. The type of medication would be a categorical variable with a domain equal to the set of possible medications. The size of the input set is the dimensionality of the problem space. It goes without saying that higher-dimensional problems are, all else being equal, more difficult to solve than lower-dimensional problems. However, low-dimensional problems with input parameters that have larger domains could be more difficult to solve than high-dimensional problems with small-domain input parameters. Moreover, the forms of the input parameters constrain the possible approaches to sampling and the algorithms that can be used to explore the problem space.

The form of the objective function—that is, how the input parameters are mapped to the output value—is necessarily unknown in a search task. However, there are qualities about the mapping function that also constrain the types of search that can be done. For instance, the mapping from potential solutions to payoffs does not have to be a deterministic process, but for any search process to be more effective than random combinations of the input parameters, there must be some signal that can be used to guide the searcher.

Another key feature of the objective function is the *smoothness* or *complexity* of the problem space. In optimization, many algorithms depend on a feature known as convexity, which means that any local minimum in the function is also a global minimum. These problems are generally solvable with "hill-climbing" algorithms that search the space by looking for any small change in the input parameters that results in a better payoff than the current solution—for instance, Newton's method for finding roots. However, this characteristic is not common to many problems in practice; on the contrary, real problems likely have many local minima that are not the globally best solution.

Finally the output value is in some ways the most important part of a search problem. Choosing the right output value to minimize (or maximize) determines whether the success of the search task translates to the actual desired outcome. This is particularly true in human computation, when the alignment of output to incentives can be critical.

Sampling the Problem Space

Exploration of a problem space involves sampling the problem space and receiving information about the value from the objective function. The actual process of selecting points in the parameter space is the subject of voluminous literature, from algorithmic optimization (Boyd and Vandenberghe 2004) to animal foraging (Kennedy and Gray 1993; Roberts and Goldstone 2006) to human collective search (Hills et al. 2008; Rendell et al. 2010; Pirolli and Card 1995). This research is mostly focused on either a *descriptive* analysis of how search actually happens—that is, what signals individuals use in practice to navigate an unknown problem space—or a *proscriptive* analysis of what methods work best, given assumptions about the objective function.

In some cases, there may also be constraints on how an individual can sample the problem space. It may be that one can only change a single parameter value at a time; the input space may be distorted, so the domain of one input depends on the value of another; there could be sequential dependencies, so the available input values for one sample may be limited by the values from the previous sample. Any of these constraints affect both the way in which one is likely to sample the problem space as well as the optimal methods for finding the global optimum (minimum or maximum) of the function.

Computer science research on optimization highlights another connection between search and general principles of computation. An optimization problem is represented as a function f that maps from set A to the real numbers—this is the objective function—with the goal of finding x' where $f(x') \leq f(x)\ \forall x \in A$. The methods developed to solve optimization problems depend on the form of the objective function. For example, objective functions that are linear versus quadratic require different approaches from functions that have a discrete solution set, which require different approaches from functions that have random variables or functions that have infinite dimensions.

Collective search can be seen as a sub-problem of optimization with the additional requirements that sampling the objective function is possible by humans (which many optimization problems are, particularly those within operations research (Heyman and Sobel 2003)), and that information about samples from the objective function can be shared between searchers. This latter component adds many interesting layers of complexity to the problem.

Because of this complexity, a substantial amount of research on collective search has leveraged agent-based models (ABMs; a type of computational model that encodes the behavior of individuals and observes the high-level outcomes of their interactions) to understand how individual decision-making can lead to different kinds of collective outcomes (Lazer and Friedman 2007; Roberts and Goldstone 2006; Rendell et al. 2010). One such influential study on the relationship between collective search and problem spaces was Levinthal (1997), who used a particular type of objective function, the *N-K problem space* (Kauffman 1993). This function has two parameters, *N* that defines the size of the problem space, and *K* that defines the complexity of the problem space. Levinthal was concerned with organizational change, but the basic features of collective search are there: sampling a problem space through specific input parameter values, an objective function that maps the input to a value, communication between collective agents, and an aggregating function for the agents' output. From the perspective of the study, the agents represented organizations, and one of the key findings is that as the number of local optima increase, so do the number of "organizational forms". This is to say, more complex problems lead to a wider variety of solutions. Additionally, in Levinthal's model the more local optima that existed, the longer it took for the organizations to converge on the global optimum. This has echoes of March's (1991) model of organizational learning which illustrated the tradeoff between exploring for new, better solutions to a problem versus exploiting the best known solution. To quote, the returns on exploitation are "positive, proximate, and predictable", while the returns on exploration are "uncertain, distant, and often negative" but are necessary to exceed the current standard and to make large strides in competitive advantage.

One approach that uses collective search for optimization is the particle swarm algorithm (Kennedy and Eberhart 1995). The particle swarm algorithm is a set of proscribed behaviors for agents searching a problem space with communication between agents, and the objective of converging on the global optimum in the problem space. The premise of the algorithm is that a collection of agents are placed randomly in a (multi-dimensional) problem space with a random direction and velocity. The agents are drawn to the best location they have found, as well as the best location found by any member of the collective, and their course and velocity in the problem space are adjusted incrementally towards these targets. The algorithm also must account for the tradeoff in exploration and exploitation, which is accomplished by the parameter that defines the strength with which the agents are drawn to the best found solutions: the stronger it is, the less likely they are to find the global optimum—though if they do, they find it quickly. Conversely, the weaker it is, the more likely the agents are to find the global maximum, but they take longer to do so. Work on this algorithm has explored many variations with the goal of improving efficiency and optimality, but there are no guarantees of convergence on the global maximum.

Nature and Structure of Communication

The factors that affect collective search discussed so far apply equally to individual search and collective search. When a *collective* is searching the problem space, however, there are two additional features that affect the performance of the collective: how the collective communicates during the search process, and how the outcomes of the collective are aggregated. For instance, consider a group of people brainstorming a solution to a product design challenge. Each individual comes up with different variations of the design given the goals and constraints (sampling the input space) and shares these potential solutions with every other member of the group. The aggregation function for brainstorming selects the best choice (where, for instance, the objective function is defined by the group). In this case the communication process can be modeled as a complete graph—all members are communicating with all other members—and the aggregation function selects the optimal solution.

Evaluating the output of the collective is also different than for a single individual. For nearly all tasks, many individuals working towards the same goal will be more effective than a single individual working towards that goal. Thus, when considering collective search the key question is whether the collective performance is greater than, equal to, or less than the sum of their individual outputs. Steiner (1972) suggested tasks could be classified in this manner: *additive* tasks are those in which the output of a group working on the task is equal to the sum of their parts, such as laying bricks; *compensatory* tasks are those in which the collective output is equivalent to the average performance of the members, such as a university being evaluated by the average number of citations of each faculty member; *disjunctive* tasks are those in which the performance of the group is as good as the best performing individual in the group, such as teams in track & field competitions that are decided by the fastest team member; *conjunctive* tasks are those in which the performance of the task is equivalent to the worst performing individual. For the brainstorming example, because the aggregating function selects the best performing individual—who came up with the best solution—it is clearly a disjunctive task. Not mentioned in Steiner's typology are tasks that have output better than the best individual (or worse than the worst individual). These types of tasks certainly exist, but are sufficiently rare as to have escaped notice.

The aggregating function is not the only determinant of the group's performance—how the members communicate with each other also affects the relationship between individual effort and group output. For instance, in brainstorming it has been shown that the all-to-all communication pattern can actually decrease the performance of the group, because of *production blocking* (Diehl and Stroebe 1987), in which the ideas of one person interferes with the thought process of another, stymieing that individual's process of sampling the problem space. On the other hand, if there is a barrier in communication between individuals, the best solution in the group may not be available to other members of the group.

Following Levinthal's study (Levinthal 1997) on collective search, Lazer and Friedman (2007) had simulated agents explore the N-K problem space, but in this case the agents were embedded in a communication network. The agents started at a random location in the problem space, and each turn could choose to sample a location within a small distance from their current position; if this location returned a higher score, the agent would move to that location, thereby executing a hill-climbing search over the parameter space (akin to convex optimization algorithms). However, the agents were also connected to a subset of other agents, and if the agent's network neighbors had a higher score, the agent would "jump" through the problem space to somewhere near the successful neighbor's location. By varying the network structure connecting the agents, Lazer and Friedman could see how information transmission affected the overall success of the group, both in terms of the speed with which the group converged on a solution and the ultimate score that the groups obtained. One of the key features of the networks they studied is the average *path length* between nodes, which is the average number of hops between nodes in the network. The famous "six degrees of separation" refers to the fact that the average path length between people is six (though on Facebook it seems to be only four (Backstrom et al. 2012)). In this model, they found that rapid information transmission, obtained through the short path lengths separating nodes, precise copying of information, or frequent copying of information, led to better performance in the short run. In contrast, slow transmission of information via long path lengths, errorful copying of information, or infrequent copying of information, led to better performance in the long run. This result arises from the fact that in complex problem spaces, insufficient exploration can lead to premature convergence on a suboptimal point, and the behavior of the agents guaranteed convergence on some point.

In subsequent experimental work, my colleagues Rob Goldstone and Andy Jones and I found similar effects (Mason et al. 2008). In our study, groups of individuals embedded in a network explored several different one-dimensional problem spaces, and the solutions were shared with each individual's network neighbors. There were three different problem spaces in this study: a simple unimodal Gaussian distribution, a multimodal Gaussian distribution with two local maxima and one global maximum, and a "needle" function with one hard-to-find global maximum and one easy-to-find local maximum (see Fig. 1). The objective function had added stochastic noise, so participants had to make repeated samples to determine the true value of a point in the problem space. The networks tested in this study were similar to those studied in Lazer and Friedman (2007): a cycle/lattice-like network, a fully-connected network, a random network, and a small-world network (Watts and Strogatz 1998) that is more clustered than a random graph but has shorter average path lengths than the cycles. In this study, we found that the full-connected network

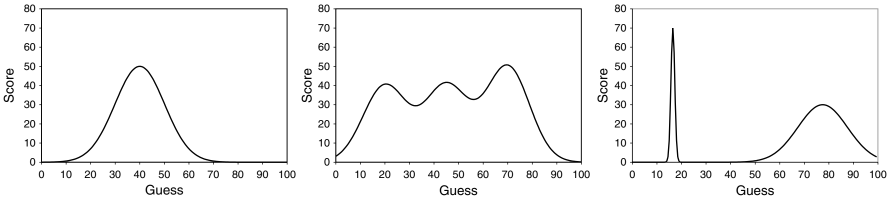


Fig. 1 Three example problem spaces explored by participants in Mason et al. (2008)

did best on the simple problem space, because rapid convergence to any maximum mean rapid convergence to the global maximum. In the multimodal problem space, however, the small-world network outperformed all other networks. In the "needle" problem space, the network with the longest path length—the lattice—outperformed the rest because the participants were more likely to find the global maximum. Again, the explanation is that the networks with the longer path lengths slowed the diffusion of suboptimal solutions long enough that other individuals weren't drawn to the local optimum before they found the global optimum.

However, in more recent work, with a problem space that qualitatively possessed all of the features hypothesized to generate this effect of long path lengths—a rugged landscape with local optima that can be found through local exploration (e.g., hill-climbing approaches)—my colleague Duncan Watts and I (Mason and Watts 2012) found the opposite pattern, that networks with small path lengths performed better. In this study, participants were playing a game called "Wildcat Wells" in which they were told they were exploring a landscape for oil. They had 15 rounds to explore, and received points on each round based on how much "oil" they found. Underlying their exploration was a two-dimensional problem space that was generated by first creating a single bivariate Gaussian distribution with fixed radius and random mean, which was then combined with a Perlin noise distribution; and example is shown in Fig. 2. As can be seen, the resulting fitness landscape had many local maxima that would be basins of attraction for purely local exploration, and one global maximum that was clearly superior to other solutions. The participants in the study were embedded in one of eight different fixed-degree networks that varied in clustering, path length, betweenness, and network constraint (Burt 1992). While there are many differences between this study and the prior work, qualitatively the set up seemed similar enough that we expected to replicate the findings, and were surprised when we did not. This suggests there must be some factor unaccounted for that changed the outcome. While it could be something as simple as a feature of the UI in the game, we believe the most likely explanation is that the complexity of the problem space—which is to say, the potential for local maxima—is not a sufficient condition for networks with longer path lengths to outperform networks with shorter path lengths in the long run. An important question for future work on collective search is determining what factors, other than the communication network, lead members of a collective exploit the currently best-known solution or continue to explore for better solutions, and what the relative effect size is for these factors.

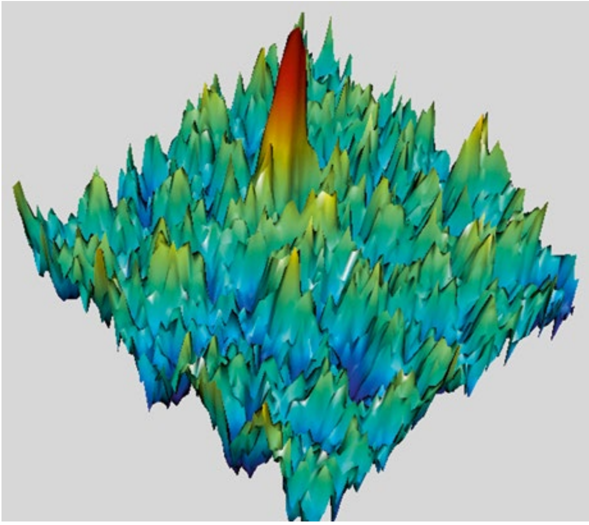


Fig. 2 Graphical representation of the problem space being explored in the "Wildcat Wells" game

Incentives for Exploration and Exploitation

One important factor that has not been explored much in collective search is the role of incentives in motivating individuals to explore new solutions versus exploit existing ones. In fact, one potential explanation for the difference between my study with Duncan Watts (Mason and Watts 2012) and previous work (Mason et al. 2008; Lazer and Friedman 2007) could be related to the participants' relative certainty or uncertainty about the opportunities for exploration (e.g., the number of rounds participants had to explore) or the relative difference in value between the local maxima and the global maximum. In other words, the perceived value of exploration versus exploitation could be critical to the effect of path length on the short-term or long-term success of the collective.

Aside from the implications for the effects of networked groups, however, it is clear that the motivations of the individuals in the collective are critical for successful search. One of the consequences of sharing information in collective search, which we observed in Mason and Watts (2012) was a social dilemma: the individuals who merely imitated others tended to outperform the individuals who did more exploration, yet groups that had fewer individuals exploring (and therefore more imitating) were less likely to find the global optimum. The same effect was observed in a competition centered around collective search (Rendell et al. 2010), in that the most successful agents were those that heavily copied other agents, but the same agents failed miserably when there were no explorers to copy.

Thus incentives must be structured appropriately to encourage sufficient exploration (to avoid early convergence on a local maximum) yet allow for exploitation of the global maximum when it is found. The incentives must also guard against free-riding behavior observed in most social dilemmas. How this is best accomplished, however, is still an open question.

Future Directions

There are a number of open questions about collective search, particularly regarding its connection to human computation. To date, most research on collective search has focused on one of two questions: what is the optimal way to coordinate collective search, and how do people actually engage in collective search? And for both of these questions there are a number of directions to explore.

A potentially productive avenue of research could connect the algorithms on optimization with collective search. Research on algorithms that converge rapidly and accurately on the global maximum that can be applied in situations in which human collectives are searching a problem space could have an enormous impact on all sorts of organizations. Re-framing team problem solving into a search task and applying the computer science research on optimization could potentially lead to great advancements in the efficiency with which groups of people solve problems. Of course, there are inevitable human elements that may confound the results in the computer science research on optimization, but that too is an interesting direction for research to explore.

All of the questions and approaches discussed so far assume that the individuals or collectives searching the problem space already know which inputs are relevant to the problem. The dimensions of the problem space being explored are defined, in most cases, by the searchers themselves. When trying to find the right combination and dosages of medications to treat a case of chronic hypertension, the doctor generally begins by choosing which medications to explore. However, this problem of feature selection is crucial to determining whether the end result is successful and how successful it is. In machine learning, the problem of feature selection was identified as a fundamental problem (Langley 1994), and although some automatic approaches have been developed, it is still a wide-open topic.

In fact, most algorithms for automatic feature selection assume there is some large set of known features, and the relevant subset must be identified. However, in real problems even just finding the potential set of features from which to select a relevant subset is its own problem. It is an interesting possibility that algorithms developed for search might be used to tackle this meta problem of finding feature candidates from an unknown (possibly infinite) set of features.

As mentioned earlier, not much research has explored the role of incentives for individuals involved in collective search. Understanding how to structure these incentives for exploration and exploitation appropriately could be an interesting avenue for research on mechanism design.

Conclusion

Many problems can be mapped into the search domain, by considering the available inputs as dimensions in a problem space. Doing this allows one to leverage a deep understanding of how people engage in search and how to best search a non-linear problem space. Moreover, this search can be improved, with the appropriate communication network and incentives, by allowing multiple individuals to search simultaneously and share information about potential solutions and outcomes. The area of collective search is therefore a promising way of exploring what is possible in collective search and human computation generally.

References

Backstrom L, Boldi P, Rosa M, Ugander J, Vigna S (2012) Four degrees of separation. In: WebSci'12: proceedings of the 3rd annual ACM web science conference, Evanston, June 2012. ACM Request Permissions

Boyd S, Vandenberghe L (2004) Convex optimization. Cambridge University Press, Cambridge/New York

Burt RS (1992) Structural holes: the social structure of competition. books.google.com

Diehl M, Stroebe W (1987) Productivity loss in brainstorming groups: toward the solution of a riddle. J Personal Soc Psychol 53(3):497–509

Heyman DP, Sobel J (2003) Stochastic models in operations research: stochastic optimizations. Dover books on Computer science series. Dover, New York

Hills TT, Todd PM, Goldstone RL (2008) Search in external and internal spaces evidence for generalized cognitive search processes. Psychol Sci 19(8):802–808

Kauffman SA (1993) The origins of order: self organization and selection in evolution. Oxford University Press, New York

Kennedy J, Eberhart R (1995) Particle swarm optimization. In: Proceedings of the IEEE international conference on neural networks, Piscataway

Kennedy M, Gray RD (1993) Can ecological theory predict the distribution of foraging animals? A critical analysis of experiments on the ideal free distribution. Oikos 68(1):158–166

Langley P (1994) Selection of relevant features in machine learning. In: AAAI fall symposium, Defense Technical Information Center. New Oreleans, LA, USA

Law E, Ahn LV (2011) Human computation, 1st edn. Morgan & Claypool Publishers, San Rafael

Lazer D, Friedman A (2007) The network structure of exploration and exploitation. Adm Sci Q 52(4):667–694

Levinthal DA (1997) Adaptation on rugged landscapes. Manag Sci 43(7):934–950

March J (1991) Exploration and exploitation in organizational learning. Organ Sci 2(1):71–87

Marquart DI (1955) Group problem solving. J Soc Psychol 41(1):103–113

Mason WA, Jones A, Goldstone RL (2008) Propagation of innovations in networked groups. J Exp Psychol Gen 137(3):422–433

Mason WA, Watts DJ (2012) Collaborative learning in networks. Proc Nat Acad Sci U S A, 109(3):764–769

Osborn AF (1957) Applied imagination: principles and procedures of creative problem-solving, Rev. edn. C. Scribner's Sons, New York

Pirolli P, Card S (1995) Information foraging in information access environments. Proceedings of the SIGCHI conference on Human factors in computing systems. ACM Press/Addison-Wesley Publishing Co. pp. 51–58

Rendell L, Boyd R, Cownden D, Enquist M, Eriksson K, Feldman MW, Fogarty L, Ghirlanda S, Lillicrap T, Laland KN (2010) Why copy others? Insights from the social learning strategies tournament. Science 328:208–213

Roberts ME, Goldstone RL (2006) EPICURE: spatial and knowledge limitations in group foraging. Adapt Behav 14(4):291–313

Stasser G, Titus W (1985) Pooling of unshared information in group decision making: biased information sampling during discussion. J Personal Soc Psychol 48(6):1467–1478
Steiner ID (1972) Group process and productivity. Social psychology. Academic Press, New York
Watts DJ, Strogatz S (1998) Collective dynamics of 'small-world' networks. Nature 393:440–442

Organismic Computing

Pietro Michelucci

Introduction

"Too many cooks spoil the broth". This is an English proverb conveying the notion that employing too many people in a collaborative effort can be detrimental to the desired outcome. In common usage, this proverb can imply that using fewer collaborators is more efficient. Of course, the verity of this proverb seems to depend on the problem space and division of labor. In a small kitchen with only two work areas, a chef and prep cook may be most effective for producing a meal in the least amount of time. Adding people would likely only introduce interference.

As for most proverbs, however, there is a proverbial counterpoint: "many hands make light work". In common usage, this proverb implies that adding collaborators actually improves efficacy. For example, painting the interior of an entire house would certainly go faster if one or more painters could be assigned to each room, working in parallel. However, even in the latter context, beyond a certain group size there may be diminishing returns and even deleterious effects. Imagine ten people trying simultaneously to paint the same eight-foot long wall: friendships could be tested under such circumstances. So it turns out that this counterexample, in the limit, leads to the same conclusion as the original proverb. Even in painting the interior of a house, it is conceivable that there could be "too many cooks". Consideration of other collaborative activities also seems to lead inevitably to the same conclusion.

The question then becomes, can engineered Human Computation (HC) produce a counterexample? In other words, can technology-mediated collaboration give rise to circumstances in which group efficacy is unbounded, that is, it actually continues to improve no matter how many proverbial cooks are added to the kitchen? Moreover, is it conceivable that HC could enable *increasing* returns with each person who is added to a collaboration?

P. Michelucci (✉)
Fairfax, VA, USA
e-mail: pem@thinksplash.com

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_36,

Collaboration Efficacy

When a group's efficacy exceeds the combined efficacy of an equally sized collection of lone contributors, the group is said to be exhibiting synergy. Consider two people working on the same jigsaw puzzle, which depicts animals in the Serengeti. Now imagine that Luca wants to work on the elephant part of the puzzle and Eva wants to build the giraffe, but they are not interested in helping each other. Thus, this is collaborative only in that there is a shared goal to which both people are contributing.

Each person selects a piece from the pile and accepts or rejects it based on the perceived likelihood that that piece belongs to the chosen animal. If a piece cannot be fitted into the puzzle it is discarded back into the pile, where it might be reselected by the other person. Clearly, there is a time cost associated with the selection and disposition of each piece. Thus, each person working independently would experience this time cost as a work pace limitation.

After a while, Luca gets frustrated because he is having trouble finding elephant pieces. So he asks Eva to give him any elephant pieces she comes across. Eva agrees and asks that Luca do the same, by providing her with any giraffe pieces he finds. In this new, enhanced collaborative mode, two evaluations are made for each piece selection instead of one. Even without a formal analysis, it is plainly evident that fewer piece selections will likely result in more correct emplacements than when the agents are operating independently. As long as the additional cognitive load associated with assessing a puzzle piece for two animals instead of one is not too great, and as long as the transaction time associated with exchanging pieces is short enough, we would expect the puzzle completion time to be faster when Luca and Eva collaborate than when they don't.

However, this collaborative effect, by itself is insufficient to conclude the existence of synergy. It is conceivable that simply working on the same puzzle can be detrimental to individual performance due to indirect effects. For example, Luca might grab the piece that Eva was about to examine, interfering with her piece selection process. Or Eva, through her own selection and replacement of pieces, could disrupt Luca's ability to keep track of the pieces he has already examined and rejected. If such interference reduces or negates beneficial collaborative effects, then perhaps working independently rather than collaboratively would be more effective. Thus, comparing two forms of collaboration is insufficient for detecting synergy. It only reveals whether one mode of collaboration is more effective than another.

So how would we go about detecting synergy? As per our earlier definition of synergy, we need to compare independent performance aggregated over individuals to group performance on the same task. In the context of our jigsaw puzzle example, one way to assess the presence of synergy would be to measure the time it takes to build two jigsaw puzzles. By employing two puzzles, each person could be assigned an entire puzzle without concern about indirect collaborative effects either positive or negative. Thus, in the non-collaborative mode, Luca would build one puzzle and Eva would build the other. We would then record their summed completion times.

In the collaborative mode, Luca and Eva would work jointly on both puzzles, and we would again sum the completion times for both puzzles. Taking this approach, if completing both puzzles took less time in the collaborative mode than in the individual assignment mode, we would have strong evidence of a synergistic effect.

With this context in mind, we will consider next why unbounded group efficacy, as well as synergy, is possible, how it might be achieved through HC, and report briefly on a recent study that supports its plausibility.

Relationships

Group intelligence is sometimes defined as a group's ability to solve problems collaboratively. One might expect that such forms of group intelligence would be highly dependent upon the intelligence of its constituent members. It turns out, however, that a small group's collective intelligence is actually not closely related to the distribution of intelligence among its individual members. Instead, a better predictor of group intelligence seems to be social perceptiveness and cooperative behavior (See Woolley and Hashmi 2014). This seminal finding (Woolley et al. 2010) supports the view that group efficacy may by highly dependent upon both the quantity and quality of relationships that exist in a group.

An interesting example of this idea exists in neurobiology. Neurons, information processing nodes in the brain, accept input from other neurons, and on the basis of those inputs compute an output (see Koch and Segev 2000). Collectively, the dynamic activation of neurons in the brain gives rise to behavior. As such, one might expect that intelligence in the animal kingdom is associated with the number of neurons in an organism's brain. Though a correlation does exist between intelligence and number of neurons, it is actually the number and type of connections among these neurons that seems to differentiate levels of intelligence among and within species (Bruer 1999). This is not entirely surprising given the propensity for the growth of new connections among neurons during learning, otherwise known as neuroplasticity. Thus, if we think of the brain as a collection of collaborating neurons, we observe a striking parallel between group intelligence as a function of connectivity among group members, and individual brain intelligence as a function of connectivity among neurons.

A similar reliance on connectivity has been observed in the collective behavior of eusocial insects (See Moses, et al., and Pavlic and Pratt 2014, both this volume), and in the role of social learning on population fitness (Smaldino & Richerson, this volume). Recent evidence (Van Raan 2013) suggests that this effect also applies at the scale of cities. The central finding is that larger cities prosper in terms of wealth and new ideas due to an increase in the number of social interactions as well as the increased likelihood that any given relationship will result in economic specialization. Indeed this effect is hypothesized to exist across our entire species in the context of a "global brain" (Heylighen 2014). This recapitulation across organizational levels and species seems to further implicate the value of relationships in group efficacy.

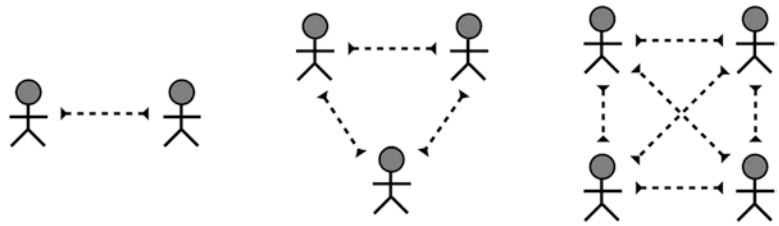


Fig. 1 Potential relationships for three different group sizes

Table 1 The progression of possible relationships as a function of group size

Group size	Relationships
1	0
2	1
3	3
4	6
5	10
10	45
30	435
100	4,950
1,000	499,500
1,000,000	499,999,500,000

Synergy Through Combinatorics

As we shall see, this view may point a way toward increasing the limits of group efficacy through HC. A simple intuition for this view derives from the following observation. As groups get larger, each additional member added to the group allows for increasingly more relationships. For example, a group with only two members is defined by a single relationship. Adding one member to form a three-person group increases the number of relationships to three. A group of size four has six possible relationships. And so on. This relationship between group size and possible relationships is depicted in Fig. 1.

This relationship between group size and the number of possible relationships in the group can be generalized. For a group of size n, the number of possible pairwise relationships in the group is given by the second-order polynomial function:

$$f(n) = (n^2 - n)/2$$

This progression is conveyed in Table 1 for ten different group sizes.

As the number of possible relationships in a group rises, so does the number of possible relationships for any single person, which is equal to the number of other members in the group. This results in a combinatorial explosion, such that with very large groups the number of possible relationships may become intractable. For example, when the group has 1,000,000 members, it seems unlikely that any given person would interact with every one of the other 999,999 members of the group. The first obvious barrier is the physical limitation. Imagine a football stadium full of people trying to manifest all possible pairwise interactions—people would be tripping over each other. But even if you obviate this physical (spatial) barrier by providing a communication infrastructure, we encounter another barrier: time. Humans are limited capacity systems. As cognitive load increases, the processing time for any individual cognitive task tends to slow down—this has been referred to as the "cognitive bottleneck" (Townsend 1990). Furthermore, we tend to be serial communicators, processing social interactions one at a time. Indeed, Kristina Lerman (this volume) has observed such limitations in the context of idea spread in the Twitter medium. "Re-tweeting", that is, receiving a message and then actively rebroadcasting it to one's network of subscribers, is an indicator of social signal processing. Analyzing such retweet data, Lerman found that attentional limitations play a role in attenuating the spread of ideas. Thus, even with a technology-mediated communication infrastructure, which overcomes physical obstacles to communication, there is a time-based cognitive constraint on manifesting social interactions.

But perhaps all possible relationships need not come to fruition. It is conceivable that certain relationships or interactions would be more useful than others, in which case the question of "who should talk to whom, and about what?" becomes central to the problem of architecting effective group behavior. Given that humans as individual processors of information have a limited capacity, optimizing group efficacy may require striking a balance between both number and *quality* of interactions.[1] To better understand what such a balance might entail, we can appeal to nature's most elaborate, and perhaps parsimonious (Hingston et al. 2008) model of useful connections: the human brain. The product of millions of years of evolutionary engineering, the human brain has been referred to (Anderson 2011) as "the most complex object in the known universe". Indeed, a large scale research effort led by the U.S. National Institutes of Health (NIH), is underway to map the "Human Connectome", that is the *patterns of connection* (see Blumberg 2013) that exist in the human brain. It is expected that this effort (Van Essen et al. 2012) will shed light on how these connection patterns give rise to intelligence and even, perhaps, consciousness (Bassett and Gazzaniga 2011).

Even without the benefit of a detailed human connectome, we can infer something about the how the brain balances quantity and quality of connections by simply applying an order-of-magnitude analysis. Linda Larson-Prior (this volume) informs us that the human brain has on the order of 100 billion information processing units, called neurons, which are interconnected. From our earlier formula, we know that the total theoretical number of connections that could exist

[1] Dunbar's number (Sutcliffe et al. 2012), represents a cognitive limit to the number of social relationships that can be maintained, and has been proposed to have a value between 100 and 230 (Hernando et al. 2009). However, these values may not generalize to non-social relationships or social relationships in constrained environments.

among these neurons is 10^{21}, which is equivalent to about 100 billion connections *per neuron*. In graph theory, the ratio of connections to nodes is referred to as *network density*. Thus, we can say that the maximum possible network density in the human brain is about 100 billion. However, according to Larson-Prior, the actual network density in the brain turns out to be closer to 10,000. This is fortunate, because if there were indeed 100 billion connections per neuron, either the human head would have to be a million times bigger to accommodate them or the neurons would have to be a million times smaller, neither of which would be morphologically tractable. There is evidence to suggest that neuron size varies across species; however, this variance is by no more than a factor of 100[2] (Herculano-Houzel 2009).

Evolutionary Tradeoffs

So what can we glean from the observation that, on average, each neuron in the human brain is connected to approximately 10,000 other neurons? If the engine of evolution is parsimonious, which it is often assumed to be, then it will optimize network density to maximize fitness (i.e. survival ability). This fitness function might reasonably involve a tradeoff between intelligence and SWaP (Size, Weight, and Power requirement). For context, the SWaP for the human brain is generally given as "the size of a large grapefruit", 1.5 kg, and 20 W, respectively. Perhaps evolutionary forces decided that this was the maximum SWaP that allowed humans to be smart enough to build spears and use fire, but not waste too many precious calories on brain support.

But what if SWaP were not a factor. By ignoring SWaP, we can engage in a couple of potentially illuminating thought experiments. So for the sake of argument, let's begin by stipulating only that the human brain, as we know it, is sized optimally. Knowing that the number of dendritic connections that will fit in such a brain is 10^{15} (Larson, this volume), we can then ask, why does the human brain have 100 billion neurons with 10,000 connections each instead of one billion neurons with a million connections each? In other words, how did the brain evolve to favor a particular network density? Since intelligence is believed to correlate to network density (Bruer 1999), neuron count, and transmission speed (Herculano-Houzel 2009), all three of these factors were likely at play in the evolutionary process,[3] yet somehow the present configuration is the one that resulted ultimately in the kind of intelligence that maximized evolutionary fitness.

[2] As an interesting digression, this degree of variance in neuron size across species does explain how elephant brains can be twice as large as human brains but contain only a quarter as many neurons.

[3] It turns out to be even a bit more complicated than that, as it has also been observed that proteomics may be a factor in intelligence. That is, the number and complexity of proteins in the synapse (the junction between neurons) tends to be greater in more intelligence species (Emes et al. 2008).

Next let us consider what would happen if we removed the size constraint. What if the human brain could be 46 ft in diameter and consume 20 million Watts of power to allow each of the 100 billion neurons to be maximally connected to each other? What kind of performance could one expect from such a brain? Stimulating a single neuron in this connection-saturated brain would result in the instantaneous activation of all 100 billion neurons. In fact, stimulating any neuron in such a brain would produce the same result (though perhaps with slightly different activation offsets due to proximity-based latencies). In practical terms, this means any signal coming into the brain, whether through visual, auditory, or other sensory channels, would give rise to the same brain state: all neurons activated. There would be no way to differentiate one experience from another, or one thought from another.[4] Indeed there is a clinical term for such a state of synchronous and excessive neuronal activity: *seizure*. This observation, that hyper-connectivity reduces information processing ability, speaks to an evolutionary rationale for network density that transcends physical considerations such as SWaP: a maximally connected brain is as useless as a completely disconnected brain. Therefore, a sparsely connected brain is required for cognition. This observation could be instructive for the design of human-based systems, suggesting that they might also be best served by limiting interconnectivity among collaborators.

Cognitive Architecture

As soon as one realizes that not all nodes in a network should be connected to each other, then a decision needs to be made. Which nodes should be connected to which other nodes, and to what effect? The form this answer takes is often referred to as a "cognitive architecture". Cognitive architectures specify representational structures and the flow of information in thinking systems.

A common distinction that arises in the design of cognitive architectures, is between associative and procedural knowledge (Bever 1992). Consider the act of learning someone's name. The light reflected off of a person's face stimulates the retina, which results in a representation of that face in the visual cortex—the part of the brain that decodes images. At the same time, the person's spoken name vibrates the tympanic membrane, stimulating the auditory cortex—the part of the brain that decodes sounds. The simultaneous stimulation of these visual and auditory patterns in those two regions of the brain causes a connection between them to be formed. The result of this connection is that, in the future, stimulating the same pattern in one of these regions will, using this newly formed connection, activate the other. For example, hearing the name will conjure a mental image, and seeing the person's

[4] Due to the existence of different types of neuronal connections (e.g., inhibitory and excitatory), this characterization of the underlying activation dynamics may not be entirely accurate, but the effective outcome still holds.

face will remind of the name. This is called forming an association, and is explained easily in the context of neurons and connections.

However, there is another type of information processing that does not lend itself so well to association-based explanation. Consider the following puzzle. I line up seven balls in a row and ask you how many there are. You count them and tell me there are seven. Now I move the balls into a circular configuration and ask you again: how many balls are there. This time you answer "seven" without even counting. You've never seen this puzzle before. How did you know that there were still seven balls? You inferred that since no balls have been removed there must still be the same number of balls. In other words, you reasoned symbolically using discrete, logical steps. Reasoning is more difficult to explain using only the concept of learned associations. At some level, of course, since a biological associative network (the brain) ultimately performs the reasoning, associations must underlie symbolic reasoning. But it is both associations and processes and the manner in which they are combined that gives rise to the myriad complex behaviors observed in human cognition. If we can duplicate this in a human computation system, perhaps we can kindle collective reasoning.

So far, we have considered evidence that group efficacy is related not so much to the intelligence of individual group members, but rather to the quality of interaction among group members. We then appealed to the simple observation that with more people in a group comes many more possible relationships, and that this effect increases with group size. This gives us hope that, under the right conditions, synergy would be possible. We also realize that very large groups have so many potential relationships that they cannot all possibly be realized. But we learn that this is not necessarily a bad thing, because in some information processing systems, such as animal brains, it would actually be detrimental to form all possible connections. We have also considered that knowledge, and indeed information processing capacity, may be stored in the patterns of connections that are formed. This suggests that in exploring methods for achieving synergy in large groups of people it may be advantageous to engineer social structures, such as hierarchy, that are designed to achieve specific information processing functionality, and more generally, to experiment with different patterns of relationships among collaborators.

Organismic Computing

Inspired by the successes of systems evolved at both the cellular (brain) and organismal (eusocial insects) grades of organization and compelled by an understanding of the factors that might influence collaboration efficacy, we now turn our attention to how one might engineer the synergies observed in these natural systems and maximize the value of each relationship. Toward that end, we propose a human computation paradigm that that seeks to use technology to enable a group of human collaborators to function simultaneously as independent agents and as tightly integrated parts of a collective "superorganism" (see Pavlic and Pratt 2014, this volume). Such an "organismic computing" system would necessarily apply tightly coupled *shared sensing* (see Lin et al. 2013 and Meier 2013), *collective reasoning* (see Blumberg 2013, Greene and Young 2013), and *coordinated action* (see Novak 2013) toward collective goals.

Shared Sensing

The notion of shared sensing is that the sensory experience of one member of the group is available to other members of the group. The challenge here is to avoid information overload. People accomplish this with attentional mechanisms, that is, by attending only to environmental inputs that are most relevant. However, there are times when even humans who are not being fed sensory data from their collaborators are still prone to information overload. Consider the difficulty of trying to attend simultaneously to a radio talk show and a person who is talking to you. So the key seems to be to increase the percentage of *relevant* information in the environment without increasing the overall amount of information. Information load can be managed in several ways: (1) information channels could be selectively turned on and off by individuals, (2) information access could be driven automatically by context, (3) sensory sharing could be confined only to subgroups that are defined functionally or hierarchically, with the interesting possibility that certain individuals could have membership in more than one subgroup, and furthermore, that subgroup membership could shift dynamically.

Collective Reasoning

Collective reasoning is intended here to mean something subtly different than distributed thinking. Distributed thinking is a direct analog to distributed computing. It means simply that an information-processing load is spread across human computational agents. Collective reasoning is more specific. It requires that information processing occurs not only within individual agents, but is encoded somehow in the interactions that occur among those agents.

Coordinated Action

Coordinated action is the notion that the product of collective reasoning, that is, the information that arises from interactive processes, exerts an influence on individual behavior toward collective goals. To develop an intuition, consider a football coach who observes patterns of interaction on the field and then uses those observations to inform the design of more effective tactics, which are used to coordinate the actions of her players. Keep in mind that in order to execute such tactical plans, each player need only be aware of the portion of the plan that concerns his own interactions with other players and, specifically, the rules that govern his behavior within those exchanges. The key concept here is that knowledge arising from interaction feeds back into individual behavior to improve it.

A Recipe for Life

It has been suggested (see Walker and Davies 2012) that a distinguishing feature of life is that, at a biologic level, the information that arises from interactions bears causal influence on the interacting elements themselves. This bears an interesting parallel to the motivation underlying organismic computing, which is that synergy arises from the influence of "higher order" information—that is, information derived from social interactions—on individual behaviors.

In a sense, higher-level information is a free lunch, because it arises without the expenditure of any additional energy, yet under the right circumstances could materially increase organismic efficacy. Indeed the rational for including collective reasoning and coordinated action in organismic computing is to leverage interactional effects to the benefit of the group so that it can function more effectively than the summed capabilities of its individual members would suggest.

Therefore, perhaps it is precisely because interactional influence affords synergy that it is a precursor for life. In other words, by drawing this parallel between organismic computing and life, we may gain insight into the sustaining role of interactional influence on life. Conversely, we may wonder if the manifestation of such interactional influence in human groups is, in turn, suggestive of superorganismic life.

An Ecological Approach to Collective Perception

In organismic computing, time scale becomes relevant. When there is an introduced latency between sensing and acting, then some tasks may become prohibitive. Imagine being blindfolded and then verbally guided by another person to reach out and grasp a glass of water on a table: "Reach forward. Slow down. Left. Nope, you went too far. Come back to the right. Forward slowly. Too fast! Oops. That's okay, we have paper towels."

Humans are active learners (Gibson and Pick 2000). That is, by simultaneously acting and perceiving in the world we develop dynamic expectation models that guide our interactions with objects. Such models describe qualities of object, called affordances (Gibson 1986), that govern how we can act upon them. Indeed, it is this tight coupling between action and perception that endows us with such models as well as the ability to use them to guide our behaviors effectively in real time. When the sense-act cycle is too slow, as in the example above, such learned models cannot be applied so we become effectively crippled.

Imagine that you are still blindfolded, but that your confederate is now using a beacon to guide you to the glass of water. He is shining a bright flashlight toward your eyes through the glass of water, which you can see as a blurry spot of light through your blindfold. Indeed that blurry spot is the only thing you can see. As you begin to reach out with your hand, your head shifts position slightly. The resultant motion parallax gives you an immediate and visceral sense of distance to the glass. As your hand gets closer to the water glass, it momentarily passes in front of it, eclipsing the spot of light perceptibly from one side to the other. You autocorrect your position and grasp the glass in the first try.

There are two remarkable aspects to this experiment. One is that it really works (try it—all you need is a glass of water, a dishtowel to serve as a blindfold, a flashlight, and a dark room with a table). And the other is that the mechanism of tightly coupling perception and action in the environment is so powerful, that a stimulus as impoverished as a blurry spot of light succeeds easily where a high-latency albeit precise natural language exchange fails miserably. Conducting such an informal experiment leads to a visceral appreciation of ecological perception, for which natural language, such as this paragraph, is similarly inadequate.

Varieties of Latency

When this ecological model is applied to *collective* perception, we observe that latency can now manifest in three ways instead of one.

- *Feedback latency*: latency in the action-perception feedback loop (previously discussed)
- *Sensory latency*: latency in the sharing of sensory information among group members (e.g., mailing a DVD vs. streaming live video)
- *Action latency*: latency that disrupts the synchronization of coordinated actions (e.g., sending instructions by courier pigeon, such that the furthest recipient is the last to act)

Without delving deeply into the implications of each of these latency types, it seems evident that only when all forms of latency are minimized can certain dynamic causal relationships be perceived and learned collectively. Consider, for example, the goal of learning invariant response behaviors in a population. Such invariances may be more or less accessible depending upon the time scale associated with coordinating probe actions, coalescing distributed sensory experiences, and/or identifying collectively observable (not individually observable) cause-and-effect relationships.

On this basis, the notion of organismic computing, insofar as it is intended to optimize group performance by maximally leveraging relationships, entails minimizing communication latencies of all varieties, or at least sustaining a pace that enables the collective to maintain currency.

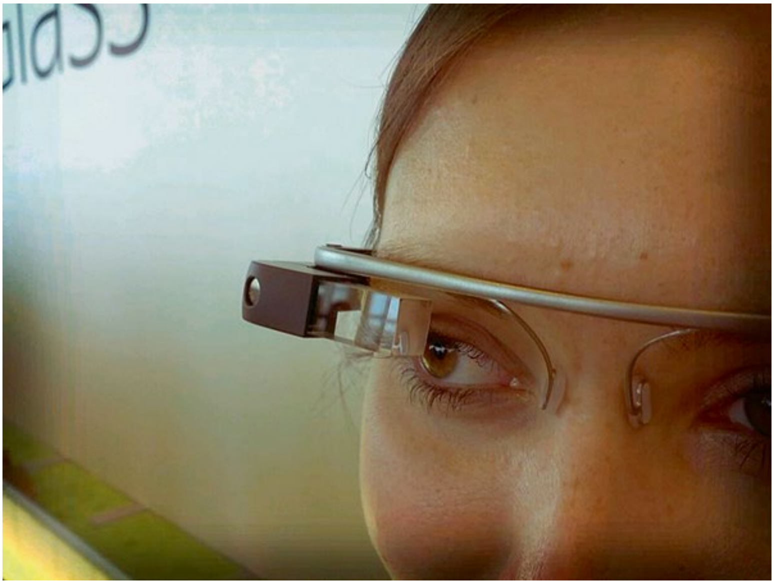


Fig. 2 Google Glass projects text and images onto the retina so the wearer can see them overlaid in the visual field

Minimizing Latency

Even if one accepts the validity of the ecological model of distributed perception described above, which implicates the role of communication latency in group efficacy, the practical matter of minimizing latency is another issue entirely. How does one enable low-latency communication on a massive scale in a manner consistent with the tenets of organismic computing (i.e., shared sensing, collective reasoning, and coordinated action)?

One avenue of recourse is to use a shared state space—a globally accessible data layer to which sensory information, directives, and high-level "processed" information can be pushed to and pulled from. Such a data layer disentangles the data availability problem from the information overload problem.

With such a data layer in place, the problem reduces to interface design. In the case of a problem space that manifests in the physical environment (e.g., capturing a fugitive), one simple approach would be to implement a heads-up display, such as Google Glass (see Fig. 2). With such an interface, textual information from the shared state space could be transmitted wirelessly and instantaneously to the wearer. Such information could be provided as needed (based upon context) or as requested, in order to minimize interference and maximize actionable information. For example, if I sent information to the shared state space indicating that a vehicle is moving north on Elm Street, automated processes could push that information only to collaborators currently situated on the north end of Elm Street. Such information might be irrelevant to

Fig. 3 Screenshot of an augmented reality application that overlays the location of real objects in the environment, such as airliners, in the parts of the image that correspond to their real-world locations

other collaborators. In addition, aggregating information from multiple sources or even recasting it into different forms (e.g., graphical) by either a human or automated system could substantially improve its utility and reduce the cognitive load.

Such an interface to a shared data layer might substantially reduce latencies, but an information-processing bottleneck remains. Any abstraction of sensory experience entails a processing load for both the producer and the consumer. For example, the communication of textual information via the shared data layer is slowed by processes of language production and perception. Even visual abstractions that seek to aggregate information require high-level cognitive processing for interpretation. For example, consider the cognitive load associated with using a street map to orient oneself to a local context. To address this, one seeks to shift from the high-latencies associated with overt cognitive translation to lower-latencies that result from the direct use of perceptual processes. Such a result can be achieved through the use of a shared augmented reality.

Augmented Reality (AR) combines virtual reality (e.g., simulation) with sensate reality (i.e. our perception of the physical world). It is a technique by which visual data is overlaid directly onto the visual environment using heads-up displays, such as Google Glass, or even smart phones that have cameras (see Fig. 3). The use of AR allows communication to occur directly within the perceptual medium of visual experience, which avoids the cognitive bottleneck. Consider the following scenario. A "smart camera" that can detect people and recognize human action (see Oltramari and Lebiere 2012) is strapped to my head, monitoring everything within my field of view. This camera happens to register a person running across the field in front of me. Because it is equipped with GPS and has access to topographic datasets corresponding to my locale, it is able to determine the location of the running person and transmit his dynamic spatial coordinates to the shared data layer. Even though my collaborators are not within line of sight of the runner, their augmented reality "smart glasses" (see Moses 2012) allow them to "see" the runner as a beacon moving across their own field of view in the actual direction of the runner. From this perceptual information, they can easily converge on the location of the runner by simply moving toward the beacon.

While communicating overtly using a heads-up display reduces communication latencies *between* members of a collaborative group, sharing sensory experience, action directives, and information processing products through an augmented reality interface reduces latencies that derive from the cognitive bottleneck that exists *within* members of the group. Thus, the use of an AR interface and supportive interaction paradigm would go along way toward meeting the low-latency requirement of organismic computing.

Recap

Given the expansive conceptual terrain that has been covered so far in this chapter, it may be worthwhile to recapitulate that briefly before describing the empirical work. The premise of this approach is that because group efficacy stems primarily from relationships, and because the ratio of possible relationships to individuals is higher in larger groups than in smaller groups, the potential for synergy seems higher in larger groups. For the most part, however, that potential does not seem to be realized in human collaboration, perhaps due to a relationship cost function that increases with group size (e.g., the coordination cost increases with group size). Thus, the thesis of this chapter is that the prospect of synergy ultimately boils down to considering both the additive value of each relationship and the cost of each relationship. On this basis, we have advanced a collaboration paradigm called "organismic computing", modeled on natural systems, that seeks to both maximize the value and minimize the cost of each relationship.

The next section reports on a pilot study that seeks to assess the relative impact of organismic computing on group efficacy.

Experiment

To explore this thesis, a pilot study was conducted to compare three collaboration paradigms in terms of their relative impact on group efficacy in a tactical "hide-and-seek" game. Each of the three paradigms utilizes a different communication modality:

1. **Radio**—uses multi-channel radio communication to permit instantaneous transfer of spoken information from one group member to one or more other group members; this is modeled loosely on standard tactical methods, such as those used by small military units
2. **Social**—uses textual and visual updates to share information that is processed cognitively via a wireless heads-up display, which minimize latencies *between* group members
3. **Organismic**—uses wireless communication via an augmented reality interface to enable direct perceptual sharing of information, which minimize latencies both *between and within* group members

Methods

The goal of this study was to produce an existence demonstration that, with the right technology-mediated collaboration paradigm, larger groups can exhibit greater synergy than smaller groups—that indeed, the more cooks there are in the kitchen, the better. Toward that we evaluated a 3D multi-user simulation environment to assess its suitability to future work and to compare tactical group performance across different levels of augmentation in a hide-and-seek game.

Testers

In order to conduct a study that is reliant on the simultaneous availability of a predetermined population of players we employed in an unconventional manner the services of uTest. uTest, Inc. is a software testing company that employs a mature process and scalable infrastructure to crowdsource software testing to a pool of over 80,000 prescreened testers from 190 countries. Because the interface used in the study software relies upon comprehension of both spoken and written English, the study population was limited to literate, native speakers of English. A significant advantage of using a service like uTest, was the ability to schedule synchronous testing to ensure the collaborative participation of groups of predetermined sizes. Additionally, the testers were accustomed to being exposed to new interfaces and also in providing useful feedback about their user experiences.

Materials

As discussed previously, smart cameras and smart glasses are emerging technologies, but not yet mature technologies. In order to inexpensively conduct this study, we developed a massively multiplayer online role-playing game (MMORPG) using a "first person shooter" (FPS) style 3D virtual environment (see Fig. 4). In this context, we simulated both heads-up displays and augmented reality. Thus, we used

Fig. 4 Screenshot of the organismic interface in the immersive 3D gaming environment that was used to conduct the study. The "floating" dots are proxies for informational beacons that might appear in an real world AR interface

virtual reality to simulate augmented reality, in what we expect will become a useful technique for conducting future related experiments.

Design

Domain

The application domain used in this study was tactical operations—specifically, hide and seek. Testers played the role of seekers, moving through the virtual world as avatars. Infiltrators (hiders) and citizens were, in gaming parlance, "non player characters" (NPCs). In other words, the computer used role-specific heuristics to determine their behaviors.

Goal

Each group of players formed a single team that played against the clock in its own instantiation of the virtual world. The goal for each team was to identify the leader of the hiders, on the basis of his distinctive behavior and the behaviors of those around him, in the least amount of time.

Avatar Roles

Each player was assigned to one of three possible roles, listed below with the corresponding functional assignments:

Fig. 5 Screenshot of the analyst interface in the Radio group

Scout—Note Behaviors & Suspicion
Analyst—Lead Scouts, Label NPCs
Director—Lead Analysts, Capture

NPC Roles

As described above, each NPC was also assigned one of three roles, each of which had a distinct behavioral profile.[5] Each NPC type is listed below with the corresponding game-based purpose:

Citizen—serves to distract the seekers from finding the Leader
Infiltrator—tends to stay near the Leader; serves as a "noisy" clue
Leader—there is only one leader, who is the target of the seekers

Modalities

There were three experimental modalities: Radio, Social, and Organismic, as described above.

Radio groups (see Fig. 5) served as a control. They used headsets connected to their computers to talk with each other over separate channels that were selectively available according to their membership in hierarchical sub-units. They were also given access to a static map, which they could annotate manually. Though they were

[5] Details concerning behavioral profiles and associated heuristics, as well as other game mechanics and design elements will be reported in a forthcoming paper that is geared more specifically experimentation platform, the empirical methods, and the repeatability of reported techniques. The present exposition is intended primarily to exemplify the core theoretical concepts and summarize key results.

Fig. 6 Screenshot of the analyst interface in the Social group

each assigned specific roles and associated instructions, as were players in other groups, they were mostly left to their own devices to implement a hierarchy and adhere to their roles. The only constraint on behavior in members of the Radio group pertained to radio channel accessibility.

Social groups communicated via textual inputs and heads-up displays (see Fig. 6). In addition, Social groups were assumed to broadcast their own positions via GPS but not possess smart cameras, so they had to manually input NPC observations through a textual interface, which simply recorded the last known location of the NPC as being the transmission location of the Avatar who reported him.

Organismic groups annotated objects directly in their environment (via point-and-click with a mouse or touchpad). These annotations were shared immersively via the AR interface elements (see Fig. 7). Unlike the Social group, members of the Organismic group were assumed to be equipped with smart cameras that would automatically geolocate and report the location of any NPCs in line of sight.

Aggregation: The Social and Organismic groups both benefitted from machine-based aggregation of the collective sensory information and both benefitted from the presentation of textual and map-based information in a simulated heads-up display. Both groups employed the same, shared state space. The key and only differences were the use smart cameras (as explained above) and augmented reality in the Organismic group.

Group Size

Performance was compared across three different group sizes, each three times larger than the next smallest size. The total population of each group as well as the distribution of roles and number of possible pairwise relationships is indicated in Table 2.

Fig. 7 Screenshot of the analyst interface in the Organismic group

Table 2 Team sizes and corresponding demographics

Size	Scouts	Analysts	Directors	*Total*	Pairwise relationships
Small	9	3	1	**13**	78
Medium	27	9	3	**39**	741
Large	81	27	9	**117**	6,786

Conditions

This study employed a 3×3 design, in which the independent variables were group size and collaboration modality. Two additional variables, game world size and NPC population, were made to co-vary with group size to support playable game dynamics.

Procedure

Each tester played in a single game to avoid learning effects. Each game was assigned a specific start time. Players received instructions in advance of the game based upon their assigned roles and collaboration modality (see Fig. 8).

The game was played until either a maximum game time was reached or the team correctly identified the Leader. In the case of a misidentification, the role of the Leader was reassigned to a new NPC.

The game server recorded event-driven client telemetry consisting of Avatar and NPC locations as well as other key events such as entering and exiting the game and making a capture.

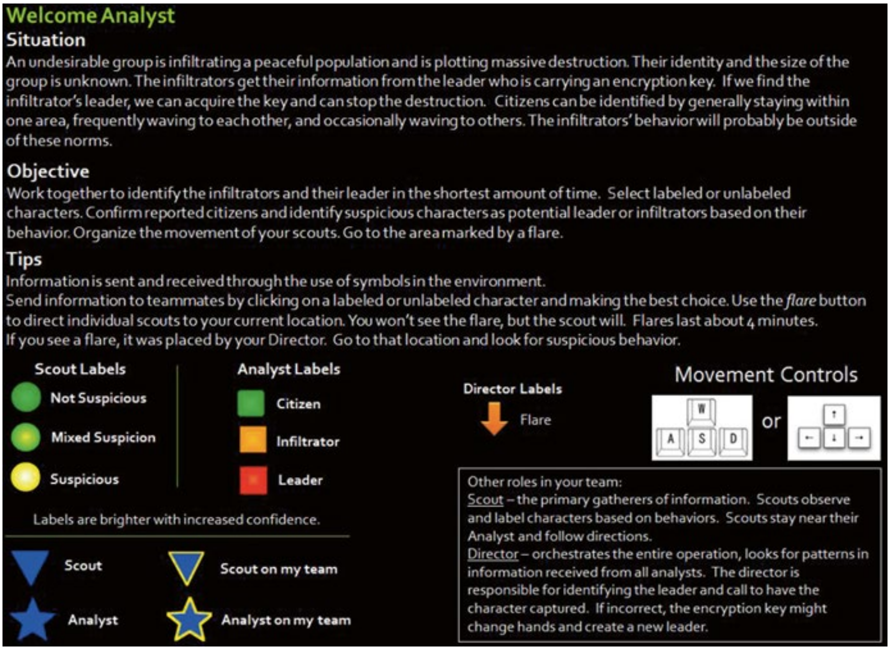


Fig. 8 Screenshot of the instructions for analysts in the Organismic group

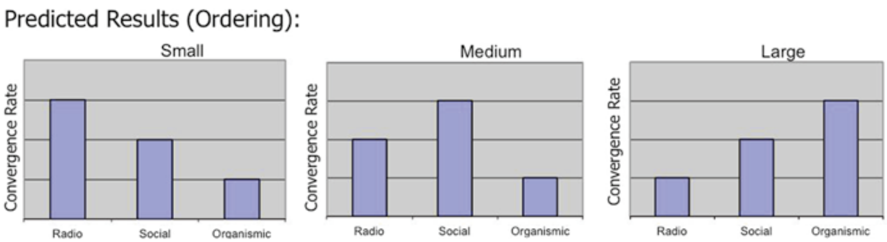


Fig. 9 Predicted ordering of convergence rate as a function of collaboration modality within each group size

Predictions

The central hypothesis was that performance in the two experimental collaboration modalities (Social and Organismic) would improve more in larger groups than in smaller groups relative to the control modality (Radio). This and a number of related hypotheses are captured succinctly by a series of three ordinal graphs of predicted performance shown in Fig. 9 with respect to a "convergence rate" metric, which represents the mean rate at which the average distance between each team member and the leader decreased over the course of the game. This metric was selected

Fig. 10 Predicted ordering of convergence rate as a function of collaboration modality within each group size

instead of completion time to overcome limitations such as serendipitous discovery, which are associated with a small sample size (i.e., one game per cell).

These ordinal predictions represent crossover effects, in which the experimental modalities (Social and Organismic) give rise to improved efficacy as group size increases. We expected a steeper learning curve for the Organismic modality than for the Social modality due to the added complexity of a visual symbology, which was used in the AR elements in the Organismic group. Thus, we anticipated the first crossover effect to occur between the Radio and Social groups when moving from a small to medium sized group. But since the synergistic effects were expected to be strongest in the Organismic group, we expected Organismic to overtake Social in the context of Large group sizes, when those effects would benefit most by manifesting within a larger number of potential interactions. In addition, relative to the Social and Organismic groups, we expected Radio performance to decrease as groups got bigger due to increasing coordination costs.

Results

Observed results are conveyed in Fig. 10, below the original predictions. The observed orderings of the convergence rates are consistent with the expected orderings described above. In particular, performance in the Radio group was strongest among small groups, performance in the Social group was strongest in the medium sized group, and Organismic performance was strongest in the large group. The latter finding lends support to the ecological model of distributed perception proposed above.

Fig. 11 Raw proximity data for the large Radio group

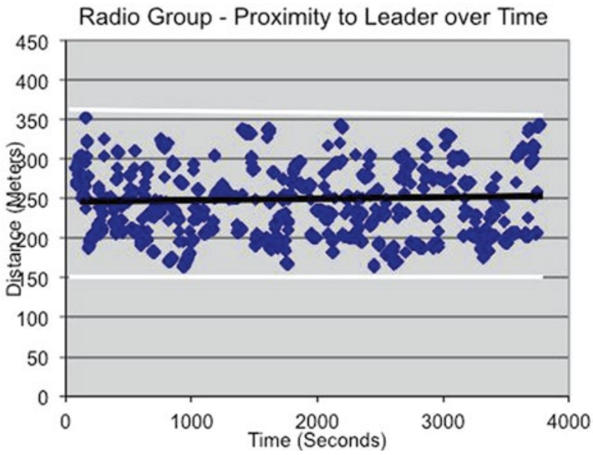
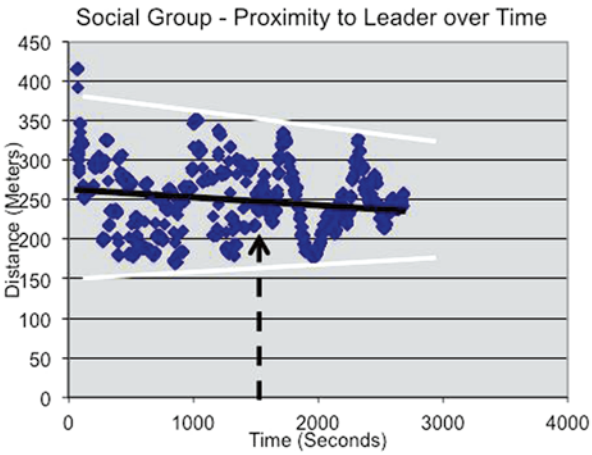


Fig. 12 Raw proximity data for the large Social group



Qualitative Analysis

In addition to these quantitative results and analysis, we consider briefly a qualitative view of the data. Figs. 11, 12 and 13 depict raw leader proximity data for each of the three large groups. It is worth noting here that leader movements in the game were cyclical. In other words, the leader would pace back and forth throughout the game, which introduces a cyclical element to the distance/time data reported here. It is important to avoid misattributing those cyclical characteristics to seeker behavior.

In order to makes sense of these graphs, three visual indicators are proposed as useful methods for qualitative analysis.

Coordination is pointed out in two of the three graphs by an arrow (no coordination behavior was observed in the Radio group). This corresponds visually to the emergence of a coherent "signal", in which there may be a cyclical trend (as explained above), but there is reduced variance. Coordination reflects a constancy of distance

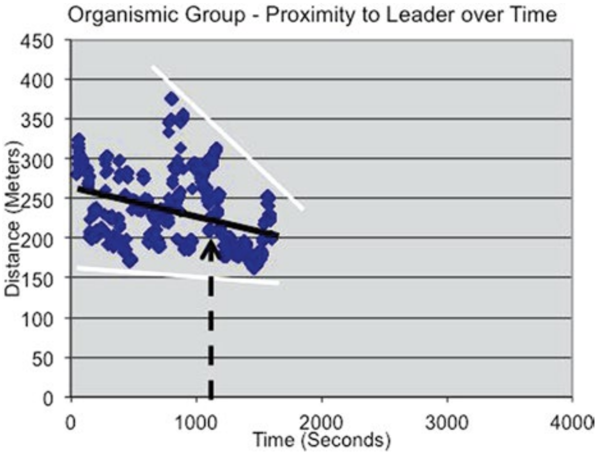


Fig. 13 Raw proximity data for the large Organismic group

among members of the Seeker group, suggesting that either their movements are coordinated or, in the absence of other indicators, that they are all stationary.

Convergence, indicated by the thick, black line, conveys the linear trend corresponding to the mean distance over time of the Seekers to the Leader. The slope of this line corresponds to the data used in the quantitative analysis above. In general, convergence indicates how quickly the Seekers converged on the Leader. The coincidence of coordination and convergence rules out the possibility that the team is stationary.

Tracking behavior manifests visually as signal amplitude. The two white lines, which form upper and lower bounds on the observed proximity data, indicate the degree to which mean distance to the leader varies over time. When the two lines are convergent (from left to right), that indicates that the distance to the leader is fluctuating less and less over time. One plausible explanation for such a trend is that more and more Seekers, irrespective of their proximity, are coordinating their movements to the Leader. An observed trend in leader-centric coordination could be construed as being indicative of knowledge spread.

With a ready understanding of these three qualitative indicators, we can proceed to analyze the raw proximity data.

Radio Analysis

In the Radio group (see Fig. 11), we observe the absence of coordination—at no point does a coherent "signal" become apparent. This suggests that in the absence of special augmentations, it is very difficult to coordinate a very large group. A flat trend line suggests that there is no change in mean proximity to the leader—in other words, no apparent progress has been made. Finally, the absence of a discernable tracking signature further supports the interpretation that the Radio team is no better off at the end of the game than at the beginning.

Social Analysis

In the Social group of seekers, we observe the distinct emergence of a “signal”, suggesting the possibility of overt synchronization of team movements. This is supported by the presence of a gradual but consistent convergence trend. Additionally, the indisputable tracking signature suggests that at least some team members are stalking the Leader.

Organismic Analysis

We observe coordination of team movements in the Organismic group earlier than in the Social group. In the Organismic group of seekers, we observe the distinct emergence of a “signal”, suggesting the possibility of overt synchronization of team movements. Rapid convergence is indisputable, even without the benefit of a linear fit. And tracking boundaries seem to indicate a rapid spread of knowledge about the identity and position of the Hider Leader for this group.

Discussion

This study was originally motivated by the goal of demonstrating, through the use of human computation, that it is possible to reverse the trend of diminishing returns in human collaboration. There was also the ancillary goal of evaluating organismic computing as a collaboration paradigm that can lead to such synergy.

Though the results are certainly tantalizing, without further investigation we cannot make any strong claims. There were methodological issues, such as a small sample size, though this was somewhat mitigated because each datum within a cell represented many samples across many seekers. And although it was valid to compare performance orderings across group size, and absolute performance numbers across modalities within a group size, it was not possible to make meaningful comparisons of absolute performance across group size because (as noted earlier) the size of the game world and NPC population co-varies with group size, which introduces a confound.

Furthermore, in the interest of meeting the stated objective we took a shotgun approach, by combining many collaborative enhancements within a single experimental condition. With so many variables at play, it would be almost impossible to establish with high confidence which augmentation(s) were responsible for the observed results.

These concerns notwithstanding, we can make several weak claims:

- We have observed a quantifiable difference in efficacy between the Social modality and Organismic modality within large groups, which could be construed as support for an ecological model of distributed perception.

- We have observed an interaction between group size and the degree of organismic augmentation (considering the discrete continuum of Radio → Social → Organismic), suggesting that as groups become larger, the collaboration benefits of such augmentation are greater.

The second claim raises the interesting possibility that if the observed trend continues indefinitely, then achieving synergy in the context of organismic computing may simply be a matter of employing a large enough group.

Future Work

The significant practical value of this empirical work was to motivate a deep list of desiderata. A first order of business would be to perform a more extensive analysis on the extant study data:

- Use the play logs to compare how much time seekers spent on overt communication between the Social and Organismic groups. This may reveal primary effects of communication latency on overall performance.
- Also compare how much information each group communicated. Perhaps both groups spent the same amount of time communicating, but the Organismic group communicated much more information in the same amount of time.
- Consider any methods that might help disentangle the effects of machine-based aggregation or collective reasoning from the effects of shared sensing and coordinated action.
- Explore methods for measuring the effects of collective reasoning on individual cognition.

A second order of business would be to repeat the study, but with a number of improvements:

- Make interface improvements that would likely further reduce the cognitive load, thereby freeing up reasoning resources.
- Implement better controls: e.g., use the same size game world and NPC populations across group sizes.
- Employ an asynchronous participation model with collective memory, such as Crowd Agents (Lasecki & Bigham 2014), which would make data collection robust to player attrition.
- Further develop the experimentation software into a reusable, generalized platform for studying organismic computing.

It is of additional interest to explore alternate approaches to enhancing collaboration in the context of many different application domains. Investigations of this type would improve our understanding of collaboration dynamics and the suitability of various human computation methods to specific applications. Indeed, such efforts may ultimately lead to generative or even adaptive models of collaboration.

Acknowledgments The author wishes to express deep gratitude to Kshanti Greene and Thomas Young of Social Logic Institute for their creative contributions and tireless execution of the present study as well as their helpful feedback on this chapter. The author would also like to acknowledge Geoffrey Bingham for his insightful comments regarding the application of ecological perception to distributed groups. Finally, the author would like to thank James Donlon for his enduring confidence and support of this speculative work. This research was funded under DARPA contract #D11AP00291.

References

Anderson A (2011, November 17) Science: brain work. The economist. Retrieved from http://www.economist.com/node/21537050

Bassett DS, Gazzaniga MS (2011) Understanding complexity in the human brain. Trends Cogn Sci 15(5):200–209. doi:10.1016/j.tics.2011.03.006

Bever T (1992) The demons and the beast: modular and nodular kinds of knowledge. In: Reilly R, Sharkey N (eds) Connectionist approaches to natural language processing. Lawrence Erlbaum Assoc, Hove, pp 213–252

Blumberg M (2013) Patterns of connection. In: Michelucci P (ed) The handbook of human computation. Springer, New York

Bruer JT (1999) Neural connections: some you use some you lose. Phi Delta Kappan 81(4):264–277

Emes RD, Pocklington AJ, Anderson CNG, Bayes A, Collins MO, Vickers CA, Grant SGN (2008) Evolutionary expansion and anatomical specialization of synapse proteome complexity. Nat Neurosci 11(7):799–806. doi:10.1038/nn.2135

Gibson JJ (1986) The ecological approach to visual perception. Psychology Press

Gibson EJ, Pick AD (2000) An ecological approach to perceptual learning and development. Oxford University Press

Greene K, Young T (2013) Building blocks for collective problem solving. In: Michelucci P (ed) The handbook of human computation. Springer, New York

Herculano-Houzel S (2009) The human brain in numbers: a linearly scaled-up primate brain. Front Hum Neurosci 3:31. doi:10.3389/neuro.09.031.2009

Hernando A, Villuendas D, Vesperinas C, Abad M, Plastino A (2009) Unravelling the size distribution of social groups with information theory on complex networks (arXiv e-print No. 0905.3704). Retrieved from http://arxiv.org/abs/0905.3704

Heylighen F (2014) From human computation to the global brain: the self-organization of distributed intelligence. In: Michelucci PE (ed) Handbook of human computation. Springer, New York

Hingston PF, Barone LC, Michalewicz Z (2008) Design by evolution: advances in evolutionary design. Springer

Koch C, Segev I (2000) The role of single neurons in information processing. Nat Neurosci 3:1171–1177. doi:10.1038/81444

Kunegis J (2011, July 25) The density of a network is independent of its size. Netw Sci. Blog. Retrieved from http://networkscience.wordpress.com/2011/07/25/the-density-of-a-network-is-independent-of-its-size/, 2 July 2013

Lasecki W, Bigham J (2014) Interactive crowds: real-time crowdsourcing and crowd agents. In: Michelucci P (ed) The handbook of human computation. Springer, New York

Lin A, Huynh A, Barrington L, Lanckriet G (2013) Search and discovery through human computation. In: Michelucci P (ed) The handbook of human computation. Springer, New York

Meier P (2013) Human computation for disaster response. In: Michelucci P (ed) The handbook of human computation. Springer, New York

Moses A (2012) A mind's eye in front of your nose. Syd Morning Her. Retrieved from http://www.smh.com.au/digital-life/digital-life-news/a-minds-eye-in-front-of-your-nose-20121123-29xhe.html, 7 July 2013

Novak J (2013) Collective action and human computation: from crowd-workers to social collectives. In: Michelucci P (ed) The handbook of human computation. Springer, New York

Oltramari A, Lebiere C (2012) Using ontologies in a cognitive-grounded system: automatic action recognition in video surveillance. In: Proceedings of the 7th international conference on semantic technology for intelligence, defense, and security, Fairfax

Pavlic T, Pratt S (2014) Superorganismic behavior via human computation. In: Michelucci P (ed) The handbook of human computation. Springer, New York

Sutcliffe A, Dunbar R, Binder J, Arrow H (2012) Relationships and the social brain: integrating psychological and evolutionary perspectives. Br J Psychol 103(2):149–168

Townsend JT (1990) Serial vs. parallel processing: sometimes they look like tweedledum and tweedledee but they can (and Should) be distinguished. Psychol Sci 1(1):46–54. doi:10.1111/j.1467-9280.1990.tb00067.x

Van Essen DC, Ugurbil K, Auerbach E, Barch D, Behrens TE, Bucholz R, Chang A, Chen L, Corbetta M, Curtiss SW, Della Penna S, Feinberg D, Glasser MF, Harel N, Heath AC, Larson-Prior L, Marcus D, Michalareas G, Moeller S, Oostenveld R, Petersen SE, Prior F, Schlaggar BL, Smith SM, Snyder AZ, Xu J, Yacoub E, for the WU-Minn HCP Consortium (2012) The human connectome project: a data acquisition perspective. NeuroImage 62(4):2222–2231

Van Raan AFJ (2013) Universities scale like cities. PLoS One 8(3):e59384. doi:10.1371/journal.pone.0059384

Walker SI, Davies PCW (2012) The algorithmic origins of life (arXiv e-print No. 1207.4803). Retrieved from http://arxiv.org/abs/1207.4803

Woolley AW, Hashmi N (2014) Cultivating collective intelligence in online groups. In: Michelucci P (ed) The handbook of human computation. Springer, New York

Woolley AW, Chabris CF, Pentland A, Hashmi N, Malone TW (2010) Evidence for a collective intelligence factor in the performance of human groups. Science 330(6004):686–688. doi:10.1126/science.1193147

Part IV
Infrastructure and Architecture

Infrastructure and Architecture for Human Computer Intelligent Collaboration

Michael Witbrock

The idea behind the traditional conception of human computation[1] has been intensely reductionist—it seeks to identify aspects of human capability that are useful in isolation, to wrap those capabilities in abstractions that allow their utility and applicability to be assessed independent of the identity of the human involved (or, indeed, of the very fact that a human is involved), and to design systems that maintain these abstractions sufficiently well for the desired computation to be performed in a way that satisfies the purposes of its consumer.

On the face of it, this description might suggest that Human Computation is anti-human—the value of humans *qua* people is sacrificed to the goal of computational effectiveness within a computation system as rigid and mechanical as, say, an accounting or airline reservation system. But this view ignores the full possibility of human computation, in which such computation takes place in the context of human-computer collaboration, in which a variety of computational agents collaboratively complete their joint goals, including both task goals and personal goals. Such a collaboration can include the complete range of computational techniques, including learning on the part of both human and non-human computational agents, and can be directed in a way that makes the tasks directed at each kind of agent both increasingly suited and increasingly satisfactory to the agent involved. This suggests, perhaps, adding to Lawhead & Estrada's characterisation of HC tasks, found elsewhere in this volume, an additional axis representing actualization c.f. disempowerment, with the latter kind of task perhaps best evoked by the vision of proletarian workers labouring in the city bowels in Fritz Lang's "Metropolis".

[1] Disruptive, Engineered Human Computation, in the Lawhead and Estrada formulation elsewhere in this volume.

M. Witbrock (✉)
Cycorp Inc./Curious Cat Company, Austin, USA
e-mail: witbrock@cyc.com

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_37,

Most probably, many human-computer collaboration systems will fall somewhere along a dystopian/utopian axis for the human components, but an inclination towards the utopian seems easily practical, and worth encouraging. And one possible means of encouragement is to engineer the infrastructure of HC in ways that facilitate computational behaviours that are as satisfying as possible for the agents engaging in them.

This chapter encompasses two major research themes that can enrich the architectural capabilities of HC systems. In one theme, the basic computational architecture is enhanced with structures like shared memory, analogues of device drivers, and reward-structure programmability that enrich the sort of tasks that can be performed by HC-based systems in a fundamental way. In the other theme, improvements to modelling (of users, of the state or results of joint human-computer computations, or both) serve to support interactions that are more natural, more appealing to human participants, or provide for the more effective reuse of (and therefore value creation from) the results of human effort.

Dominic DiFranzo and James Hendler open the section within the second theme, advocating the use of Linked Open Data (LOD) methodologies to support the creation, via Human Computation, of content that is more persistent, reusable, and reused, and therefore more valuable. In the other direction, they advocate the increased and more widespread use of curation in an HC framework in efforts to mitigate the severe data-quality issues in current LOD datasets. The section editor then expands on this theme, illustrating how the sophisticated knowledge representation capabilities of the Cyc knowledge-representation and reasoning system can be applied in a practical setting to drive machine interactions with human contributors that are more natural and conversational, and less like data entry and curation. In a broad sense, both these chapters envision using richer formal knowledge representations to bring the underlying computational system closer to the ability of the human "workers" (who may also be users), enabling the computer to support more sophisticated and sustained human computations.

Schall uses formal descriptions of HC processes for an orthogonal purpose that moves HC software towards industrial reliability, showing how both human and machine computation can be formalised within a service description framework enabling integration into service-description-based Business Process Management (BMP) systems. The chapter describes how characteristics that might seem unique to Human Computation, such as matching tasks with qualified performers, and dealing with intermittent availability of both tasks and human workers, can be readily mapped onto formal service description concepts such as task descriptions and service delegation.

The other major theme: fundamental architectural enhancement of HC systems, is represented in (Lasecki et al. 2012) Chorus system. In Chorus, simple agents provide temporary unplanned teams of human Mechanical Turk workers with the connectivity and memory infrastructure required to allow them to collaboratively produce a coherent, useful conversation. The agents also manage worker staging, to ensure that the system is highly responsive. These fundamental architectural enhancements of the HC system (latency control and shared memory) support a qualitatively different kind of HC application, assembling a responsive, conversational, unitary agent (in this case, a location-based information agent) out of constantly forming and re-forming teams of workers.

The observation that these spontaneous teams often volunteer useful information beyond what is asked for is at least suggestive that the team members find it a satisfying process at an individual level. We might, then, imagine extending this system to one in which collaborating teams of human and computer agents perform the task while simultaneously trying to increase the performance and satisfaction of other agents. For example, human agents could be offered the choice of contributing directly to a particular conversational interaction or of providing training input to a software agent that produced a near-miss response. Similarly, software agents could be offered a choice of offering a potential interaction response, or of offering input (from, for example, a tedious interaction on the web, or a mathematical or otherwise involved computation such as image transcoding, or from the interaction history of the user) to a human agent to inform the human's contribution. As such a collaborative system evolved, we might expect that the interactions it could support would become richer and richer, while becoming increasingly interesting to the human computation components.

Sun and Dance, in their chapter, consider the division of labour between humans and computers in human-computer collaborative systems, and more broadly consider humans as components in computational systems. They discuss both the advantages of humans (ability to partially solve NP-complete problems, and to solve AI complete problems) and their weaknesses as computational resources (cognitive biases, memory limitations). Finally they outline how use of humans as computational resources might be improved, by understanding more clearly how the API should be addressed (i.e. how the instructions should be formulated), by formulating a coherent notion of intuitiveness, and by beginning to characterise computational complexity with respect to HC implementations.

Lathia broadens the section's view of HC by considering humans as distributed sensing and computing platforms, both with on-board sensing and computation, and with "peripheral systems" such as telephones, activity trackers and smartcards. The chapter characterises sensing both in terms of its effect on the user (e.g. obtrusiveness) and the nature of the data produced (e.g. the granularity of the spatial region over which it is gathered). The issues and potential benefits of sensing-at-the-human are illustrated by the use of smart-card (Transport for London "Oyster" cards) data to identify a need for fare selection optimisation and possible means to achieve it.

Pulling back to the abstract level a little, Morishima adds to Schall's work on process representation and Laesecki's introduction of persistent storage and inter-human communication, both of which add to the programming methodology of HC systems. Morishima's work revolves around developing a programming language (CyLog) that abstracts humans as data sources, and HC operations, formalised as rules, as partial means to satisfy database queries. As an admittedly imperfect but useful abstraction, CyLog models humans as rational agents, and uses the Aspect Oriented Programming paradigm (Kiczales et al. 1997) to cleanly compartmentalise the reward structure of the HC system from the task formalisation itself. Morishima also introduces Crowd4U, a platform based on CyLog, which enables researchers to experiment with HC. Finally, Castelli, Mamei, Rosi & Zambonelli increase the scope of abstraction one level further, by introducing SAPERE, which considers HC as an element in a self-aware pervasive computational ecosystem which describes all of its agent elements uniformly using semantic annotations. SAPERE uses ecological abstractions, such as spread and decay, to combine agent semantic annotations with topographic and other environmental criteria to model agent interactions. The SAPERE abstraction is illustrated by using it to model a HC task of activity recognition based on closed-circuit camera sensing.

All the work in this section represents exciting progress in extending from the very limited "systems level" design facilities of "micro-task" platforms such as Mechanical Turk or CrowdFlower, or "macro-task" platforms such as Task Rabbit, eLance, 99 Designs and VoiceBunny by adding what are, in effect, higher level programming abstractions designed around the capabilities of humans. These richer models, in turn enable the creation of systems offering HC tasks with far more complex structure than those—perhaps problematic with respect to worker satisfaction—tasks that are currently widespread (such as "Captchas", sentiment analysis, or photograph content filtering). This evolution offers the very real possibility of building systems that are highly effective and efficient, that can produce complex results of enduring value, and that provide an enjoyable and fulfilling experience for the agents involved.

References

Kiczales G, Lamping J, Mehdhekar A, Maeda C, Lopes CV, Loingtier J, Irwin J (1997) Aspect-oriented programming. In: Proceedings of the European conference on object-oriented programming (ECOOP), Springer LNCS 1241, June 1997

Lasecki WS, Kulkarni A, Wesley R, Nihols J, Hu C, Allen JS, Bigham JP (2012) Chorus: letting the crowd speak with one voice. Technical report #983, University of Rochester

Interactive Crowds: Real-Time Crowdsourcing and Crowd Agents

Walter S. Lasecki and Jeffrey P. Bigham

Introduction

People can now connect instantly via nearly ubiquitous broadband connections, enabling interactive computational systems that are no longer constrained to machine-based automation alone. Instead, they can work in concert with the on-demand labor of people available on the web (the crowd), recruited on-demand and working synchronously together to complete tasks in real-time. The resulting model resembles a Distributed AI (discussed in Chapter <Dist AI>), but with a mix of human and machine agents composing the network.

Crowdsourcing workflows typically involve dividing tasks into smaller, separable tasks (Little et al. 2010; Dai et al. 2010). Using independent tasks provides an effective means of leveraging human intelligence to solve discretized problems, such as image labeling, offline transcription, handwriting recognition, and more. However, this model cannot handle acquiring consistent input for an ongoing task from workers. In order to expand the power of crowd algorithms, new models have been introduced that present approaches for continuous real-time crowdsourcing. This allows the crowd to be used to generate responses within the few-second time window needed for interactive tasks (Nielsen 1993). In these workflows, workers are engaged for longer periods of time, allowing them to receive feedback from the system as the task evolves due to their own input, as well as the input of others.

W.S. Lasecki (✉)
Department of Computer Science, University of Rochester, Rochester, USA
e-mail: wlasecki@cs.rochester.edu

J.P. Bigham
Human-Computer Interaction Institute, Carnegie Mellon University, Pittsburgh, USA

Department of Computer Science, University of Rochester, Rochester, USA
e-mail: jbigham@cmu.edu

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_38,

We describe how models of continuous crowdsourcing can be used to enable task completion using both synchronous and asynchronous groups of workers. We then explore a new model of continuous real-time crowdsourcing called a "crowd agent" that allows groups of workers to interact with both users of crowd-powered systems and their environment, as if they were a single individual. This model provides a means of abstracting away the collective in crowdsourcing by making the crowd appear as a single intelligent entity.

Next, we discuss a set of crowd agents that have been developed based on this new model that are capable of a variety of different functions and actions that were not previously possible to complete using the crowd. Finally, we conclude with a discussion of the potential future advances that this approach enables.

Background

Human computation was introduced to integrate people into computational processes to solve problems too difficult for computers to solve alone, but has not been applied to real-time control problems. Human computation has been shown useful in writing and editing (Bernstein et al. 2010), image description and interpretation (Bigham et al. 2010; von Ahn and Dabbish 2004), and protein folding (Cooper et al. 2010), among many other areas.

Most abstractions for human computation focus on increasing quality, and generally introduce redundancy into tasks so that multiple workers contribute and verify the results at each stage. For instance, guaranteeing reliability through answer agreement (von Ahn and Dabbish 2004) or the find-fix-verify pattern of Soylent (Bernstein et al. 2010). Unfortunately, this takes time, making these approaches poorly suited for real-time domains. For a more in-depth discussion of several of the most widely used crowdsourcing workflows, see Chapter <Workflows>.

Several systems have previously explored how to make human computation interactive. For example, VizWiz (Bigham et al. 2010) answers visual questions for blind people quickly. It uses quikTurkit to pre-queue workers on Amazon's Mechanical Turk so that they will be available when needed. Crowd agents need multiple users to be available at the same time in order for its input mediators to work correctly. Prior systems have also needed multiple workers to be available. For instance, the ESP Game encouraged accurate image labels by pairing players together and requiring them both to enter the same label, although ESP Game players could also be paired with simulated players (von Ahn and Dabbish 2004). Seaweed reliably got Mechanical Turk workers to be available at the same time to play economic games by requiring the first worker to arrive to wait (generally for a few seconds) (Chilton 2009). Crowd agents similarly utilize the input of multiple workers and ask workers to wait until other workers have arrived, but engages them for longer control tasks. Specialized remote control systems even allow aircraft to be piloted remotely. The main difference between these prior systems and general real-time crowd control systems such as Legion, which allows multiple workers to control a single interface, is the idea that multiple workers can collectively solve a problem as a single, more reliable worker.

Assistive Crowds

Prior work has shown how crowds can be used to assist users in their daily lives. Systems such as VizWiz (Bigham et al. 2010), which provides blind users with answers to visual questions in nearly real-time, shows that crowds can provide vital aid to users. Soylent (Bernstein et al. 2010) introduced the idea that crowd work could be made accessible from inside our existing applications—in that case, inside an existing word processor. Similarly, EmailValet (Kokkalis et al. 2013) uses the crowd to generate to-do lists from a partial view of a user's inbox. Mobi (Zhang et al. 2012) helps a user by generating a travel itinerary offline. But to truly work *with* the crowd, these systems need to be able to be recruited quickly and work synchronously with the end user. Soylent's reported delays of tens of minutes make the difference between collaborative cooperative work and iteration.

Overview

In this chapter, we present a discussion of the following:

- Real-time crowdsourcing, which provides responses to users within seconds
- Continuous crowdsourcing, which allows workers to engage in longer individual sessions to complete tasks which require workers to maintain context
- Crowd Agents, which combine the input of multiple workers contributing to continuous real-time tasks into a single output that retains the properties of a single reliable individual

Real-time crowdsourcing grew from the need for assistive systems, but allows a greater range of capabilities and interaction than was previously possible. These crowd-powered systems provide a useful service to end users, as well as insight into how users would interact with intelligent systems if they worked robustly.

Real-Time Crowdsourcing

Crowdsourcing has been shown to be an effective means of leveraging human computation to compute solutions to difficult problems; however, existing models only support usage in offline cases. Enabling systems that are able to use human computation to quickly and intelligently respond to user input can support a key aspect of nearly all systems: interaction, but requires latencies of only a few seconds. Current platforms such as Mechanical Turk typically require workers to browse large lists of tasks, making it difficult to recruit workers within such a small time window. To recruit crowd workers to answer immediately, work on real-time and nearly real-time crowdsourcing has looked at pre-recruiting workers for a given task, then having them remain partially engaged with a task until they are prompted to switch to a new one (Bigham et al. 2010; Bernstein et al. 2011). Using this approach, it is possible to get workers to join a task in less than 1 s (Bernstein et al. 2012).

Applications

With such low response times possible, we can begin to think of interactive systems powered by the crowd. Bernstein et al. used these quick-acting crowds to enable a camera application that filters a short video down to a single best frame within seconds. Unlike Soylent, real-time crowds allow Adrenaline (Bernstein et al. 2011) to work behind the scenes the same way an automated would, without an explicit request step by the user, and VizWiz allows users to ask time-relevant questions and get responses within a minute. This responsiveness enables a new style of interaction with the crowd: seamless integration of new functions within the paradigm of traditional interfaces.

Continuous Crowdsourcing

Even with synchronous real-time systems, there are tasks that cannot be completed using one-off responses. For instance: instead of selecting a video frame, what if we wanted workers to help caption the video in real-time? Traditional approaches would divide the task into multiple pieces, and ask workers to accept a short task transcribing one of them. However, this means these approaches do not allow workers to maintain context of the topic or the terms being used, often divide words over two different pieces, and require workers to immediately recognize their place in the task and begin working from there in order to work properly. All of these factors reduce workers' ability to complete the task quickly and correctly.

Furthermore, issuing discrete tasks presents a model in which workers are frequently interrupted by either being asked to change topics or delayed before continuing to a subsequent task. Industrial and organizational psychology and cognitive science have looked at modeling the effects of these interruptions. For instance, prospective memory measures the ability to remember an ongoing task when interrupted by another. This ability to remember context is most strongly affected by the length and magnitude of the topic difference in the interrupting task. Unfortunately, on many crowd platforms, it would not be surprising to have a multi-part college course transcription task interleaved with a flower-labeling task.

In order to maintain context and allow workers to interact more robustly with the task at hand (for instance, learning new words that are used later in a conversation, or reacting to an object falling in the path of a robot being driven by the crowd), Lasecki et al. introduced the idea of continuous crowdsourcing (Lasecki et al. 2011). In continuous crowdsourcing tasks, workers are engaged for longer periods of time in order to allow them to maintain this context. In the most general case, they are asked to connect to a task for as long as they choose, and will be able to seamlessly continue working on the same job until they choose to leave. In this model, workers must be compensated for their input on the fly in order to make the payments scale with the size of the task. In the next section, we will see how the idea of continuous crowdsourcing can be combined with real-time and synchronous tasks to enable the crowd to provide accurate, highly generalizable feedback.

Crowd Agent Model

Real-time, synchronous, and continuous tasks each individually present means of providing functionality in a fundamentally different way. However, in order to leverage human computation in this way, we are forced to develop one-off systems that use this collective input to accomplish some task. For instance, continuous crowdsourcing offers many advantages, but presents issues with how to provide users or systems with reliable responses (those verified by agreement between workers) in real-time. Allowing for repeated or multi-stage interaction, using the crowd requires a framework for merging collective input in real-time.

To address these issues, Lasecki et al. (2011) introduced the Crowd Agent model of crowdsourcing. This model recruits workers to complete synchronous continuous tasks in real-time, then uses a task-specific input mediator to merge their results into a single output. This allows the system to harness the power of collective intelligence, while abstracting away the multitude of responses and allowing the user or system to interact with a single, highly skilled, intelligent agent. Furthermore, using this model presents the ability to easily partially automate tasks by using existing automated systems as individual contributors, allowing systems to benefit from a synthesis of human and machine intelligence, and to scale gracefully towards becoming fully automated in the future.

Advantages

This method not only allows for new types of repeated-interaction tasks to be completed, but also strives to enable the crowd to retain many of the beneficial aspects of a single human user. Some of these properties that endow the crowd with a sense of agency include:

- **Unified Output:** By merging the input of workers in real-time, these systems create a single output stream, similar to that of a single skilled user. This is a critical aspect for combining crowd-powered systems with existing single-user systems (i.e. GUIs).

- **Collective Memory:** Organizational memory refers to a process in which groups of individuals collectively remember and pass down information from one generation to the next. In the context of crowdsourcing, part actions and decisions can be passed down to the current set of workers both explicitly via messages or labels (Lasecki et al. 2013b) and implicitly via behaviors (Lasecki et al. 2012b), even when no workers from the current session were present when the memory was created.
- **Consistent Output:** We can leverage the idea of collective memory to support consistent actions by the crowd. That is, actions or behaviors that are in line with the crowd's previous actions. This is important in systems that users engage in repeated interactions with, such as intelligent assistants (Lasecki et al. 2012c), where the prior interactions must be reliably recalled to facilitate the interaction.

Crowd Agents

In this section, we briefly describe a few recent systems that have been developed using the crowd agent model, and explore how they each demonstrate new potential uses for and capabilities of the crowd (Fig. 1).

Legion

Legion (Lasecki et al. 2011) is a system that enables the crowd to control existing single-user interfaces. It was the first work to introduce the idea of crowd agents. It leveraged the ability of this model to create a single combined output using an input mediator to control existing interfaces without the need to modify them. Legion's input mediator selected a single 'leader' at any given time step (usually

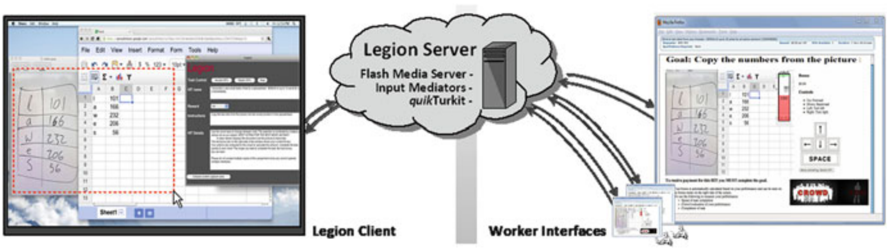


Fig. 1 Legion system architecture. In this example, a user has outsourced a spreadsheet text-entry task. The Legion client allows end users to choose a portion of their screen to send to crowd workers (outlined in *red* on the *left*), sends a video stream of the interface to the server, and simulates key presses and mouse clicks when instructed by the server. The server recruits workers, aggregates the input of multiple workers using flexible input mediators, and forwards the video stream from the client to the crowd workers. The web interface presents the streaming video, collects worker input (key presses and mouse clicks), and gives workers feedback

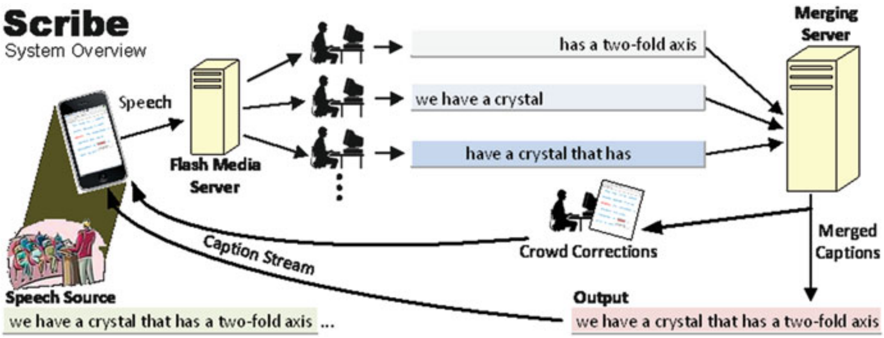


Fig. 2 Legion:Scribe system. Audio is sent to multiple non-expert captionists who use Scribe's web-based interface to caption as much of the audio as they can in real-time. These partial captions are sent to a server to be merged into a final output stream, which is then forwarded back to the user's mobile device

~1 s in length) to control the system, instead of averaging inputs, taking a vote, or letting a single individual or series of individuals control an interface. This leader is selected from the set of workers based on their past agreement with the collective as a whole, and at each time step workers are re-ranked, and the best leader at that point is selected.

Legion was demonstrated effective on a variety of tasks that ranged from robot navigation to controlling word processing applications, performing OCR, and enabling assistive keyboards. Further work demonstrated the crowd's ability to remember information over time, even in the presence of complete worker turnover, on a virtual navigation task (Lasecki et al. 2012b).

Legion:Scribe

Legion:Scribe (aka 'Scribe') (Lasecki et al. 2012a) is a system that enables groups of non-expert typists, such as those available from the crowd, to caption audio in real-time. To accomplish this, Scribe extended the underlying input mediator used in Legion to synthesis inputs from the whole set of workers instead of deciding on a single worker to listen to at any given time. This merger is performed using Multiple Sequence Alignment (MSA), a process most commonly associated with genome sequencing in computational biology (Naim et al. 2013). Using this approach, Scribe is able to use the partial captions of multiple workers to generate a single complete final transcript (Fig. 2).

Real-time captioning converts speech into text with a per-word latency of less than 5 s. Real-time captions are a vital accommodation for deaf and hard of hearing people that provides access to spoken language. Prior to Scribe, the only viable option for providing real-time captions were expensive and hard-to-schedule professionals who required 2–3 years of training and cost \$100–\$300 per hour or more (depending on skill set). Since Scribe is able to use anyone who can hear and type,

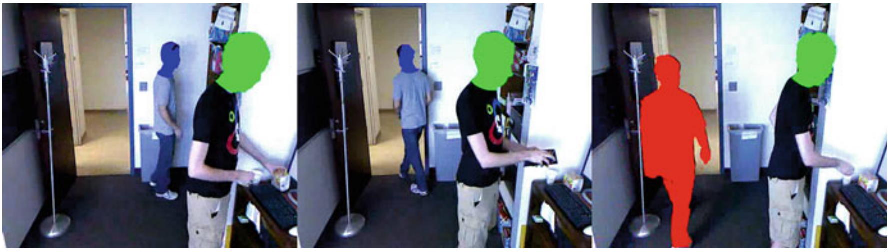


Fig. 3 Example of the video steam (with automatically generated privacy veils) that Legion:AR presents to workers during the activity labeling process. Each actor in the scene is color-coded to allow workers to identify which person's actions are being labeled

without the need for prior training, workers can easily be recruited for \$8–10 per hour. Scribe can use as few as 3–6 workers to reach professional-level quality, meaning the same accommodation can be provided for a fraction of the cost. Furthermore, by recruiting crowd workers on-demand, the captioning services can be made far more available than is possible using professionals.

Legion:AR

Legion:AR (Lasecki et al. 2013b) is a system that uses the crowd to generate activity recognition labels in real-time. Activity recognition is important because it allows systems to provide context-relevant responses to users. Legion:AR focuses primarily on two domains: (i) an assistive living domain in which prompting and monitoring systems help cognitively impaired and older users live more independently for longer, and (ii) a public monitoring domain in which systems provide timely assistance by observing actions in a public space, such as calling an ambulance when a car accident is observed, or calling the police when an armed robbery or other crime is observed (Fig. 3).

Automated systems have struggled with these types of tasks because people perform actions in very different ways and in very different settings. In most cases these variations require explicit training in advance, meaning that systems are brittle and unable to handle change or new actions well. Legion:AR allows an automated system to call on the crowd in real-time to provide labels and training data. Unlike even experienced labelers, Legion:AR is able to produce these labels with extremely low latency, and keep up with live video. To do this, it uses an input merging approach similar to Legion:Scribe, in which labels are aligned and merged into a single stream. In order to provide consistent labels, it also presents workers with a display of what the other workers are currently suggesting, and labels those that have previously been used for similar activities (Fig. 4).

Once these labels are generated, they are used to train an automatic system. This means that over time, the system is able to gracefully transition from being fully crowd-powered to being fully automated.

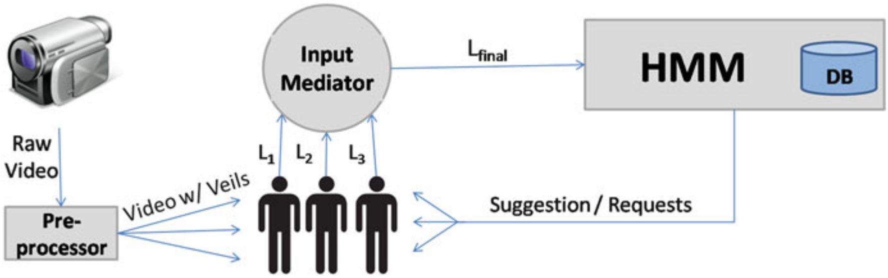


Fig. 4 Legion:AR system. Workers submit labels to Legion:AR, which forwards a final label and segment to train the HMM. The HMM can then add this to a database and use the information to identify the activity later and forward that prediction back to the crowd

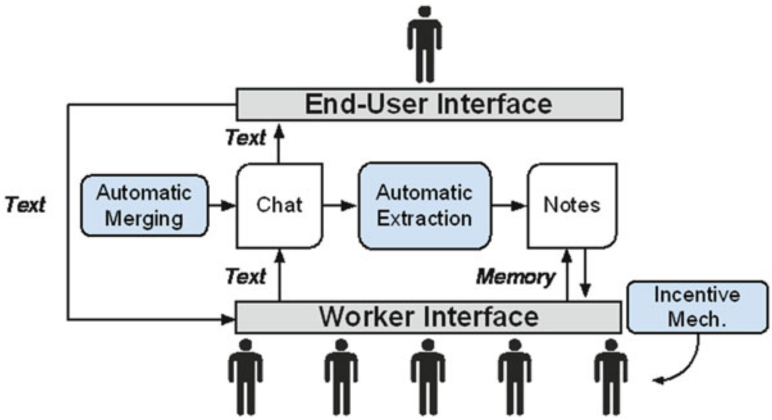


Fig. 5 Chorus system. Chorus is a framework for conversational assistance that combines human and machine intelligence. To end users, Chorus operates like a typical instant messaging client—workers type or speak messages, and receive responses in natural language. Crowd workers propose and vote on candidate responses, and are motivated to contribute via the incentive mechanism

Chorus

Chorus (Lasecki et al. 2013c) is a system that provides a conversational interface to a crowd-powered personal assistant. Conversational interaction has been a goal of natural language processing (NLP) researchers for several decades because it enables fluid interactions that more closely resemble those between people. Chorus leverages people's understanding of the state of the conversation, the context it occurs in, and common knowledge facts, to complete a requested information-finding task (Fig. 5).

Fig. 6 An example of a multi-part question that required accurate framing in the image. This task took over 10 min to complete using VizWiz, while similar tasks take less than a minute using Chorus:View because workers can help the user iteratively refine the framing and use prior knowledge to answer questions

In order to respond with the same consistency as a single assistant would, the Chorus uses an explicit propose-and-vote scheme combined with an incentive mechanism that encourages workers to contribute useful information, and then agree with others if their response would be more appropriate. It also provides a 'notes' window to explicitly allow a second synchronous crowd to curate memories from throughout the conversation. This allows workers to not only be consistent within a given discourse, but also between multiple sessions in which different workers may be present. Experiments showed that the crowd was able to answer over 84 % of user questions (including follow-up and clarification questions) accurately, and successfully remember information from prior interactions.

Chorus:View

Chorus:View (Zhong et al. 2012) combines the ability to hold conversations with the crowd with a visual question answering service similar to VizWiz. By using streaming video instead of single images, and allowing users to engage in an ongoing conversation with the crowd about what is shown, View enables a much more fluid support of visual tasks, similar to how a single person assist if they were collocated with the user. This is particularly beneficial for tasks that involve consecutive questions, such as finding a given type of food, locating the instructions, then finding the cook time, or even just framing an item in the camera's view. All of which turn out to be very common tasks that blind users need answers to (Brady et al. 2013) (Fig. 6).

Future Directions

Using the models of crowd work discussed here allows for a class of systems which can interact with users in a natural way, understand the context of the surroundings, and respond in a manner consistent with prior interactions. In the future, we expect to be able to expand the scope of the human intellect that such systems are able to harness in order to understand longer term collective beliefs, desires, and intents. We can also use these crowds as a means of training existing artificial intelligence systems on-demand.

Novel Workflows

Using multiple synchronous workers instead of just one allows for novel workflows that would not otherwise be feasible using a single individual. For example, Lasecki et al. (2013a) showed that using multiple automatically coordinated workers, it is possible to slow down the playback rate of audio to make the captioning task in Scribe easier, while reducing the per-word latency. Slowing down the audio to be captioned presents workers with an easier motor task, resulting in increased performance. This also allows workers to keep up with each word as they hear it instead of first listening to a segment of audio, memorizing it, and then typing it, meaning latency is also reduced. However, while effective, this approach is not possible to use in real-time with a single worker, who would necessarily fall behind. In contrast, by using the crowd, it is possible to automatically interleave workers so that someone is always captioning live content. In the future we expect to see more such workflows that improve on what is possible for a single user to accomplish alone.

Improved Synthesis with Automatic Approaches

For many of the problem domains presented above, automated systems exist that try to generate solutions with varying degrees of success. One of the greatest benefits to the crowd agent model is that it treats each contributor as a noisy input, while being agnostic to exactly how it has generated its answer. This allows automated systems to be used as individual contributors that can learn and grow over time, just as human workers do. Using multiple inputs simultaneously also presents new opportunities to train systems which are still being explored. As seen above in Legion:AR, it also provides a means of smoothly transitioning between a system that is fully crowd powered to one that is fully automated, all without ever needing to expose end-users to an unreliable system during training.

Conclusion

In this chapter, we have presented an overview of the current work in interactive crowd systems, as well as some possible directions of future work. Work on continuous real-time crowdsourcing systems promises to enable interactive intelligent systems that can both operate using human intelligence and train automated systems in novel ways. They also allow human workers to work jointly on tasks in novel and efficient ways. The potential of these models is just beginning to be explored, but these systems lay the groundwork for interactive intelligent systems, powered by the crowd.

References

Bernstein MS, Little G, Miller RC, Hartmann B, Ackerman MS, Karger DR, Crowell D, Panovich K (2010) Soylent: a word processor with a crowd inside. In: Proceedings of the 23nd annual ACM symposium on user interface software and technology, UIST'10, ACM, New York, pp 313–322

Bernstein MS, Brandt JR, Miller RC, Karger DR (2011) Crowds in two seconds: enabling real-time crowd-powered interfaces. In: Proceedings of the 24th annual ACM symposium on user interface software and technology, UIST'11, ACM, New York, pp 33–42

Bernstein MS, Karger DR, Miller RC, Brandt J (2012) Analytic methods for optimizing real-time crowdsourcing. In: Proceedings of the collective intelligence, Boston, MA

Bigham JP, Jayant C, Ji H, Little G, Miller A, Miller RC, Miller R, Tatarowicz A, White B, White S, Yeh T (2010) Vizwiz: nearly real-time answers to visual questions. In: Proceedings of the 23nd annual ACM symposium on user interface software and technology, UIST'10, ACM, New York, pp 333–342

Brady E, Morris MR, Zhong Y, Bigham JP (2013) Visual challenges in the everyday lives of blind people. In: Proceedings of the ACM SIGCHI conference on human factors in computing systems (CHI 2013), Paris

Chilton L (2009) Seaweed: a web application for designing economic games. Master's thesis, MIT

Cooper S, Khatib F, Treuille A, Barbero J, Lee J, Beenen M, Leaver-Fay A, Baker D, Popovic Z, Players F (2010) Predicting protein structures with a multiplayer online game. Nature 466(7307):756–760

Dai P, Mausam, Weld DS (2010) Decision-theoretic control of crowd-sourced workflows. In: Twenty-fourth AAAI conference on artificial intelligence (AAAI 2010), Atlanta

Kokkalis N, Köhn T, Pfeiffer C, Chornyi D, Bernstein MS, Klemmer SR (2013) EmailValet: managing email overload through private, accountable crowdsourcing. In: Proceedings of the 2013 conference on computer supported cooperative work (CSCW '13), ACM, New York

Lasecki WS, Murray K, White S, Miller RC, Bigham JP (2011) Real-time crowd control of existing interfaces. In: Proceedings of the ACM symposium on User Interface Software and Technology, UIST'11, ACM, New York, pp 23–32

Lasecki WS, Miller CD, Sadilek A, Abumoussa A, Borrello D, Kushalnagar R, Bigham JP (2012a) Real-time captioning by groups of non-experts. In: Proceedings of the ACM symposium on User Interface Software and Technology (UIST 2012), Boston, MA, pp 23–34

Lasecki WS, White S, Murray K, Bigham JP (2012b) Crowd memory: learning in the collective. In: Proceedings of collective intelligence (CI'12), Boston

Lasecki WS, Miller CD, Bigham JP (2013a) Warping time for more effective real-time crowd-sourcing. In: Proceedings of the international ACM conference on human factors in computing systems, CHI'13, page to appear, Paris, France

Lasecki WS, Song Y, Kautz H, Bigham J (2013b) Real-time crowd labeling for deployable activity recognition. In: Proceedings of the international ACM conference on computer supported cooperative work and social computing (CSCW 2013), San Antonio, TX, pp 1203–1212

Lasecki WS, Wesley R, Nichols J, Kulkarni A, Allen JF, Bigham J (2013c) Chorus: a crowd-powered conversational assistant. In: Proceedings of the 23nd annual ACM symposium on user interface software and technology (UIST'13), St. Andrews, UK. UIST 2013. PP 151-162

Little G, Chilton LB, Goldman M, Miller RC (2010) Turkit: human computation algorithms on mechanical Turk. In: Proceedings of the 23nd annual ACM symposium on user interface software and technology, UIST'10, ACM, New York, pp 57–66

Naim I, Lasecki WS, Bigham JP, Gildea D (2013) Text alignment for real-time crowd captioning. In: Proceedings of the North American Chapter of the association for computational linguistics (NAACL 2013), To appear, Atlanta, GA

Nielsen J (1993) Usability engineering. Morgan Kaufmann, San Francisco

von Ahn L, Dabbish L (2004) Labeling images with a computer game. In: Proceedings of the SIGCHI conference on human factors in computing systems, CHI'04, ACM, New York, pp 319–326

Zhang H, Law E, Miller RC, Gajos K, Parkes DC, Horvitz E (2012) Human computation tasks with global constraints. In: Proceedings of the SIGCHI conference on human factors in computing systems (CHI'12), pp 217–226, Austin, TX

Zhong Y, Thiha P, He G, Lasecki WS, Bigham JP (2012) In: Proceedings of the ACM conference on human factors in computing systems work-in-progress (CHI 2012), Austin

The Semantic Web and the Next Generation of Human Computation

Dominic DiFranzo and James Hendler

Introduction

Human computation has come a long way in the pursuit of joining the abilities of humans and machines to solve interesting problems that neither could have done alone. While we have seen great success in human computation in everything from annotating images (Von Ahn 2006) to building global knowledge repositories (Yuen et al. 2009), these systems are still confided to the platforms and data models they were built from. They are silos of services and data, disconnected from each other at both the data and application level. These systems are usually single purpose, and not general enough to be used with new or different data, or in situations different than they were built for. The datasets and data models these systems produce and use can also be difficult to repurpose or reuse. Users of these systems also can't easily move from one system to the next, instead having to create completely new accounts, having a new identity and different reputation. Their past work and reputation do not transfer over to different human computation systems.

Semantic Web and Linked Data

In the past decade, Semantic Web technologies and linked open data have worked to solve many of the challenges of linking heterogeneous data and services together (Shadbolt et al. 2006). The World Wide Web Consortium (W3C) defines the Semantic Web as "a common framework that allows data to be shared and reused across

D. DiFranzo (✉) • J. Hendler
Tetherless World Constellation, Rensselaer Polytechnic Institute, Troy, NY, USA
e-mail: difrad@rpi.edu

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_39,

application, enterprise, and community boundaries". Where as the World Wide Web is a collection of documents linked together, the Semantic Web is a web of data, concepts, ideas and the relationships that connect them together. This technology allows for data to be integrated in a seamless and schema-less fashion and moves us beyond the unstructured "throw-away" data, often produced by human computation systems, to a more sophisticated system that attempts to bridge the worlds of knowledge representation (the study of how best to represent and encode knowledge) and information presentation (how that information should be displayed) (van Ossenbruggen and Hardman 2002). This structuring and connecting of raw data allows the implicit semantics to become more explicit, allowing for both machines and humans to gain real and new information that may have been hidden before.

The Linked Open Data Cloud is a realization of this web of data. It is a network of over 295 different heterogeneous datasets (covering domains in Life Science and Health Care, Government, Finance, Academic Publications, Geographic, Media and User-generated Content) linked together. A visualization of this data cloud along with statistics can be seen at http://lod-cloud.net/.

While the LOD cloud has grown and expanded greatly in recent years due to the 'publish first, refine later' philosophy of the Linking Open Data movement, groups like the Pedantic Web[1] have revealed issues in data quality and consistence in Linked Open Data (Simperl et al. 2011).

Take for example DBPedia, a semantic version of information found in Wikipedia articles and one of the central nodes in the Linked Open Data Cloud (Auer et al. 2007). DBPedia has an entry for every county in the US, but it is impossible to automatically query or lists all of these entries as the labeling of these US counties are not consistent, or in many cases even exist. This makes the data difficult if not impossible to use. These data errors are difficult to find using automated methods and is often left for the community at large to resolve (Simperl et al. 2011). However the incentive structures for this community to help improve the quality and consistence of Linked Data (assuming there even exists procedures and interfaces that allow this) are usually flawed as there is a disconnect between the effort involved, and the benefit received. Human computation may be able to fix the broken incentive scheme of Linked Open Data (Siorpaes and Hepp 2008a).

It's also not just in data curation that human computation can be used in Linked Data, but in all parts of the Linked Data life cycle. Link Data creation, alignment and annotation all reflect the interest and expertise of many people, and as such are community efforts that can be facilitated by human computation (Siorpaes and Hepp 2008a). Linked Data also needs better ways to create and manage meta-data and provenance, which is currently missing in many important Linked Datasets. In all of these areas, human abilities are indispensable for the resolution of those particular tasks that are acknowledged to be hardly approachable in a systematic, automated fashion. Human computation can help provide, at scale, these human abilities needed to improve and expand Linked Data (Simperl et al. 2011).

[1] http://pedantic-web.org/

By combining Semantic Web and Human Computation, we get the best of both worlds. We can build the next generation human computation system that reach beyond the platform they were built upon. Their data can be reused repurposed in ways never intended by their creators (thereby answering the call of the Long Tail, in that a large percentage of requests made to a web system are new, or unintended (Anderson 2007)). Likewise, Linked Open Data can be improved by using human computation strategies to help curate and improve data quality, data linking, vocabulary creation, meta-data (data that defines and explains data) and provenance (history of how data has been created, modified and used). In other words, Human Computation can be used to help curate Semantic Web data in the Linked Open Data cloud while the Semantic Web can be used to provide better user continuity and platform consistency across Human Computation systems.

In rest of this chapter, we will consider how Human Computation systems have been used to curate Linked Open Data. We will then discuss some of the large questions and challenges facing promise of Semantic Web technologies being used to connect and expand Human Computation systems.

Examples of Semantic Web Curation Using Human Computation Systems

Many researchers have already explored combining human computation strategies with Semantic Web technologies with the purpose of improving and curating Linked Data. The OntoGames series covers what they list as the complete Semantic Web life cycle in their series of games (Siorpaes and Hepp 2008b). They list this life cycle to be ontology creation, entity linking, and semantic annotation of data. In this section we list the games that fit into these categories (both in and out of the OntoGames series) along with some semantic human computation systems that fall into categories outside these three.

Ontology Creation

This first section deals with applications that use human computation to help collaboratively build ontologies. OntoPronto (a part of the OntoGame Series) is a "Game with a Purpose" to help build the Proto ontology (general interest ontology). In it, two players try to map randomly chosen Wikipedia articles to the most specific class of the Proton ontology. If they agree on a Proton class for their article, they get points and proceed with to next specific level (Siorpaes and Hepp 2008b).

Conceptnet is another system in ontology building. Conceptnet is a semantic network of common knowledge from it's users. The system works by having users ask other users questions related to interesting topics. The answers from these other users are then evaluated by a majority of the users and taken as basis for constructing the ontology (Havasi et al. 2007).

Entity Linking

One of the most important and difficult parts of Linked Data is linking similar entities together to create the connections between datasets. Different games and human computation system have been built to try to address this issue, allowing from better linking in the Linked Open Data cloud. One of the first of these systems was SpotTheLink (a game in the OntoGames Series). SpotTheLink is a collaborative game to link DBPeida to the Pronto Ontology (as created in the OntoPronto game). In the game, players have to agree on which Pronto ontology class fits best with a randomly selected DBPedia entity. This agreement creates a mapping between these concepts, which is encoded in SKOS (Simple Knowledge Organization System). While the sample size was small, the users that did play the game would play a second or third round over 40 % of the time, showing the potential for replay value and more work done (Thaler et al. 2011).

ZenCrowd is another system that uses both automated techniques and crowdsourcing to link entities in unstructured text to the Linked Open Data cloud. It uses a probabilistic framework to create a list of candidate links from entities it mines out of unstructured text. It then dynamically generates micro-tasks from these entities and candidate link list for use in a mechanical Turk platform, where users complete these tasks to make or rate the links. This unique system allowed for larger amount of links to be created and verified than non semi-automated systems. Issues arose though with payment incentives for the mechanical Turk platform (Demartini et al. 2012).

UrbanMatch is a game used to help link photos on the web to the geographic points of interest they are from. It's a mobile game, played from a smartphone, that makes players actually visit these points of interest and verify that photos found online actual depict the landmark. The more links a player creates, the more points they gain. While accuracy and completeness of linking data using this game improved, it has yet to be seen if the game mechanics can motivate players to continue play, or invite other to play with them (Celino et al. 2012).

Data Annotation

A lot of the data and information on the web is in an unstructured format (free-text, images, videos, audio files, etc.). Lots of work on systems to bring semantic annotation and structure to this data has been taking place to help make this data more easily available and useable. OntoTube and OntoBay are two such systems that attempt to do this by allowing users to semantically annotate data in YouTube and eBay. In OntoTube, users are shown videos from YouTube, and asked from questions about the video. The answers to these questions correspond to semantic annotations that will be attached to the video, provided that a majority of the users agree in their answers. OntoBay works the same way, with a very similar interface, but instead of videos on YouTube, it works with listings on eBay. These simple games allow for the creation and annotation of new data into the Linked Open Data Cloud, and make the data in YouTube and eBay available to the greater Semantic Web community (Siorpaes and Hepp 2008a).

Data Quality

As stated before, data quality and consistency is a huge issue for Linked Data. Finding and correcting these errors are hard if not impossible to do in a completely automated way. Some researchers have looking into building games to try to incentivize people to help curate Linked Data. WhoKnows? is a quiz game styled after the “Who Want to Be a Millionaire?” game. While playing this game, users can report errors in the answers they get from the quiz. If enough player agree on these errors, a correction, or patch, to this dataset is created, using a patch ontology created by the developers. This allows the correction of the data to be taken in not only by the original publisher of the data, but also by other data consumers who may have their own local copy of the published Linked Data (Waitelonis et al. 2011).

Challenges

Although we have seen some of the current work in using Semantic Web technology with human computation strategies, there is still a lot of work to be done to bring out the full potential. As we discussed before, one of the hopes of merging Semantic Web technology with human computation is the ability to create new cross platform, linked and general-purpose human computation systems. To be able to build systems where the applications and not just the data could be linked together. The systems we reviewed used and contributed back to Linked Data, thus allowing their data to be linked, reused and repurposed by others. However, the applications themselves were standalone and usually single purpose. New work in semantic integration of services, like the RDFAgents protocol (Shinavier 2011) and SADI (Semantic Automated Discovery and Integration) services (Wilkinson et al. 2009) can help bring this along. Efforts to experiment with this technology in human computation systems are greatly needed.

The services we explored all had a wide array of interfaces presented to users. In fact many of the papers reported that developing an interface for their system to be one of the biggest challenges they faced (Siorpaes and Hepp 2008a; Simperl et al. 2011; Kochhar et al. 2010). Semantic Web applications are still very new and experimental, and designing user interfaces that best represent and produce semantic data need to be explored and tested more. How exactly a layperson will read and write to the Semantic Web is still poorly understood. New visualizations and useable designs will be needed.

Building interfaces to semantic systems is not the only challenge to their development. Many issues exist for building systems on top of services and datasets that are constantly in flux, have up-time and reliability constraints, issues in latency, and problems with data quality and consistency (Knuth et al. 2012). Some of these issues will dissipate as these technologies grow and mature, but some are inherent to the very nature of the Semantic Web. New methods in application development to address these challenges must be better researched and explored.

Inline with this challenge, but deserving of it's own mention, is scalability. Human computation on the Semantic Web will require scalability in both reading and writing semantic data. None of the services previously mentioned attained anything close to web scale (most had number of users in the hundreds). It is yet to be seen what a truly web scale semantic human computation system will look like, and what parts of it architecture will serve as the bottleneck.

New ontologies and vocabularies will need to be developed to help manage and link these human computation systems together. Some of the projects experimented a little with building some ontologies to help patch data corrections to Linked Data (Knuth et al. 2012), but more needs to be explored.

Semantic user management was something that was largely ignored in these examples. One of the hopes of using semantic web technologies in these systems is that users can easily sign on into new systems, and have their points and reputation follow them. There exist protocols in place to handle these use cases, like WebID.[2] This could create new avenues in linking these systems together. These technologies must be integrated and explored more.

Although the Linked Open Data Cloud is a collection of datasets that are linked together as one big graph, querying or building applications on top of it as if it were one connected graph is incredibly difficult if not impossible. Technologies like SPARQL Federated Query[3] and LinkedDataSail (Shinavier 2007) have come about as a response to this challenge. Semantic human computation system need to explore this space more to pull from all of the Linked Open Data Cloud, and not just a node inside it.

As stated before, the Linked Open Data Cloud suffers from a lack of meta-data and provenance information. According to http://lod-cloud.net/, over 63 % of the datasets in the LOD Cloud do not provide provenance meta-data. While many of the systems explored in this chapter worked on increasing accuracy and consistency in Linked Data, none of them had a focus on meta-data and provenance. This is a big gap in the LOD Cloud and human computation methods should be leveraged to help address it.

Many different types of human computation systems have been explored in the past (Yuen et al. 2009). It has yet to be seen what type of incentives, platform, games, rules, systems and architectures will work best for human computation on the Semantic Web (Siorpaes and Hepp 2008a). More experiments and tests must be done to explore this further.

[2] http://www.w3.org/2005/Incubator/webid/spec/

[3] http://www.w3.org/TR/2013/REC-sparql11-federated-query-20130321/

Finally, another area that the Semantic Web could be explored to improve Human Computation systems is by providing a more global, consistent "State Space". State Space in a Human Computation system is the collection of knowledge, artifacts and skills of both the human users and the computer system that help define the state, or current stage at a given time, of the Human Computation system. This State Space can often be very messy, disjointed, incomplete or inconsistent. The Semantic Web could provide this common platform and medium for representing knowledge that persists despite the asynchronous behaviors of the human participants. More research is needed to explore how this could work, how best to represent this knowledge, and what advantages this could bring to future Human Computation systems.

Conclusion

The future of human computation and the Semantic Web holds great promise. Although there have been some great experiments in building human computation system for the Semantic Web, there are still many challenges and questions left. More work on leveraging the best of semantic web technologies and human computation is greatly needed to bring forth this next generation.

References

Anderson C (2007) The long tail: how endless choice is creating unlimited demand. (Random House Business Books, New York, 2007).

Auer S, Bizer C, Kobilarov G, Lehmann J, Cyganiak R, Ives Z (2007) Dbpedia: a nucleus for a web of open data. In: The semantic web. Springer Berlin Heidelberg, pp 722–735

Celino I, Contessa S, Corubolo M, Dell'Aglio D, Della Valle E, Fumeo S, Krüger T (2012) Urbanmatch–linking and improving smart cities data. In: Linked data on the web workshop, LDOW

Demartini G, Difallah DE, Cudré-Mauroux P (2012) ZenCrowd: leveraging probabilistic reasoning and crowdsourcing techniques for large-scale entity linking. In: Proceedings of the 21st international conference on world wide web, ACM, pp 469–478

Havasi C, Speer R, Alonso J (2007) ConceptNet 3: a flexible, multilingual semantic network for common sense knowledge. In: Recent advances in natural language processing, pp 27–29

Knuth M, Hercher J, Sack H (2012) Collaboratively patching linked data.arXiv preprint arXiv:1204.2715

Kochhar S, Mazzocchi S, Paritosh P (2010) The anatomy of a large-scale human computation engine. In: Proceedings of the ACM SIGKDD workshop on human computation, ACM, Washington, DC, USA, pp 10–17

Shadbolt N, Hall W, Berners-Lee T (2006) The semantic web revisited. Intell Syst IEEE 21(3):96–101

Shinavier J (2007) Functional programs as linked data. In: 3rd workshop on scripting for the semantic web, Innsbruck

Shinavier J (2011) RDFAgents specification. Technical report 20110603, Rensselaer Polytechnic Institute

Simperl E, Norton B, Vrandecic D (2011) Crowdsourcing tasks in linked data management. In: Proceedings of the 2nd workshop on consuming linked data COLD2011 co-located with the 10th international semantic web conference ISWC 2011, Bonn, Germany

Siorpaes K, Hepp M (2008a) Games with a purpose for the semantic web. Intell Syst IEEE 23(3):50–60

Siorpaes K, Hepp M (2008) Ontogame: weaving the semantic web by online games. In: The semantic web: research and applications, Springer Berlin Heidelberg, pp 751–766

Thaler S, Simperl E, Siorpaes K (2011) Spotthelink: a game for ontology alignment. In: Proceedings of the 6th conference for professional knowledge management, Innsbruck, Austria

van Ossenbruggen JR, Hardman HL (2002) Smart style on the semantic web. Centrum voor Wiskunde en Informatica

Von Ahn L (2006) Games with a purpose. Computer 39(6):92–94

Waitelonis J, Ludwig N, Knuth M, Sack H (2011) WhoKnows? Evaluating linked data heuristics with a quiz that cleans up DBpedia. Interact Smart Edu 8(4):236–248

Wilkinson MD, Vandervalk B, McCarthy L (2009) SADI semantic web services-,cause you can't always GET what you want! In: Services computing conference. APSCC 2009. IEEE Asia-Pacific. IEEE, pp 13–18

Yuen MC, Chen LJ, King I (2009) A survey of human computation systems. In: Computational science and engineering. CSE'09. International conference on, vol 4, Vancouver, Canada. IEEE, pp 723–728

Conversational Computation

Michael Witbrock and Luka Bradeško

Introduction

Systems that crowdsource data to support everyday human activities such as listening to music, watching movies and going out to eat have become ubiquitous. The data-collection ranges from relatively unobtrusive (thumbs up in Pandora, stars in Netflix) to relatively intrusive (check-ins, writing reviews on Yelp or eBay or Amazon). A common characteristic of these systems, though, is that they use an interface that captures either only very simple data (item selections, integer ratings) if the data is to be machine understandable or they capture data whose meaning is opaque to the machine (e.g. blocks of text constituting a review). A second characteristic, which results in part from this limitation in the complexity of the crowd-sourced content that is actually understood by the machine, is that the behaviour of the interface changes relatively little, if at all, in response to the information the user has previously provided.

This stands in stark contrast to natural human interaction, in which responses are at worst partially understood, and in which both the content and the form of the interaction can change very dynamically based on what has happened previously.

We are interested in building AI systems that are true companions (see also Forbus and Hinrichs 2004), who care about the humans they work with, and who retain and use detailed information about those humans and their worlds to help them build lives that are more enjoyable and more enriching. Such systems will be characterised by heterogeneity—it is likely, for example, that knitting, or hang-gliding, or telenovelas are as important to some users as restaurants or coffee shops. To work, the systems will have to be able to elicit detailed knowledge of the entities,

M. Witbrock (✉)
Cycorp Inc. and Curious Cat Company, 7718 Wood Hollow Dt. Austin, TX, USA
e-mail: witbrock@cyc.com

L. Bradeško
Institut Jožef Stefan and Curious Cat Company, Jamova 39, Lubljana 1000, Slovenia

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_40,

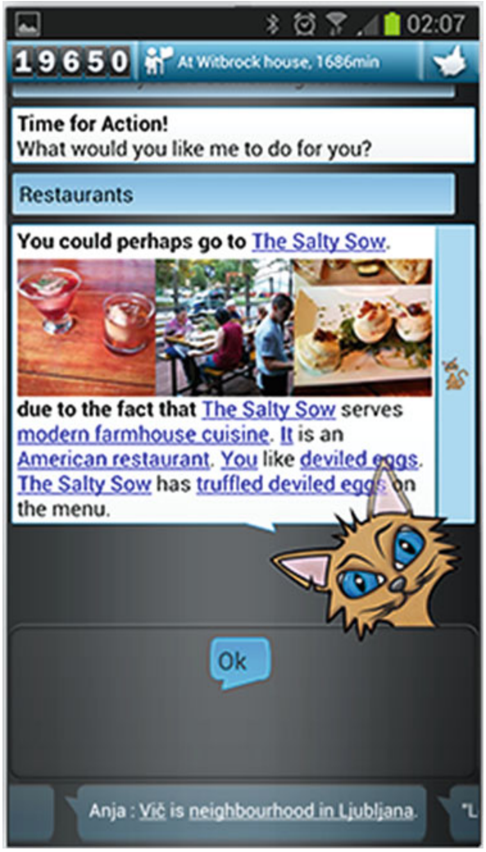


Fig. 1 Curious Cat is an AI based application that makes specific recommendations to its users based on detailed world knowledge and knowledge about the user. The recommendations are synthesised from a logical representation, which, in turn, is synthesised from knowledge entered conversationally by users. Uniform, formal knowledge representation makes it straightforward, e.g. to associate the picture of devilled eggs in the photo-strip with the content of the user-specific synthetic review

processes and situations involved in these heterogeneous topics, and of the way they interact to affect a user's life, and be able to apply that knowledge appropriately. To enrich and improve the experience of information gathering and to connect the specific data provided with the needs of other users of the system, crowd-taught companion systems will need to have a good amount of prior knowledge and intelligence. Only then will they be able to ask the right questions at the right time, situate what is learned within existing knowledge, and combined that knowledge to deliver the right advice to the right persons when needed (Fig. 1).

One tempting approach might be to provide this increased intelligence itself through human-computation, achieving adaptability either by platform enhancements that support complex tasks (e.g. Lasecki 2013) or by using human agents to decompose the task (e.g. Kittue et al. 2011, and Dai's chapter in this volume), but this is unlikely to scale to supporting systems that can cost-effectively serve millions of users, each of whom needs to be understood in detail. Ultimately, the flexibility will have to come from the use of rich knowledge representations that can cover complex heterogeneous domains, powerful inference systems that can operate over them to combine knowledge to produce personalised content, and Natural Language systems that can render these representations into human comprehensible form. It is likely that, in the future, probabilistic logics, deep learning (e.g. Mirowski and LeCun 2010) and similar techniques will be able to support such operations, but at present, predicate calculus representations, supported by efficient forward inference and natural language generation provide a good balance of power and practicality.

Curious Cat

Curious Cat (CC) is a commercial Android 2.2 (and above)[1] companion application that simultaneously learns about the world and about its users by conversational interaction based on user activities. Its goal is to use what it learns to produce activity suggestions for users that have a high probability of being deemed worth acting on. Its reward is acquiring knowledge from those users (who are acting as a source of human computation), both conversationally and unobtrusively, as they go about their lives.

The underlying operation of Curious Cat is based on representing knowledge in a uniform logical representation (CycL—see below for details) allowing background knowledge, specific knowledge about locations, goods and activities, and knowledge about the user and his or her context to be combined by inference. Recommendations are produced using natural language synthesis from an underlying logical representation, and the same representation and NL generation capability is used to drive conversational knowledge capture from humans.

Curious Cat was in closed Beta at the time of writing, and at that point had been used by more than 250 active users from among more than 500 downloads. Collectively those users were responsible for more than 200,000 assertions into the agent's knowledge base.

Curious Cat is implemented on an AI infrastructure platform that makes heavy use of Cyc. The licensed version of Cyc that is being used is substantially similar to ResearchCyc, which is available to researchers under a zero cost licence.[2]

The Cyc Platform

Cyc is a common sense artificial intelligence project started in 1984 and under continuous development since (Lenat 1995; Matuszek et al. 2006; Lenat et al. 2010). It comprises three main components: a knowledge base, an inference engine, and a lexicon and natural language generation system. It also has support for compositional parsing to logic, but that capability is not yet incorporated into CC.

[1] Downloadable at the time of writing from: https://play.google.com/store/apps/details?id=cc.curiouscat

[2] See http://research.cyc.com

The Cyc ontology and Knowledge Base (KB) is a formalized representation of a vast quantity of fundamental human knowledge: facts, rules of thumb, and heuristics for reasoning about the objects and events of everyday life. It consists of terms and assertions which relate those terms. These assertions include both simple facts (i.e., ground assertions) and rules. The version of Cyc used by CC consists of more than 17 k predicates, 550 k concepts and 6 m assertions. It is divided into many (currently thousands of) "contexts" (or "micro theories"), each of which is essentially a collection of assertions that share a common set of assumptions. This division can be based on the domain of knowledge, a particular interval in time, etc, and allows the system to independently maintain assertions which would be contradictory otherwise. The use of contexts makes it possible to speed inference by focusing it on a relevant set of theories, and in the case of CC, to separate possibly contradictory input from multiple users. The KB is under continuous development, with new knowledge being added through a combination of automated and manual means. Concepts can be expressed as individual logical symbols (e.g. CityOfNewYorkNY is a logical term representing New York City) or they can be expressed functionally, enabling the compositional creation of terms in much the same way as is permitted by a natural language but with a precise, machine-understandable semantics. "(LiquidFn Nitrogen)", for example, applies the logical function, LiquidFn (which denotes the liquid form of whatever tangible stuff is its argument) to Nitrogen, and this entails, *inter alia*, that the represented substance, liquid nitrogen, can flow.

Cyc can add assertions to the KB autonomously by forward inference, in which rule application is trigged by the addition of new knowledge. The added assertions may, in turn, trigger other rules, in a cascade. Truth maintenance retracts these added assertions if the supports on which they were based become unavailable. This forward inference and truth-maintenance mechanism can be used to implement agent-like behaviour, and is used extensively by Curious Cat.

The Cyc Inference Engine is tightly integrated with the KB and it supports both this multi-step forward and goal-driven backward inference over very large rule sets (and large data accessed from databases or triple stores). It performs general logical deduction (including modus ponens, modus tollens, and universal and existential quantification), with specific inference mechanisms (such as inheritance, automatic classification, etc.) as special cases. It has effectively applied for query answering in multiple domains (Panton et al. 2006; Deaton et al. 2005; Lenat et al. 2010), and has previously been adapted for lower expressivity reasoning (Ramachandran et al. 2005). Curious Cat extends on those previous efforts in terms of anticipated scale, the degree of user modelling and consequent ability to proactively make suggestions, in the heterogeneity of the covered domains, and in the fact that it is intended for completely general audiences with no training whatsoever.

Curious Cat Implementation

The ability to support human computation via rich interactions relies (inter alia) on three AI capabilities: (1) Efficient inference from knowledge base state and user context to knowledge capture goals; (2) Carefully designed interactions that enable the user who is providing the human computation to understand what is being asked of her or him; and (3) Sufficiently natural generation of Natural Language text from logic.

The typical knowledge capture process using inference consists of at least four steps, which all depend on the specific logical vocabulary. These steps consist of (1) Deciding when and what to ask, (2) Converting a logical question into a natural language, (3) Checking the consistency and existence of the answer, (4) Asserting the answer into the ontology and/or creating new concepts supporting it. For the ease of understanding in this paper we will explain the process on a simplified example of answering a few questions about the imaginary place where CC and its user just went to eat. Each element of the operation depends on a combination of task independent and task dependent logical vocabulary that is used to describe the rules and assertions that control system operation.

Interaction Design for Successful Knowledge Capture

Although computers have completely reliable means for keeping track of state, human computation providers do not. A human-computation-based system like CC must, therefore accommodate state-loss and other features of human cognition. One form this accommodation can take is providing supplemental information, that is not directly relevant to a knowledge capture goal, but which provides a context that allow the person satisfying the goal to do so more accurately or with less effort.

For example: Curious Cat often wants to capture several similar kinds of knowledge about an entity in the world. The knowledge base contains complete knowledge of all previously captured facts expressed using a (for example) a particular predicate[3] and complete knowledge of how the resulting knowledge is embedded in the rest of the KB. From the AI system's point of view, all that is needed to gather more facts using that predicate is to repeatedly ask the same question. However, from a human's point of view repeatedly asking e.g. "What kind of venue is the Salty Sow" would be irritating and confusing. It is much more helpful to provide context, e.g. "Besides being a modern American restaurant, what kind of venue is the Salty Sow?" which is far more likely to elicit new information (e.g. "patio bar").

Inference to Knowledge Capture Goals

To enable inference to be used to decide when and what to ask, vocabulary is needed for storing and reasoning about the user's context and to specify to the system how to formulate questions. For this paper we will describe a representative subset of this

[3] A predicate is a relation between logical terms; for example the predicate typeOfPlace may be used to express a relation between particular business venues and their types: (typeOfPlace SpiderHouseCafe Coffeehouse-Organization) is true only if Spider House (a café in Austin TX) is a coffee house, which it is.

Table 1 The predicate lastVenue allows the system to reason about the users' most recent location. Here it is defined as a binary predicate (it relates two concepts) and it is defined to be a relationship between a particular user of CC and a particular venue that CC knows about. These definitions, using "isa" and "arg1Isa" and "arg2Isa" are directly analogous to the description of the argument signature of a method in a programming language like Java. The important thing to understand, though, is that this predicate allows the system to know where its users are or have recently been, and to combine that knowledge automatically with other knowledge, using rules

Concept: lastVenue
(isa lastVenue BinaryPredicate)
(arg1Isa lastVenue CuriousCatUser)
(arg2Isa lastVenue CuriousCatVenue)

vocabulary, excluding the upper ontology[4] and other peripheral concepts. The process of conversational knowledge capture will be explained through an example in which the user of the system goes to a new place for lunch and Curious Cat asks what type of place it is.[5]

Definition of the Vocabulary

The salient vocabulary describing the user context is defined as follows (Table 1):

The lastVenue predicate is used to create assertions like: (lastVenue CuriousCatUser0 Venue1) representing the most recent location of a user, which can be combined by inference with information about that venue to e.g. infer likely activities. An important support for such reasoning is, of course, knowledge of what kind of place some venue is (Table 2):

With this vocabulary we can describe venues using logical statements like this: (typeOfPlace HoodBurger-TheBurgerPlace FoodTruck-Organization), i.e. that "Hood Burger" is a food truck.[6]

The second predicate "secondaryTypeOfPlace" allows us to state additional type information (e.g. that, in addition to being a food truck, Hood Burger is a Burger Joint). This predicate is, from a logical point of view, redundant. One could equally well have multiple typeOfPlace assertions with the same first argument and different second arguments representing multiple types assigned to and further describing the same place. However, as explained above, this second predicate supports an

[4] The upper ontology is Cyc's basic division of the world into e.g. Abstract and Concrete things, and the abstract descriptions of the relationships of those things to, for example, events; this upper ontology is very important computationally, but is not generally readily intelligible by users.

[5] This type information is, of course, extremely valuable in setting the future course of the system-human conversation, and, since there are many, many types in the world, illustrates a kind of interaction that is expansive—it increases the range of information the system may have to deal with in the future by introducing new concepts.

[6] This is functionally equivalent to (isa HoodBurger-TheBurgerPlace FoodTruck-Organization). The representational choice here is for reasons of efficiency.

Table 2 Predicates that describe places in terms of their types. This information can be combined with information about the user's location to infer what sorts of questions are relevant. Both predicates take instance of the Place as the first argument and a subclass of the Place as the second (and, in the second case, third argument. In Cyc's logical language, is a relates a particular individual thing (such as a café) to a class of which it is a member, so the arg1Isa assertion says that the first argument of typeOfPlace must be some individual place (such as the restaurant "the Salty Sow"). On the other hand, genl relates subtypes to supertypes, so the arg2Genl assertion says that the second argument of typeOfPlace must be a kind of place (such as "modern American restaurant" or "coffee house")

Concept: typeOfPlace
(is a typeOfPlace BinaryPredicate)
(arg1Isa typeOfPlace **Place)**
(arg2Genl typeOfPlace **Place)**
Concept: secondaryTypeOfPlace
(is a secondaryTypeOfPlace TernaryPredicate)
(arg1Isa secondaryTypeOfPlace **Place)**
(arg2Genl secondaryTypeOfPlace **Place)**
(arg3Genl secondaryTypeOfPlace **Place)**

interaction design that partially reflects the agent's current state of knowledge to the user, making that user's computation substantially easier. We believe that this is an important design principle for Human Computation systems: that they should, to the extent possible, remove unnecessary cognitive burdens or constraints on the humans with whom they are collaborating. The second predicate, by allowing the system to talk about two types of places at once, as in (secondaryTypeOfPlace HoodBurger-TheBurgerPlace FoodTruck-Organization BurgerJoint-Restaurant) allows the system to say "Besides a kind of food truck, Hood Burger is also a burger joint". Even more helpfully, when eliciting knowledge, as in (secondaryTypeOfPlace HoodBurger-TheBurgerPlace FoodTruck-Organization ?WHAT), where ?WHAT is unknown, it enables asking, helpfully "Besides a kind of food truck, what other kind of place is Hood Burger". Some detail on how this NL is generated is given below.

The third, and perhaps most interesting, predicate needed for our example is used by the system to make it the case that the system intends to ask a particular question (which, if nothing intervenes to cause the system to retract the inference, will happen).[7]

Table 3 defines vocabulary that represents the system's intention to ask questions—an intention that is acted upon by the Android interface. The predicate defined takes three arguments, where the first one is the user whom to ask, the

[7] The Cyc platform that underlies much of Curious Cat has extensive truth maintenance support for forward inference, so if one of the supports of a conclusion (for example a conclusion that the system intends to ask a question), is removed, the conclusion will automatically be withdrawn. One example of this is when the system changes the (lastVenue <User> <Venue>) assertion for a user, which may cause it to reconsider its intention to ask the user a subset of its questions about that venue.

Table 3 Definition of the predicate that triggers the process of asking a user question in the CC system. This predicate includes functional information (the user to ask, and the logical sentence defining the question they should be asked), but it also includes information that is simply intended to improve usability (a list of possible answers that the UI can offer the user)

Concept: curiousCatWantsToAskUser
(isa curiousCatWantsToAskUser TernaryPredicate)
(arg1Isa curiousCatWantsToAskUser CuriousCatUser)
(arg2QuotedIsa curiousCatWantsToAskUser CycLOpenExpression)
(arg3Isa curiousCatWantsToAskUser List)

second one is the actual logical sentence that it will ask[8] and the third one is the list of possible suggested answers which can be either predefined or inferred from the context. An example of such assertion is (curiousCatWantsToAskUser CuriousCatUser0 (secondaryTypeOfPlace Venue1 FoodTruck-Organization ?ANSWER) ((TheList Restaurant-Organization Cafe-Organization Bar-DrinkingEstablishment)) which means that the system intends that CuriousCatUser0 be asked what, besides a food truck, Venue1 might be. The list "restaurant, café, bar" serves both to clarify what is being asked, and to provide a quick means of answering for the common cases.

Rules That Drive Conversation

In previous paragraph we defined the vocabulary that supports inference to decide when to ask specific questions. Now we define forward inference rules which will actually do the work when the logical statement like (lastVenue CuriousCatUser0 HoodBurger-TheBurgerPLace) appears in the KB (it is asserted by the phone using its GPS location and either CC's own mapping of places to locations, or Factual Inc's open location identifiers).

From Table 4 we can see the rule that is used by forward inference to trigger Curious Cat's desire to ask the question about the user's venue.[9] This rule states that for all users, if the type of their last venue is unknown, the system should ask about the type of the venue with a suggestion list that is defined for this type of question. This suggestion list is defined with other rules or direct assertions, but the details are beyond the scope of this explanation.

[8] Putting the representation outside the range of merely first-order logic.

[9] Whether that desire is acted on depends on the state of the user interface. It is also worth noting that this rule is simple for expository purposes—CC desires to ask this question whenever a user is in a place *and* it does not already have an answer. In the full system, other information, including the number of similar questions the user has asked, their level of enthusiasm for answering, and the length of time they have been at the venue, can all readily be used to limit how often the rule triggers.

Table 4 Rule that triggers the type of place question. The meta-requirement prevents the question from being asked, if the answer is known for the current user (keeping track of users in this way is done using the Cyc context mechanism)

Rule:
(implies (and
(suggestionsForCuriousCatQuestionType
VenueTypeOfPlace-CuriousCatQuestion
?SUGGESTIONLIST)
(lastVenue ?USER ?VENUE))
curiousCatWantsToAskUser ?USER (typeOfPlace
?VENUE ?SUGGESTIONLIST))
Meta Requirement:
(unknownSentence (thereExists ?PLACETYPE (typeOf-
Place ?VENUE ?PLACETYPE)))

Table 5 Rule that triggers the more specific question about the type of food truck place. This rule is specific to food-trucks, and may appear overly specific. However, it is possible in Cyc to write forward rules that add specific rules like this to the KB. This form of "self programming" behaviour is used extensively in CC

Rule:
(implies (and
(isa ?VENUE FoodTruck-Organization)
(lastVenue ?USER ?VENUE)
(suggestionsForCuriousCatQuestionType FoodTruckSecondaryTypeOfPlace-
CuriousCatQuestion ?SUGGESTIONLIST))
(curiousCatWantsToAskUser ?USER
(secondaryTypeOfPlace ?VENUE FoodTruck-Organization ?TYPE)
?SUGGESTIONLIST))

In order to get more specific knowledge about the place following our assistive interaction design we need another rule which takes into consideration the previous answer and produces a more detailed question:

Once the rule for "typeOfPlace" has furnished some information about a venue, this "secondaryTypeOfPlace" rule only triggers if that (or some other) interaction has revealed the venue to be an instance of the FoodTruck-Organization and if it is part of the user's venue history. Some kinds of places, and notably food trucks, are more defined by a secondary type (e.g. more questions are unleashed by the knowledge that a food truck is also a cupcake bakery, than by its identity as a mobile food vendor). Forward inference on triggering this rule produces assertions like (curiousCatWantsToAskUser CuriousCatUser0 (secondaryTypeOfPlace HoodBurger-TheBurgerPlace FoodTruck-Organization ?TYPE) (TheList TacoStand IceCreamTruck)) which the system understands as an intention to ask the CuriousCatUser0 the following question: (secondaryTypeOfPlace HoodBurger-TheBurgerPlace FoodTruck-Organization ?TYPE), after it has been converted to English, as "Besides a kind of food truck, what other kind of place is Hood Burger". The interaction resulting from the use of the vocabulary, and of the rules defined in Tables 4 and 5 can be seen on images 1, 2. How the logic is actually converted is described in the next paragraph.

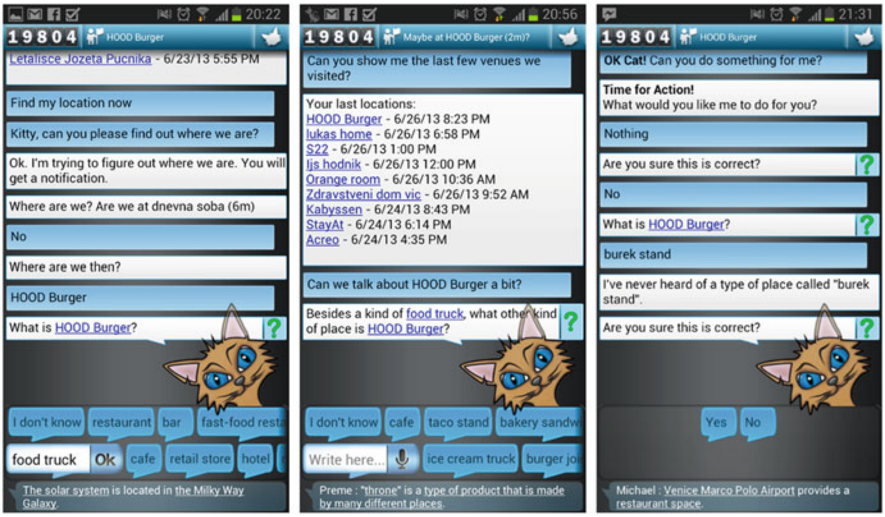


Fig. 2 The inferred intention to ask a questions results in a conversational dialogue supporting knowledge capture. The image to the right illustrates the creation of an entirely new type—the large pre-existing knowledge base makes this an increasingly rare occurrence, and makes entering new knowledge about particular places, for example, more straightforward

Natural Language Generation (NLG)

Of course, inferring logical sentences that represent conversational intentions is not sufficient. Those intentions must result in interactions in natural language. We have given some examples of that in the text, and in Fig. 2; Table 6 shows how this is conversion is done. Because logic is quite far from a human language, a uniform logic-language system is infeasible—these templates represent a compromise—they are specific enough to represent the mapping from logic precisely, but they are "linguistic" enough to permit composition into larger discourse units.

Checking consistency and asserting the answer

When the user answers the question presented (Fig. 2), the CC system has to check whether the answer is something it knows already or it is completely new concept, and if the answer is consistent with existing knowledge. This is done with the help of the definitions of the vocabulary, including the type constraints on predicate arguments. Other consistency constraints, including attempting to prove the contrary of a prospective assertion, may also be used.

It's worth noting, of course, that the answers from humans come to the system as NL strings, not as precise logical terms. So the first thing the system must do is to find out whether the answer has a logical interpretation that satisfies the vocabulary

Table 6 The system is supported by Natural Language Generation, allowing the synthesis of both declarative and interrogative forms, according to patterns illustrated here. The patterns support e.g. correct number and tense generation depending on arguments and context

As a question	As a statement
(genTemplate-QuerySentence (secondaryTypeOfPlace :PLACE :PRIMARY ?SECONDARY) (ConcatenatePhrasesFn (ConcatenatePhrasesFn (PhraseFromStringFn "besides a kind") (ConcatenatePhrasesFn-NoSpaces (Pp-PNpFn Of-TheWord (ParaphraseFn-Constrained nonPlural-Generic :PRIMARY)) (PhraseFromStringFn ", what other kind of place is"))) (ParaphraseFn :PLACE)))	**(genTemplate secondaryTypeOfPlace (ConcatenatePhrasesFn (ConcatenatePhrasesFn (PhraseFromStringFn "besides a kind of") (ConcatenatePhrasesFn-NoSpaces (Pp-PNpFn Of-TheWord (ParaphraseFn-Constrained nonPlural-Generic :ARG2)) (PhraseFromStringFn ","))) (PhraseFn-Tensed :ARG1 (Sentence-NpIsXpFn (ParaphraseFn :ARG1) (ConcatenatePhrasesFn (PhraseFromStringFn "also") (Np-DetNbarFn-Indefinite (PhraseFn-Constrained nonPlural-Generic (ParaphraseFn :ARG3))))))))**
Input: (secondaryTypeOfPlace HoodBurger-TheBurgerPlace FoodTruck-Organization ?WHAT)	**Input: (secondaryTypeOfPlace HoodBurger-TheBurgerPlace FoodTruck-Organization BurgerJoint)**
Result: "Besides a kind of food truck, what other kind of place is Hood Burger?"	**Result: "Besides a kind of food truck, Hood Burger is also a burger joint."**

Table 7 Query to check the ontology constraints of the answer. The system knows that it is asking a question about "typeOfPlace". This query looks up the argument constraints on the second argument for typeOfPlace (the argument for which the user has supplied the string "food truck" as a possible answer), it looks up all the possible meanings of "food truck" (including FoodTruck-Organization), and it checks whether any of those meanings meet the argument constraints it found for typeOfPlace

Query:
(#$and (#$termStrings ?TERM "food truck") (#$or (#$and (#$unknownSentence (#$thereExists ?ARGENL (#$argGenl typeOfPlace 2 ?ARGENL))) (#$equalSymbols ?ARGENL #$Nothing)) (#$and (#$argGenl typeOfPlace 2 ?ARGENL) (#$genls ?TERM ?ARGENL))) (#$argIsa typeOfPlace 2 ?ARGISA) (#$isa ?TERM ?ARGISA))

(ontology) requirements. If we continue with our example based on Hood Burger, the query looks as shown in Table 7.

The query in Table 7 returns the concept (class) of FoodTruck-Organization, which can be immediately used as an answer without more disambiguation since it is the only result and it fits the ontology constraints.

The final results of this process are the assertions: (secondaryTypeOfPlace Venue1 FoodTruck-Organization BurgerJoint) (typeOfPlace Venue1 FoodTruck-Organization)

In cases when the query returns more than one answer (because there are many possible types with the same name, due to polysemy), the user is presented with options to pick one among the possible answers or to define a new one—the NLG system attempts to describe the choices as differentially as possible.

If the above query returns no results, then the CC system searches the full KB to find an appropriate type based on its English name string. If it finds one then it uses inference to check whether that answer is provably disjoint with the constraints or not. If they are, a new "type" concept is created with the name. When neither the constrained search nor full KB search returns any answers, the CC system tries to create a new (previously unknown) concept which fits the ontology constraints. For example, if instead of "food truck" the answer had been "burek stand"[10], a concept previously unknown to the system, the system would try to create a new type of place called just that (illustrated to the right of Fig. 2). The result would be the new concept

Concept: BurekStand
(isa BurekStand Collection)
(genls BurekStand Place)

And the assertion (typeOfPlace Venue1 BurekStand), meaning that "Venue1 is a burek stand", our new kind of place.

Conclusions

Logical representations, forward inference, and natural language term recognition and generation provide a practical means to describe interactions with users that allow them to provide knowledge and other computational output to a collaborative system, and provide a convenient mechanism for making use of that knowledge, as it is captured, to drive the capture process itself.

An important component of making these interactions satisfactory to users is to design them so that machine computation accommodates human abilities, for example by feeding back contextual detail that a machine would not need, but that helps humans to understand and track the status of their task.

In our running example, we have illustrated how capturing knowledge in a representation that supports automatic knowledge-combination for suggestions, and is suitable for natural language generation, can be employed in a way that allows human cognitive preferences and limits to be respected, in this case by making the existing knowledge context explicit, in part.

[10] A burek is a kind of savoury filled pastry popular in Balkan countries and elsewhere.

Future Work

While the system described here is already able to interact with mobile users to elicit knowledge that is useful for highly tailored recommendations, it is somewhat inflexible in its interaction style, and does not put sufficient initiative for the order and content of interactions with the human. One important step in supporting increased human initiative in the collaboration will be incorporating a compositional language to logic system, Semantic Construction Grammar, currently under development at Cycorp. It may also, in some cases, be useful to incorporate humans in the loop, especially if, while satisfying a user's needs, these humans are training the system as a side effect.

To fully realise the goal of producing companion systems, it will also be necessary to reduce the task focus of the current system somewhat (it currently concentrates heavily on getting answers and providing suggestions). The same mechanisms that currently produce intentions to ask questions could be adapted to drive other discourse elements, such as comments and greetings.

References

Deaton C, Shepard B, Klein C, Mayans C, Summers B, Brusseau A, Witbrock M (2005) The comprehensive terrorism knowledge base in Cyc. In: Proceedings of the 2005 international conference on intelligence analysis, McLean

Forbus K, Hinrichs T (2004) Companion cognitive systems: a step towards human-level AI. In: AAAI fall symposium on achieving human-level intelligence through integrated systems and research, Washington, DC

Kittue A, Smus B, Kraut R (2011) CrowdForge: crowdsourcing complex work, technical report, CMU-HCII-11-100, Human-Computer Interaction Institute, Carnegie Mellon University

Lasecki WS (2013) Real-time conversational crowd assistants. CHI extended abstracts 2013, pp 2725–2730

Lenat DB (1995) Cyc: a large-scale investment in knowledge infrastructure. Commun ACM 38:32

Lenat D, Witbrock M, Baxter D, Blackstone E, Deaton C, Schneider D, Scott J, Shepard B (2010) Harnessing cyc to answer clinical researchers' ad hoc queries. AI Mag 31:3

Matuszek C, Cabral J, Witbrock M, DeOliveira J (2006) An introduction to the syntax and content of cyc. In: 2006 AAAI spring symposium on formalizing and compiling background knowledge and its applications to knowledge representation and question answering

Mirowski P, LeCun Y (2010) Dynamic auto-encoders for semantic indexing. NIPS workshop on deep learning

Panton K, Matuszek C, Lenat DB, Schneider D, Witbrock M, Siegel N, Shepard B (2006) Common sense reasoning—from cyc to intelligent assistant. In: Cai Y, Abascal J (eds) Ambient intelligence in everyday life. LNAI 3864. Springer, pp 1–31

Ramachandran D, Reagan P, Goolsbey K (2005) First-orderized researchCyc: expressivity and efficiency in a common-sense ontology. In: Papers from the AAAI workshop on contexts and ontologies: theory, practice and applications. Pittsburgh

Modeling Humans as Computing Resources

Yu-An Sun and Christopher Dance

To understand the unique characteristics of humans as computation resources, we have to understand how humans perform tasks, instead of treating them as black boxes with inconsistent efficiency. If we view humans as computing nodes, it is clear that improving the computing efficiency of each node will result in an increase of the total efficiency of the human computation procedure or workflow. Since 1971, Tversky and Kahneman (1971) conducted a series of experiments to test their hypotheses and developed a dual-mind theory to demonstrate that humans have two types of cognitive operations—intuitive and reflective. The intuitive mind processes tasks quickly and automatically, but is prone to heuristic and biased judgments. The reflective mind operates more slowly, but in a rule-based fashion. The two systems also process different types of content. The intuitive system acts on content that is affective, concrete, specific, and based on casual propensities and prototypes. The reflective system, on the other hand, processes content that is neutral, based on statistics, abstractions and sets. From the two system perspective, computer algorithms can be viewed as mimicking the reflective operations of human minds. This research field is far richer and larger than what's discussed so far; this article is merely an attempt to scratch the tip of an iceberg. By identifying several unique characteristics of humans as computing resources, we further point out open problems that human computation researchers should pay attention to in order to design better instructions and institutive algorithms.

Y.-A. Sun (✉)
Xerox Research Center Webster, Webster, NY, USA
e-mail: Yuan.sun@xerox.com

C. Dance
Xerox Research Centre Europe, Grenoble, France

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_41,

By cross-analyzing the state of art research in human computation field (Law and von Ahn 2011) and the two system theory (Tversky and Kahneman 1971), we may observe that a human is unique as a computation resources in the following ways:

Humans Can Solve Computer Hard Problems

A rich body of literature, summarized in the ‘Human Computation’ (Law and von Ahn 2011) has shown human computation can be effectively used in the computer vision, translation, transcription, and OCR domains to produce quality results that computers still cannot. Cornety et al. studied the Irregular Strip Packing problem, which is known to be NP(non-deterministic polynomial time) hard, and showed that a crowd outperformed the best algorithms in the literature (Corney et al. 2010). Solutions of NP hard problems, in computational complexity theory, cannot be located in Non-deterministic Polynomial time. An famous example of NP hard problems is traveling salesman problem. It asks for a solution of the shortest possible route that goes through each city exactly once given a list of cities and distance between them. However, not all NP hard problems are solvable by human. A new classification of computational problems, ‘AI complete’, is developed to describe those with have polynomial time solutions with the help of ‘Human Oracle’(Shahaf and Amir 2007).

Humans Are Very Good at Exception Handling

As Little and Sun pointed out in their recent work (Little and Sun 2011), humans are particularly good at exception handling. When processing an OCR task on a medical form, a human knows to look for the other address field to find answers when the address field is filled with the phrase ‘Same’ without additional instructions to specify all the possible scenarios.

Humans Have Creativity

This unique characteristic extends humans from a computing resource to a creative resource. Crowdsourcing platforms such as 99 Designs take full advantage of humans’ creativity. In addition to graphic design tasks, there are other examples such as Soylent, a Word Macro (Bernstein et al. 2010) asking the crowd to help shorten a paragraph, and Adrenaline (Bernstein et al. 2011), using mobile phone cameras to capture video and asking the crowd to identify a single photo as the best moment of the video.

On the flip side, the two minds theory also suggests that humans have the following disadvantages when used as computing resources:

Humans Have Cognitive Load Limitation

Human have short-term memory limitations, which constrain the size of input that humans can accept. For example, if a human is asked to identify two identical images from a set of 1,000 images which are all different except for the two, it is obvious that the human cannot remember the index numbers of those two identical images after reviewing all 1,000 of them. As Kahneman and Tversky pointed out in 1971 (Tversky and Kahneman 1971), humans are also poor at estimating probabilities. Humans tend to make judgements based on prototypes. Ariely (2008) also pointed out that human are predictably irrational when making decisions. This is unique disadvantage of humans as computing resources.

Humans Are Vulnerable to Psychological Manipulation

The anchoring effect (Tversky and Kahneman 1992), the sunk-cost fallacy (Arkes and Blumer 1985) and the crowding-out effect (Frey and Jegen 2001) are all psychological manipulations that can change the result of a human computation task. They can either change the quality, cost or time that associated with a human computation task. When asked to perform tasks related to decision making, humans can have biases from the irrelevant information that is also presented to them, thus the accuracy of the judgment is affected. Sunk-cost fallacy and crowding-out effects are related to how to motivate humans to perform the same task in different efficiency. Clearly, a better motivated human performs better.

Humans Are Prone to Errors, Especially for Reflective Tasks

As Kahneman and Frederick pointed out, the reflective mind operates more slowly than the intuitive mind, and is more likely to cause errors. When questions get more difficult to analyze, humans tend to switch from reflective operation to intuitive operation. This is called ‘attribute substitution’ in Kahneman and Frederick’s work (2002).

After understanding the above unique characteristics of using humans as computing resources, the question we ask is, how to evaluate the efficiency of various ‘algorithms’ when we give humans instructions to perform a task? Big-O notation is used to analyze time complexity for computer algorithms, and input size and output size are part of the notation. Currently when people conduct human computation tasks, instructions are written in natural language. Is there a need to construct corresponding 'instruction sets' analogous to those used for computer programs? How do we write intuitive instructions in order to take advantage of the intuitive mind? There are quite a few open questions yet to be answered and this article does not aim to provide answers, but rather to point out directions towards a better human computation instruction design.

Are Instructions Better Given in Natural Language?

In current work in the human computation research area, when experiments are conducted on Amazon Mechanical Turk or Crowdflower, people are given instructions written in natural language instead of pseudo code. One reason to give instructions in natural language is that humans are going to read the instructions. However, a question that remains open is "What constitutes a readable set of instructions?" Does the length of the instructions matter? Would humans operate better with several-pages-long instructions or with clear examples considering the short term memory limitation we pointed out earlier? Writing computer programs via examples is not a new approach. However, instead of giving human computation workers instructions, showing them examples could result in quicker and better understanding, especially for human's intuitive operation. In addition, are video or audio instructions better than graphical or written instructions? Clearly, the answer to the last question is situation-specific. For example, when assembling furniture, it is better to have diagrams, but when asked to perform a play, it is better to have a script. Frequently, a mix is better than one individual modality, thus the question may become, what is the better mix of different type of instructions for a specific type of task?

How to Design Intuitive Human Computation Instructions/ Algorithms?

When two sets of instructions are equally understandable for humans, how do we measure intuitiveness? In the following example, we are going to use a *sorting* operation, which is a classical problem with multiple algorithms with similar average-case time complexity, to illustrate the difference in intuitiveness. We are not suggesting using human as computing resources for sorting operations. When writing instructions for humans asked to perform a *sorting* operation, is a bucket sort more intuitive than quick sort? One might measure the "intuitiveness" of an algorithm by the time for humans to understand the instructions. This measure is particularly important in situations where understanding the instructions can become the bottleneck for humans performing the task. In the sorting example, if a human takes

10 s to understand a bucket sort and 5 min to understand quick sort, on first encountering it, with clear instructions, and a human takes 30 s to perform a bucket sort and 28 s to perform quick sort, we can still conclude that bucket sort is more intuitive than quick sort hence it performs better in terms of time. However, we should also consider quality and cost as performance measurements. We might hypothesize that:

Hypothesis: The Intuitiveness (measured by function *I*) of a human computation algorithm predicts the efficiency (for quality, time AND cost) with which humans will implement <tasks A, B and C>.

The question is, what might functions I and X plus tasks A, B and C actually be?

This is pretty complex because performance related psychological effects such as the anchoring effect, the Yerkes-Dodson law (1908), and learning effect could all play a role. As we discussed earlier, the anchoring effect affects accuracy of the human computation. Yerkes-Dodson law states that performance increases with motivation, but only to a point. Performance decreases after that point even with higher motivation. Humans also perform better with familiar tasks because they learn from the past experience (Ebbinghaus 1885). With all the psychological effects in play, it is obviously a difficult hypothesis to validate. In addition, the presentation of the task to test this hypothesis is also absolutely critical so that readability becomes a controlled variable.

Once an Intuitive Algorithm Is Designed, How Will It Scale with Input Size?

In computational complexity analysis, different complexity affects the ability to scale. . In other words, even if we could directly compare input size, it may well be that some polynomial-time problem families take hours to solve when the size is n for some small n, yet, some NP families take only seconds for the same size n. Then scaling comes in because the polynomial-time family still only takes hours for size 1,000 n yet the NP family takes centuries for 1,000 n.

What about scaling for humans taking short term memory limitation in consideration? One idea could be adding computer help to reduce the input size. For example, when asking human to compare 10,000 images and answer whether they are all taken at the same city, an image clustering algorithm can pre-process the 10, 000 images to reduce it to a reasonable number within human's short term memory limitations. However, in the recent attempt of formalizing complexity theory and classification of AI-complete problems (Shahaf and Amir 2007), the short term memory limitation was not addressed. The same complexity notion is used to analyze 'Human Assisted Turing Machines'. This might not be a problem when input size is relatively small, but certainly does not incorporate the cognitive limitation of humans.

In conclusion, there is still a long way to go for human computation researchers to fully explore the richness of human's cognitive operations. A lot of open questions remain to be answered regarding the design of intuitive yet scalable human computation algorithms.

References

Ariely D (2008) Predictably irrational, Harper Collins

Arkes H, Blumer C (1985) The psychology of sunk cost. Organ Behav Hum Decis Process 35: 124–140

Bernstein M, Greg L, Miller R, Hartmann B, Ackerman M, Karger D, Crowell D, Panovich K (2010) Soylent: a word processor with a crowd inside. UIST: ACM symposium on user interface software and technology

Bernstein M, Brandt J, Miller R, Karger D (2011) Crowds in two seconds: enabling real-time crowd-powered interfaces. UIST: ACM symposium on user interface software and technology

Corney JR, Kowalska I, Jagadeesan AP, Lyn A, Medellin H, Regli W (2010) Crowdsourcing human problem solving strategy. CrowdConf 2010

Ebbinghaus H (1885) Memory: a contribution to experimental psychology. Dover, New York

Frey BS, Jegen R (2001) Motivation crowding theory. J Econ Surv 15(5):589–611

Kahneman D, Frederick S (2002) Heuristics of intuitive judgment: extensions and applications. Cambridge University Press, New York

Law E, von Ahn L (2011) Human computation. Morgan & Claypool synthesis lectures on artificial intelligence and machine learning

Little G, Sun Y (2011) Human OCR. ACM CHI 2011 workshop on crowdsourcing and human computation

Shahaf D, Amir E (2007) Towards a theory of AI completeness. In: 8th international symposium on logical formalizations of commonsense reasoning (Commonsense 2007), California

Tversky A, Kahneman D (1971) Belief in the law of small numbers. Psychol Bull 76:105–110

Tversky A, Kahneman D (1992) Advances in prospect theory: cumulative representation of uncertainty. J Risk Uncertain 5:297–323

Yerkes RM, Dodson JD (1908) The relation of strength of stimulus to rapidity of habit-formation. J Comp Neurol Psychol 18:459–482

Service Oriented Protocols for Human Computation

Daniel Schall

Introduction

Human computation is increasingly becoming mainstream and commonplace in collaborative computing environments. This is partially due to the success of applications such as 'games with a purpose' (von Ahn 2006) (a game wherein two people need to agree on a common set of keywords which they associate with images) and the success of commercially available crowdsourcing platforms such as Amazon Mechanical Turk (AMT) (Amazon 2013). Human computation and crowdsourcing are often used interchangeably to address problems that computers cannot yet tackle on their own in an efficient manner (von Ahn 2006). In a similar spirit, crowdsourcing is commonly defined as *'the act of taking a job traditionally performed by a designated agent and outsourcing it to an undefined, generally large group of people in the form of an open call'* (Howe 2006). A recent taxonomy and survey (Quinn and Bederson 2011) overviews existing literature in the context of human computation and crowdsourcing.

Many crowdsourcing systems have recently emerged on the World Wide Web (Doan et al. 2011) including CrowdFlower (2013), oDesk (2013), ClickWorker (2013), SmartSheet (2013), and SpeechInk (2013). These platforms allow people to work on tasks such as transcription of spoken language into text, translation of text, tagging of images, and coding as well as integration of scripts and APIs. By nature, platforms on the World Wide Web are under constant flux and change. Also, human computation and crowdsourcing platforms are dynamic with people around the globe joining and leaving communities (Ipeirotis 2010). Thus, it is essential to account for the dynamics in the availability of a large-scale human

D. Schall (✉)
Siemens Corporate Technology, Vienna, Austria
e-mail: daniel.schall@siemens.com; daniel.schall@gmail.com

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_42, © Springer Science+Business Media New York 2013

workforce, changing skills of crowd workers, and changing requirements with regards to the underlying communication protocol.

- Protocols for human computation and crowdsourcing have to account for these dynamic aspects by supporting adaptive (flexible) interactions that may span numerous human and software services as well as a range of devices,
- Monitoring and logging mechanisms to observe the environment and ongoing interactions and
- A crowd worker discovery mechanism based on workers' capabilities, evolving skills and availability constraints.

In previous work we proposed service-oriented computing principles to address the challenges of human computation and crowdsourcing in large-scale environments. *Mixed service-oriented systems* (Schall 2011) consist of Human-Provided Services (HPS) (Schall et al. 2008) and Software-Based Services (SBS) that can be composed to jointly solve crowdsourcing tasks (Schall 2012). The novelty of mixed service-oriented computing environments is the application of social principles to coordinate the execution of human tasks within open Web-based systems. Existing XML-based industry standards such as WS-HumanTask (WS-HT) (Amend et al. 2007) and Bpel4People (B4P) (Agrawal et al. 2007) can be integrated into mixed service-oriented systems to support the coordination of a set of distributed human tasks. *Non-functional requirements* play an essential role for the integration of these standards in open computing environments because of the inherent dynamic nature of Web-based systems. Our prior work (Schall 2012) provides the basis for the following discussions on service-oriented protocols for human computation.

We start with an overview of the system context in section "System Context Overview" detailing the environment in which human computation is performed. In this work we discuss the basic functional and non-functional protocol requirements with regards to human computation (see section "Basic Protocol Requirements"). Based on the basic protocol requirements we describe the mapping of human interactions onto a service-oriented infrastructure in section "Service Oriented Protocol". This is detailed at a technical level by providing an actual XML-based description of a Human-Provided Services interface.

System Context Overview

A general system context overview is provided by Fig. 1. Fig. 1 shows various actors and their roles and a set of architectural building blocks. The essential roles within a human computing or crowdsourcing environment are listed and described in the following.

- **Platform Provider**: the platform provider (not shown in Fig. 1) is responsible for providing and maintaining the crowdsourcing platform. The platform provider

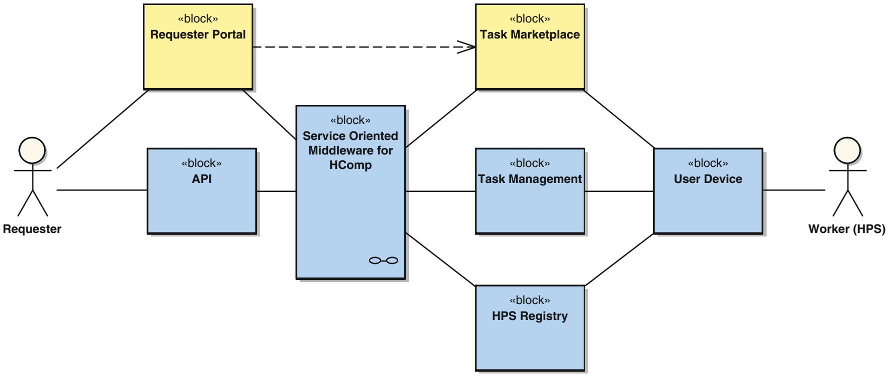


Fig. 1 Main architectural building blocks

must also ensure that enough workers are available for task processing (e.g., by providing incentive mechanisms) and that requesters post enough tasks so that workers are able to obtain rewards.

- **Requester**: the requester sends human tasks to the platform. The requester may be a person who wishes to crowdsource a set of tasks or a software agent that needs to outsource certain steps in a computational process to a human. The requester typically pays a monetary reward to the workers as well as the platform provider.
- **Worker (HPS)**: the worker or Human-Provided Service (HPS) (Schall et al. 2008) either claims or receives a human task. The HPS concept enables the seamless integration of human capabilities into a service-oriented human computation or crowdsourcing environment (Schall 2012). An HPS can be discovered like an SBS. However, interactions with an HPS need to be performed in a manner suitable for a human. The result of the human task is returned by the HPS and delivered to the requester.

From a technical point of view, various building blocks are needed to realize a service-oriented platform for human computation. First, we briefly discuss two building blocks that common systems such as AMT implement to realize a marketplace for task-based crowdsourcing. However, market-based crowdsourcing is not focus of this work.

- **Requester Portal**: the requester has the ability to create new tasks and monitor the progress of human tasks using a Web-based portal.
- **Task Marketplace**: human tasks may be presented in the task marketplace. Workers (HPS) have the ability to discover and claim new tasks that are available through the marketplace.

The following building blocks can be found in a service-oriented human computation environment:

- **API**: The requester has the ability to create human tasks through the API programmatically. This allows to integrate human computation techniques into existing infrastructures.
- **Service Oriented Middleware for HComp**: The middleware implements features such as HPS discovery, asynchronous interaction support, monitoring and mining of interactions.
- **Task Management**: The task management provides a standardized interface to create and manipulate task instances. For that purpose, existing standards such as WS-HT (Amend et al. 2007) can be adopted that already define the various states of generic human tasks.
- **HPS Registry**: A Human-Provided Service can offer the capability to, for example, 'translate documents', 'perform document review', or 'provide help and support'. The HPS Registry helps to discover the demanded service by performing a lookup procedure.
- **User Device**: Workers have the ability to use their preferred device to interact with the system. This may demand the adaptation of the user interface.

Basic Protocol Requirements

In this section we discuss the most important functional and non-functional requirements of a service-oriented protocol for human computation.

Functional requirements include:

1. The protocol must support *asynchronous interactions* between requesters and workers (HPS). Humans operate at a different speed than software services and thus all interactions should be performed asynchronously. This is true for most interactions where human input is needed.
2. The protocol must support seamless *interactions and composition of human and software services*. A seamless infrastructure integrates the capabilities of people and software services to perform computation in a hybrid service-oriented system.
3. The protocol must support *well defined interfaces* that define the possible interactions. Well defined interfaces help to discover the appropriate service (HPS).

Non-functional requirements must be defined to provide assurance with regards to different qualities:

1. The requester must be able to explicitly state non-functional requirements in form of *service-level agreements (SLAs)*. SLAs must be stated so that the middleware platform is able to interpret and enforce the negotiated agreements.

2. All *interactions must be monitored* to provide historical information that can be used for later analysis. Indeed, monitored information is only used by the platform itself to perform, for example, ranking and selection of the most suitable HPS. Selecting skilled workers helps to improve or guarantee the quality of delivered task results.
3. *Interactions must support flexibility* to enable load balancing and delegation. Since human tasks may arrive in an unpredictable manner (bursts), HPSs must be able to perform task delegation to balance their workload. This should be assisted by the platform to help finding the appropriate delegation receiver.

The following section details the protocol and describes how the presented functional and non-functional requirements are satisfied by the protocol.

Service Oriented Protocol

We propose the application and extension of existing Web service standards for human computation and crowdsourcing. Here we present a concrete HPS interface example and discuss how the discussed requirements are addressed by the proposed protocol and Web service standards.

Technical description. Technical service interfaces are typically described by using the well-established Web Services Description Language (WSDL).[1] WSDL interfaces help to discover and invoke services through a late binding mechanism. The very same description language can be used to describe an HPS.

Listing 1 shows the definition of a WSDL interface of a *translation service* (to translate a document from one language to another) that is used to interact with people in a service-oriented manner. Using WSDL as the interface description language brings the important advantage that the same standard is used to describe both HPS and software services (SBS). The WSDL in Listing 1 is automatically generated by the Service Oriented Middleware for HComp (see Fig. 1). It shows also the structure of the complex data type that is passed to the HPS to perform the actual task. Type information is used to automatically generate XML-based graphical user interfaces using forms technologies (see XForms[2]).

Notice, human task related information in managed by a separate task management service and is not depicted by the presented HPS interface. Task management can be implemented as a WS-HumanTask infrastructure (see Amend et al. 2007; Schall 2012). This allows for a clear separation of the generic task model (task states, transitions, etc.)

[1] http://www.w3.org/TR/wsdl

[2] http://www.w3.org/MarkUp/Forms/

```
<wsdl:definitions name="TranslationService" ...>
 <wsdl:types>
  <xs:schema elementFormDefault="unqualified" tns="http://...">
   <xs:element name="assignProcRequest"
    type="tns:assignProcRequest" />
   <xs:element name="getStatus" type="tns:getStatus" />
   <xs:element name="getProcResult" type="tns:getProcResult" />
   <!-- responses omitted-->
   <xs:complexType name="desc">
    <xs:element name="docTitle" type="xs:string" />
    <xs:element name="docUri" type="xs:string" />
    <xs:element name="length" type="xs:string" />
    <xs:element name="language" type="xs:string" />
    <xs:element name="translation" type="xs:string" />
    <xs:element name="translationUri" type="xs:string" />
    <xs:element name="mimeType" type="xs:string" />
    <!-- further details omitted-->
   </xs:complexType>
   <!-- other types... -->
  </xs:schema>
 </wsdl:types>
 <wsdl:message name="assignProcRequest">
  <wsdl:part element="tns:assignProcRequest" name="params" />
 </wsdl:message>
 <!-- messages... -->
 <wsdl:portType name="TS">
  <wsdl:operation name="assignProcRequest">
   <!-- in-/output... -->
  </wsdl:operation>
 </wsdl:portType>
 <wsdl:binding name="TSSoapBinding" type="tns:TSService">
  <soap:binding style="document" transport="http://schemas..."/>
   <!-- operations... -->
 </wsdl:binding>
 <wsdl:service name="TSService">
  <wsdl:port binding="tns:TSSoapBinding" name="TSPort">
   <soap:address location="http://somehost:8080/..."/>
  </wsdl:port>
 </wsdl:service>
</wsdl:definitions>
```

Listing 1 Interface description of human translation service: description is used to define complex data types and to support discovery of HPS

and the actual application specific service (HPS) model. Thus, HPSs can be designed at any time and registered with the HPS registry (see Fig. 1).

The elements in Listing 1 are used for the following purpose:

- `docTitle`: defines the name of the document to be translated by a human.
- `docUri`: contains the location (e.g., link to a document repository) where the document can be downloaded from.
- `length`: the length (number of words) of the document.

```
<soap:Envelope
 xmlns:soap="http://schemas.xmlsoap.org/soap/envelope/"
 xmlns:hps="http://www.danielschall.at/hps/">
 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 xmlns:wsa="http://schemas.xmlsoap.org/ws/.../addressing>"
 <soap:Header>
  <hps:Timestamp value="2013-03-07T17:24:18"/>
  <hps:TaskUri="http://.../HumanTask#42"/>
  <wsa:MessageID>uuid</wsa:MessageID>
  <wsa:From>http://.../Actor#Actor1</wsa:From>
  <wsa:To>http://.../Actor#Actor2</wsa:To>
  <wsa:ReplyTo>http://.../Actor#Actor3</wsa:ReplyTo>
  <wsa:Action>http://.../Type/Translate</wsa:Action>
 </soap:Header>
 <soap:Body>
  <hps:Request>
   <!-- request omitted-->
   <hps:keywords>document, translation</hps:keywords>
  </hps:Request>
 </soap:Body>
</soap:Envelope>
```

Listing 2 HPS message log example: logs are used to perform analysis of interactions

- `language`: specifies the language in which the document is written.
- `translation`: the target language to be translated to (e.g., provide translation from German to English).
- `translationUri`: the location (repository) of the translated document. The person translating the document may provide an alternative location that can be requested via the Web service operation `getProcResult`.
- `mimeType`: states the acceptable format of the translated document (e.g., PDF, Word document, or plain text).

To support monitoring and logging, interactions are captured through XML message interceptors deployed within the service runtime environment. Messages are saved in a log database for analysis. An example interaction log is shown by Listing 2, which includes various message header extensions for message correlation and context-aware interaction analysis.

The purpose of the most important extensions is outlined in the following:

- `Timestamp` captures the actual creation of the message and is used to calculate temporal interaction metrics, such as average response time.
- `TaskUri` describes the context of interactions based on the task performed by the user. The `TaskUri` helps to correlate messages and task context.
- `MessageID` enables message correlation, i.e., to properly match requests and responses.
- `WS-Addressing` extensions, besides `MessageID`, are used to route requests through the collaborative (social) network. Routing is performed through delegation but can also be assisted by the middleware through a rule based system.

Addressing functional and non-functional requirements. The functional and non-functional requirements are satisfied as follows:

- Asynchronous interactions (see functional requirement 1) are supported through the operations `assignProcRequest` and `getProcResult`.
- Interactions and composition of human and software services (see functional requirement 2) are supported because the same technical standards and framework are used.
- Well defined interfaces (see functional requirement 3) are supported through the use of well defined XML-based WSDL interfaces.
- Service-level agreements (see non-functional requirement 1) are technically supported through the Web services stack. A detailed discussion is provided in Schall (2012).
- Interactions are monitored (see non-functional requirement 2) to provide mechanisms for temporal analysis, message and task correlation, and fine-grained expertise analysis.
- Interactions can be performed in a flexible manner (see non-functional requirement 3) through delegation patterns (at the technical level, WS-Addressing[3] mechanisms help to route messages).

Conclusions

In this chapter we have discussed service oriented protocols for human computation. Service oriented protocols help implementing applications for human computation such as the presented translation service and other possible applications such as GWAP. The main advantage of service oriented protocols is the potential integration of crowdsourcing into business environments that are based on Web services technologies and related BPM standards. Interface design of human-based services is an important issue. Using the Web service description language in combination with SOAP is an effort to standardize interface descriptions for human computation. These formal XML-based standards help defining domain concepts as data types in a rigorous manner. Other more Web-centric data formats such as JSON based data types may also be used to exchange task requests with human-based services. However, these formats currently lack a standardized approach for describing service interfaces. Web-centric data formats in the context of human computation will be analyzed in our future work.

[3] http://www.w3.org/Submission/ws-addressing/

References

Agrawal et al. A (2007) Ws-bpel extension for people (bpel4people), version 1.0
Amazon (2013) mturk.com. Accessed 05-Mar-2013
Amend M et al. (2007) Web services human task (ws-humantask), version 1.0
ClickWorker (2013) clickworker.com. Accessed 05-Mar-2013
CrowdFlower (2013) crowdflower.com. Accessed 05-Mar-2013
Doan A, Ramakrishnan R, Halevy AY (2011) Crowdsourcing systems on the world-wide web. Commun ACM 54(4):86–96
Howe J (2006) The rise of crowdsourcing. Wired 14(14):1–5
oDesk (2013) odesk.com. Accessed 05-Mar-2013
Ipeirotis PG (2010) Analyzing the amazon mechanical turk marketplace. XRDS 17:16–21
Quinn AJ, Bederson BB (2011) Human computation: a survey and taxonomy of a growing field. In: Proceedings of the 2011 annual conference on human factors in computing systems, CHI '11, pp 1403–1412, New York. ACM
Schall D (2011) A human-centric runtime framework for mixed service-oriented systems. Distrib Parallel Databases 29(5–6):333–360
Schall D (2012) Service Oriented Crowdsourcing: architecture, protocols and algorithms. Springer Briefs in Computer Science. Springer New York, New York
Schall D et al. (2008) Unifying human and software services in web-scale collaborations. IEEE Internet Comput 12(3):62–68
SmartSheet (2013) smartsheet.com. Accessed 05-Mar-2013
SpeechInk (2013) speechink.com. Accessed 05-Mar-2013
von Ahn L (2006) Games with a purpose. IEEE Comput 39(6):92–94

CyLog/Crowd4U: A Case Study of a Computing Platform for Cybernetic Dataspaces

Atsuyuki Morishima

Introduction

Many emerging crowdsourcing/human-computation applications, including GWAPs (von Ahn and Dabbish 2008) (e.g., the ESP game) and Q&A services (e.g., Yahoo! Answers), are *data-centric*. They need humans to input, manage, and process data. In addition, computation is not necessarily closed in machines in most of data-centric applications today. To appropriately design such applications, it is important to explicitly handle data-centric computations of *both* humans and machines.

This chapter addresses a computing platform for cybernetic dataspaces. Here, we use the term *cybernetic dataspace* to emphasize that we explicitly deal with data-centric computations of both of humans and machines.

The chapter is organized as follows. First, it discusses a database-oriented view of cybernetic dataspaces, in which we consider humans to be *data sources* in the space. Second, we introduce a database abstraction, called CyLog (Morishima 2010; Morishima et al. 2011), which is one of the first languages to handle humans as data sources. CyLog is unique in that it models humans as *rational data sources* that behave rationally in the given incentive structure so that CyLog can naturally incorporate human intelligence into data-centric computation without losing a well-defined semantics. Finally, we explain Crowd4U, a computing platform for cybernetic dataspaces. Crowd4U was developed as a crowdsourcing platform for academic purposes and can run applications written in CyLog. CyLog and Crowd4U are being developed as part of the FusionCOMP project that started in 2009. This chapter describes a snapshot of this project as of March 2013. Because FusionCOMP is an ongoing project, the designs of CyLog and Crowd4U are subject to change in the future.

A. Morishima (✉)
University of Tsukuba, Tsukuba, Japan
e-mail: mori@slis.tsukuba.ac.jp

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_43,

A Database-Oriented View of Cybernetic Dataspaces

To explicitly handle human/machine computation, we need *abstractions* to describe such computation. Examples of such abstractions range from the formal models of computation to executable programming languages, and a good abstraction serves as a powerful tool both in theoretical research and software development. However, existing abstractions for computation have been designed only to describe the behavior of computers and do not offer tools for modeling people as components of computation. Human computation is out of the scope of languages; the logic of interaction with people needs to be implemented from scratch using primitive functions (e.g., GUIs or command-line interfaces) or crowdsourcing APIs (e.g., Amazon Mechanical Turk). More importantly, it is difficult to analyze or predict the behavior of the entire system that involves machine and human activities.

A fundamental question is whether we can develop a *principled* abstraction for data-centric human/machine computations. There have been some previous attempts to develop abstractions for human/machine computations. For example, CrowdForge (Kittur et al. 2011) uses the map-reduce abstraction to describe crowdsourcing applications. CrowdLang (Minder and Bernstein 2012) is a language that uses control and data flows to describe human/machine computation. The development of such abstractions, however, is still in its infancy.

An approach to develop such abstractions is to adopt and extend database abstractions and consider humans as *data sources*. There are at least two types of approaches. The first type, which we call the pure database approach here, is to allow humans to participate in the process of database queries, generally focusing on *data independence* in the presence of human data sources. Most existing work, including CrowdDB (Franklin et al. 2011), Qurk (Marcus et al. 2011), and Deco (Parameswaran et al. 2012), allows the humans to join the execution of SQL-like queries while there exists a study that uses Datalog-like formalism for discussing query processing involving humans (Parameswaran and Polyzotis 2011). In short, they try to construct the traditional database layer on top of human/machine computing resources, as shown in the left-hand side of Fig. 1. The database abstraction layer tries to give programmers the view in which the humans in the dataspace are considered to be a data source as a whole, so that they can submit queries to the system in the same way as to the traditional database management systems. Then, the processor delegates the operations that are difficult for a machine to process (e.g., subjective comparisons) to humans.

The second type, which we call the extended database approach, is to *extend* database abstractions for explicitly providing the means to utilize the power of human data sources, in order to design a wider range of data-centric applications that are more complex than traditional database queries (see the right-hand side of Fig. 1). CyLog is such an example; it has several extensions for utilizing human intelligence to implement a wider range of data-centric applications. As demonstrated in Morishima et al. (2012), CyLog can implement (1) a program for extracting structured data from tweets, in which the main contributor of data extraction is gradually changed from the crowd to the machine, and (2) a program for collecting data, in which humans participate in the process of identifying manageable small tasks to solve a bigger problem.

CyLog: A Programming Language for Cybernetic Dataspaces

In this section, we provide a brief introduction to CyLog, a programming language designed to provide a principled abstraction for describing, analyzing, and executing programs in cybernetic dataspaces. CyLog is a rule-based language with syntax similar to that of Datalog (Ceri et al. 1989), a data-oriented variant of Prolog. However, an essential difference in CyLog and existing languages is that CyLog handles human computation as a first-class component and allows us to *design and analyze* the behavior of users, whereas others give no hints on whether the users will behave in an expected manner.

CyLog assumes that a relational database exists in the cybernetic dataspace, as shown in the right-hand side of Fig. 1. The database has a set of *relations* (tables) each of which stores a set of *tuples* (rows) that conform to the *schema* of the relation. For example, as shown in Fig. 2, a database can have a relation called *Image*; the schema is identified as `Image(filename, year)` and has an example set of tuples `{ (file1.jpg, 2010), (file2.jpg, 2013)}`. Figure 2 illustrates the relation in a tabular form.

A key idea of CyLog is that it provides the means to partly abandon the *closed world assumption*, which is adopted by most existing database and logic-based programming languages. With this assumption, we assume that a fact does not hold in the real world if the database has no data to represent that fact. For example, if there is no information on a student in the student database of a university, we assume that there is no such student in the university. In contrast, CyLog allows data to be *open* in that when the data is not stored (or cannot be derived) in the database, it tries to

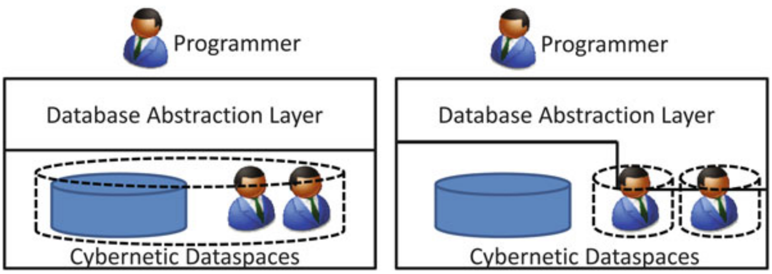


Fig. 1 Approaches based on database abstractions to deal with human data sources. The pure database approach (in the *left-hand side*) tries to give the programmer a view to which he can submit database queries, and tries to realize data independence in the presence of human data sources. The extended database approach (in the *right-hand side*) extends the traditional database abstraction to explicitly provide the means to utilize the power of human data sources to design a wider range of data-centric applications

Fig. 2 Relation `Image`

filename	year
file1.jpg	2010
file2.jpg	2013

extend the world by asking people whether or not a fact holds true. Therefore, it can naturally incorporate the processes of interactions with people in the language design.

However, it quickly becomes non-trivial to provide well-defined semantics in the presence of human computations. CyLog borrows concepts from game theory to model humans as novel *rational data sources*, which behave rationally within the given incentive structure. Because incorporating a feedback system is essential to the use of game theory in defining the semantics of computation not closed in machines, CyLog has a built-in reward system at the language level. In addition, CyLog allows programmers to describe the incentive structure as the *game aspect* so that the code to implement the incentive structure is separated from that for the other logics in the program. This makes it easy to implement, analyze, and maintain the incentive structures.

Rule-Based Language with Open Predicates

The main component of a CyLog program is the set of *rules*, each of which has the form *head* `<-` *body*`;` Each rule specifies that, for each combination of tuples satisfying the conditions specified in the *body*, we insert some other tuple into the relation specified in the *head*. For example, assume that we have the `Image` relation shown in Fig. 2. Then, CyLog rule

```
Image2013(filename) <- Image(filename, year), year=201 3;
```

inserts a tuple `(file2.jpg)` into the relation having schema `Image2013(filename)` because we have a tuple `(file2.jpg, 2013)` in the `Image` relation.

In CyLog rules, these relations (e.g., `Image`) are called *predicates*, and a predicate followed by its arguments (e.g., `Image(filename, year)`) is called an *atom*. A rule without a *body* is called a *fact*, which means that the tuple specified in the *head* is inserted into the database without any condition. For example, the CyLog fact

```
Image(filename:"File3.jpg", year:1999);
```

inserts the tuple `(File3.jpg, 1999)` into the `Image` relation.

CyLog allows predicates to be *open*, which means that the decision as to whether a tuple exists in the relation is performed by humans when the data cannot be derived (computed) from the data in the database. For example, to solicit humans to provide keywords to label the images stored in the `Image` relation, we can write the following rule:

```
Label(filename, keyword)/open <- Image(filename);
```

Fig. 3 Payoff matrix for the sESP game

Player A/Player B	cricket	baseball
cricket	(1,1)	(0,0)
baseball	(0,0)	(1,1)

Here the value of the attribute `keyword`, which does not appear in the rule body, will be given by humans. As explained later, Crowd4U, which is a crowdsourcing platform that can run CyLog programs, *crowdsources* providing values for open predicates.

Note that human resources are not always available, and the computing power of humans widely varies. One of the advantages of the rule-based language is that it has an affinity toward the asynchronous and parallel nature of human computation and crowdsourcing. For example, the above rule naturally describes parallel executions to compute keywords for more than one image.

Rational Datasources

In this section, we explore how the semantics of open predicates are defined. The key problem is that the human factors affect program executions. Because people might lie and need motivation to participate in the computation, it is difficult to predict the execution results. One possible approach is to consider people as *rational* data sources. By "rational," we mean that people provide data in a way that is consistent with the expected rewards. CyLog adopts terms and concepts from game theory to not only implement but also *design and analyze* the appropriate behavior of rational data sources. CyLog is therefore unique because other languages provide no hints as to whether users will behave in the expected manner.

Games are abstract concepts that have been studied well in the literature from both theoretical and practical aspects, and game theory is known to be useful when discussing not just real "games" but any system that involves incentive structures, such as networks, auctions, and GWAPs (Jain and Parkes 2009; Shoham 2008).

As an example, we discuss a simplified version of the ESP game (von Ahn and Dabbish 2008), which we call the sESP game. In the game, an image is shown to two players, and each player is required to predict keywords that the other would give for the same image. If the given keywords match with each other, the players are rewarded, and the matched keyword will be used for the label of the image.

A game is often written as a *payoff matrix*; Fig. 3 shows a part of the payoff matrix of the sESP game (only two terms are shown in the matrix). The Y and X axis show the possible actions of Player A and B, respectively. The matrix shows that each player can enter `cricket` or `baseball` for a given image. It also describes how payoffs are given to players. In each cell, (v_1, v_2) means that Players A and B receive v_1 and v_2 as their payoffs when they choose certain actions. In the sESP game, if they give the same term, they receive the payoffs.

Figure 4 illustrates the same sESP game in a tree style called the *extensive form* (Vega-Redondo 2003). Each *path* from the root to a terminal node corresponds to

Fig. 4 Extensive form

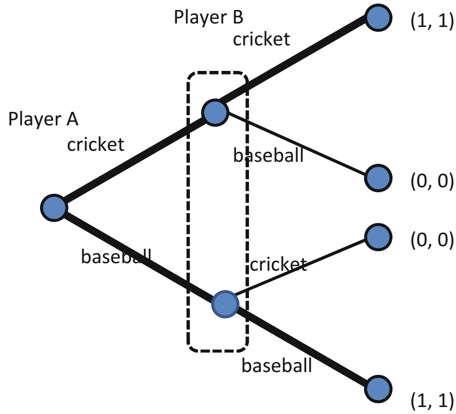
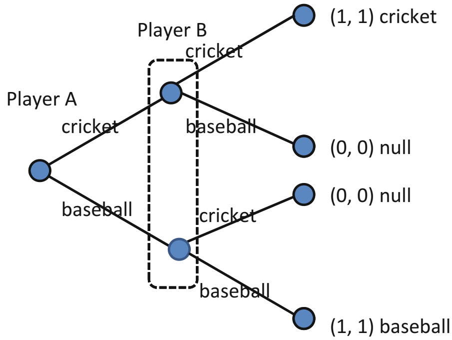


Fig. 5 Data game



each cell in the matrix shown in Fig. 3 and represents a possible play of the game. The leaf nodes are associated with payoffs to the players. The dotted circle means that the player B does not know the choice Player A took for his/her action.

To model humans as data sources, we consider the actions that the players chose to be values for open predicates. Then, we define the semantics of open values as actions in the *solutions* of the game (Vega-Redondo 2003), which are the paths taken by rational workers. For example, the solutions of the sESP game are the paths in which the players provide the same term (bold line).[1] Therefore, it is important to maintain the information on the path in each game play. We also extend the game to associate an *output value* with each terminal node. For example, Fig. 5 shows the sESP game using the extensive form associated with output values. We call such a game with output values a *data game*. The output value of a data game is defined as the output value of the solution.

[1] In reality, some terms are more likely to be chosen and therefore the expected payoffs are different.

```
Rule:
  LabelInput(filename, keyword)/open <- Image(filename);
  Label(filename, keyword:g(filename)) <-Image(filename), g(filename)@end;
Game:
  g(filename)@time(10): duplicate, {LabelInput}
```

Fig. 6 A simplified ESP game

Game Aspects

Programmers implementing a game need to write code to maintain paths of the game and compute output values and payoffs to players; however, if we use other general-purpose programming languages to implement games, the code tends to become more difficult to read, analyze, and tune. The reason here is two-fold. First, the code fragments related to games are implicitly encoded in many different parts of the program. In general, programmers can implement more than one incentive structure in a program. In some cases, the code for an interaction with humans may be associated with more than one game (i.e., the code for different games overlap in the program). Even in the case that a program implements only one game structure, the code to maintain paths and compute payoffs are not localized. Second, the functions to implement games are repeatedly implemented in different programs.

An important principle in software development is the *separation of concerns*. CyLog is unique in that it introduces the *game aspect* that separates the code for the incentive structures from the other set of logic encoded in the program. The game aspect describes the incentive structures at one particular place in the code. Then, the codes to maintain the path of each game play and to compute payoffs and values are automatically generated from the aspect description so that the programmer does not have to manually write the code for implementing the incentive structures. This allows programmers to write, analyze, and tune the incentive structures for rational data sources more easily.

As an example, Fig. 6 shows a fragment of a CyLog program that implements the sESP game. The code consists of two parts. The `Rule` part contains a set of CyLog rules. The first rule solicits humans to input keywords for the given image, and the second uses the inputs to compute a label for the image. The `Game` part describes the game aspect of the sESP game. In the description of the game aspect, each game is identified by a function called a *Skolem function*. For example, `g(filename)` is the Skolem function for the sESP game, which means that a game is created for each specified parameters (namely, a game is created for each image specified by `filename`). We call each game a *game instance*. In the description, `g(filename)@time(10)` specifies that the game instance ends in ten seconds.

For each game instance, a special table called a *path table* is automatically constructed. The path table maintains the *path* of the play of the game instance to show how the game reached the last state. The relations specified in `{ ..}` (e.g., `LabelInput` in Fig. 6) supply tuples inserted to the path table. Given the description of the game aspect in Fig. 6, (1) a game instance is created for each image file,

Fig. 7 Path table

Order	Date	Player	Rel	Action
1	10:10am	Kate	LabelInput	cricket
2	10:11am	Ann	LabelInput	cricket
3	10:12am	Pam	LabelInput	baseball

(2) a path table is constructed for the game instance, and (3) the inputs to the `LabelInput` are inserted into the path table to be recorded as the choices of players. Fig. 7 shows an example of the path table. It has the schema P(order, date, player, rel, action), and each tuple in the table records when and who entered the open values for `LabelInput`.

In the `Game` part of Fig. 6, `duplicate` is the *game aggregation* used in the game. A game aggregation is a function that implements a data game and computes the payoffs to players and the output values. Some game aggregations are predefined in the CyLog library. The `duplicate` implements the same data game as that in Fig. 5 except that it extends the table with an infinite number of players and terms (values). Assuming that people behave rationally, it is expected that the value is computed by the solutions with the aggregations. Assume that we have the path table shown in Fig. 7. With `duplicates`, the payoff values for `Kate`, `Ann`, and `Pam` are 1, 1, and 0, respectively, because `Kate` and `Ann` agreed on the values. The output value is `cricket`, which is the value given by both `Kate` and `Ann`.

The payoffs and output values produced by the game aggregations are consumed as follows. First, payoffs are given to the players at a specified time. By default, the payoffs are given to them when the game ends. Second, output values are consumed by rules in the rule part. For example, the output value of game `g(filename)` is consumed by the second rule in Fig. 6. The "`keyword:g(filename)`" in the head specifies that the value of the `keyword` attribute is the output value of the game identified by `g(filename)`.

Crowd4U: A Platform for Cybernetic Dataspaces

Crowd4U is a crowdsourcing platform that has an engine for executing CyLog code, and is designed to harness the power of people in academia. Crowd4U *crowdsources* input values for open predicates through the scripts placed primarily on the Web sites of universities. This section explains the highlights and the architecture of Crowd4U.

Highlights of Crowd4U

1. **Platform being Developed by Universities:** To our knowledge, Crowd4U is one of the first non-commercial, microtask-based crowdsourcing platforms

being developed by universities. It is unique in several ways compared to other systems (e.g., Bossa official website). First, Crowd4U provides a high-level abstraction for complex human/machine computation. Second, as explained later, it supports various task assignment and incentive structures including push/pull-style task assignments. Third, Crowd4U microtasks can be performed at different Web sites via embedded scripts. As of March 2013, the collaborators are from 15 universities and the related scripts are embedded in many Web sites. Because many workers voluntarily perform tasks on Crowd4U, they are called *contributors*. Many of these contributors are university students. Although contributors are not required to create accounts on Crowd4U, the number of the accounts created is more than 230 as of March 2013. The estimated number of anonymous workers on Crowd4U is more than 1,000.

2. **Microtask-based Platform:** Crowd4U supports microtasks that contributors can perform in a short period of time. The microtasks are similar to human intelligence tasks of Amazon's Mechanical Turk. Programmers write CyLog codes to define microtasks that contain open predicates and register them to the task pool of Crowd4U. Contributors perform the microtasks registered in the task pool.
3. **Non-commercial Ongoing Projects:** Crowd4U is hosting several non-commercial crowdsourcing projects. For example, L-Crowd is a project started by active LIS researchers in Japan to apply crowdsourcing technologies to library problems. In 2012, they designed microtasks to identify different books that have the same ISBN in an effort to clean the bibliography database of the National Diet Library (L-Crowd; National Diet Library). Another example is a project to develop a crowdsourcing solution for quickly understanding what happened in disaster scenes. Today, many people upload photos and movies on the Web. The project is trying to design microtasks to infer how a tornado moved in Tsukuba City in 2012, using the photos and movies uploaded on the Web.
4. **Various Incentive Structure:** Although Crowd4U is a microtask-based platform, it is not a microtask market. Rather, Crowd4U microtasks are performed within a variety of incentive structures. For example, some contributors are university students who perform microtasks when they try to download PDF files used in courses. Some of them perform microtasks driven by gamefication mechanisms (e.g., ranking of their contributions). Some of them happen to find and perform microtasks when visiting Web sites that embed scripts to host microtasks. Some of them perform microtasks by using "Crowd4U terminals" located on campus. In these ways, Crowd4U can serve as a testbed for various incentive structures.
5. **Advanced Functions Embedded in the Engine:** One of the purposes for which Crowd4U was developed was to implement advanced functions that are not incorporated in commercial crowdsourcing platforms. Crowd4U provides a suite of built-in functions to support CyLog programming. It provides functions to implement common game situations including majority votes, duplicate (coordination) games, and other data games with various types of incentive structures. Crowd4U also provides a RAD (rapid application development) tool for crowdsourcing applications. It has a facility to easily define Web views that are associated to

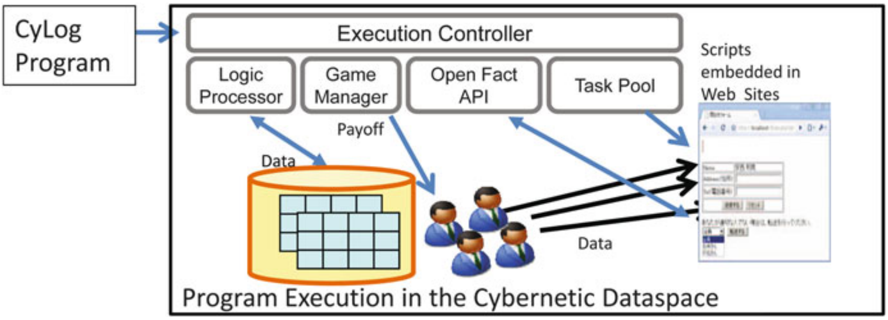


Fig. 8 CyLog/Crowd4U architecture. The logic processor evaluates rules without open predicates. The game manager manages data (e.g., payoffs) related to games. The open fact API provides the means for external codes to tell Crowd4U that some facts hold. The task pool manages the set of microtasks to be assigned to workers (contributors)

open predicates and event invocations, making it easy to implement interactions with humans. In the future, we also plan to embed mechanisms to improve data quality.

Crowd4U Architecture

Figure 8 shows the architecture of Crowd4U, which is also described below. **Query Processing for Data-centric Complex Human/Machine Computations.** To execute complex data-centric crowdsourcing applications, Crowd4U supports the execution of CyLog codes by both machines *and* humans, interleaved with the execution of programs in general-purpose languages. It provides an API (Open Fact API) to invoke event signals to indicate that a new fact holds (i.e., a tuple exists in a relation), which allows us to naturally combine CyLog programs with other programs written in procedural languages.

The CyLog processor adopts a semi-naïve event-driven evaluation strategy in which the rules are evaluated in a bottom up manner; Crowd4U knows that a fact holds when it receives an event signal that indicates that the fact holds. Rules are processed as follows: when all of the atoms in a rule body hold, we need to determine whether the head of the rule holds. When the head is not open, the logic processor evaluates the rule based on the first-order predicate logic. Internally, it invokes an event signal to indicate that the head holds, and the result tuple is inserted into the database. When the head is open, humans are assumed to perform microtasks, which will call the open-fact API in order to indicate whether the fact holds or not.

Handling Rational Data Sources. According to the behavior of workers during the program execution, the game manager provides values that represent payoffs. Currently, payoff values in Crowd4U are not monetary, but are visible to workers, showing the contribution of each worker. From our experience, non-monetary indicators to inform the workers of the value of their contributions work well on Crowd4U, and we have found that it is very important to make their contributions visible. Therefore, Crowd4U explicitly acknowledges the contributions of workers in different ways. We plan to make Crowd4U able to support monetary incentives in the future.

Lessons Learned and Discussions

Our experience shows that an integrated abstraction of human/machine computation is indispensable to achieve the non-adhoc development of software in cybernetic dataspaces. This section addresses some of the lessons learned.

There is an affinity between rule-based languages and integration of human/machine computations. We have found that a rule-based language allows us to concisely describe data-centric applications. In particular, the rule-based language is appropriate to handle human computation in that (1) human computation is often asynchronous and the rule-based code does not impose unnecessary timing constraints, and that (2) complex human/machine interactions often require event-driven executions that rule-based languages can implement in a straightforward way. In addition, describing the code as a set of rules brings flexibility to data-centric human/machine computation. For example, allowing code to dynamically add and delete rules naturally achieves *higher-order crowdsourcing*, in that humans not only contribute to data input, but also participate in program evolution.

The integrated abstraction has great potential. We found that not only is the integrated abstraction indispensable for appropriately designing applications in cybernetic dataspaces, but also that it raises novel opportunities and challenges. The "open" predicates allow us to execute the same program in flexible ways with different mixtures of human/machine computation. The feature is interesting in various scenarios. For example, it is often the case that we need to quickly develop a software tool but there is no time to fully implement it. In such a case, the integrated abstraction allows us the seamlessly transfer from the human-powered tool to the fully automatic tool, by first allowing humans to execute most of the program with open predicates, and then gradually replacing those open predicates with implemented automatic functions.

Modeling humans as a data source is a challenge. CyLog uses game theory to define the semantics of rational data sources; however, games are not a magic wand. In some applications, it may be difficult to provide real benefits (e.g., points, money, and evaluation scores) to be modeled by payoff values. Humans are not necessarily rational; however, we believe that modeling humans as rational data sources is a good starting point because we have a theoretical background, and for some applications the concept of rational data sources works well. An interesting open question is whether we can apply the results from other fields such as cognitive and behavioral sciences in mitigating the limitations of this rationality assumption.

Acknowledgements The author is grateful to the members and collaborators of the FusionCOMP project, and the contributors who are performing microtasks on Crowd4U. Their names are listed at http://crowd4u.org. Note that the list contains only the names of contributors who have accounts on Crowd4U, and there are many more anonymous contributors who perform microtasks on Crowd4U. The FusionCOMP project is partially supported by PRESTO from the Japan Science and Technology Agency, and by the Grant-in-Aid for Scientific Research (#25240012) from MEXT, Japan.

References

Bossa official website. http://boinc.berkeley.edu/trac/wiki/BossaIntro.

Ceri S, Gottlob G, Tanca L (1989) What you always wanted to know about datalog (and never dared to ask). IEEE Trans Knowl Data Eng 1(1):146–166

Crowd4U. http://crowd4u.org

Franklin MJ, Kossmann D, Kraska T, Ramesh S, Xin R (2011) CrowdDB: answering queries with crowdsourcing. In: SIGMOD Conference, Athens, pp 61–72

FusionCOMP project. http://www.kc.tsukuba.ac.jp/~fusioncomp/

Jain S, Parkes DC (2009) The role of game theory in human computation systems. In: KDD workshop on human computation, Paris, pp 58–61

Kittur A, Smus B, Kraut R (2011) CrowdForge: crowdsourcing complex work. In: CHI extended abstracts, Vancouver, pp 1801–1806

L-Crowd: applying crowdsourcing technology to library problems. http://crowd4u.org/lcrowd/

Marcus A, Wu E, Madden S, Miller RC (2011) Crowdsourced databases: query processing with people. In: CIDR 2011, Asilomar, pp 211–214

Minder P, Bernstein A (2012) CrowdLang: a programming language for the systematic exploration of human computation systems. In: SocInfo 2012, Lausanne. Springer, Berlin/New York, pp 124–137

Morishima A (2010) A database abstraction for data-intensive social applications. In: The 5th Korea-Japan database workshop 2010 (KJDB2010), Jeju Island, 28–29 May

Morishima A, Shinagawa N, Mochizuki S (2011) The power of integrated abstraction for data-centric human/machine computations. In: VLDS 2011 held at VLDB2011. Seattle, Please see http://www.wikicfp.com/cfp/servlet/event.showcfp?eventid=16021&copyownerid=24065

Morishima A, Shinagawa N, Mitsuishi T, Aoki H, Fukusumi S (2012) CyLog/Crowd4U: a declarative platform for complex data-centric crowdsourcing. PVLDB 5(12):1918–1921

National Diet Library. http://www.ndl.go.jp/en/index.html

Parameswaran AG, Polyzotis N (2011) Answering queries using humans, algorithms and databases. In: CIDR 2011, Asilomar, pp 160–166

Parameswaran AG, Park H, Garcia-Molina H, Polyzotis N, Widom J (2012) Deco: declarative crowdsourcing. In: CIKM 2012, Maui, pp 1203–1212

Shoham Y (2008) Computer science and game theory. Commun ACM 51(8):74–79

Vega-Redondo F (2003) Economics and theory of games. Cambridge University Press in England, Please see http://en.wikipedia.org/wiki/Cambridge_University_Press

von Ahn L, Dabbish L (2008) Designing games with a purpose. Commun ACM 51(8):58–67

Multiagent Environment Design for Pervasive Human-ICT Systems: The SAPERE Approach

Gabriella Castelli, Marco Mamei, Alberto Rosi, and Franco Zambonelli

Introduction

A wide range of novel pervasive computing scenarios can be best modeled via a set of autonomous software components interacting on a locality scope. For example, an application related to monitoring an environment via ICT sensors and devices requires software components (i) deployed on sensors and autonomously adapting to environmental conditions, (ii) locally interacting to calibrate, verify each other findings, and diffuse information.

A Multiagent systems is a software system naturally fitting such a model. A multiagent system is composed of multiple autonomous entities (agents) interacting with each other to realize a given application. Interactions in a multiagent system are modeled as taking place within an environment representing the space where agents'activities take place (Weyns et al. 2004; Platon et al. 2006). In such a system, the environment is defined as a first-class abstraction with the role of providing the surrounding conditions for agents to exist, the mediation of agent interactions, and access to resources (Platon et al. 2006). In a number of research proposals, especially in the context of pervasive computing research, the environment is modeled, realized and implemented by means of a software infrastructure for the provisioning of a number of digital services comprising interaction and communication facilities, discovery, and life-cycle management. In pervasive computing, such an infrastructure is used by the agents to ubiquitously access services to support better interaction with the surrounding physical world and with the activities occurring in it. The environment infrastructure also support users in deploying customized services, making the overall infrastructure as open as the Web currently is

G. Castelli (✉) • M. Mamei • A. Rosi • F. Zambonelli
DISMI, University of Modena and Reggio Emilia, Modena, Italy
e-mail: gabriella.castelli@unimore.it; marco.mamei@unimore.it; alberto.rosi@unimore.it; franco.zambonelli@unimore.it

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_44,

Zambonelli (2012). Unfortunately, most of the solutions so far are proposed as "add-ons" to be integrated in existing frameworks (Babaoglu et al. 2006; Kari and Rozenberg 2008). The result is often increased complexity of current frameworks and the emergence of a contrasting trade-off between different solutions.

In our opinion, there is need to tackle the problem of modeling the environment of a MAS at the foundation, answering the following ambitious question: is it possible to conceive a radically new way of modeling integrated MAS and their execution environments, such that the apparently diverse issues of context-awareness, dependability, openness, flexible and robust evolution, can all be uniformly addressed once, and for all, via a sound and programmable self-organization approach? The overall goal of the SAPERE (*Self Aware Pervasive Service Ecosystems*) project is to show that a novel positive answer to the above question exists, and it will try to go even further, by defining an innovative framework in which all the identified issues can be solved via a limited set of "laws" embedded in the framework to support and rule its self-organizing activities.

The rest of this chapter is structured as follows: in section "The SAPERE Approach" we overview the main ideas at the base of SAPERE. In section "The SAPERE Software Infrastructure" we present the software infrastructure managing the environment associated with SAPERE applications. In section "Application Scenario: Environment Design" we present the modeling of an exemplary pervasive computing application focusing on crowdsourcing activities. In particular, we highlight how the SAPERE environmental abstractions effectively support the design of such an application. Finally, section "Conclusions" provides some conclusions.

The SAPERE Approach

SAPERE takes its primary inspiration from natural ecosystems, and starts from the consideration that the dynamics and decentralization of future MAS will make it appropriate to model the overall world of services, data, and devices as a sort of distributed computational *ecosystem* (Viroli and Zambonelli 2010).

Specifically (see Fig. 1), SAPERE considers modeling and architecting a MAS environment (Weyns et al. 2004) as a non-layered *spatial substrate*, laid above the actual pervasive network infrastructure. The substrate embeds the basic laws of nature (or *eco-laws*) that rule the activities of the system. It represents the environment on which individuals of different species (i.e., the agents) interact and combine with each other (in respect of the eco-laws and typically based on their spatial relationships), so as to serve their own individual needs as well as the sustainability of the overall ecology. Users can access the ecology in a decentralized way to use and consume data and services, and they can also act as "prosumers" by injecting new data and services.

For the *agents* living in the ecosystem, SAPERE adopts a common model and a common treatment. All agents in the ecosystem (whether sensors, actuators, services, users, data, or resources in general) have an associated semantic

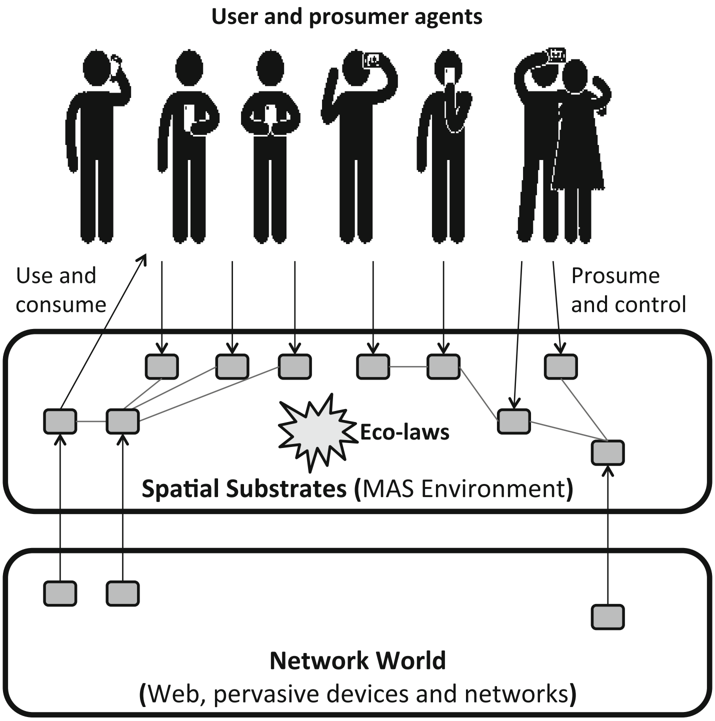


Fig. 1 The SAPERE Reference Architecture. The environment abstractions support agents' activities and interactions, easing MAS design

representation, which is a basic ingredient for enabling dynamic unsupervised interactions between components. To account for the high dynamics of the scenario and for its need of continuous adaptation, SAPERE defines such annotations as "living", active entities, tightly associated to the agent they describe, and capable of reflecting its current situation and context. Such *Live Semantic Annotations* (LSAs) thus act as observable interfaces of resources, as well as the basis for enforcing semantic forms of dynamic interactions (both for service aggregation/composition and for data/knowledge management).

The *eco-laws* define the basic policies driving *reactions* among the LSAs of the various agents of the ecology. In particular, the idea is to enforce, on a spatial basis, and possibly relying on diffusive mechanisms, dynamic networking and composition of data and services. Eco-laws, in particular, support agent discovery and interaction by connecting (bonding) their LSAs; they support distributed operations by allowing LSAs to be spread and aggregated across the network; they also allow the deletion of unused LSAs for garbage collection.

Following the SAPERE approach, MAS design proceeds by coding the agents' computation activities (business logic) and by specifying agents' LSAs. The LSA of an agent comprises both a description of the agent current situation and capabilities, and a description of its needs and requests.

Once this specification is realized, the SAPERE environment takes care of running the overall MAS. In particular, eco-laws will trigger reactions among the agents allowing them to execute their tasks.

The SAPERE Software Infrastructure

To turn the above described reference architecture into an operational one, a software substrate should proactively mediate interactions between components (i.e., in general terms, all those active agents that participate to the ecosystems). That is, it should act as an active environment in which to store the continuously updating LSAs of agents, so as to adaptively support the matching process triggering eco-laws in dependence of the current conditions of the overall ecosystem.

From the distribution viewpoint, the SAPERE infrastructure is formed by a network of nodes, each hosting a local LSA-space, with neighbor relations typically shaped according to some spatial or network relations. The LSA-space is a sort of local tuple space, which hosts LSAs in the form of tuples. The shape of the actual network of connection is determined by a reconfigurable component, which can be based on, e.g., a strategy that connects nodes based on spatial proximity or rather one relying on social proximity (Zambonelli et al. 2011). The shape of such network determines the paths along which LSAs on a node can propagate and diffuse to other nodes.

Whenever an agent (whether corresponding to a device, a sensor, a service, or to an application agent) approaches a node, its own LSA is automatically injected and stored into the LSA-space of that node, making the component part of that space and of its local coordination dynamics (i.e., subject to eco-laws operations). The LSA can also propagate and diffuse in the network of LSA spaces to enable distributed operations. When a component moves away from a node, its LSA is automatically removed from that space.

From the viewpoint of individual agents, the environmental infrastructure provides them (via a simple API) with the possibility of advertising themselves via an LSA, and supporting the continuous updating of the LSAs. In addition, such an API enables agents to detect local events as the modifications of some LSAs or the enactment of some eco-laws on available LSAs.

- From the viewpoint of the underlying network, the environment accounts for transparently absorbing dynamic changes at the arrival/dismissing of the supporting devices, without affecting the perception of the spatial environment by individuals, and is able to detect events on LSAs and to trigger the necessary eco-laws. Eco-laws are realized as a set of rules embedded in SAPERE nodes. For each node, the same set of eco-laws applies to rule the dynamics between LSAs. In particular, we identified four basic eco-laws that can fully support MAS activities:The **Bonding eco-law** enables the interaction between components that live in the same SAPERE node. The Bonding eco-law realizes a bond between two components, i.e., a virtual link between their LSAs. Such a bond is established as a result of a pattern matching mechanism on the set of LSAs. Once a bond is established the agent holding the LSA is notified of the new bond and can trigger actions accordingly. In SAPERE we do not have a standard “read” operation. Agents express in their LSA the fact that they wish to bind with other LSAs. Once eco-laws create the bond, the agent can then read the bonded LSA.

- The ability to aggregate sparse information into an higher level summary of the system’s context is another fundamental requirements for SAPERE application. The **Aggregation eco-law** is intended to aggregate LSAs together so as to compute summaries of the current system’s context. An agent can express in its LSA a request for aggregation. The aggregation eco-law is triggered by such a request (via pattern matching) and creates the summary of the system’s context.
- **The Decay eco-law** enables the elimination of components from the SAPERE environment. The Decay eco-laws applies to all LSAs that specify a decay property to update the remaining time-to-live according to a specific decay function, or allows actually removing LSAs that, based on their decay property, are expired. The Decay acts, therefore, as a kind of garbage collector capable of removing LSAs that are no longer needed in the ecosystem or a no longer maintained by a component.
- Since the SAPERE model is based on a set of networked interaction spaces, it is of course fundamental to provide a mechanism to send information to remote spaces. In SAPERE we designed a **Spread eco-law** capable of diffusing LSAs to remote spaces. One of the primary usages of the spread eco-law is to enable searches for components that are not available locally, and to enable the remote advertisement of services that available in the local node. An agent can in fact express in its LSA a request for spreading. The spreading eco-law is triggered by such a request (via pattern matching) and sends the LSA to other spaces.

The presented SAPERE architecture and the identified eco-laws can support software engineers to the design of multiagent applications. The proposed environment abstractions, in fact, provide a framework to manage and orchestrate the agents’ activities. In the next section, we sketch the design of an exemplary multiagent application illustrating how the above abstractions support a complex human computation activity, namely software design for multiagent systems.

Application Scenario: Environment Design

In this section we present the design in SAPERE terms of an exemplary application involving human and ICT computation. The application aims at monitoring an environment by using a network of Closed Circuit Cameras (CCCs) and crowd sourced people who will be dynamically recruited to classify and judge situations happening in the environment on the basis of “human computation capabilities” (Alt et al. 2010; Das et al. 2010; Greene et al. 2011).

In SAPERE, the elements (users, sensors and software services) composing the application are all modeled as components of the ecosystem, all with a common associated semantic representation (the LSA) expressing their identities, their necessities and their objectives. SAPERE is also based on a notable separation of concerns between an application's computation and interaction. Computation (i.e., the main application business logic) is coded in the SAPERE agents using agent-oriented software engineering methodologies (Weyns et al. 2009). Interaction consists in writing agents' LSAs and managing their evolution over time.

On this basis, users' smartphones and camera sensors should have installed an application collecting information (e.g., their current location and their availability in being possibly recruited) and enabling interactions. The core application will be coded via standard software engineering methodologies, but its interactions are mediated by the SAPERE space. In particular, the application will expose components information and availability in a LSA injected in the SAPERE node embedded in the device. The user's LSA will be a tuple in a form assimilable to:

LSA=[type="human" id="user_id",location=(x,y)]

while, for example, the LSA of a CCC sensor will assume the following aspect:

LSA=[type="sensor" id="CCC_id",location=(x,y),resolution=1920x1080].

The LSAs are "live" in the sense that their fields are constantly updated to reflect the current agent's location and availability. The application willing to monitor the environment will also expose via a proper LSA the tasks it requires. The LSA will be a tuple in a form assimilable to:

LSA=[user=?,location=(x,y),task="take a picture of the environment"].

SAPERE *spread eco-law* will selectively forward requests to interested components of the ecosystem, whatever they are users, sensors, software services, etc. Upon a positive match the *bond eco-law* will connect the request to an appropriate component. Once this bond takes place, the component is notified of the requests, can take the picture and send it back to the requesting Web service. At this point, the Web service can decide to dynamically recruit a human user who will be requested to use human judgment capabilities to classify the situation. Following an approach similar to the one before, it will expose an LSA in the form:

LSA=[type="human",user=?,location=(x,y),task="classify a situation"].

Users who will accept the task will later report their findings.

There are a number of SAPERE characteristics that can efficiently support the development of such an application.

1. SAPERE adopts an unified treatment and representation of ecosystem components, such as devices, users, software have a common associated semantic representation expressing their identities, their necessities, their objectives.
2. SAPERE architecture allows the programmers to focus on the agents' business logic while leaving to the SAPERE infrastructure (environment) the task of enabling interactions and coordination among the agents.

3. Networking mechanisms integrated in the SAPERE infrastructure allow contacting possibly interested users in an efficient way.
4. LSA's flexible structure and the use of semantic technologies allow users to express in a flexible way the activities requested to the users and viceversa for the user to flexibly express what they are willing to do.
5. The uncoupled nature of shared space interaction coupled with the locality scope enforces by the SAPERE nodes allows enforcing powerful and flexible communication patterns. The requesting service can leave a request LSA in the environment for interested passers-by users.

Conclusions

From our perspective, the abstractions and mechanisms proposed by SAPERE are effective in supporting the design and development of pervasive multiagent systems involving ICT devices and human actors.

Currently, we are working towards improving some implementation aspects of the SAPERE middleware, in particular with regard to optimizing LSAs storing and access, and at extending its support for semantic data representation. In addition, it will be important to develop a suitable IDE (Integrated Development Environment) supporting application designers in applying the SAPERE abstractions and to understand how to realize LSAs and eco-laws compatible with their needs.

On this basis, as plan for future work, we intend to experience the SAPERE approach with a number of innovative services in the area of crowd management and urban computing (Alt et al. 2010; Zambonelli 2012). The final goal would be to have a SAPERE-based "standard" methodology to code such applications instead of relying on one-off custom solutions.

Acknowledgements Work supported by the FET Proactive Initiative on "Self-awareness in Autonomic Systems", under grant No. 256873.

References

Alt F, Shirazi AS, Schmidt A, Kramer U, Nawax Z (2010) Location-based crowdsourcing: extending crowdsourcing to the real world. In: NordiCHI conference, Reykjavik

Babaoglu O et al. (2006) Design patterns from biology for distributed computing. ACM Trans Auton Adapt Syst 1(1):26–66

Das T, Mohan P, Padmanabhan V, Ramjee R, Sharma A (2010) Prism: platform for remote sensing using smartphones. In: MobiSys conference, San Francisco

Greene K, Thomsen D, Michelucci P (2011) Exploration in massively collaborative problem solving. In: IEEE international conference on privacy, security, risk and trust IEEE international conference on social computing, Boston

Kari L Rozenberg G (2008) The many facets of natural computing. Commun ACM 51:72–83

Platon E, Mamei M, Sabouret N, Honiden S, Parunak V (2006) Mechanisms for environments in multi-agent systems: survey and opportunities. J Auton Agents Multi-Agent Syst 1(14):31–47

Viroli M, Zambonelli F (2010) A biochemical approach to adaptive service ecosystems. Inf Sci 10(180):1876–1892

Weyns D, Parunak H, Michel F, Holvoet T, Ferber J (2004) Environments for multiagent systems, state-of-the-art and research challenges. In: Environments for multi-agent systems. International workshop on environments for multi-agent systems, New York, pp 47–52

Weyns D, Parunak H, Shehory O (2009) The future of software engineering and multi-agent systems. Int J Agent-Oriented Softw Eng 4(3):31–47

Zambonelli F (2012) Toward sociotechnical urban superorganisms. IEEE Comput, 47(8)

Zambonelli F, Castelli G, Mamei M, Rosi A (2011) Integrating pervasive middleware with social networks in sapere. In: International conference on selected topics in mobile and wireless networking, pp 145–150

The "Human Sensor:" Bridging Between Human Data and Services

Neal Lathia

Introduction

The technology that we encounter in our daily lives is increasingly characterised by its ability to collect from, store, and process growing amounts of diverse data streams about human life. These sources of data range from small to large and offer varying levels of insight into our behaviours, thoughts, decisions, and actions. Researchers' ability to source from these data streams gives rise to the idea of *humans as sensors*: existing and technologically feasible (social, sensing, smartphone, or otherwise) systems digitise a growing number of facets of life, and provide access to data that historically remained quantitatively unavailable.

While the existence of human data sources continues to grow, an ongoing challenge that researchers face is how to go beyond analysis, and towards designing and building systems that put these datasets to use. The overarching question is: if we can source data from humans (who, in doing so, act as our 'sensors'), how do we fit these people into the design or broader framework of systems that leverage the data's value?

In this chapter, we consider how researchers may reason about building the bridge between human data and systems. We do so by discussing three broad questions. First, where does human-sensor data come from? The literature in this area is diverse and fragmented, yet crowd-sourcing, participatory sensing, and database mining all share remarkable similarities in terms of the data that they provide. We therefore characterise these data sources in terms of their original purpose, the obtrusiveness of collecting them, and their underlying structure. Second, what system models exist where this data is applicable? In particular, we discuss recommendation, retrieval, and emerging behaviour-mediating technologies; each of these

N. Lathia (✉)
Computer Laboratory, University of Cambridge, Cambridge, UK
e-mail: neal.lathia@cl.cam.ac.uk

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_45, © Springer Science+Business Media New York 2013

provide a framework for applying human data to serve different needs. In doing so, we expose a set of examples of systems that have been built using data from crowds and sensors. We highlight how these system models are applied to human mobility data by describing a case study of using public transport access records to build automated alert and fare-recommendation systems. Finally, we discuss three open research questions in this space, which encompass the spectrum of system building: understanding data representativity, uncovering how system design affects learning about humans, and the challenge of reasoning about relevance and effectiveness across different application domains.

Generating Data Streams

We begin by discussing the context of using humans as sensors. Broadly speaking, a *sensor* is a means of translating a physical phenomenon into a digital signal; for example, a gyroscope measures its orientation. Referring to a “human sensor” acts as an umbrella term for our growing ability to capture data streams about human life—activities, movements, thoughts, behaviours—and encompasses the variety of both explicitly and implicitly available means of collecting data from or about people, via digital systems.

For example, the widespread adoption of smartphones (as well as other sensor-enhanced devices) now means that people regularly carry items that are instrumented with means for collecting data streams that tell us about their behaviour; moreover, smartphones are, naturally, built as interactive devices and thus ideal for collecting data that is manually input while on the go. Researchers can therefore leverage these devices to collect data via both sensors (Eagle and Pentland 2006) and participant’s direct input (Intille et al. 2003).

How does all this data collection come to fruition? The research literature that has begun to emerge in this space often discusses contributions as broadly related to *participatory sensing* or *crowd-sourcing*. The narrative behind the former (Burke et al. 2006) is that the increasing ubiquity of sensor-enhanced devices gives rise to the potential of collecting data from a community of, for example, participating smartphones. The focus here tends to be on the availability of sensors (Lane et al. 2008): the spread of smartphones throughout the world has translated the historical problem of *deploying* sensor networks into one of *harnessing* volunteers’ sensors for data collection. Crowd-sourcing (Doan et al. 2011), instead, focuses on the opportunity that arises to tackle large problems by means of groups of volunteers, where each individual’s contribution may be small (e.g., writing an article or, indeed, appending a couple of sentence to one) but the group’s output is significant (e.g., creating Wikipedia). Much like participatory sensing, the key factor that determines the success of crowd-sourced data collection is the ability to harness, engage, and maintain a community of contributors. Delineating the nuances between these two groups therefore goes beyond the scope of this chapter: instead, we focus on similarities and where the two methods are beginning to meet. Both crowd-sourcing and participatory sensing are active means of generating data: advances in mobile devices now means that data from “crowds” is often accompanied by sensor data (e.g., geo-located tweets Quercia et al. (2012a)), crowds’ smartphones can be used to work on sensor-related tasks (Yan et al. 2009), and sensor data can be harnessed by means of social systems (e.g., fine-grained GPS mapping from Foursquare check-ins (Shaw et al. 2013) and merging social activity sharing with sensor sampling (Hossmann et al. 2012)). In fact, some research now blends the terms together into *crowd-sensing* (Ra et al. 2012).

An alternative source of human data arises, instead, from those repositories of information that are automatically created as we use different systems: many of these databases have presented the opportunity for research that comes as a consequence of the pre-existence of data, unlike the more traditional research that solicits defining questions and hypotheses prior to data collection. Increasingly, a variety of daily tasks that we engage in create digital “footprints” including: renting a shared bicycle (Froehlich et al. 2009b), making a telephone call (Ratti et al. 2010), clicking on web links (Radlinski and Joachims 2007), taking a photograph (Girardin et al. 2008), and entering a public transport system (Bryan and Blythe 2007) are amongst a growing family of actions that are automatically logged. More importantly, while individual entries may, at face value, seem meaningless, the aggregation of large samples of these datasets convert them into invaluable resources for insight into human life, even though they were originally created to serve other purposes.

The landscape of human-driven data that is available to modern-day researchers is, therefore, seemingly limitless. How can it be characterised? In the following section, we review how many of these data sources are being put to use; we close here by broadly characterising the varying qualities of human-sourced data:

1. **Obtrusiveness.** A hallmark characteristic that differentiates different sources of “human sensor” data is the effort and commitment that is required by participants in order to serve the researcher’s purposes. Consider, for example, the difference between participatory sensing to gauge transportation modalities (Reddy et al. 2010) vs. sensor-augmented experience sampling to measure the geography of happiness (Mackerron. 2012). In the former, all that is required is for participants to contribute samples from their smartphone sensors, in the latter, participants must manually complete momentary mood assessments.
2. **Original Purpose.** Human-sourced data is further differentiated by considering why it was originally created. For example, the original intent behind tweets and Foursquare check-ins is to participate in those services’ social functions by, for example, sharing your location with your friends. The fact that these data sources are now used to study mobility (Noulas et al. 2012a) and mood (Quercia et al. 2012b) is divorced from the data’s original purpose; on the other hand, sensor samples collected after manually inputting an activity (Hossmann et al. 2012) and location traces from participants’ cars (Froehlich and Krumm 2008) construct datasets that directly respond to a research question at hand.

3. **Perspective Hierarchy.** A common theme amongst human data sets is that, by being sourced from individuals, they contain a hierarchy of perspectives. For example, tweets may be used to study and build for individuals (Gupta et al. 2013) or cities (Quercia et al. 2012b). Similarly, sensors can reflect on both individuals' (Consolvo et al. 2008) or city-wide (Lathia and Capra 2011a) behaviours. Navigating and building from these data sources often means picking one level of this hierarchy for analysis.
4. **Structural or Itemised.** A final means of characterising human-sourced data comes from asking how each source encodes human behaviour. Drawing from the previous example (Hossmann et al. 2012), some systems directly associate behaviour with data; in this case, sensor streams are 'labelled' with the user's current activity. On the other hand, alternative sources provide a means of analysing behaviour via the *relationships* that emerge within the data. For example, web clicks on search results encode an intent for information (Radlinski and Joachims 2007), and mobility traces encode underlying communities and social relationships (Brown et al. 2012).

In the above, we introduced the first step required to build with "human sensors:" using humans to source data about daily life. In the following section, we review how these sources of data are being translated into new insight—where empirical measurement has historically been elusive—and how the web has become the primary example of using this data as a foundation for information services.

Putting Data Streams to Use

As introduced above, the state-of-the-art facilitates collecting data from humans more so than putting that data to use within systems designed to include facets of human computation. Arguably, while data collection methodologies may suffice for scientific enquiry, the value of human data has not been fully reaped until it is part of an ecosystem that supports those behaviours that it was measured from. In this section, we discuss a number of examples where visions of how human data can be integrated into systems have appeared. In particular, we consider scenarios where the human data is not simply used as a body of knowledge that can be accessed (e.g., Wikipedia), but fully enables the existence of new systems. Broadly speaking, we decompose these systems into three categories: those that support *recommendation*, *information retrieval*, and *behaviour change*.

Recommending Items

The idea of giving users personalised, automated recommendations pervades the online world. Hallmark examples of these systems include Amazon.com's product recommendation (Linden et al. 2003) and Netflix's movie recommender

(Amatriain 2012) systems; however, social (Gupta et al. 2013) and search (Das et al. 2007) systems are now also characterised by personalisation algorithms. These systems are grounded in a view of the world where the number of available 'items' (e.g., movies, e-commerce products) far outnumber each user's ability to sort through, evaluate, and find the subset that best matches their preferences (Ricci et al. 2010). These systems therefore tend to use *collaborative filtering*: an algorithmic approach that takes as input a sparse set of ratings or representations of preference, and produces personalised rankings for each user (Sarwar et al. 2001). In doing so, they close the loop between human computation (since humans provide ratings that 'evaluate' items) and machine learning (which predict values for unrated items) to serve behavioural interests (finding new items).

The value of the model underpinning the design of recommender systems is that they are fully agnostic of *what* is being recommended. In principle, this means that *any* scenario that can be described as a set of 'items,' a set of 'users' with preferences, and required a mapping between the two may suit systems that leverage both humans' relevance judgements and machine learning.

A growing set of examples show how this model is being applied outside of the domain of the web. A notable example relates to discovering places and social events in the physical world: for example, cell phone data sourced from a city can be used to infer those social events that people are attending, and provide them with recommendations about others that they may be interested in Quercia et al. (2010). Similarly, GPS data from people's smartphones can be used to find venues of interest (Takeuchi and Sugimoto 2006; Shaw et al. 2013), or even recommend where to go after visiting their current location (Noulas et al. 2012b). Beyond location-discovery, data from human mobility has been applied to, for example, helping taxi drivers find their next fare (Yuan et al. 2011). Similarly, the photos that we take have been shown to uncover the world's interesting locations (Crandall et al. 2009): this data fits into recommender-style applications by, for example, supporting tourists who are navigating an unfamiliar place (Girardin et al. 2008).

Mediating Behaviour

While recommendation systems seek to support those contexts where users are navigating large 'item' repositories, there are many domains beyond the user-item model where behaviours may nonetheless be mediated by the data that they produce. These systems are often referred to as *persuasive* (Fogg 2002) or behaviour-change (Hekler et al. 2013) technologies, which merge the data that can be sensed or collected about human behaviour with behavioural theory about how habits are formed, changed, or maintained. In many instances, the loop between data collection and system design is closed by using the data to provide feedback to a user.

Such systems have already been applied to a host of domains. For example, users' own data can be fed back to them in the context of sustainable travel choices (Froehlich et al. 2009a) and physical activity (Consolvo et al. 2008). In these cases, the data that is collected by directly measuring users' behaviours is then returned via feedback interfaces which may nudge participants' choices (Kalnikaite et al. 2011). We further note that–rather than using the data to provide feedback to users–data about human activities can also be used to mediate the control of systems: for example, room occupancy prediction can be used to automatically tailor the management of household heating systems (Krumm and Brush 2011).

The key design issue in these domains seems to revolve around how to best implement behavioural theory in systems and, in parallel, how systems may augment the potential to extend behavioural theory (Hekler et al. 2013). In these cases, individuals' data plays two roles: first, it provides a source for self-measurement and understanding (Li et al. 2010). Moreover, it supports automating the extent that people can track their commitment and progress towards accomplishing their goals, representing their identities, and uncovering their own inconsistent behaviours (Consolvo et al. 2009).

Monitoring and Retrieving Knowledge

A final application scenario for 'human sensor' data is to use the information that is collected to support retrieval-type contexts which were previously inaccessible to users. The application scenarios here are far-reaching; unlike the above, the model here relies on supporting information needs that can be expressed via a query of some type. For example, participatory sensing indicates that systems that what once relied on time tables to provide public transport information (Ferris et al. 2010) can now use passengers' smartphone sensors to crowd-source bus inter-arrival times (Zhou et al. 2012): the answer to the question "how long do I have to wait at this bus stop?" can be driven by sensor data, rather than time tables.

Beyond individuals' mobility within a city, the data available from movements of crowds has led researchers to advocate for using these sources to guide future urban planning and leadership (Soto and Frias-Martinez 2011; Ratti et al. 2006): city "leaders and service providers are looking to base decisions on data" (Amini et al. 2011). In these cases, crowd-sourced data replaces the kinds of queries (for example, "how do people use these urban spaces?") that were previously answered by arduous field studies. For example, consider the challenge of defining urban neighbourhoods: the data sourced from people's movements now supports dynamically defining and visualising communities (Cranshaw et al. 2012). It is worth noting that, in some cases, similar datasets may be applicable to multiple domains. For example, the previous section mentioned using sensor-enhanced taxis to help drivers find their next fare. Similar placements of sensors on taxis can help local governments track their city's pollution (Yu et al. 2013).

Even the behaviours that accompany retrieval-systems can themselves become informative. For example, search query behaviour related to mental health has been shown to be seasonal (Ayers et al. 2013); similarly, queries about medication reveals the side-effects of combined drug usage (White et al. 2013). Such data, that emerges from people using a search engine, could thus become the foundation for monitoring and retrieval tools that support medical practitioners’ work.

The examples above decompose the usage of human-sourced data into three generic application scenarios: those that support recommendation, mediate (and aide to change) behaviours, and those that create new forms of retrieval systems. All of these application scenarios share a common vision: that is, taking the collection of “human sensor” data beyond mere analysis, and using it as a foundation for building systems. As above, in many of these cases the data that was collected was not originally intended to support the design of such systems. In the following section, we consider a particular use case: translating data that was originally intended to log public transport financial transactions into a basis for personalised transport information systems.

A Case Study: Computation with Smart Cards

In this section, we focus on one particular use case: turning transport smart cards into sources of information for travel services, with a particular focus on how it may be implemented in London, England (Lathia et al. 2012b). Smart cards are increasingly being adopted by public transport authorities across the world: they are typically personal RFID-enabled cards that store passengers’ fares or tickets. In doing so, they facilitate the process of paying for and accessing public transport and remove the need to carry paper-based tickets. However, a consequence of automating the billing process is that detailed records about millions of passengers’ movements and fare purchases throughout a public transport network are created.

For example, the London public transport system uses the Oyster card: a personal contact-less smart card that allows passengers to access all of the city’s multimodal transport systems, which includes underground trains (11 interconnected lines with 270 stations), overground trains (5 lines with 78 stations), and buses (about 8,000 buses service 19,000 stops). The Oyster card itself is used to store fares, which come as both credit/pay-per-journey or travel passes, and is then used to enter and exit train stations and when boarding buses. By 2009, this system accounted for approximately 80 % of all public transport trips in the city (Weinstein 2009).

What do these records tell us about the city and its public transport passengers? Recent work has uncovered that this data contains a hierarchy of information, that ranges from patterns about individuals and communities, to city-wide behaviours: navigating between these levels demonstrates the granularity of analysis that this data enables. At the grandest scale, Oyster data reflects the overriding week-day commuting pattern of the city and shows that the metropolitan area of London, when considered based on passenger flows, has a polycentric structure (Roth et al. 2011). Similarly, analysis of the large-scale features of the data shows how passengers travel choices relate to the financial behavioural incentives that are delineated by the transport authority (Lathia and Capra 2011a); for example, peak-fares seem to guide travellers' fare purchase decisions (rather than their choice to travel), students do not buy fares that they would be eligible to purchase at a discount, and the availability of free travel radically alters peoples' likelihood of taking a bus. By stepping down to the community-level, and using the data about where humans travel between to model how communities interact with one another, the same records show that mobility patterns correlate with social deprivation (Lathia et al. 2012a). Finally, individual patterns of mobility measured via these smart cards uncovers the variance between different individuals' travel choices (Lathia et al. 2010a) and the extent that passengers overspend on public transport by failing to relate their travel behaviours to the fare most suited to them (Lathia and Capra 2011b).

How can this researchers 'close the loop,' and turn this insight into human-data driven systems? We consider two examples, which leverage the latter analytic results: namely, that smart card data provides a means for measuring differences between *individuals'* behaviours. Therefore, a first step into this domain could entail diversifying the output of transport information services, in order to cater for personal differences. Historically, public transport information systems have been centred on the system itself (by providing, for example, the location of a transit service or, scheduled and estimated arrival times of trains or buses) and has not automatically tailored its output to individual travellers. Consider, for example, the Transport for London Travel Alerts[1]; this system requires passengers to manually set their travel choices and times. Replacing this manual input with the automated smart card data would not only alleviate users from this task, but allows for data that can predict travel times more accurately than time tables (Lathia et al. 2012b) and automatically rank the importance of station alerts in such a way to even capture the importance of places that travellers have not historically visited (Lathia et al. 2010a).

Similarly, the individual-level analysis uncovered that passengers often make the incorrect fare choices when making purchases: at a city-wide level, this overspending was estimated to be approximately GBP 200 million per year (Lathia and Capra 2011b). Part of the problem emerges from the difficulty that people have in (a) estimating their own travel needs, and (b) linking their own forecasts with the optimal fare, particularly since the nuances between fares may not be apparent. Smart cards, however, act as an implicit diary for all public transport usage, and reveal that mobility is consistent enough that simple, moving-average based techniques are sufficient for accurately predicting those features of mobility that are relevant to fare purchases. Moreover, supervised learning techniques were then shown to be able to accurately predict between 77 % (Naive Bayes) to 98 % (Decision Trees) of the 'cheapest' fares that passengers need.

Both of these results demonstrate how data that has been sourced from human behaviour can be used and contribute back towards guiding it. By borrowing techniques that have been widely applied in the online world (personalisation and recommender systems) and applying them to pervasive data (from smart cards), this

[1] http://alerts.tfl.gov.uk/

research demonstrates how many of the tools, techniques, and data sources to build future systems are already available today: the biggest challenge being how to build an appropriate bridge between them. In the following section, we consider a range of open problems that are hidden within this cycle of system building, and discuss how research is required to address them.

Looking Forward: Three Research Challenges

In the previous sections, we broadly characterised the what (data sources) and how (system models) of building systems with humans in the loop. As we move forward, it is likely that the applicability of these techniques will pervade many more facets of daily life; to date, many interesting datasets continue to remain behind closed doors. However, many open challenges remain; these range from technical to ethical challenges associated with computation with humans. In this section, we consider those open challenges that directly relate to building systems: a discussion of the broader issues of privacy and informed consent, while relevant, goes beyond the points that we enumerate here. Instead, we focus on three questions: (a) is the data itself valid? (b) does tailoring a system's design affect its inferences? and (c) how can the loop between data and user be closed most appropriately?

A prominent issue that arises when building the kind of (recommendation, retrieval, behaviour-related) systems discussed above is that understanding what is *not* represented in the data is often overlooked (boyd and Crawford 2011). In this case, it is worth differentiating between data *sparsity*, which undermines the predictive power of machine learning algorithms, and data *representativity*, which is more about considering the inherent bias in the collection of *any* human dataset. Participation in publicly deployed applications, whether online or offline and with or without sensors, will be limited to self-selecting users; moreover, the typical distribution of participation between those who do tends to be highly non-uniform. Naturally, data derived from social media is 'skewed' by the extent that its users are a fair sample of the general population. While it is arguable that analysing data from millions trumps the same analysis on dozens, this newfound scale is not, in itself, a problem, but understanding the demographics of these sources is Mislove et al. (2011). This issue gains importance if we consider that the way people use systems often breaks from our assumptions: notable examples include that online accounts (which, historically, researchers have assumed to belong to individuals) are shared between household members (Bellogin et al. 2012), and twitter data (which researchers assume is sourced from humans) is actually rife with automated bots (Chu et al. 2010).

From a system perspective, the interplay between the quantity and granularity of data collection continues to stand at odds with the obtrusiveness and energy consumption required on participants and their devices. For example, smartphone sensing applications will always need to trade off between sensor sampling rates and the battery usage; while recent work (Rachuri et al. 2013) shows how off-loading to nearby sensors may alleviate this problem, the question of how this design should vary between different contexts remains open. To what extent does optimising for battery life or, more generally, technical-related facets of a system (storage, connectivity, etc.) influence a system's ability to make reasonable inferences about a person's behaviour? Moreover, one of the most challenging tasks of system design is that of translating a high-level requirement (e.g., "I would like to give my users recommendations") into a task that can be suitably addressed by algorithms (i.e., "predicted preference ratings can be used to rank unrated content"). This process critically defines how systems reason about the data at hand and, more broadly, begins to put boundaries on the system's behaviour. In continuing with the recommendation example: defining an algorithmic approach to recommendations as one of generating a static set of predictions from user rating data will, for example, explicitly ignore the system's temporal behaviour (Lathia et al. 2010b).

Finally, the bridge itself between the data and systems continues to pose open research questions. Architecting a system to feed back any kind of information from human-sourced data requires researchers to continuously revisit the concept of relevance (Saracevic 1975), not only in order to effectively close the loop, but also as a means of evaluating whether systems are worthwhile or achieving their goals (Hekler et al. 2013). This becomes particularly challenging in the context of the kind of potential systems described above: many of the methodologies for most appropriately evaluating the quality of the system (e.g., recommendation quality, whether/to what extent behaviour has been changed) remain elusive and the subject of active research.

Conclusion

In this chapter, we have provided a broad overview of how the *human sensor* may fit into the design of future systems. The main question that we discussed was, given the growing diversity and availability of data that encodes human behaviour: how can this data become an integral part of systems that support our every day life?

To do so, this chapter discussed the data itself, by considering the similarities in the different techniques that have historically been used to source it. Most notably, whether crowd-sourcing, participatory-sensing, or database mining techniques are adopted, the result is a representation of human behaviour which can be characterised in terms of its collection obtrusiveness (manual or automated), its original purpose (e.g., social vs. sensor), and its structure (both implicit hierarchy and whether behaviour is directly encoded or emerges from relationships in the data). We further broadly characterised three kinds of systems (recommendation, retrieval, behaviour-mediating) that can leverage these sources, and touched on three issues that remain unsolved within the context of system design. By attempting to thread a high-level narrative that demonstrates how human data can be put to use, this chapter naturally did not delve into the details of recommender, retrieval, and persuasive systems, although many suitable resources exist for further reading (Baeza-Yates and

Ribeiro-Neto 1999; Ricci et al. 2010; Fogg 2002). Finally, we discussed how these considerations motivate to future research: as future systems are designed, researchers and practitioners will need to tackle open questions about the input data itself, the potentially biasing role a system plays once high-level requirements are translated to algorithmic solutions, and how to most appropriately draw the link back to the humans that each system is designed for.

References

Amatriain X (2012) Mining large streams of user data for personalized recommendations. SIGKDD Explorations Newsletter 14(2):37–48

Amini L, Bouillet E, Calabrese F, Gasparini L, Verscheure O (2011) Challenges and Results in City-Scale Sensing. In: IEEE Sensors 2011, Limerick Ireland

Ayers J, Althouse B, Allem J, Rosenquist J, Ford D (2013) Seasonality in seeking mental health information on google. Am J Prev Med 44(5):520–525

Baeza-Yates R, Ribeiro-Neto B (1999) Modern information retrieval. Addison Wesley, Harlow

Bellogin A, Diez F, Cantador I (2012) Time feature selection for identifying active household members. In: ACM CIKM, Maui, Hawaii

boyd d, Crawford K (2011) Six provocations for big data. In: A decade in internet time: symposium on the dynamics of the internet and society, Oxford

Brown C, Nicosia V, Scellato S, Noulas A, Mascolo C (2012) Where online friends meet: social communities in location-based networks. In: ICWSM, Dublin

Bryan H, Blythe P (2007) Understanding behaviour through smartcard data analysis. Proc Inst Civ Eng Transp 160(4):173–178

Burke J, Estrin D, Hansen M, Parker A, Ramanathan N, Reddy S, Srivastava M (2006) Participatory sensing. In: Workshop on world-sensor-web: mobile device centric sensory networks and applications, Boulder

Chu Z, Gianvecchio S, Wang H, Jajodia S (2010) Who is tweeting on twitter: human, bot, or cyborg? In: Proceedings of the 26th annual computer security applications conference, New Orleans

Consolvo S, McDonald D, Toscos T, Chen M, Froehlich J, Harrison B, Klasnja P, LaMarca A, LeGrand L, Libby R, Smith I, Landay J (2008) Activity sensing in the wild: a field trial of UbiFit garden. In: ACM CHI, Florence

Consolvo S, McDonald D, Landay J (2009) Theory-driven design strategies for technologies that support behavior change in everyday life. In: ACM CHI, Boston

Crandall D, Backstrom L, Huttenlocher D, Kleinberg J (2009) Mapping the world’s photos. In: WWW’09: proceeding of the 18th international conference on world wide web, Madrid

Cranshaw J, Schwartz R, Hong J, Sade N (2012) The livehoods project: utilizing social media to understand the dynamics of a city. In: ICWSM, Dublin

Das A, Datar M, Garg A, Rajaram S (2007) Google news personalization: scalable online collaborative filtering. In: WWW, Alberta

Doan A, Ramakrishnan R, Halevy A (2011) Crowdsourcing systems on the world wide web. Commun ACM 54:86–96

Eagle N, Pentland A (2006) Reality mining: sensing complex social systems. Pers Ubiquitous Comput 10:255–268

Ferris B, Watkins K, Borning A (2010) OneBusAway: results from providing real-time arrival information for public transit. In: Proceedings of CHI, Atlanta

Fogg BJ (2002) Persuasive technology: using computers to change what we think and do. Ubiquity 2002:2

Froehlich J, Krumm J (2008) Route prediction from trip observations. In: Intelligent vehicle initiative, SAE world congress, Detroit

Froehlich J, Consolvo S, Dillahunt T, Harrison B, Klasnja P, Mankoff J, Landay J (2009a) UbiGreen: investigating a mobile tool for tracking and supporting green transportation habits. In: ACM CHI, Boston

Froehlich J, Neumann J, Oliver N (2009b) Sensing and predicting the pulse of the city through shared bicycling. In: 21st international joint conference on artificial intelligence, Pasadena

Girardin F, Calabrese F, Fiore FD, Ratti C, Blat J (2008) Digital footprinting: uncovering tourists with user-generated content. IEEE Pervasive Comput 7:36–43

Gupta P, Goel A, Lin J, Sharma A, Wang D, Zadeh R (2013) WTF: the who to follow service at twitter. In: WWW, Rio de Janeiro

Hekler E, Klasnja P, Froehlich J, Buman M (2013) Mind the theoretical gap: interpreting, using, and developing behavioral theory in hci research. In: ACM CHI, Paris

Hossmann T, Efstratiou C, Mascolo C (2012) Collecting big datasets of human activity one checkin at a time. In: Workshop on hot topics in planet-scale measurement, Lake District

Intille S, Rondoni J, Kukla C, Ancona I, Bao L (2003) Context-aware experience sampling. In: ACM CHI extended abstracts, Ft. Lauderdale

Kalnikaite V, Rogers Y, Bird J, Villar N, Bachour K, Payne S, Todd PM, Schoning J, Kruger A, Kreitmayer S (2011) How to nudge in situ: designing lambent devices to deliver salience information in supermarkets. In: ACM ubicomp, Beijing

Krumm J, Brush A (2011) Learning time-based presence probabilities. In: Pervasive, San Francisco

Lane N, Eisenman S, Musolesi M, Miluzzo E, Campbell A (2008) Urban sensing systems: opportunistic or participatory? In: Workshop on mobile computing systems and applications (HotMobile), New York

Lathia N, Capra L (2011a) How smart is your smartcard? measuring travel behaviours, perceptions, and incentives. In: ACM international conference on ubiquitous computing, Beijing

Lathia N, Capra L (2011b) Mining mobility data to minimise travellers' spending on public transport. In: ACM SIGKDD 2011 conference on knowledge discovery and data mining, San Diego

Lathia N, Froehlich J, Capra L (2010a) Mining public transport usage for personalised intelligent transport systems. In: IEEE international conference on data mining, Sydney

Lathia N, Hailes S, Capra L, Amatriain X (2010b) Temporal diversity in recommender systems. In: ACM SIGIR, Geneva

Lathia N, Quercia D, Crowcroft J (2012a) The hidden image of the city: sensing community well-being from urban mobility. In: Pervasive, Newcastle

Lathia N, Smith C, Froehlich J, Capra L (2012b). Individuals Among Commuters: Builder Personalised Transport Information Services from Fare Collection Systems. Elsevier Pervasive and Mobile Computing: Special Issue on Pervasive Urban Applications. 9(5):643–664.

Li I, Forlizzi J, Dey A (2010) Know thyself: monitoring and reflecting on facets of one's life. In: ACM CHI workshops, Atlanta

Linden G, Smith B, York J (2003) Amazon.com recommendations: item-to-item collaborative filtering. IEE Internet Comput 7:76–80

Mackerron G (2012) Happiness and environmental quality. PhD thesis, The london school of economics and political science

Mislove A, Lehmann S, Ahn Y, Onnela J, Rosenquist J (2011) Understanding the demographics of twitter users. In: AAAI ICWSM, Barcelona

Noulas A, Scellato S, Lambiotte R, Pontil M, Mascolo C (2012a) A tale of many cities: universal patterns in human mobility modelling. PLoS ONE 7(5):e37027

Noulas A, Scellato S, Lathia N, Mascolo C (2012b) Mining user mobility features for next place prediction in location-based services. In: IEEE ICDM, Brussels

Quercia D, Lathia N, Calabrese F, Lorenzo GD, Crowcroft J (2010) Recommending social events from mobile phone location data. In: IEEE ICDM, Sydney

Quercia D, Capra L, Crowcroft J (2012a) The social world of twitter: topics, geography, and emotions. In: ICWSM, Dublin

Quercia D, Ellis J, Capra L, Crowcroft J (2012b) Tracking gross community happiness from tweets. In: AAAI CSCW, Seattle

Ra M, Liu B, Porta TL, Govindan R (2012) Medusa: a programming framework for crowd-sensing applications. In: ACM MobiSys, Lake District
Rachuri K, Efstratiou C, Leontiadis I, Mascolo C, Rentfrow P (2013) METIS: exploring mobile phone sensing offloading for efficiently supporting social sensing applications. In: IEEE PerCom, San Diego
Radlinski F, Joachims T (2007) Active exploration for learning rankings from clickthrough data. In: In proceedings of KDD, San Jose
Ratti C, Pulselli RM, Williams S, Frenchman D (2006) Mobile landscapes: using location data from cell phones for urban analysis. Environ Plan B Plan Des 33(5):727–748
Ratti C, Sobolevsky S, Calabrese F, Andris C, Reades J, et al. (2010) Redrawing the Map of Great Britain from a Network of Human Interactions. PLoS ONE 5(12): e14248.
Reddy S, Mun M, Burke J, Estrin D, Hansen M, Srivastava M (2010) Using mobile phones to determine transportation modes. ACM Trans Sens Netw 6(13)
Ricci F, Rokach L, Shapira B, Kantor P (eds) (2010) Recommender system handbook. Springer
Roth C, Kang S, Batty M, Barthelemy M (2011) Structure of urban movements: polycentric activity and entangled hierarchical flows. PLOS ONE 6:e15923
Saracevic T (1975) Relevance: a review of and a framework for the thinking on the notion in information science. J Am Soc Inf Sci 26(6):321–343
Sarwar B, Karypis G, Konstan J, Reidl J (2001) Item-based collaborative filtering recommendation algorithms. In: WWW, Hong Kong
Shaw B, Shea J, Sinha S, Hogue A (2013) Learning to rank for spatiotemporal search. In: Web search and data mining (WSDM), Rome
Soto V, Frias-Martinez E (2011) Robust land use characterization of urban landscapes using cell phone data. In: Workshop on pervasive and urban applications, San Francisco
Takeuchi Y, Sugimoto M (2006) An outdoor recommendation system based on user location history. In: ACM ubicomp, Orange, California, USA
Weinstein L (2009) TfL's Contactless ticketing: oyster and beyond. In: Transport for London, London
White R, Tatonetti N, Shah N, Altman R, Horvitz E (2013) Web-scale pharmacovigilance: listening to signals from the crowd. J Am Med Inform Assoc 1 May 2013 vol. 20 no. 3 404–408
Yan T, Marzilli M, Holmes R, Ganesan D, Corner M (2009) mCrowd: a platform for mobile crowdsourcing. In: ACM conference on embedded networked sensor systems (SenSys), Berkeley
Yu X, Fu Q, Zhang L, Zhang W, Li V, Guibas L (2013) CabSense: creating high-resolution urban pollution maps with taxi fleets. In: ACM MobiSys, Taipei
Yuan J, Zheng Y, Zhang L, Xie X, Sun G (2011) Where to find my next passenger? In: ACM Ubicomp, Beijing
Zhou P, Zheng Y, Li M (2012) How long to wait? predicting bus arrival time with mobile phone based participatory sensing. In: ACM MobiSys, Lake District

Part V
Algorithms

Algorithms: Introduction

Remco Chang and Caroline Ziemkiewicz

The idea of treating humans as computational units has challenged and redefined our understanding of what computing is. Since Luis von Ahn introduced CAPTCHA a decade ago,[1] a fast-rising number of crowdsourcing games have used human computation to solve a wide range of problems. The rapid development of crowdsourcing games has outpaced our understanding of the theory and algorithms that are common to them. Indeed, in a 2008 article in Communications of the ACM, Jeannette Wing notes that one of the five unsolved problems in Computer Science is to define computing when it can be performed by both humans and machines.[2]

The purpose of this section is to explore this unsolved problem. The authors represented here examine systematic and general ways to treat humans as computational units. These approaches necessarily take a multidisciplinary approach and consider cognitive and societal factors as well as the mathematical foundations of human and machine computation. With a better understanding of which computational jobs are best suited to humans and which are best suited to machines, solutions can begin to optimize the combination of human and machine computation.

The six chapters in this section represent a diverse set of perspectives and approaches. The first half of the section examines the background of human computation algorithms from three distinct viewpoints: philosophical and societal implications, computational complexity theory, and a survey of existing crowdsourcing methods.

The chapter by Aidan Lyon and Eric Pacuit examines three methods to aggregate human judgment from a historical and philosophical perspective. Using an example

[1] Von Ahn, L., Blum, M., Hopper, N. J., & Langford, J. 2003. CAPTCHA: Using hard AI problems for security. Advances in Cryptology – EUROCRYPT 2003. 294–311.

[2] Wing, J. 2008. Five deep questions in computing. Communications of the ACM 51(1):58–60.

R. Chang (✉)
Tufts University, Somerville, MA, USA
e-mail: remco@cs.tufts.edu

C. Ziemkiewicz
Aptima, USA

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_46, © Springer Science+Business Media New York 2013

dating back to the days of Aristotle, the authors show that through simple methods of aggregation, groups of humans have the ability to make predictions with uncanny accuracy.

Jordan Crouser, Alvitta Ottley, and Remco Chang propose that human computation can be seen as an extension of human-computer collaboration. Based on Fitts' 1951 HABA-MABA (Humans are better at/Machines are better at) list, the authors give examples of how the distribution of work between humans and computers can be quantified and optimized.

The chapter by Peng Dai provides a comprehensive survey of crowdsourcing algorithms by organizing them into six different workflow models. In these workflows, human and machine computations have defined inputs and outputs that feed into each other to solve complex problems. This categorization of crowdsourcing algorithms as workflows provides insight in how to engineer these algorithms.

While the above three chapters provide perspectives on human computation in general, the following chapters examine the role of humans in existing computational algorithms. The first approach sees humans as agents in distributed intelligent agent algorithms, the second as elements of evolutionary computing, and the last as input for social recommendation engines.

The chapter by Edmund Durfee compares the similarities and differences between Distributed Artificial Intelligence (DAI) algorithms for intelligent artificial agents and cooperation between intelligent human agents in real life. The findings suggest that there are rich and promising opportunities to use concepts and methods from the DAI community to develop efficient human computation systems.

The chapter by Jeffrey Nickerson examines evolutionary algorithms in the context of human computation. In some evolutionary algorithms, human input can be used to affect selection, modification, and evaluation. Integrating humans into this typically expensive computational process can reduce overall computation costs, because humans provide domain knowledge that is crucial to finding an optimal solution quickly, but is typically inaccessible to machines.

The final chapter by Ido Guy presents algorithms for leveraging social content (for example, data found on social network websites) to provide social recommendations. In particular, the author provides algorithms for analyzing information from social websites to recommend media content and friends. This chapter summarizes years of research by the author in this domain and demonstrates that data provided by humans contain a wealth of information. With the appropriate algorithms, these data can be used to greatly enhance our understanding of social networks and relationships.

The Algorithms Section is not exhaustive, but it does include a range of ways to think about the problem of human-machine computation. In fact, this section presents six unique approaches to integrate human input in existing computational algorithms and to develop new algorithms that distribute work between humans and machines. While they represent novel and creative perspectives to human computation, all authors note that the field of human computation is still in its infancy, and there are many open problems that need to be addressed. Together, these chapters represent a step towards answering Wing's question of how human computation changes our understanding of what is computable.

The Wisdom of Crowds: Methods of Human Judgement Aggregation

Aidan Lyon and Eric Pacuit

Introduction

The Wisdom of Crowds is all the rage in these heady Web 2.0 days. But the idea is an old one, and one that goes back to the philosophers of antiquity:

> For the many, of whom each individual is but an ordinary person, when they meet together may very likely be better than the few good, if regarded not individually but collectively, just as a feast to which many contribute is better than a dinner provided out of a single purse.
>
> —Aristotle, *Politics*, Book III, §XI.

The basic idea is also a simple and familiar one: two heads are often better than one, and more are even better. A classic example comes from a contest of some 800 people at a county fair in Plymouth, 1906. The contest was to guess the weight of an ox, slaughtered and dressed. Francis Galton found that the *average* of the crowd's guesses was within 1% of the true weight of the ox, despite huge errors in most of the individual guesses (see Galton 1907a, b). Somehow, the crowd knew more as a collective than many of its individuals.

Although the idea is an old one, there has been a recent boom in research into the Wisdom of Crowds, and this appears to be at least partly due to the now widespread availability of the Internet, and the advent of social media and Web 2.0 applications. Never before has it been so easy to get a crowd and leverage their collective wisdom

A. Lyon (✉)
Philosophy, University of Maryland, College Park, MD, USA

Munich Centre for Mathematical Philosophy, Ludwig Maximilian University of Munich, Germany
e-mail: alyon@umd.edu

E. Pacuit
Philosophy, University of Maryland, College Park, MD, USA
e-mail: epacuit@umd.edu

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_47,

for some task. There are now many well-documented and contemporary examples of the so-called Wisdom of Crowds[1]:

- Amazon's product recommendations.
- Wikipedia and Intellipedia.
- Netflix's movie recommendation algorithm.
- Prediction markets.
- Online citizen science.
- Google's PageRank algorithm.

Discussions of these examples (and many others) can be found in Surowiecki (2005), Page (2008), Nielsen (2011), and Landemore and Elster (2012). This paper will provide an overview of some of the theory behind all of the examples. By thinking carefully about what they have in common and how they differ from each other, we can find new ways to make these applications better. Sometimes such research will simply result in better movie recommendation services, but sometimes it will have much more serious consequences. For example, there are now many Web 2.0 tools being designed to help track and predict the outbreaks of emerging infectious diseases (cf. Collier et al. 2006; Brownstein et al. 2009; Keller et al. 2009; Lyon et al. 2012b, 2013) and even to diagnose rare diseases (e.g., Nuwer 2013). By developing a better understanding of the Wisdom of Crowds, we should be able to improve upon such tools, and thereby make better forecasts of disease outbreaks (among other things).

To begin, in section "Thinking About the Wisdom of Crowds", we'll lay out a simple conceptual framework for thinking about the Wisdom of Crowds. We'll identify six core aspects that are part of any instance of the Wisdom of Crowds. One of these aspects is called *aggregation*, and this will be our primary focus for the remainder of the paper. An aggregation method is the method of bringing the many contributions of a crowd together into a collective output. In the example of the crowd at Plymouth guessing the ox's weight, the aggregation method was the *averaging* of the crowd's individual guesses. This, however, is not the only method of aggregation available. In the next three sections, we'll discuss three broad kinds of aggregation methods: mathematical aggregation, group deliberation, and prediction markets.

Thinking About the Wisdom of Crowds

A good way to start thinking systematically about the Wisdom of Crowds is to think about what you would do if you had a burning desire to use the Wisdom of Crowds to do something—because, say, it just seems like a fun thing to do.

[1] We say "so-called", because examples of the Wisdom of Crowds often have little to do with the notion of wisdom that philosophers care about (see Andler (2012) for further discussion), and they often involve only a group of people—even just a handful—and not a *crowd* in the usual sense of the word. Nevertheless, we will stick with the words that seem to have stuck.

The very first thing you have to do is decide what you want to achieve. Do you want to predict the outcome of an election? Recommend products to customers? Decide if someone is guilty of a crime? Write an academic paper? Solve a murder mystery? Predict disease outbreaks? Whatever it is you want to do, we will call this the desired *output* of your endeavour; it's what you want to get out of the Wisdom of Crowds. As we'll soon see, it's important to be clear about your desired output, because this can have a big impact on how you use your crowd.

Speaking of which: you need to get yourself a crowd. Perhaps you have one already, because you have some willing friends. Or perhaps you don't have any friends, but you have some cash to rent a crowd. Or perhaps you see a free crowd—e.g., there could be people on Twitter regularly tweeting information that you could use. We'll call the process of getting a crowd *recruitment*—even if you don't recruit anyone in the usual sense of the word. This recruitment process is very important, for there are many things to consider. Does your crowd need to consist of experts on some topic? Or can they just all be regular folk? How large does your crowd have to be?[2] Does your crowd have to be diverse? Will members of the crowd talk to each other? And so on. These are all important and complicated issues to deal with, and we'll put them aside for now; we simply flag them here because they are important.

The next thing to do is decide how your crowd will contribute to your output. For example, if you want to determine someone's guilt or innocence, perhaps your crowd can contribute by giving their own judgements of guilt or innocence. You might then judge the person to be guilty if and only if everyone in your crowd judges the person to be guilty. However, maybe you need to be more nuanced: instead of your crowd giving outright "guilty" or "innocent" verdicts, perhaps you want to know how *confident* they are in their verdicts. If everyone judges the person to be guilty, but they are only 70 % confident in their judgements, then you might be reluctant to, say, send the person to death row. When you've decided whether you want outright judgements or probabilities—or something else—we'll say that you've decided your *inputs*; you've decided what input the members of your crowd are going to have in your endeavour to achieve your desired output. Note that everyone needn't give the same kind of input. For example, you may want one half of your group to give product reviews, and the other half to rate the qualities of those reviews. Also note that the inputs needn't be of the same kind as your desired output—e.g., there are ways to turn probabilities (inputs) into an outright judgment (output); and there ways to turn outright judgements (inputs) into a final probability (output). We've mentioned a few kinds of inputs, but there are many others. To name just a few, inputs could be: votes, preferences, sentences, arguments, probability distributions, lines of computer code, quality ratings, translations, relevance rankings, or text transcriptions through services like reCAPTCHA (Von Ahn et al. 2008).

[2] Psychologists have found that even just single person can function as a crowd of individuals (see e.g., Vul and Pashler 2008; Herzog and Hertwig 2009; Hourihan and Benjamin 2010).

Once you've decided what kind of inputs you want to get out of your crowd, you have to work out how to *get them out* of your crowd. This is really important for all sorts of reasons. For example, some members of your crowd may have an incentive to lie to you (perhaps the person on trial seduced the wife of someone in your crowd). Or maybe their contributions are valuable to them, and so you need to pay for their contributions in some way. Maybe some of your members are shy while others are overbearing, and so you may need to make sure everyone has equal opportunity to make their contribution. We call this process of getting the inputs out of your crowd *elicitation*. Your method of elicitation can be crucial for getting the most out of your crowd. For example, psychologists have shown that how you ask for probability assignments from people can have a dramatic effect on how overconfident they are (see e.g., Klayman et al. 1999).

Let's say you've decided what you want to do (output), got yourself a crowd (recruitment), worked out how they will contribute to your endeavour (inputs), and how you will get those contributions (elicitation). The next step is called *aggregation*: you need to convert the contributions of your crowd into your desired output. We've touched on an aggregation method already: judge the person on trial to be guilty if and only if everyone in your crowd judges them guilty. Another aggregation method is: judge them guilty if and only if a *majority* of your crowd judges them guilty. Yet another: judge the person guilty if and only if the average probability assigned to the guilty verdict by your crowd is above 90 %. As you can probably tell, there are a lot of aggregation methods to choose from, and different aggregation methods will have different properties. Much of the rest of this paper is devoted to the topic of aggregation, so we'll leave further discussion of these matters to later sections.

There is one final aspect, and we call it *evaluation*, and it is how you evaluate the output of your endeavor. Sometimes evaluation will be straightforward. For example, if your crowd judged the person to be guilty, and they are in fact guilty, then your crowd got it *right*, and maybe that's all you care about. But you might also be concerned that your crowd will mistakenly judge an innocent person to be guilty, and that being wrong in this way (a false positive) is much worse than judging a guilty person to be innocent (a false negative). If so, you may have to decide how to balance these different kinds of error against each other. There are plenty of other standards of evaluation. If you're guessing the weight of an ox, you might want to *minimise the error* of your crowd's judgement. If you're forecasting the weather, you might want your announced "chances of rain" to be *well calibrated*.[3] If your crowd is writing encyclopedia articles, you might want the articles to have *few grammatical errors*, or to have few *factual inaccuracies*, or to have a *unified style*—or, probably, some combination of all of these virtues. How you choose to evaluate the output will have a big impact on your choices regarding the other five aspects we've identified. For example, some aggregation methods can be good at producing a collective judgement with the appropriate level of confidence but not very good at producing accurate judgements (cf., Lyon et al. 2012a).

[3] For example, it should rain on 90 % of the days that your crowd says there is a 90 % chance of rain.

To summarise, we've identified six core aspects to the Wisdom of Crowds:

1. The Output
2. The Recruitment
3. The Inputs
4. The Elicitation Method
5. The Aggregation Method
6. The Standard of Evaluation

There are two important qualifications that we now need to make. The first is that we presented these components as steps in a chronological process: decide what you want to, get your crowd, decide on your inputs, work out how to elicit them, work out how to aggregate them, and then work out how to evaluate them. However, it should be clear by now that there is no set chronological order to these aspects. For example, perhaps your most important criterion is that the output is a *fair* one. If so, this will put heavy constraints on the how you settle the other issues—e.g., your aggregation method may have to give equal weight to everyone's input, rather than unequal weight (cf. section "Mathematical Aggregation"). So instead of thinking of the aspects as steps in a chronological process, they should thought of as components of a reflective equilibrium.

The second important qualification is that these aspects can overlap with each other and that their borders are blurry. In fact, two of the main kinds of aggregation methods discussed in this paper—discussion groups and prediction markets—can also be thought of as elicitation methods. For example, a prediction market works by getting people to place bets with each other on whether some event will occur or not—e.g., whether Hilary Clinton will win the 2016 US Presidential Election. The "market price" of a prediction market is an aggregate of all of the individual bets, and, when interpreted as the probability of the event in question happening, can be highly effective in forecasting whether the event will happen. However, the market price is determined by the individual bets being made, and those bets can be used to infer the people's individual subjective probabilities of the event happening. So the prediction market both elicits and aggregates the human judgement inputs. Although the above six aspects overlap with each other, we believe they nevertheless provide a convenient conceptual framework for thinking about the Wisdom of Crowds.

All of the aspects are extremely important, but due to limitations on space, in this paper we will restrict our focus to the aggregation aspect of the Wisdom of Crowds. In fact, we will need to restrict our focus even further: we'll limit our discussion to aggregation methods that take only simple kinds of human judgements as input: votes, estimates, probabilities, etc., and these will always be *epistemic* judgements—that is, we won't the discuss the aggregation of inputs such as preferences, judgements of fairness, etc. And we won't discuss methods for aggregating more complex kinds of inputs, such as sentences to wikipedia articles, product reviews, text translations, contributions to legislation, etc.

The aggregation methods that we will focus on fall roughly into three broad categories: Mathematical Aggregation (section "Mathematical Aggregation"), Deliberation Methods (section "Deliberation Groups"), and Prediction Markets (section "Prediction Markets").

Mathematical Aggregation

Perhaps the most common aggregation method is averaging, specifically, *unweighted linear averaging*. Suppose there are N people in your crowd, and we number each individual, $i = 1, 2, 3, \ldots, N$. Let j_i be the elicited judgement of person i (e.g., the number of jelly beans in a jar). The unweighted linear average of your crowd's judgements is defined as:

$$\text{Unweighted Linear Average} = \frac{1}{N}\sum_{i=1}^{N} j_i$$

This simple method of averaging is considered by many as a standard benchmark of aggregation. For example, Armstrong (2001b) recommends it as a good default option, especially if you don't know anything about the abilities of the individuals in the group. If you do have such information, you may want to use some kind of weighted average (see below).

Averaging has its drawbacks. It can make sense when the individual judgements are clustering around a central value, but it can have undesirable consequences when the distribution of judgements takes on another shape. For example, consider the following hypothetical estimates of the effect that Obama's economic policies will have on US GDP. Average growth in GDP for the next decade will be:

(i) $-0.1, 0.1, 0.2, -0.3, 0.1, 0.3, -0.3, 0.2, -0.1, -0.1\%$
(ii) $-19.1, 5.1, 5.2, 4.7, -20.5, 5.4, 4.7, 4.6, 4.8, 5.1\%$

In both cases, the average of the estimates is 0 %, but the distributions of the guesses differ in an important way. The first set of estimates cluster around 0 %, but the second tend to cluster more around 5 % than they do around 0 %. The only reason why the average of the second set of estimates is 0 % is because of the two extremely negative estimates. In the first case, the individuals could agree to 0 % as the collective judgement as a compromise—perhaps because 0 % is so close to each individual estimate. However, in the second set, no one believes that the effect will be about 0 %, so to take 0 % as the collective judgement seems like a rather odd thing to do. For this sort of reason, a better strategy may be to take the *mode* of the estimates. In this way, the mode can be a more democratic aggregation method than the average. The mode is just one statistical property of the distribution of guesses we could use as an alternative to the mean. Other options include the median, the mean with outliers removed, the geometric mean, the maximum entropy expectation, and so on. In short, any of the tools of statistics can be used to construct a more sophisticated aggregation method.

Another way to move beyond simple averaging is to use a *weighted* average. A weighted average gives more weight to some of the estimates over than others. Using the same notation as before, but where w_i is the weight given to judgement j_i, the weighted linear average of the crowd's judgements is defined as:

$$\text{Weighted Linear Average} = \frac{1}{N}\sum_{i=1}^{N} w_i j_i$$

Using a weighted average can make sense when, say, you know some members of your crowd are more reliable than others. For example, if you know from past experience that Ann is twice as good at guessing the number of jelly beans in a jar than Bob is, then you might want to take the average of their guesses, but give twice as much weight to Ann's guess than to Bob's. This is a variant of a method known as *Cooke's method* (*cf.* Cooke 1991). The core idea is that you should use the past performance of the members of your crowd to determine how much weight you should give to their current judgements (see Clemen (2008) for a study of the method's performance). This is not the only way to use a weighted average. It may make sense to weight the judgements by how confident the individuals are in their judgements. If someone is not very confident in their judgement, then perhaps their judgement shouldn't contribute much to the collective judgement. Various results in psychology suggest that, at least in some cases, confidence in a judgement correlates with the accuracy of that judgement (e.g., Koriat 2012).

A more complicated way to take a weighted average is to elicit degrees of *peer respect* along with the judgements (thus making the inputs slightly more complex). Suppose you find yourself in a group of people who all give judgements about some issue, but you think some members of the group are experts on the issue at hand and others are not. You would probably be unhappy with any collective judgement that gave equal weight to everyone—you'd prefer a collective judgement that gave more weight to the experts than to the fools. Similarly, everyone else will feel the same way—although they may have different opinions as to who are the experts. For any individual k, if they respect each person i to degree $w_{k\,i}$, it looks like they should average as follows:

$$\text{Respect Weighted Average} = \frac{1}{N}\sum_{i=1}^{N} w_{ki}\, j_i$$

(where the $w_{k\,i}$ are all between 0 and 1, and for each fixed i, the $w_{k\,i}$ sum to 1; so there is no need for a normalisation term). This aggregation method will produce a new judgement j_i' for each person i. Lehrer and Wagner (1981) argue that there is nothing special about these new judgements, and so if they vary, then everyone should now average again, using the new judgements and original weights of respect. Lehrer and Wagner prove that if everyone continues to average in this way, they will reach a group *consensus*: all of the averaged judgements will approach a unique consensus judgement j_c. Lehrer and Wagner argue that this consensus judgement has a number of virtues—both pragmatic and epistemic. One potential drawback to this method of aggregation, however, is that people's judgements of each other's level of expertise do not track the accuracies of their judgements. Burgman et al. (2011) found that such ratings of expertise were poor guides to judgement accuracy. There can also be practical difficulties in getting people to rate each other's expertise—especially if those ratings are to be made public (*cf.* Regan et al. 2006). For an extensive discussion of the Lehrer–Wagner consensus model, see Loewer and Laddaga (1985) and the other papers in the same special issue of *Synthese*.

So far, we have only discussed examples where the inputs and outputs are quantity estimates and so can be represented with real numbers. If the inputs are not like this, we have to choose a different kind of aggregation method. Another common sort of judgement are outright judgements of the form "guilty"/"innocent", "yea"/"nay", "black"/"white", and so on. Such judgements cannot be averaged, but there are, nonetheless, ways to aggregate them. Perhaps the most natural and common is what is known as the *majority rule*: the collective judgement is "guilty" ("innocent") if and only if more than 50 % of the individual judgements are "guilty" ("innocent"). A famous theorem, known as the Condorcet 1785 jury theorem (rediscovered by Black (1963)), shows that as you add more and more people to the crowd and aggregate their judgements using the majority rule, then if each person has a greater than 50 % chance of being right, and if they make their judgements independently of one another, then the probability that the collective judgement is correct will approach certainty. The theorem requires that the people in the crowd make their judgements independently of each other, which is a somewhat implausible of real life situations. However, Ladha (1992) generalised the theorem to allow for there to be some dependencies between the crowd's judgements. And there have now been a number of other generalisations of the theorem to make its application to real life situations more plausible. List and Goodin (2002) generalised the theorem to cover other aggregation methods and Grofman et al. (1983) generalised the theorem to allow for people who don't have a greater than 50 % chance of judging correctly.

Things get tricky if the inputs and outputs are more complex than single all-or-nothing judgements. Suppose that instead of simply judging whether someone is guilty, we want our crowd to provide some reasoning for this judgement. For example, suppose that G means the person is *guilty*, N means they were *nearby* when the crime was committed, and $N \to G$ means that *if they were nearby, then they are guilty*. Now suppose we have 30 people in our crowd, and they make the following judgements on the three propositions, N, $N \to G$, and G:

	N	$N \to G$	G
10 people say	True	True	True
Another 10 people say	True	False	False
The remaining 10 people say	False	True	False
So, the greater-than-50 % majority rule says	True	True	False

If the collective judgement is defined using the greater-than-50 % majority rule, then the collective judgement on the three propositions will be *logically inconsistent*,[4] even though every individual in the group is perfectly consistent. This paradox has come to be known as the *doctrinal paradox*, and it has generated a large literature (see e.g., List 2012; Dietrich 2012; Cariani 2011). This sort of inconsistency result

[4] This is because N and $N \to G$ entail G, so if the former two propositions are true, the latter has to be true.

shows that your choice of inputs and outputs can be incredibly important. Keep them simple, and you can get a result like the Condorcet Jury Theorem which says your crowd will probably do good things. But make the inputs and outputs a little more complex, and all of a sudden your crowd can be logically inconsistent. (Note that if the inputs are only judgements on N and $N \rightarrow G$, and the output is simply a judgement on G, then there is no inconsistency.)

In the above discussion, the collective output is evaluated in terms of its accuracy. However, as we explained in section "Thinking About the Wisdom of Crowds", there are other standards of evaluation. May (1952), for example, identified four *procedural* constraints, which he called *Universal Domain*, *Anonymity*, *Neutrality*, and *Positive Responsiveness*. These are (arguably) plausible procedural constraints that an aggregation method should satisfy (with outright judgements as inputs and outputs). For example, Neutrality requires that if everyone changes their judgement, then the collective judgement should change accordingly. May proved that the majority rule is the only aggregation method that satisfies all four constraints. For further discussion of these issues see e.g., Maskin (1995), Woeginger (2003), and Asan and Sanver (2002).

Much more could be said on the topic of mathematical aggregation, and we have only discussed simple kinds of aggregation methods on fairly simple kinds of inputs and outputs. For further discussion see Armstrong (2001a), List and Pettit (2002), Grofman et al. (1983), and Pacuit (2012). We now turn to another way in which judgements can be aggregated: through group deliberation.

Deliberation Groups

The aggregation method that most readers will have had direct experience with is a deliberation group: the "crowd" meets to discuss the problem at hand, and after a period of discussion, they arrive at a collective judgement.[5] The group discussion can be structured or unstructured. In an ideal situation, the discussion will elicit from each member of the group not only their judgements, but also their reasons, arguments and evidence that back up these judgements. Through discussion and debate, the group can sort through all of the evidence and arguments leading to a more informed solution.

A common criticism of unstructured group discussion is that it *enhances* cognitive errors rather than mitigates them. In addition, there are many social phenomena that hinder a group's ability to reach a correct judgement, even if, in principle, the group has all the pieces needed to solve the problem. We note the following three

[5] As we noted in section "Thinking About the Wisdom of Crowds", not all deliberation groups are instances of judgement aggregation. For example, the crowd could simply meet to share information and then still give different individual judgements, which could then be aggregated using one of the methods described in sections "Thinking About the Wisdom of Crowds" or "Prediction Markets".

issues. *Bias against the minority*: There is a tendency for groups to ignore isolated, minority or lower-status members. *Anchoring effect*: There is a tendency to rely too heavily, or "anchor", a judgement on one piece of information (for example, the first announced judgement, the judgement of the most senior person in the group, or the judgement of the loudest person in the group). *Common knowledge effect*: Information held by all members of the group has more influence on the final decision than information held by only a few members of the group (see Gigone and Hastie 1993). See Sunstein (2011) for a discussion of other problems with group deliberation.

Despite its many flaws, unstructured deliberation can be fruitful in certain circumstances. For instance, the unstructured discussion in the comments section of the polymathblog led to a new proof of the Hales-Jewett Theorem (see Polymath 2012). Other examples that may benefit from unstructured debate and discussion include writing a novel or finding the correct wording of a piece of legislation. Indeed, group brainstorming sessions are often used to generate new ideas and creative solutions to a variety of problems. However, some research shows that interacting brainstorming groups come up with fewer new ideas than does aggregating the collective ideas from a group of non-interacting individuals (Diehl and Stroebe 1987). The social dynamics of the group also often interferes with the group's ability to achieve its intended goal. Therefore, it is important to develop methods to keep the group focused on the task at hand (e.g., see Gerber (2009) and Bao et al. (2010) for methods aimed at improving brainstorming sessions).

One way to diminish the effect of the psychological phenomena mentioned above is to *structure* the deliberation. A method that has been widely used is the *Delphi method* (Linstone and Turoff 1975). This actually refers to a whole range of methods. What is common among the different implementations is that the members of the group provide their initial judgement *before any discussion takes place*, then there are a number of rounds in which the group members can discuss and revise their judgements.

After the group members give their initial judgments to the moderator, there are a number of ways to proceed. A sample session may run as follows: The moderator shows everyone in the group the initial judgements (making public the judgement of each member of the group). Members of the group are encouraged to discuss their reasons for their initial judgements. After a round of discussion, each person in the group is asked (either privately or publicly) if they want to revise their initial judgement. The second round judgements are then given to the moderator who aggregates them using one of the methods from section "Mathematical Aggregation". There are many ways to vary the group interactions: (i) The initial judgements are kept anonymous. (ii) Members of the group are asked to judge how confident they are in their judgements. (iii) Rather than taking part in an unstructured discussion, the members of the group are given time to do their own research in light of receiving each other's judgements. (iv) Members of the group are asked to judge how confident they are in another (randomly selected) person in the group's judgements. (v) The process continues until consensus is achieved (or some large subgroup achieves consensus). There is mounting evidence that structuring group deliberation in this way leads to more accurate predictions (Amrstrong (2006); Graefe and Armstrong (2011)).

Sometimes no amount of discussion will lead the group to a consensus opinion. This means that group deliberation may only be a partial solution to an aggregation problem, and consequently, the moderator may have to use an additional aggregation method to form the final group judgement (e.g., the moderator might average the final estimates). However, one must be careful with how these aggregation methods are combined, for it is possible for group deliberation to improve the individual judgements in a group, while making the collective judgement worse. For example, suppose that there are 10 people estimating a parameter whose true value is 40 with the following initial estimates: 15, 18, 20, 22, 30, 45, 50, 55, 60, and 61. Using an unweighted average, the group estimate is 37.6. If the new estimates after the discussion period are: 16, 25, 21, 23, 31, 41, 41, 40, 41, and 45, then each individual improved their estimate. However, the average of these estimates is 32.4, and so the collective judgement (understood as the average) is worse after discussion. Nevertheless, there is data to show that discussion, in an appropriately structured deliberation group, can improve the group estimate—see e.g., Burgman et al. (2011).

For a much more detailed overview of deliberative groups and collective group judgements in general, see Fidler et al. (2013).

Prediction Markets

Recently, there has been quite a lot of interest in the use of *prediction markets* as a method for aggregating individuals opinions about future events. Suppose that we are interested in whether an event will take place at some time in the future (for example, will Hillary Clinton run for president in 2016?). Rather than gathering experts to discuss their opinions about this event, the approach we highlight in this section is to create a market in which individuals trade contracts whose payoffs are tied to the future event. For instance, suppose that there is an option that pays \$10 if Hillary runs for president in 2016 and \$0 otherwise. Ignoring any transaction costs, if an investor pays \$7 for the event "Hillary Clinton will run for president in 2016", then she earns \$3 dollars if Clinton runs and loses \$7 if she does not run. Under standard decision-theoretic assumptions (such as that investors are *risk-neutral*), investors should be willing to pay up to a price that equals their estimated probability that an event will happen. The market price, or equilibrium price, is the value such that if an investor were willing to sell below the price, the other investors would buy the stock driving the price back up (similarly for anyone willing to buy above the market price). The market price has been interpreted as the aggregate probability of the investors (Manski 2006; Wolfers and Zitzewitz 2004) and has been shown to be remarkably accurate in predicting events (Arrow et al. 2008; Rothschild 2009).

Prediction markets have many advantages over deliberation as a method for aggregating individual judgements (see also Sunstein (2011) for a discussion of this). The primary advantage is that prediction markets provide the right incentives for a diverse population to disclose the information that they privately hold. Furthermore, the economic incentive in a market encourages traders to search for the best available information. Moreover, even if the investors are unsophisticated or not well-informed, the *efficient market hypothesis* states that markets are good aggregators of information (see Lo (2007) for an overview). See Wolfers and Zitzewitz (2004) and Arrow et al. (2008) for an extensive discussion of predication markets, including an overview of the experimental evidence and case studies that demonstrate the benefits of using markets to predict future events.

Markets work well when there is a large and diverse group and each person is likely to get different types of information. This suggests that implementing a prediction market may not always be feasible. There are two central problems that can make a prediction market infeasible. The first is that there must be a large enough group of people that are interested and engaged with the market. The second is that in order to use a prediction market, you must be interested in predicting whether or not some event will happen at some specific moment in the future. This is important since It must be perfectly clear which bets to payoff. In addition to problems of feasibility, prediction markets face a number of other challenges.

The economic incentives provided by a prediction market do a good job mitigating many of the biases that infect group deliberation discussed in the previous section. Still, prediction markets are influenced by investor biases. The most well-known is the *favorite long-shot bias*. This bias causes investors to undervalue events with probabilities close to 1. Similarly, investors tend to over-value events that have probabilities close to 0. This type of bias is well-documented and can have an effect on the market price (Thaler and Ziemba 1988).

Recent work has questioned the relationship between the market price and the distribution of beliefs of the investors in a prediction market. Othman and Sandholm (2010) study the behavior of simple agents that sequentially interact with the market. They show that by varying the order of participation in a market, the market price can converge to an arbitrary value (see Frongillo et al. (2012) for a generalization of this result).

A final challenge for the use of prediction markets as an aggregation method is the possibility of manipulation. Since most prediction markets have a relatively low volume, it would be relatively inexpensive for an investor to take losses in order to affect the market price. An example of this type of manipulation was recently observed in the Intrade market to predict the outcome of the 2012 election. According to a Washington Post article (Plumer 2012), a few months before election day, there was a huge swing towards Romney in the market which appears to have been driven by someone spending about \$17,800 to push up Romney's chances of winning. The surge only lasted about 6 minutes before other traders brought the price back down. It is still unclear whether this was a manipulation by an investor attempting to sway perceptions of the race or simply an example of a trader who made an expensive trade. There is evidence that attempts to manipulate prediction markets tend to fail (Hanson et al. 2006).

In fact, Hanson and Oprea (2009) offer a model in which attempts to manipulate the market actually *increases* the accuracy of the market price (see also Chen et al. (2010) for a general study of manipulation in prediction markets).

We conclude this short section with a few brief comments about some computational aspects of prediction markets. Virtually all the prediction markets currently in use restrict trade to "simple" events that can all be explicitly listed and monitored. So, for example, bets are made on events of the form "horse *A* will win" rather than more complex events such as "horse *A* will beat horse B which will beat horse C", "horse A will win and horse B will come in third" or "horse A will win if horse B comes in second". Initial research shows that allowing individuals to trade on more complex and/or conditional events has significant advantages (*see* Chap. 29), but it raises many difficult computational challenges. For example, it is no longer feasible to explicitly list all the possible bets—e.g., in a horse race with 10 horses, there are $10! = 3,628,800$ many different possible permutations that would need to be listed. Therefore, it is important to develop *combinatorial betting mechanisms* that allow investors to succinctly express their bets. Another computational challenge is that allowing bets on more complex events makes it much more difficult to match buyers with sellers. In general, it may be necessary to look beyond bilateral trades and consider complex multilateral trades. There is much more to say about the computational aspects of prediction markets. See Pennock and Sami (2007) and Chen and Pennock (2010) for a discussion of these issues and further references to the relevant literature.

Conclusion

There is a growing literature focused on the Wisdom of Crowds spanning many disciplines such as philosophy, computer science, management science, social psychology and social choice theory. And it can be difficult to pin down exactly what the class of phenomena is that is loosely called the "Wisdom of Crowds" by these diverse research communities.

In section "Thinking About the Wisdom of Crowds", we outlined out a simple conceptual framework for thinking systematically about the Wisdom of Crowds. We did this by taking the perspective of someone interested in using the Wisdom of Crowds to solve a problem. For example, suppose that you are interested in finding the answer to some question (e.g., Is the defendant guilty?), a prediction about a future event (e.g., Will Hillary Clinton run for president in 2016?), or an estimation of some parameter (e.g., How many jelly beans are in the jar?).

Once you identify the group of people that will make up your crowd, you must decide how to leverage the "wisdom" of the crowd to solve your problem. This involves eliciting useful information from each member of the group and deciding how to aggregate this information. There are many different methods that can be used to aggregate the judgements of a group of people. We discussed three broad categories of aggregation: Mathematical Aggregation, Group Deliberation and Prediction Markets.

All of the aggregation methods we discussed in this paper accept *human* judgments as inputs, and we primarily focused on *epistemic* judgements that were simple in form—e.g., they are about the true value of a parameter, or the answer to a yes/no question. Moving beyond this limited focus would allow us to examine a wider variety of examples of the Wisdom of Crowds. (See, for example, Nielsen (2011) for a discussion of collective judgements in situations where there are no objective facts against which the judgements can be evaluated.)

Clearly, there is much more to say about the Wisdom of Crowds. It is certainly going to take a diverse group of researchers to fully understand this phenomena.

References

Amrstrong JS (2006) Should the forecasting process eliminate face-to-face meetings? Int J Appl Forecast 5:3–8

Andler D (2012) What has Collective Wisdom to do with Wisdom? In: Landemore H, Elster J (eds) Collective wisdom. Cambridge University Press, Cambridge, pp 72–84

Armstrong J (2001a) Principles of forecasting. Kluwer Academic, Boston

Armstrong S (2001b) Combining forecasts. In: Armstrong S (ed) Principles of forecasting: a handbook for researchers and practitioners. Kluwer Academic, Norwell

Arrow KJ, Forsythe R, Gorham M, Hahn R, Hanson R, Ledyard JO, Levmore S, Litan R, Milgrom P, Nelson FD, Neumann GR, Ottaviani M, Schelling TC, Shiller RJ, Smith VL, Snowberg E, Sunstein CR, Tetlock PC, Tetlock PE, Varian HR, Wolfers J, Zitzewitz E (2008) The promise of prediciton markets. Science 320:877–878

Asan G, Sanver R (2002) Another characterization of majority rule. Econ Lett 75(3):409–413

Bao P, Gerber E, Gergle D, Hoffman D (2010) Momentum: getting and staying on topic during a brainstorm. In: Proceedings of the SIGCHI conference on human factors in computing systems, Atlanta, pp 1233–1236

Black D (1963) The theory of committees and elections. Springer, Norwell, MA

Brownstein JS, Freifeld CC, Madoff LC (2009) Digital disease detection—harnessing the web for public health surveillance. N Engl J Med 360(21):2153–2157

Burgman M, McBride M, Ashton R, Speirs-Bridge A, Flander L, Wintle B, Fidler F, Rumpff L, Twardy C (2011) Expert status and performance. PLoS One 6(7):e22998

Cariani F (2011) Judgment aggregation. Philos Compass 6(1):22–32

Chen Y, Pennock D (2010) Designing markets for prediction. AI Mag 31(4):42–52

Chen Y, Dimitrov S, Sami R, Reeves D, Pennock D, Hanson R, Fortnow L, Gonen R (2010) Gaming prediction markets: equilibrium strategies with a market maker. Algorithmica 58(4): 930–969

Clemen RT (2008) Comment on cooke's classical method. Reliab Eng Syst Saf 93(5):760–765

Collier N, Kawazoe A, Jin L, Shigematsu M, Dien D, Barrero R, Takeuchi K, Kawtrakul A (2006) A multilingual ontology for infectious disease surveillance: rationale, design and challenges. Lang Resour Eval 40(3):405–413

Condorcet M (1785) Essai sur l'application de l'analyse á la probabilité des décisions rendues á la pluralité des voix. Paris: l'Imprimerie Royale. (Reprint, 1972, Chelsea, New York)

Cooke RM (1991) Experts in uncertainty: opinion and subjective probability in science. Oxford University Press, New York

Diehl M, Stroebe W (1987) Productivity loss in brainstorming groups: toward the solution of a riddle. J Personal Soc Psychol 53(3):497–509

Dietrich F (2012) Judgment aggregation and the discursive dilemma. In: Encyclopedia of philosophy and the social sciences. Sage

Fidler F, Wintle B, Thomason N (2013) Groups making wise judgements. Final report to IARPA (Intelligence Advanced Research Projects Activity)
Frongillo R, Della Penna N, Reid M (2012) "Interpreting prediction markets: a stochastic approach." In: Proceedings of neural information processing systems
Galton F (1907a) Letters to the editor: the ballot-box. Nature 75:900–1
Galton F (1907b) Vox Populi. Nature 75:450–1
Gerber E (2009) Using improvisation to enhance the effectiveness of brainstorming. In: Proceedings of the SIGCHI conference on human factors in computing systems, Boston, pp 97–104
Gigone D, Hastie R (1993) The common knowledge effect: informaiton sharing and group judgments. J Personal Soc Psychol 65(5):959–974
Graefe A, Armstrong JS (2011) Comparing face-to-face meetings, nominal groups, delphi and prediction markets on an estimation task. Int J Forecast 27(1):183–195
Grofman B, Owen G, Feld SL (1983) Thirteen theorems in search of the truth. Theory Decis 15(3):261–278
Hanson R, Oprea R (2009) A manipulator can aid prediction market accuracy. Economica 76(302): 304–314
Hanson R, Oprea R, Porter D (2006) Information aggregation and manipulation in an experimental market. J Econ Behav Organ 60(4):449–459
Herzog SM, Hertwig R (2009) The wisdom of many in one mind improving individual judgments with dialectical bootstrapping. Psychol Sci 20(2):231–237
Hourihan KL, Benjamin AS (2010) Smaller is better (when sampling from the crowd within): low memory-span individuals benefit more from multiple opportunities for estimation. J Exp Psychol Learn Mem Cogn 36(4):1068
Keller M, Blench M, Tolentino H, Freifeld C, Mandl K, Mawudeku A, Eysenbach G, Brownstein J (2009) Use of unstructured event-based reports for global infectious disease surveillance. Emerg Infect Dis 15(5):689
Klayman J, Soll J, González-Vallejo C, Barlas S (1999) Overconfidence: it depends on how, what, and whom you ask. Organ behav Hum Decis Process 79(3):216–247
Koriat A (2012) When are two heads better than one and why? Science 336:360–2
Ladha KK (1992) The condorcet Jury theorem, free speech, and correlated votes. Am J Political Sci 36(3):617–634
Landemore H, Elster J (2012) Collective wisdom: principes and mechanisms. Cambridge University Press, New York
Lehrer K, Wagner C (1981) Rational consensus in science and society: a philosophical and mathematical study, Vol 24. D Reidel, Dordrecht
Linstone HA, Turoff M (1975) The Delphi method: techniques and applications. Addison-Wesley, Reading
List C (2012) The theory of judgment aggregation: an introductory review. Synthese 187(1): 179–207
List C, Goodin RE (2002) Epistemic democracy: generalizing the condorcet Jury theorem. J Political Philos 9(3):277–306
List C, Pettit P (2002) Aggregating sets of judgments: an impossibility result. Econ Philos 18(01):89–110
Lo, A (2007) Efficient market hypothesis, in The New Palgrave: A Dictionary of Economics, L. Blume, S. Durlauf (eds), Palgrave Macmillan
Loewer B, Laddaga R (1985) Destroying the consensus. Synthese 62(1):79–95
Lyon A, Fidler F, Burgman M (2012a) Judgement swapping and aggregation. In: 2012 AAAI fall symposium series
Lyon A, Nunn M, Grossel G, Burgman M (2012b) Comparison of web-based biosecurity intelligence systems: biocaster, epispider and healthmap. Transboundary Emerg Dis 59(3):223–232
Lyon A, Grossel G, Nunn M, Burgman M (2013) Using internet intelligence to manage biosecurity risks: a case study for aquatic animal health. Divers Distrib 19(5–6):640–650
Manski C (2006) Interpreting the predictions of prediction markets. Econ Lett 91(3):425–429

Maskin E (1995) Majority rule, social welfare functions and game forms. In: Choice, welfare and development: a festschrift in honour of Amartya K. Sen. Oxford University Press, Oxford, pp 100–109

May K (1952) A set of independent necessary and sufficient conditions for simply majority decision. Econometrica 20(4):680–684

Nielsen M (2011) Reinventing discovery: the new era of networked science. Princeton University Press, Princeton

Nuwer R (2013) Software could make rare diseases easier to spot. New Sci 218(2913):21

Othman A, Sandholm T (2010) When do markets with simple agents fail? In: Proceedings of the 9th international conference on autonomous agents and multiagent systems, AAMAS '10, Toronto, vol 1, pp 865–872

Pacuit E (2012) Voting methods. In: Zalta EN (ed) The stanford encyclopedia of philosophy. (Winter 2012 edn)

Page S (2008) The difference: how the power of diversity creates better groups, firms, schools, and societies (new edn). Princeton University Press, Princeton

Pennock DM, Sami R (2007) Computational aspects of prediction markets. In: Nisan N, Roughgarden T, Tardos E, Vazirani V (eds) Algorithmic game theory. Cambridge University Press, Cambridge/New York

Plummer B (2012) How to swing the prediction markets and boost Mitt Romney's fortunes, The Washington Post Wonkblog, October 23

Polymath DHJ (2012) A new proof of the density Hales-Jewett theorem. Ann Math 175(3): 1283–1327

Regan HM, Colyvan M, Markovchick-Nicholls L (2006) A formal model for consensus and negotiation in environmental management. J Environ Manag 80(2):167–176

Rothschild D (2009) Forecasting elections: comparing prediction markets, polls and their biases. Public Opin Q 73(5):895–916

Sunstein C (2011) Deliberating groups versus prediction markets or Hayek's challenge to habermas. In: Social epistemology: essential readings. Oxford University Press, Oxford/New York, pp 314–337

Surowiecki J (2005) The wisdom of crowds: why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations. Doubleday, New York

Thaler R, Ziemba W (1988) Anomalies: parimutuel betting markets: racetracks and lotteries. J Econ Perspect 2(2):161–174

Von Ahn L, Maurer B, McMillen C, Abraham D, Blum M (2008) Recaptcha: human-based character recognition via web security measures. Science 321(5895):1465–1468

Vul E, Pashler H (2008) Measuring the crowd within probabilistic representations within individuals. Psychol Sci 19(7):645–647

Woeginger G (2003) A new characterizaiton of the majority rule. Econ Lett 81(1):89–94

Wolfers J, Zitzewitz E (2004) Prediction markets. Journal of Economic Perspectives, 18(2):107–126

Balancing Human and Machine Contributions in Human Computation Systems

R. Jordan Crouser, Alvitta Ottley, and Remco Chang

Motivation

As we enter an age of increasingly larger and noisier data, dynamic interplay between human and machine analysis grows ever more important. Researchers and toolbuilders work to better understand and support the analytical process through systems that employ novel visual interactive interfaces along with computational support. These systems leverage the acuity of the human visual system as well as our capacity to understand and reason about complex data, nuanced relationships, and changing situations. In designing and building these systems, we rely on the intuition that the lived experience, perceptual advantage, and adaptability of the human analyst may prove crucial in areas where purely computational analyses fail. Similarly, by pairing the human analyst with a machine collaborator we hope to overcome some of the limitations imposed by the human brain such as limited working memory, bias, and fatigue.

With many promising examples of human-machine collaboration in the literature and everyday life, how do we tell if a new problem would benefit from human-computer collaboration and how should we allocate computational tasks? At present, balancing the cost of building and deploying a collaborative system with the benefits afforded by its use is precarious at best. We rely heavily on researcher intuition and current field-wide trends to decide which problems to approach using collaborative techniques. While this has led to many successes, it has also led to the investment of significant time and energy into inefficient collaborative solutions for problems that might better have been (or have already been) solved by human or machine alone.

While we have come a long way from listing tasks best assigned to human or machine (Fitts 1951), appropriate function allocation in collaborative systems is still far from a perfect science (Sheridan 2000). However, the effectiveness of any

R.J. Crouser (✉) • A. Ottley • R. Chang
Department of Computer Science, Tufts University, Medford, MA 02155, USA
e-mail: rcrous01@cs.tufts.edu; alvittao@cs.tufts.edu; remco@cs.tufts.edu

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_48,

collaborative system is heavily dependent on how well it leverages the skills that humans and machines have to offer while minimizing waste. In the absence of a secret formula to prescribe this interplay, how do we balance the expected contributions of human and machine during the design process, and how do we evaluate the effectiveness of systems once we've built them? Herein, we seek to address these questions.

Toward that end, this chapter is organized into three main sections. Section 1 provides a broad definition of human computation and human-computer collaboration, as well as comparisons between a few well-known problems and their human-computer collaborative solutions. Section 2 considers the relative strengths of human and machine collaborators. Section 3 discusses the open problem of function allocation in human-computer collaborative systems, and will provide some insight on applying this knowledge during the design process. We hope that this chapter will leave the reader with an improved understanding of the complementary strengths of human and machine, as well as actionable information about best practices for real world design.

1 - Beyond Crowdsourcing: Human Computation as Human-Computer Collaboration

In this section, we develop a working definition for the term *human computation*. It is important to note that human computation is not synonymous with terms such as *collective intelligence*, *crowdsourcing*, and *social computing,* although they are related. Before we continue, we will first define a few of these terms in the interest of developing a context for defining human computation. For a more detailed discussion on definitions of Human Computation, please see Chapter 9 - "Synthesis and Taxonomy of Human Computation".

Crowdsourcing is the practice of obtaining services, ideas, or content by soliciting contributions from a large group of people.

Collective intelligence is the notion that groups of individuals working together can display intelligent behavior that transcends individual contributions.

Social computing is the intersection between people's social behaviors and their interactions with technology.

In many cases, a single system could be classified under more than one of these headings. At the same time, none of them fully captures the notion of human computation. As such, there are many working definitions of human computation in the literature:

> …using human effort to perform tasks that computers cannot yet perform, … (Law and von Ahn 2009)
>
> …a technique that makes use of human abilities for computation to solve problems. (Chan et al. 2009)

A computational process that involves humans in certain steps… (Yang et al. 2008)

…systems of computers and large numbers of humans that work together in order to solve problems that could not be solved by either computers or humans alone. (Quinn and Bederson 2009)

Working from these definitions, we can begin to come to consensus regarding what constitutes human computation. First, the problem must involve some form of *information processing*. This may occur as part of an algorithmic process, or may emerge through the observation and analysis of technology-mediated human behavior. Second, human participation must be *integral to the computational system or process*. In this work, we do not consider systems with only superficial human involvement to fall under the umbrella of Human Computation.

We can think of human-computation as a kind of human-computer collaboration. In this context, collaboration as defined as *a process in which two or more agents work together to achieve shared goals,* and human-computer collaboration as *collaboration involving at least one human and at least one computational agent* (Terveen 1995). This has also been called *mixed-initiative systems* (Horvitz 1999). In a mixed-initiative system, both the human and the machine can initiate action, access information and suggest or enact responses (Thomas and Cook 2005). The field of Visual Analytics is a perfect example of human-computer collaboration, as Visual Analytics systems leverage both analyst intelligence and machine computation in a collaborative effort to solve complex problems.

Under this definition, we see that crowdsourced computation is just the tip of the HC iceberg. Along a continuum between human-heavy and machine-heavy collaboration such as the one posed by Bertini and Lalanne in 2009, crowdsourced computation falls at one extreme:

With few exceptions, the computational burden falls almost entirely on the human collaborators in typical crowdsourcing applications such as image labeling and text translation. Human-based genetic algorithms also fall on the human-heavy end of the continuum, as the human agents determine both population fitness and genetic variation. In these systems, the primary role of the machine collaborator is to distribute tasks and collect results, a role with relatively trivial computational requirements. On the other extreme, algorithms for unsupervised learning functions with near autonomy from the human collaborator. Here, the human's role is to set the parameters of the algorithms and to verify the results. While less common, there are an increasing number of algorithmic approaches that attempt to maximize the contributions from both collaborators.

Without question, the term *human computation* spans a wide range of possible applications and computational distributions. Among all these, many of the most interesting and successful human computation systems not only balance the contribution of human and machine, but also leverage the complementary computational strengths of both parties. In the following sections, we will explore some of these strengths and how they can impact the distribution of labor in a human computation system.

2 - Complementary Computation

Both human and machine bring to the partnership varying strengths and opportunities for action, and during collaboration, each must be able to perceive and access these opportunities in order for them to be effectively leveraged. These affordances define the interaction possibilities of the team, and determine the degree to which each party's skills can be utilized during collaborative problem solving. The set of problems warranting a collaborative technique is equivalent to the set problems where there is an opportunity to effectively leverage affordances on both sides of the partnership in pursuit of the solution.

Instead of deciding who gets (stuck with) which task, we can begin to reason about which party can best contribute to the collective goal at each stage. The answer may not be only the human, or only the machine, but could in fact be both. By designing such that both human and machine are aware of the affordances made available to them by their collaborators, we encourage the development of more flexible procedures for collective problem solving.

Relative Strengths of Human and Computers

Fitts made the first published attempt in 1951 categorize tasks when he created a list of tasks the *humans are better at* and *machines are better at* (see Table 1). This is often abbreviated in the literature as *HABA-MABA*. While for many years this list was viewed as mantra for the division of labor, frequent and consistent technological advances in computation, automation and robotics make function allocation and the HABA-MABA list a moving target. The distinction between human and machine

Table 1 An outdated comparison of the relative strengths of humans and machines

Humans are better at:	Machines are better at:
Detecting small amounts of visual and auditory energy	Responding quickly and applying great force smoothly and precisely
Perceiving patterns of light or sound	Performing repetitive, routine tasks
Improvising/using flexible procedures	Storing information briefly, then erasing it completely
Storing large amounts of information and performing selective recall	Reasoning deductively
Reasoning inductively	Multitasking
Exercising judgment	

Table 2 A modern comparison of human vs machine affordances

Human Affordances	Machine Affordances
Visual perception	Computational Power
Spatial ability	Scalable, persistent storage
Linguistic ability	Efficient/reliable data transfer
Creativity	Freedom from bias
Adaptability	Precision
Sociocultural awareness	Parahuman sensing
Expertise/lived experience	Environmental Tolerance

is now less clear. For example, while in the 1950s humans were indeed better at storing large amounts of information, today's machines far exceed the capacity previously imagined, and the advent of distributed storage is rapidly enabling the outpacing of human memory by machines.

While Fitts' list was aimed at simply comparing the two for basic labor division, for many years it was incorrectly interpreted as gospel for function allocation for human-machine collaborative systems. Jordan (1963) criticized this approach stating that the underlying foundation should be that humans and machines are complementary rather than antithetical. Price (1985) also supported this view, arguing that function allocation could be better viewed as an interactive process rather than a divisive listing and that there may exist several optimal solutions for a given problem. Although the comparative approach to division of labor in human-computer collaborative systems was unwarranted, Fitts' list laid the foundation for thinking about the respective strengths of humans and machines.

Affordances: A Changing Perspective

In recent years, researchers have argued that the original perception of function allocation and Fitts' list no longer makes sense (Sheridan 2000; Dekker and Woods 2002). Dekker and Woods (2002) also provided counterargument to the validity of Fitts' list. They discussed how human-machine collaboration transforms human practice and forces analysts to adapt theirs skills and analytics processes. They argue for a shift in attention, moving away from allocation of tasks to a focus centered on how to design for harmonious human-machine cooperation. That is, how do we get humans and machines to play nicely, and work effectively? In response, a more recent framework (Crouser and Chang 2012) categorizes tasks based on relative strengths or affordances—opportunities for action. For a listing of some of the affordances they examined, see Table 2.

Human computation is an ideal approach to problems where there is an opportunity to leverage both human and machine affordances in pursuit of the solution. By framing potential collaboration in terms of the affordances at our disposal, we can then consider which of these affordances could be used to approach a problem and construct a solution.

3 - Effectively Leveraging Human and Machine Affordances

The success of human-computer collaborative systems hinges on leveraging the skills of both the human and the computer. That said, in order to address the problem of balancing and allocating workload in a human-computer collaborative system, it is first necessary to explore the space of problem difficulty relative to human and machine.

While problem difficulty for a machine can be defined as space and time complexity, for the human we propose that problem difficulty is attributed to two main sources: knowledge necessary to solve the problem and time investment required to solve a problem. The level of difficulty for one party may not necessarily transfer to the other. For instance, some problems such as character recognition are inherently easy for a human and can sometimes be performed in constant time but can be computationally expensive or unsolvable for the machine. The inverse can also be true. We can think about the problem space as having two orthogonal dimensions: human difficulty and machine difficulty. Figure 1 depicts some well-known sample problems within in this space.

In this diagram, problems appearing in the lower left region are trivial; that is, they are comparatively easy for both humans and machines. These problems, such

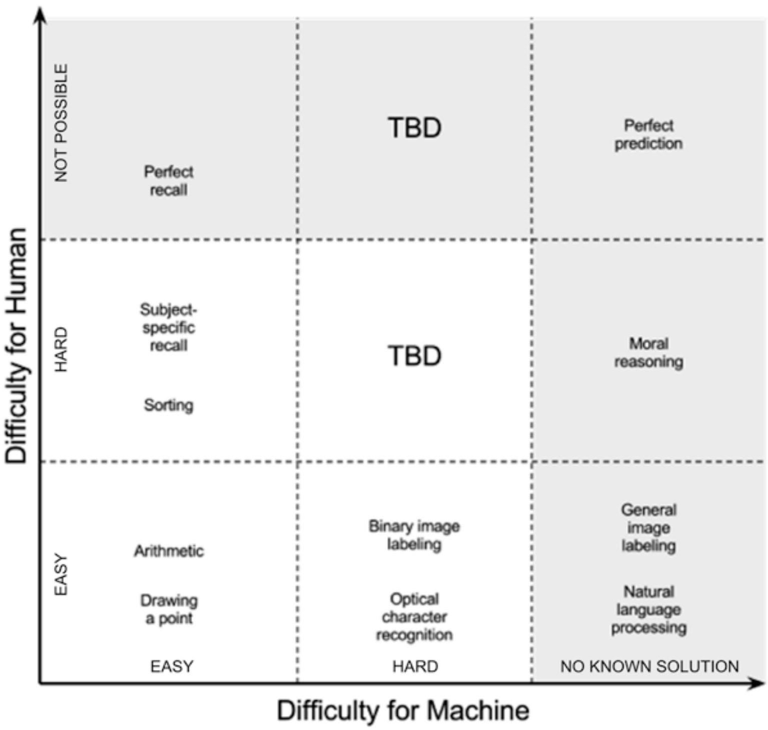


Fig. 1 A matrix of example problems based on respective difficulty levels for humans and machines

as arithmetic or simple shape rendering, generally do not warrant a human-computer collaborative solution. As we move to the right along the x axis, we encounter many of the problems addressed in early human computation systems: image labeling, character recognition, language processing, etc. These problems are difficult for machines, but relatively straightforward for humans. Here, the overhead cost incurred by involving human processing power is minimal compared with the resources required to achieve comparable performance using a machine.

As the field of human computation progresses, we are becoming more invested in applying collaborative techniques to solve problems that are difficult or impossible for *either* humans or machines alone, but which may be solvable through collaboration. In these problems, we are especially interested in how to best allocate the computational resources of the human and machine collaborators, allowing each party to play to its strengths.

Technically speaking, a human can simulate any process the machine can execute. After all, we designed the algorithms in the first place. Given an understanding of the process, enough paper and a sufficient supply of pencils, a human operator could write out the contents of each register, perform each bitwise operation, and record each result by hand. However, the time and resources required to compute exactly the same result are exorbitant. In addition, humans are susceptible to fatigue, and we are arguably limited by the capacity of our working memory and unreliable recall. In this sense, human operations are *expensive*, and there are cases where it is possible to reduce the number of human operations while maintaining optimal performance.

Consider the following example (Shahaf and Amir 2009): Imagine that we are given n randomly selected samples that we wish to classify. We know that the classifiers are simple threshold functions:

$$h_w(x) = \begin{cases} 1, & | \; x > w \\ 0, & | \; x \le w \end{cases}$$

with the value of w depending on the input. Assume that we do not know w, but a human can classify the data correctly without too much trouble. We can think of the human as an *oracle*, a black box which is able to correctly answer specific questions in a single operation. Using the human as an oracle, there are several ways to approach this problem, each with benefits and drawbacks:

1. We could ignore the human and use a pure machine computational approach, first sorting the set of samples according to their x values and then choosing a random threshold value that falls between the lowest and highest values. This requires $O(n \log n)$ time to sort the list, and guarantees that at least two of the samples will be classified correctly. This is not a very promising bound on accuracy.
2. We could use a pure human computational approach, asking the human to classify each sample in the dataset. Because as we assumed that the human can always classify samples correctly, this method guarantees 100 % accuracy. In addition, this method requires $c \times n = O(n)$ operations, where c corresponds to the amount of time it takes the human to classify one sample. Under the usual metrics for evaluating algorithmic complexity, the method is technically “faster”.

However, the value of the constant c may be enormous, which means that for all reasonably-sized input sets, this approach really isn't much better.

3. Finally, we could try a collaborative solution. First, the machine sorts the set of samples according to their x values, requiring $n \log n$ operations. Next, the human is asked to classify the sample that falls in the middle of the sorted list. If she answers "1", we know that all the samples above should also be labeled "1". Similarly, if she answers "0", we know that all the samples below should also be labeled "0". From here, the human is recursively asked about the middle point in the remaining half of the list that remains unlabeled. This is simply binary search, which implies that the human will be asked to classify at most $\log n$ samples for a cost of $c \times \log n$. Using this algorithm, we are able to dramatically reduce the workload for the human operator while maintaining 100 % accuracy simply by being clever regarding which samples to ask her about.

In this example, the third approach is clearly superior to the other two in terms of maximizing accuracy and efficiency. However, there are several key assumptions that need to be dressed. For example, what is the scale of the constant c? In human computation, we argue that this scale depends on the affordance being leveraged. This is perhaps most readily apparent in the field of information visualization. Through visualization, we transform the task of assessing abstract numerical information to evaluating visual information, leveraging the human visual processing system and thereby decreasing the per-operation cost c. As designers, it is important to consider the implications of leveraging various combinations of affordances between human and machine.

A few caveats: under this model, there is an explicit assumption the human oracle will always be able to provide the correct answer at a fixed (albeit large) cost. In reality, humans don't work this way. Intuition and experience indicate that humans eventually get tired or bored, and as a consequence their speed and accuracy suffer. Even under optimal working conditions, humans are fallible; instead of modeling the human as omniscient, we may wish to model human oracles as accurate with some probability $p < 1$. This fallibility may require measures to ensure that the overall probability of correctness is higher than that of any single oracle. In addition, the human may wish to request more information from the system before she can make a determination. This mutual querying behavior cannot be captured by an oracle model; we assume that the oracle takes input, and gives a correct answer using only that input. Because of this, it is important to continue to develop more nuanced models human of behavior in human computation systems, and to design metrics by which we can more robustly evaluate algorithmic complexity and performance in human-machine collaborative systems.

Conclusion

While history has often depicted human and machine as antithetical, we argue that in many cases their relative strengths prove complementary. Though originally designed simply to relieve human operators of tedious tasks, today it is perhaps more fitting to view machines as collaborators in the pursuit of solutions to challenging problems. We hope that this chapter has left the reader with a better understanding of some of the intricacies of balancing human and machine contributions in human computation systems as well as candidate methods for evaluating the optimality of that balance. We believe that work in the emerging field of human computation will help us to expand our understanding of what is computable, and that human-computer collaboration could lead to significant advances in tackling currently intractable problems.

References

Bertini E, Lalanne D (2010) Investigating and reflecting on the integration of. ACM SIGKDD Explor Newslett 11(2):9–18

Chan KT, King I, Yuen M (2009) Mathematical modeling of social games. Proc CSE 1205–1210

Crouser RJ, Chang R (2012) An affordance-based framework for human computation and human-computer collaboration. IEEE Trans Vis Comput Graph 18(12):2859–2868

Dekker SW, Woods DD (2002) Maba-maba or abracadabra? Progress on human–automation co-ordination. Cogn Technol Work 4(4):240–244

Ferguson G, Allen J (1998) Trips: an integrated intelligent problem- solving assistant. In: Proceedings of the national conference on artificial intelligence, Wiley, pp 567–573

Fitts PM (1951) Human engineering for an effective air-navigation and traffic-control system. National Research Council, Division of Anthropology and Psychology, Committee on Aviation Psychology, Washington

Gibson J (1977) The theory of affordances. In: Perceiving, acting, and knowing, pp 67–82

Horvitz E (1999) Principles of mixed-initiative user interfaces. In: Proceedings of the SIGCHI conference on human factors in computing systems: the CHI is the limit, ACM, pp 159–166

Jordan N (1963) Allocation of functions between man and machines in automated systems. J Appl Psychol 47(3):161

Law E, von Ahn L (2009) Input-agreement: a new mechanism for collecting data using human computation games. Proc CHI 3:1197–1206

Price H (1985) The allocation of functions in systems. Human factors. J Hum Factors Ergon Soc 27(1):33–45

Quinn A, Bederson BB (2009) A taxonomy of distributed human computation. Tech report HCIL-2009-2023, University of Maryland

Rickel J, Johnson W (1999) Animated agents for procedural training in virtual reality: perception, cognition, and motor control. Appl Artif Intell 13(4–5):343–382

Shahaf D, Amir E (2007) Towards a theory of AI completeness. In: 8th international symposium on logic formalizations of commonsense reasoning

Sheridan TB (2000) Function allocation: algorithm, alchemy or apostasy? Int J Hum-Comput Stud 52(2):203–216

Terveen L (1995) Overview of human-computer collaboration. Knowl Based Syst 8(2–3):67–81

Thomas J, Cook K (2005) Illuminating the path: the research and development agenda for visual analytics, vol 54. IEEE

Valdés-Pérez R (1999) Principles of human-computer collaboration for knowledge discovery in science. Artif Intell 107(2):335–346

Yang Y, Zhu BB, Guo R, Yang L, Li S, Yu N (2008) A comprehensive human computation framework: with application to image labeling. Proc MM

Constructing Crowdsourced Workflows

Peng Dai

Overview

With the success of several major crowdsourcing platforms such as Amazon's Mechanical Turk, oDesk, and Crowdflower, crowdsourcing has becoming more and more popular. Numerous employers have started to use crowdsourcing to complete their tasks. These tasks range from simple, micro tasks, such as image tagging, OCR, and question answering, to ones that are traditionally completed by domain experts, such as audio transcription, natural language processing, and translation. To yield high quality results, crowdsourcing employers usually decompose a complex task into multiple, pipelined tasks and build workflows over these tasks where workers may contribute to the end result collaboratively. This chapter provides an overview of crowdsourcing workflows and algorithms of controlling such workflows. To build pipelined tasks where the output of a previous task is treated as the input of a later task, people have innovated iterative workflows. To optimize iterative and other workflows, artificial intelligence and decision theory have been applied to automate the construction of such a workflow. Hybridized workflows help combine the efforts of both computer programs and human work and integrate results from various resources in a smart way. Unlike computers, human workers do not perform consistently over time. To overcome this limitation, people have come up with ideas on how to leverage human efforts and temporal information. Beyond the efforts, people also customize various workflows for specific application domains to better serve their needs.

P. Dai (✉)
Google Inc., 1600 Amphitheater Pkwy, Mountain View, CA 94043, USA
e-mail: daipeng@cs.washington.edu

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_49,

Iterative Workflows

Find-fix-verify (Bernstein et al. 2010) is a workflow introduced to solve complex tasks by workers who have little domain knowledge. The idea is to decompose the task into several smaller, sequential tasks, where the input of a later task is dependent on the output of a previous task. The motivating example is word processing. The task input is an article on a specific topic, and the output is an improved version of it. The task is complex since workers only have limited professional editing skills. Using find-fix-verify, one such task is divided into three steps (sub-tasks). In the find step, workers are asked to highlight places where a modification maybe needed (such as typos, grammar errors, unclear descriptions, verbose sentences, etc.). In the fix step, a different group of workers are instructed to make edits around places that are highlighted in the find step. Finally, the new edits are presented to another set of workers to do proofreading. By decomposition, each worker concentrates only on a very small aspect of a complex task, with the results sufficiently validated. As a result, crowdsourcing workers with little domain knowledge achieve results of the same quality as by domain experts.

Little et al. (2009, 2010) first introduced the idea of iterative improvement. Given a complex task such as transcribing a handwritten document, the idea is to have workers sequentially improve a current best solution, based upon others' work. At the beginning of each iteration, an *improvement task*, which presents the handwriting image with the current transcription (initially empty), is assigned to a worker, and the worker is asked to improve the transcription, such as transcribing new words, checking transcribed words, language, and grammar. Then a fixed number of different workers are recruited to take part in a *ballot task*, with the image shown and the old and new transcriptions presented side-by-side. The workers are supposed to vote which one is more accurate. The transcription that receives the majority vote is chosen as the starting point of the new improvement iteration. To execute such a workflow, a budget is usually set up ahead of time, and the workflow stops when it runs out of budget. Iterative improvement is a very powerful workflow as it lets workers work collaboratively and improve based on each others' work . See Fig. 1 for a transcription task and the ultimate crowdsourced results by applying iterative improvement.

Automated Workflows

Dai et al. (2010, 2011) observed there are several places where an iterative workflow can be optimized. For example, the beginning of the first few iterations are typically easy, therefore the new solutions are typically better than existing solutions, so voting is not as important as when improvements are getting hard in later iterations. Also, it is hard to decide when to stop the improvement process, as tasks can vary massively in difficulty level. Intuitively speaking, a few continuous, unsuccessful iterations probably indicates that the current solution is already in good

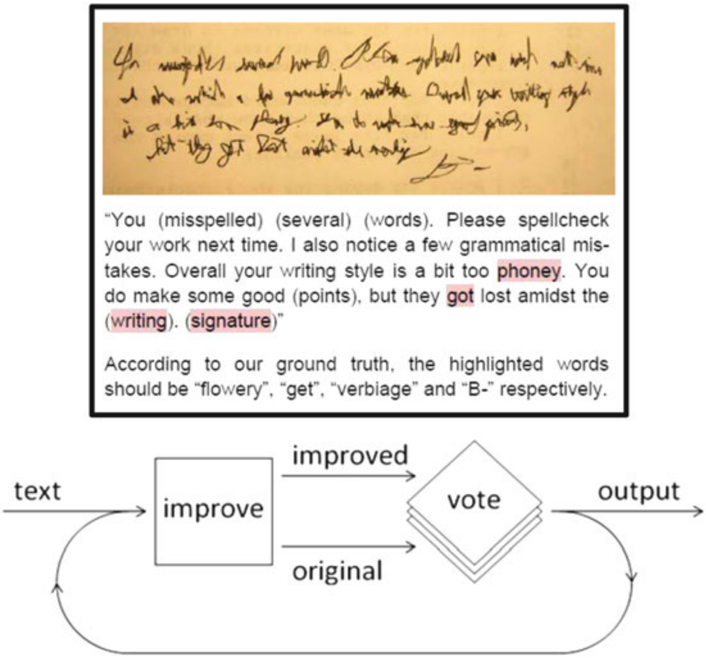


Fig. 1 A handwriting transcribing task solved by Mechanical Turk workers using an iterative improvement workflow. The tasks show the text written by a human and the existing transcription. After several iterations, the transcription is very close to ground truth (with errors highlighted) (Figures adapted and reprinted from Little et al. (2009))

quality; a good point to terminate. However, a non-adaptive workflow is unable to catch these subtleties and signals, thus works sub-optimally. The conclusions are reminiscent of the findings on human recognition behavior using psychometric data (Michelucci 2000). As humans are trying to make sense of the world, their perceptual systems employ a similar model of continuing to iterate and aggregate information until there are diminishing returns.

Following the observations, they model controlling such a workflow as a *partially-observable Markov decision process*, a generally-used statistically framework for decision making problems. Consider generating an artifact, such as an image description, a handwriting transcription, an audio transcription, etc., a workflow typically consists of two general tasks: an improvement task, i.e., improving the current artifact, and a ballot task, or making a judgment call among multiple artifacts. The model on the performance of workers on these tasks are generated and refined by machine learning techniques. An agent, called TurKontrol, is assisted by a decision-theoretic engine that makes automated decisions on when and which task is crowdsourced, given the worker performance model. One typically workflow can be illustrated in Fig. 2. The input to the workflow is an initial artifact, α. TurKontrol first decides whether an improvement is needed, based on a prior distribution of the quality of α. If so, an improvement task is generated and some

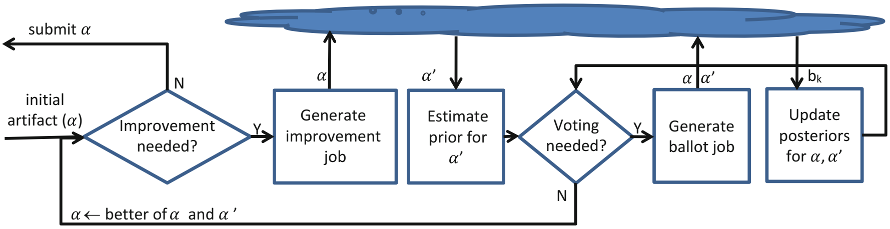


Fig. 2 A typical crowdsourced workflow automatically controlled and optimized by agent TurKontrol

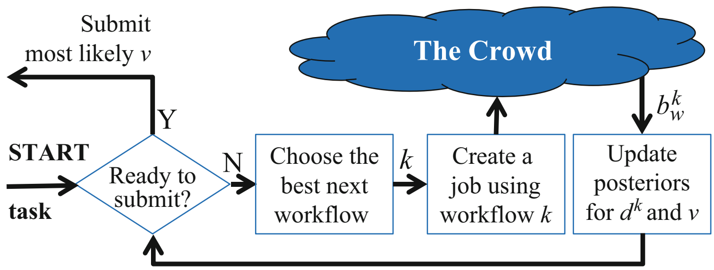


Fig. 3 AgentHunt's decisions on multiple crowdsourced workflow exploitation (Reprinted from Dai et al. (2013))

worker outputs a new artifact, α'. Given the worker identity information, TurKontrol computes its belief over the quality of α' and decides whether the next task should be a ballot, an improvement, or to terminate the workflow. On a ballot task, TurKontrol updates its quality beliefs given the ballot answer. Otherwise, the agent chooses a better artifact and either uses it as the starting point on an improvement task, or submits it on workflow termination.

Dai et al. (2010, 2011) demonstrated that applying the decision-theoretic agent significantly improves the ultimate result of the solution for the image description domain. To achieve results of the same quality, a non-adaptive workflow spends 28.7 % more money on average. Note that TurKontrol can be applied to control and optimize any type of workflows.

Optimizing a workflow has opened the door to solving many previously-unsolvable tasks. Indeed, for many cases, a single, optimized workflow is useful for most of the tasks. Yet, the best workflow can sometimes be domain-specific, and various workflows may provide independent evidence, thus increases confidence in the results. Therefore exploiting multiple workflows and applying a mixed workflow can further improve task solving efficiency and the quality of the results. Lin et al. (2012a,b) then generalized the crowdsourced workflow control problem into managing multiple workflows. See Fig. 3 for the decisions made by the agent, AgentHunt. To accomplish a task, there are k workflows, possibly with overlapping task types, that the agent can choose from. In addition to the decisions made by TurKontrol, a workflow evaluation phase is completed before a new task is posted. The best workflow is picked up and the next task along that workflow is chosen to be the new task. Results show that optimizing multiple workflows significantly improves the accuracy of named-entity tagging results over optimizing a single workflow.

Hybridized Workflows

Previously we discuss workflows where tasks are completed purely by crowdsourcing workers. In order to leverage human and computer resource, people have come up with workflows where part of them are executed by computer programs. Cascade (Chilton et al. 2013) is one automated such workflow that can be applied in generating taxonomy of data items, such as answers to a question, colors, concepts, etc. Generating taxonomy is challenging for workflows such as iterative improvement, since taxonomy is dynamic and can grow very quickly, and the task has many divergent ways to tackle, so a focused workflow with merely improvement and voting tasks is not generally enough. The workflow constructed by cascade goes along the pipeline of three crowdsourcing tasks as well as an automated result aggregation process. Initially, a *generate task* asks workers to suggest a category for multiple data items. Given candidate categories for each data item from the first step, a *selectbest task* lets workers choose the most suitable category for each item. Categories with sufficient votes from the previous step go through a *categorize task*, where workers vote whether an item fits a category. Subsequently, an automated structural inference algorithm generates the taxonomy of all the data items, based on human results. Empirical study on Mechanical Turk shows that the cascade algorithm performs close to expert agreement with competitive costs.

Skierarchy is a technology, innovated by SteuSert.com, that utilizes crowdsourcing resources hierarchically based on the qualities of each. It has been applied to solve tasks from domains that typically require human expertise, such as natural language processing. The hierarchy maintained is illustrated in Fig. 4. The bottom level is composed of machine learning algorithms trained from expert labels. Computer programs are less expensive compared to human resource, yet yield lower quality results. The results of machine learning algorithms are used as initial references by higher hierarchies. The middle level consists of crowdsourcing workers who have little to median training over the task domain. They produce higher quality results than computer programs, but are not as good as experts. Their job is to answer questions, with answer references generated by computer programs, and escalate ones they feel uncertain of on the top level resides domain experts, who make judgment calls over tasks where crowdsourcing workers escalated. By building such a hybridized workflow, not only results are sufficiently cross validated by various sources, but also human experts, the most demanding resource, can concentrate on the hardest tasks.

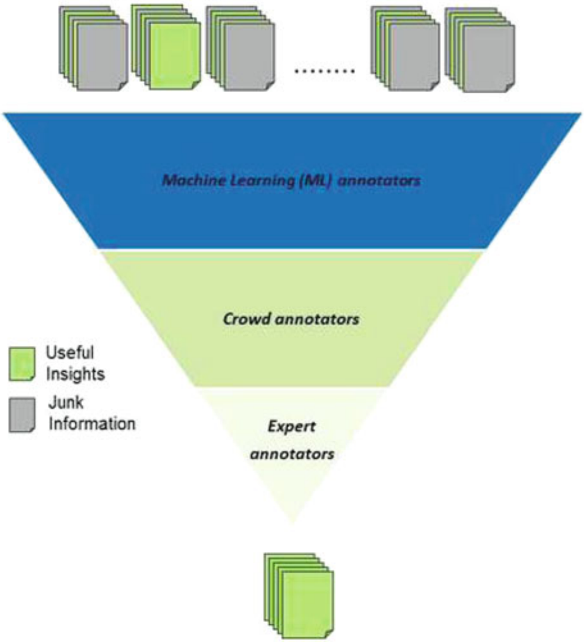


Fig. 4 Hierarchical structure of how SetuServ manages and exploit crowdsourcing resources (Figure reprinted from www.setuserv.com)

Temporal Workflows

Unlike computer programs, workers' accuracies often change over time, e.g., due to fatigue, mood, task familiarity, etc. One refined temporal worker model defines the accuracy of a worker given the time of day, day of the week, as well as the amount of time a worker has been on a particular task. Donmez et al. (2010) propose to learn temporal worker accuracy models and use them to predict the best workers to assign tasks to optimize the overall performance. Sharing the same observations, Rzeszotarski et al. (2013) take a step further to help workers recover from accuracy declines by inserting micro-breaks, such as reading a comic or playing a game, into each worker's task sequence. They also propose to build an AI-assisted agent that uses temporal worker models to predict when a worker may feel tired and automatically suggest the most effective micro-breaks.

Other Workflows

In addition to the general-purpose workflows, many other workflows have been designed to handle specific applications.

Mobi (Zhang et al. 2012) is a crowdsourcing system where workers collaboratively complete large tasks such as generating an itinerary with time and budget constraints, represented in natural language. Mobi allows workers to see and iterate ideas online, and check and edit the details of a running itinerary through a visualization tool. One useful feature of the system is a todo list automatically generated out of unsatisfied constraints, which helps provide directions and hints for workers to improve the current solution. For speech-to-text transcription tasks, a workflow that breaks clips into small segments, and later merges their transcriptions has been proved effective (Liem et al. 2011). Calculating calories on a food plate (Noronha et al. 2011) can be decomposed into several sequential tasks, such as tagging food items, matching items with nutrition database, and measuring individual item's calories, etc., to achieve reasonably accurate results. Kulkarni et al. (2012) show that the crowd can even help design and execute complex workflows, with and without the help from crowdsourcing employers. Lasecki et al. (2011) design a workflow that enables employers to easily crowdsource tasks with an integrated user interface, such as a robot navigation tool, and allow multiple workers to control the same interface in real-time.

Future Work

Building crowdsouced workflows are a new area and has already demonstrated its usefulness. We believe there are many new directions to explore in the future. Existing workflows usually have limited types of crowdsourcing tasks, most of which belong to two categories: content generation and content verification. To achieve more complicated and specialized tasks through crowdsourcing, such as software development, business consulting, financing, etc., one tends to crowdsource more important decision making tasks. Those new tasks will challenge workers' management and coordination skills, such as resource and budget management, task decomposition, or even workflow control and execution.

Many opportunities exist on building more complex, even multi-level workflows, where the task of a higher-level workflow can be a lower-level workflow. For example, small teams of humans in social contact online may collaborate and complete one such task – their own collaboration forms an independent workflow with its interface to the high-level workflow structured. Crowdsoucing may be applied on several layers, such that the worker of a bigger task can be the employer of smaller tasks.

References

Bernstein MS, Little G, Miller RC, Hartmann B, Ackerman MS, Karger DR, Crowell D, Panovich K (2010) Soylent: a word processor with a crowd inside. In: UIST, New York, pp 313–322

Chilton L, Little G, Edge D, Weld DS, Landay JA (2013) Cascade: Crowdsourcing taxonomy creation In: CHI, Paris

Dai P, Lin CH, Mausam, Weld DS (2013) POMDP-based control of workflows for crowdsourcing. Artif Intell 202:52-85

Dai P, Mausam, Weld DS (2010) Decision-theoretic control of crowd-sourced workflows. In: AAAI, Atlanta

Dai P, Mausam, Weld DS (2011) Artificial intelligence for artificial artificial intelligence. In: AAAI, San Francisco

Donmez P, Carbonell JG, Schneider J (2010) A probabilistic framework to learn from multiple annotators with time-varying accuracy. In: SIAM international conference on data mining (SDM), Columbus, pp 826–837

Kulkarni A, Can M, Hartmann B (2012) Collaboratively crowdsourcing workflows with turkomatic. In: Proceedings of CSCW, Seattle

Lasecki WS, Murray KI, White S, Miller RC, Bigham JP (2011) Real-time crowd control of existing interfaces. In: Proceedings of UIST, Santa Barbara

Liem B, Zhang H, Chen Y (2011) An iterative dual pathway structure for speech-to-text transcription. In: HCOMP, San Francisco

Lin CH, Mausam, Weld DS (2012a) Crowdsourcing control: moving beyond multiple choice. In: UAI, Toronto

Lin CH, Mausam, Weld DS (2012b) Dynamically switching between synergistic workflows for crowdsourcing. In: AAAI, Toronto

Little G, Chilton LB, Goldman M, Miller RC (2009) Turkit: tools for iterative tasks on mechanical turk. In: KDD workshop on human computation, Paris, pp 29–30

Little G, Chilton LB, Goldman M, Miller RC (2010) Turkit: human computation algorithms on mechanical turk. In: UIST, New York, pp 57–66

Michelucci PE (2000) A quantum model of recognition. Dissertation, Indiana University

Noronha J, Hysen E, Zhang H, Gajos KZ (2011) Platemate: crowdsourcing nutrition analysis from food photographs. In: UIST, Santa Barbara

Rzeszotarski J, Chi E, Paratosh P, Dai P (2013) And now for something completely different: introducing micro-breaks into crowdsourcing workflows. In: HCOMP WiP track

Zhang H, Law E, Miller R, Gajos K, Parkes DC, Horvitz E (2012) Human computation tasks with global constraints. In: CHI, Austin, pp 217–226

Distributed Intelligent Agent Algorithms in Human Computation

Edmund H. Durfee

Motivation

Just as Artificial Intelligence algorithms have often been patterned after how an individual person reasons, so too have Distributed Artificial Intelligence (DAI) algorithms drawn inspiration from social-scientific insights about how groups of people reason collectively in a coordinated fashion. While interactions among self-interested agents has been one important avenue of research (leading to computational algorithms informed by game theory, auctions, negotiation, etc.), in keeping with the themes of this book, the discussion in this chapter is limited to distributed algorithms that support and promote cooperation among artificial agents. (For a broad overview of DAI, see the book by Weiss (2013).) This chapter first provides a categorization of different families of algorithms for cooperative problem solving by computational agents. It then highlights how lessons learned in DAI could inform the development of human computational systems, and concludes by illustrating examples of combining human computation and networks of intelligent agents, suggesting directions for future efforts.

Cooperating Intelligent Agents

At the risk of oversimplifying, research into using distributed computation to realize intelligent behavior falls along a spectrum. One extreme sees intelligence as emerging from interactions among computational elements that are not individually intelligent, such as how neurons collectively implement cognition (see the Foundations

E.H. Durfee (✉)
Computer Science and Engineering, University of Michigan, Ann Arbor, MI 48104, USA
e-mail: durfee@umich.edu

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_50,

chapter "From Neural to Human Communication" by Larson-Prior). At the other extreme, each computational element is independently intelligent, but by cooperating such elements can achieve more than any can alone. For example, in the early years of DAI, the practical successes of expert systems raised the question of whether and how such systems could cooperate, much like how human medical specialists could collaboratively solve a complicated case. The belief then, as now, was that systems comprised of well-constructed agents with complementary awareness and abilities would be more versatile, extensible, and verifiable, presuming that they could cooperate well to solve problems.

An underlying assumption that has colored the development of cooperative problem-solving (CPS) algorithms bears stating: that communication is slow relative to computation. This perspective reinforced the emphasis on viewing each agent as a sophisticated problem solver in itself, where agents would invest a lot of "thought" into what information is important to share with others (and when to share it), and into how to interpret and utilize received information. In turn, like a human expert, each agent's time was extremely valuable, making it a priority that each agent was applying its expertise to the right problems at any given time.

Task-Sharing Algorithms. An early and continuing research problem in the field has thus been to develop algorithms to allocate problem-solving tasks to the "right" agents. An early, influential distributed algorithm for this was the Contract-Net Protocol (Davis and Smith 1983). In brief, the algorithm was: (1) an agent with a task that needs to be done broadcasts an Announcement that describes the task, indicates the eligibility criteria for agents to do it, and specifies the content of a bid for the task; (2) eligible agents submit Bids indicating their suitability for being awarded the task; (3) the original agent sends an Award message to one or more of the bidders; (4) the awardee(s) send(s) Result message(s) with task outcome details when done.

Since the Contract-Net Protocol, a variety of related bidding techniques have been developed, such as using computational auctions as a means for collecting bids and determining assignments (Wellman 1993). Furthermore, allocation presumes that agents can describe tasks and capabilities in ways that enable matching, and that agents know where to advertise their capabilities or announce their tasks, etc., and correspondingly there has been considerable emphasis on description languages for agents and the services they provide (Sycara et al. 2002), and for machinery for matchmaking and registration (Klusch et al. 2009).

Result-Sharing Algorithms. While task-sharing emphasizes moving tasks to agents suited for them, in some domains the "tasks" are inherently distributed, such as in distributed sensor networks where agents' local tasks are to interpret sensory inputs, with a collective objective of understanding global phenomena across the sensed region. The challenge agents face, then, is deciding which local results to share, with whom, and what to do with received results, in order to jointly solve the entire problem.

Algorithms for making such decisions range from more brute-force strategies of sharing everything with everyone, to protocols that, for example, pass a tentative global solution from agent to agent, where each agent augments it in turn. An early approach called Functionally-Accurate Cooperation (Lesser and Corkill 1981) exploited completeness to decide what results to share (those that couldn't be further improved locally) and locality to decide where to send them (to agents working on neighboring tasks).

Another important family of algorithms solve distributed constraint satisfaction problems (DCSPs) (Yokoo et al. 1998) and distributed constraint optimization problems (DCOPs) (Modi et al. 2006). Agents in this context are solving local problems whose solutions are constrained by solutions to other agents' problems, such as designing components of a complex artifact (like a car or computer) where the components must "fit" together. The algorithms are typically distributed versions of systematic and local search techniques (Russell and Norvig 2010).

Finally, a variation of this type of problem is the computational version of a social choice problem, where instead of collectively constructing a joint solution the agents are collectively selecting a particular complete solution, where each agent's expertise might lead it to favor a different solution. Computational strategies follow patterns familiar in human systems, including voting techniques, aggregating rankings among alternatives, or prediction markets (see the Techniques and Modalities chapter on Prediction Markets by Berea) where degrees of confidence are conveyed with agents' preferences (Brandt et al. 2013).

Action-Coordination Algorithms. A final category of algorithms focuses on the problem of coordination itself. Task-Sharing and Result-Sharing techniques largely assume that, while agents' local problems are related, the actions an agent takes to solve its problem(s) are not influenced by, and do not influence, the actions of others. When this is not the case, agents need to plan their joint activities to ensure that action choices by each agent do not interfere with other agents' contemporaneous actions, and enable (or do not prevent) other agents from taking useful actions downstream.

One method for coordination is obviously centralization, where a single decision-maker orchestrates the actions of all agents in concert, but such a strategy fails to scale to large problems and is slow to respond to dynamics that introduce localized disruptions to planned activities. Distributed algorithms typically sequence coordination and planning in some order. For example, when the space of conceivable interactions is large but few of them are likely to be realized, the agents should formulate their local plans independently and then use them to identify the small number of realizable interactions and coordinate over them (Cox and Durfee 2009). On the other hand, when the possible interactions are limited to a few pre-identifiable cases (e.g., roadway intersections), then coordination can precede local planning: agents agree on how to resolve interactions ahead of time (e.g., defining "right-of-way" rules for robots on collision courses) and then independently build local plans that respect them (Shoham and Tennenholtz 1995). Hybrid versions of these interleave planning and coordination by, for example, allowing agents to form abstract plans, using these plans as the basis of coordination decisions, and then allowing each agent to flexibly elaborate its abstract coordinated plan into details that fit its evolving local context (Clement et al. 2007).

Swarm Intelligence

When agents are individually intelligent, a small number of agents can accomplish a lot, but they need to be coordinated well. A counterpoint to this mindset instead utilizes much larger numbers of simpler agents, where miscoordination of a particular agent is less injurious to the performance of the collective. That is, agents are at a granularity more like insects than humans, and agent systems (hives, colonies) are orders of magnitude larger. Communication is simpler and more local, where agents might influence only a few of their neighbors, but over time such influences ripple through the collective as agents interact with agents who interact with other agents, and so on. The idea is that, while at any given time some agents might be failing to contribute, statistically enough agents are doing useful work for the system as a whole to function successfully.

This type of behavior appears, for example, in evolutionary and swarm algorithms. Genetic algorithms (Holland 1992) are a long-standing example of an evolutionary approach, where each generation by chance produces some less fit individuals, but favoring the fitter individuals in stochastically producing offspring by crossing their features means that the population's fitness as a whole climbs over successive generations. In this context, agents embody tentative solutions and use survival/reproduction to conduct a parallel stochastic local search.

An example swarm algorithm is Ant Colony Optimization (ACO) (Dorigo and DiCaro 1999), which manifests parallel stochastic local search patterned after how ants find good solutions to foraging problems (see also the Foundations chapter on biological networks by Moses and the Impact chapter on Superorganisms by Pavlic and Pratt). In ACO, a better solution corresponds to a more efficient path in a state space, where agents leave a trace behind at a visited state that fades over time, but that will encourage other agents to also visit that state. In a physical (ant) environment, this means that paths between desirable points (e.g., home and a food source) might initially be found through random exploration, but over time locations along shorter paths will be visited more frequently, increasing the intensity of trace in those states to further draw agents to that path. This same idea can be applied to conceptual spaces, where agents can use the ACO algorithm to find the best way to "connect the dots" through such a space (Dorigo et al. 2006).

Implications for Humans in Computation

Not surprisingly, because it has adopted and adapted algorithms inspired by how people work together, traditional cooperative problem-solving techniques correspond well with viewing humans as computational components, because those techniques are indeed patterned on human-human collaboration. Thus, at a high level, the strategies of task- and result-sharing, and of cooperative planning and organization, borrow from rather than contribute to insights for devising human computation systems.

Looking deeper, however, reveals potentially valuable insights because computational agents require specifications not needed by people. Research in cooperative problem solving has had to answer questions that are typically assumed away due to "commonsense" in human systems, such as how should tasks be decomposed into subtasks that will match "common" capabilities, when should a partial hypothesis be shared, and which actions need to be coordinated with actions of other agents? People often work in a context where answers to such questions are established and understood, but such boundaries are increasingly blurry and permeable. As humans become enlisted to engage in increasingly impromptu, virtual, and ephemeral collaborations, such previously-implicit decisions will require explicit reasoning of the kind already in computational systems, to bootstrap cooperative decisions until experience within a context (if it persists long enough) can engender appropriate conventions and expectations. This viewpoint adheres to traditional DAI assumptions about coarse granularity of problems, where agents perform sophisticated reasoning for considerable periods of time between communications, but where the combinations of agents working together evolve rapidly.

Communication is not the bottleneck it once was, however. Swarm-like behavior, involving frequent interactions between finer-grained reasoning activities, to statistically converge on emergent performance, has become an increasingly attractive model for cooperative problem solving among artificial agents, as well as among humans. Crowdsourcing falls into this sphere, where people are asked to perform brief cognitive tasks, with communication interleaved between those tasks. The tasks could involve labeling images (von Ahn and Dabbish 2004), folding proteins (Cooper et al. 2010), or transcribing a small piece of text (Parent and Eskenazi 2010), for example. Swarm algorithms no doubt can play a role in informing the development of mechanisms for guiding such forms of human computation, at least when that human computation aims to converge on a single "optimal" solution.

Arguably, however, traditional DAI concepts for task-sharing, result-sharing, and action coordination can also be gainfully mined for strategies to make such human computation systems more versatile. At the risk of oversimplifying both literatures, crowdsourcing has focused less, or at least less explicitly, on matching tasks to agents (people) than the DAI literature has. That is in some ways intentional: one of the benefits of crowdsourcing is that, by casting a wide net, contributions from those beyond "the usual suspects" can enrich the space of solutions explored. Yet, there are reasons to narrow the participants for some problems. The ESP game (von Ahn and Dabbish 2004), for example, implicitly wants participants who label images in "typical" ways (as reinforced by rewarding participants whose labels match each other); those who see things differently will not receive positive reinforcement. A DAI approach to such a problem would put the definition of capabilities required by participating agents at the forefront, and allocate tasks to such agents, whereas the art of creating a good crowdsourcing strategy is embodying such knowledge implicitly in the interface to most strongly attract the human computation that is expected to be most effective. Casting such human computation problems in DAI terms could complement existing crowdsourcing methodologies, to reveal implied expectations about desirable participants, which could in turn inform the development of an interface that differentially appeals to the desired subpopulation.

Such methodological advances appear crucial for moving human computation forward. To date, crowdsourcing emphasizes parallel, largely independent reasoning tasks, whether these are very fine-grained (e.g., labeling different images) or coarser (e.g., editing Wikipedia pages). Challenging cooperative problem solving applications involve *interdependencies*, where the quality of computation in one part of the network depends highly on the computational decisions in other parts. It remains to be seen how such problems scale to human computational models, where either the complexity of problems posed to humans must grow to include the non-local context of activities in remote parts of the system, or the complications due to interdependencies can be managed transparently to the humans by the computational infrastructure.

Examples of the kind of interaction envisaged by the previous paragraph have begun appearing in a variety of cooperative human-agent problem-solving technologies. For example, the DARPA Coordinators program (Hiatt et al. 2009) developed techniques to simplify human decision making by handling interdependencies between different (groups of) people through networked computational assistants (called coordinators). Because their coordinators would propagate critical coordination information behind the scenes, each human (group) would have constraints from other groups implicitly captured in its current local problem. For example, if two groups were to rendezvous, and one group was delayed, the other group's coordinator would convey the necessary information (e.g., to delay further movement) without unnecessary details. Hence, each group could focus only on its current problem locally and (seemingly) independently, because dependency constraints were reflected in the local problem specification.

A second example is in human-agent collaboration for on-demand scheduling of people to participate in virtual teams (Chen et al. 2010; Durfee et al. 2013), which assumes that some relevant human knowledge (about preferences that are evolving and/or private) could only be applied by involving humans in the scheduling process. Yet, the complexity of temporal and compatibility constraints are beyond human reasoning capabilities, so there is also a natural role for machines. The result is a system that allows computational agents and human agents to share initiative to converge on feasible and preferable schedules that neither group of agents could do alone. For example, the (machine) computational agents representing human specialists for a medical consultation could rapidly and comprehensively search for possible specialist teams that could hold an online consultation, and could optimize the timing of such a consultation. However, humans might be aware of friction among particular people that would be too sensitive to explicitly provide to the system but nonetheless should disqualify some of the feasible teams. Thus, human-agent collaboration allows all of the considerations to be brought to bear when converging on a consultation team.

In its brevity, this chapter could only scratch the surface both of the richness of ideas from the DAI field that could be mined to enhance and extend the development of human computational systems, and of the challenges and opportunities for doing so. The principal argument put forward can be summarized as follows: As systems involving humans increasingly view a human as a networked computational resource disembodied from a defining local social context, the techniques developed for characterizing and harnessing computational agents will become increasingly pertinent to such human computational systems.

Acknowledgements I would like to thank the editors for their helpful suggestions. This work was supported, in part, by the NSF under grant IIS-0964512.

References

Brandt F, Conitzer V, Endriss U (2013) Computational social choice. In: Weiss G (ed) Chapter 6 of Multiagent systems, pp 213–283

Chen W, Tang K, Mihalcik D, Tang Y, Durfee E, Dumas M (2010) Employing human knowledge to solve integrated coordination problems. In: International symposium on Collaborative Technologies and Systems (CTS), Chicago, IL, pp 285–294

Clement BJ, Durfee EH, Barrett AC (2007) Abstract reasoning for planning and coordination. J Artif Intell Res 28:453–515

Cooper S, Khatib F, Treuille A, Barbero J, Lee J, Beenen M, Leaver-Fay A, Baker D, Popović Z, Players F (2010) Predicting protein structures with a multiplayer online game. Nature 466:756–760

Cox JS, Durfee EH (2009) Efficient and distributable methods for solving the multiagent plan coordination problem. Multiagent Grid Syst 5(4):373–408

Davis R, Smith R (1983) Negotiation as a metaphor for distributed problem solving. Artif Intell 20:63–109

Dorigo M, DiCaro G (1999) Ant colony optimization: A new meta-heuristic. In: Proceedings of the 1999 conference on evolutionary computation, Washington, DC, pp 1470–1477

Dorigo M, Birattari M, Stutzle T (2006) Ant colony optimization: artificial ants as a computational intelligence technique. IEEE Comput Intell Mag 1(4):28–39

Durfee EH, Boerkoel JC Jr., Sleight J (2013) Using hybrid scheduling for the semi-autonomous formation of expert teams. Future Gener Comput Syst

Hiatt LM, Zimmerman TL, Smith SF, Simmons R (2009) Strengthening schedules through uncertainty analysis. In: Proceedings of the twenty-first international joint conference on artificial intelligence (IJCAI-09), Pasadena, CA, pp 175–180

Holland J (1992) Adaptation in natural and artificial systems. MIT Press, Cambridge

Klusch M, Fries B, Sycara KP (2009) OWLS-MX: a hybrid semantic web service matchmaker for OWL-S services. J Web Sem 7(2):121–133

Lesser VR, Corkill DD (1981) Functionally accurate, cooperative distributed systems. IEEE Trans Syst Man Cybern 11:81–96

Modi PJ, Shen W-M, Tambe M, Yokoo M (2006) ADOPT: asynchronous distributed constraint optimization with quality guarantees. Artif Intell 161:149–180

Parent G, Eskenazi M (2010) Toward better crowdsourced transcription: transcription of a year of the *Let's Go* bus information system data. IEEE Spok Lang Technol Workshop. Proceedings of the 2010 IEEE Workshop on Spoken Language Technology, Berkeley, CA, 312–317

Russell S, Norvig P (2010) Artificial intelligence: a modern approach (3rd edn). Chapter 6. Prentice Hall, Boston

Shoham Y, Tennenholtz M (1995) On social laws for artificial agent societies: off-line design. Artif Intell 73:231–252

Sycara K, Widoff S, Klusch M, Lu J (2002) Larks: dynamic matchmaking among heterogeneous software agents in cyberspace. Auton Agents Multi-Agent Syst 5(2):173–203 von Ahn L, Dabbish L (2004) Labeling images with a computer game. In: Proceedings of the SIGCHI conference on human factors in computing systems (CHI'04), Vienna, Austria, pp 319–326

Weiss G (ed) (2013) Multiagent systems: A modern approach to distributed artificial intelligence, 2nd edn. MIT Press, Cambridge

Wellman MP (1993) A market-oriented programming environment and its application to distributed multicommodity flow problems. J Artif Intell Res 1:1–23

Yokoo M, Durfee EH, Ishida T, Kuwabara K (1998) The distributed constraint satisfaction problem: formalization and algorithms. IEEE Trans Knowledge Data Eng 10:673–685

Human-Based Evolutionary Computing

Jeffrey V. Nickerson

Introduction

Evolution explains the way the natural world changes over time. It can also explain the way the artificial world changes, the way ideas replicate, alter, and merge (Campbell 1960; Dawkins 1983). This has led to a family of related computer procedures called *evolutionary algorithms* (Fogel 1994; Holland 1975). These algorithms are being used to design products, generate art, and solve mathematical problems (Bentley and Corne 2002; Eiben and Smith 2003; Goldberg 1989).

While these algorithms run on computers, they also can be performed by people: members of a crowd can create designs, modify and combine designs, and evaluate designs. Such *human-based evolutionary algorithms* are useful when many different ideas are needed, and human cognition is called for.

Overview of Evolutionary Algorithms

Evolutionary algorithms rely on several concepts. The algorithms search a large space of possible solutions that together form a *population*. Each solution is represented as a *genotype*: a bit string or a more complex data structure. These solutions are evaluated with respect to their *fitness*. Some members of a population live and some die through *selection.*

New members of a population, the next generation, are born through replication and modification. Modification takes two forms. One is *mutation*: a single solution is copied, and then altered. The other is *recombination*: parts of two solutions, the parents, are swapped to create offspring. The parents can be chosen through a

J.V. Nickerson (✉)
Stevens Institute of Technology, Hoboken, USA
e-mail: jnickerson@stevens.edu

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_51,

```
create a first generation of solutions
evaluate the solutions
while generations continue to improve
   do as many matings as are needed
            choose a pair of solutions
            combine them to yield offspring
            mutate the offspring
            evaluate the offspring
   select the next generation from the parents and offspring
output the best solutions
```

Fig. 1 The loop of an evolutionary algorithm

process called *tournament selection.* Two random potential suitors are compared. Another two suitors are also compared. The winners of the comparisons mate (Goldberg 1989).

These concepts have been explored by a large community of researchers over the last several decades (De Jong 2006; Eiben and Smith 2003; Reeves 2003; Zhou et al. 2011). Figure 1 shows the pseudocode of a typical implementation.

More details on evaluation, modification and selection processes follow.

Evaluation Processes

Each solution is evaluated. The evaluations are used for choosing the solutions to be recombined, and for deciding which solutions will survive into the next generation. When a problem has one primary objective the evaluation is straightforward. For example, if the lowest cost design of a product is sought, then the costs of each of its components are added together.

Evolutionary algorithms are particularly suited to solving problems with multiple objectives. For a product design, the objectives might include minimizing cost, maximizing speed, and maximizing quality. These goals involve trade-offs, but there are some designs that are better than others along all three dimensions. These dominating solutions constitute the *Pareto front*. Evolutionary algorithms run in such a way that each generation presses solutions toward this front. When little progress is made on improving along the chosen objectives, the algorithm then focuses on evenly distributing solutions along the front (Deb 2001). For example, this might yield viable products at different levels of price, quality, and speed.

Modification Processes

In evolutionary computing, modifications take place using mutation and recombination. How are solutions represented in the first place? They can be mapped onto bit strings. But sometimes they are represented at a higher level as parameters, or as trees, or grammars, or graphs. Depending on the representation, mutation and recombination are programmed in different ways.

When bit strings are used, mutation will flip random bits on and off. Recombination will take subsequences of bits from each of two solutions and swap them; this is called a *crossover*. When graph structures are used, mutations will delete edges, and recombinations will swap sub-graphs.

The ratios of mutation and recombination are adjusted according to the problem domain. For some domains, pure mutation provides the fastest optimization. For other problems, recombination works faster, because it preserves desirable sets of features in parent solutions (Spears 1992; Stadler and Wagner 1997).

Selection Processes

Modifications may be destructive. Flipping a bit or deleting an edge may degrade the solution: children are not necessarily fitter than parents.

A ratchet—a tool that secures already accomplished work—is needed, so that every generation is at least as good as its predecessor. Better modifications can easily be kept and worse modifications can be dropped from the population. But what if all the children are worse than a particular parent? The selection process can replicate the parent unchanged to the next generation, insuring that good solutions will not be forgotten. This process of promoting the best solutions to the next generation is called *elitism* (De Jong 1975; Deb 2001). With this mechanism in place, mutation and recombination can focus on generating variety.

Memetic Techniques

In the typical evolutionary algorithm, mutation and recombination are blind, a random shift of a number, or a random mixing of two parents. Memetic algorithms inject domain knowledge into the algorithm. Heuristics are used to insure children are better than parents. For example, once parents are picked, the space of possible children might be systematically explored. Or an effort might be made to augment the combination of two parents with other solutions. Or the objectives themselves might be changed to surface fresh solutions (Moscato and Cotta 2010).

When people rather than computers implement evolutionary algorithms, memetic processes occur naturally. For example, when people are asked to combine things, they will make use of past knowledge, and therefore introduce external ideas into the process. And when people are asked to modify an idea, it is unlikely the modification will be random. In this way, human-based evolutionary algorithms encompass many ideas described in the memetic algorithms literature. The role of humans is discussed next in more detail: in particular, how people evaluate and modify solutions.

Evaluation by Humans

An evolutionary algorithm might be used to generate a floor plan for a building. But programming a function that reacts to plans the way people do is difficult. Better to let people perform the evaluation of the computer-generated solutions (Gero and Kazakov 1997). Such a process is called an *interactive genetic algorithm* (Quiroz et al. 2009; Takagi 2001).

Many systems that use this approach were created more than a decade ago, well before the advent of crowd work (Gero 1996). Now such systems can be run using crowds: evaluation can be completed faster by using more people. However, some tasks, such as evaluating floor plans, demand either special expertise (say, understanding flow in a building) or particular values (say, the preferences of a client with respect to window placement).

Preferences can also be used to drive the generation of art: In a system called PicBreeder images are morphed and the user steers the evolution according to their own aesthetic (Secretan et al. 2011).

More generally, evaluation can be performed without an explicit objective by asking people to choose between two solutions. But asking for comparisons versus asking for absolute ratings may yield different results (Füller et al. 2010). And experts will rate things differently than novices (Welsh 2012).

Modification by Humans

Instead of directing a computer to mutate a string, we can ask a person to improve an idea. Instead of directing a computer to swap the bits of two strings, we can ask a person to combine the best elements of two ideas. While computers will often generate infeasible solutions, people will tend to generate feasible solutions because they have a sense of context, background knowledge, and cognitive skills.

Combining concepts leads to innovation, claim many (Fleming et al. 2007; Kijkuit and Van Den Ende 2007; Perry-Smith and Shalley 2003). Original brainstorming techniques encouraged idea combination (Osborn 1953), and in some experimental contexts combination generates better alternatives than simple modification (Kohn et al. 2011). In the wild, combination also plays a role: digital media makes it easy to remix ideas and designs (Kyriakou et al. 2012; Lessig 2008; Seneviratne and Monroy-Hernandez 2010; Tuite et al. 2012). Invention is arguably driven by context (Gabora 2005; Perkins 2000). A new idea can occur because of an experience. Then why not use people to perform modification as part of an evolutionary system?

The earliest expression of this idea (Kosorukoff 2001) was followed by an experiment showing that, for a particular problem, an evolutionary algorithm with human evaluation and modification converged in fewer generations than an algorithm using computer modification (Cheng and Kosorukoff 2004). The research also suggested the potential of virtual communities organized by evolutionary algorithms (Kosorukoff and Goldberg 2001).

Since the advent of crowd work (Kittur et al. 2013), researchers have performed experiments on a range of open ended problems: fixing an oil spill (Nickerson and Sakamoto 2010) design chairs (Yu 2011; Yu and Nickerson 2011), designing clocks (Yu and Nickerson 2013). In these experiments, the crowd performed most of the functions of evolutionary algorithms. The crowd first generated the initial population of ideas. For the oil spill problem, the crowd generated ideas by writing; for the product designs, by sketching. Then another crowd evaluated the ideas. A computer used these evaluations to select ideas for combination, which were presented to yet another crowd for combination. The algorithm was run for three generations.

Ideas improved, as measured by the crowd. Recombination created new ideas. Many of these ideas were worse than the parents. But the selection process culled such results and kept the successful combinations, leading to a better set of designs at the end than at the beginning.

How does the crowd choose features of original ideas in the combination process? In one experiment involving combinations of sketches the crowd members were attracted to novel but practical features (Yu and Sakamoto 2011). They were effectively performing evaluation even as they generated variety through recombination. This process is similar to a memetic algorithm, in that recombination is not done blindly, but makes use of domain knowledge. Unlike computers, though, people can add ideas collected over a lifetime to the system. The introduction of new ideas can happen in a variety of ways: for example, members of the crowd can be asked to critique solutions, and then other members asked to modify the ideas in response to the critique (Tanaka et al. 2011).

There are many different ways that crowds and computers can cooperate in evolutionary computing. Crowds can perform all functions, or just a subset of functions, leaving the rest to computers. The modification steps of the algorithm can include only mutation, or only recombination, or both. Each of the algorithm's steps, because they are run many times, can benefit from both machine learning and human learning. In sum, there is a family of possible techniques ranging from the fully computer-based evolutionary algorithms to fully human-based evolutionary algorithms, with the many parts of this spectrum still unexplored (Yu and Nickerson 2013).

Discussion and Final Thoughts

Evolutionary algorithms, as implemented by computers, have been used to solve a wide range of problems. But they can be computationally expensive, because many solutions need to be generated and assessed for hundreds of generations (Deb 2001).

And they are not a cure-all: they are essentially equivalent to random search if they don't use specific domain knowledge (Wolpert and Macready 1997). Like many metaheuristic algorithms they work well when objectives can be simply formulated and solutions can computationally evaluated (Kochenberger 2003).

The solution representations need to be conducive toward automated mutation and recombination. Some solutions are harder to evaluate than others, and harder to combine. In these cases, human-based evolutionary algorithms provide an alternative. People can bring knowledge and experience to the modification and evaluation processes, and thereby reduce the number of generations necessary to find feasible solutions.

There is much we don't know. While systems in which people perform the evaluation step have been extensively explored, systems in which people perform modification are rare, because such systems require an available, scalable human resource pool. In the past such pools were difficult to assemble, but now crowd work platforms are growing. This presents an opportunity for further experiments.

How do people combine things? People do not perform crossover operations like computers do. Instead, people create hybrids by selectively combining features (Wisniewski 1997). The result is not merely the sum of its constituents (Hampton 1988; Osherson and Smith 1981). One interesting stream of research argues that the idea recombination process can be modeled using quantum theory, which can represent many aspects of concept combination—for example, the emergence of new features—that classical probability theory cannot (Aerts et al. 2013).

People bring to the combination process a sense of context, which is determined by their expertise. Crowd members will not necessarily be expert in a particular problem domain: expertise has to be found, or created through training, or synthesized (Page 2012).

Training may be explicit, or it may happen naturally through the repetition of idea generation and evaluation steps. Once trained, crowd workers will bring their increasing expertise to related problems (Nickerson 2013). Such processes take place naturally in the traditional work place, as skilled workers increase the productivity of individual companies and industries overall. We do not know yet how crowd workforces will develop, but it is possible these workforces will build knowledge as crowd members modify and combine their ideas, as they share techniques and tools with each other. Then these human and machine collectives will become more creative. They will evolve.

Acknowledgements This material is based upon work supported by the National Science Foundation under grants IIS-0855995 and IIS-0968561.

References

Aerts D, Gabora L, Sozzo S (2013) Concepts and Their Dynamics: A Quantum-Theoretic Modeling of Human Thought, Topics in Cognitive Science 5(5):737–772.

Bentley PJ, Corne DW (2002) Creative evolutionary systems. Morgan Kaufmann, San Francisco

Campbell DT (1960) Blind variation and selective retention in creative thought as in other knowledge processes. Psychol Rev 67:380

Cheng CD, Kosorukoff A (2004) Interactive one-max problem allows to compare the performance of interactive and human-based genetic algorithms. In: Proceedings of genetic and evolutionary computation–GECCO 2004, Seattle, Springer

Dawkins R (1983) Universal Darwinism. In: Bendall DS (ed) Evolution from molecules to men. Cambridge University Press, Cambridge, pp 403–425
De Jong KA (1975) Analysis of the behavior of a class of genetic adaptive systems. (Ph.D.), University of Michigan
De Jong KA (2006) Evolutionary computation: a unified approach. MIT press, Cambridge
Deb K (2001) Multi-objective optimization using evolutionary algorithms, 1st edn. Wiley, Chichester/New York
Eiben AE, Smith JE (2003) Introduction to evolutionary computing. Springer, New York
Fleming L, Mingo S, Chen D (2007) Collaborative brokerage, generative creativity, and creative success. Adm Sci Q 52(3):443–475
Fogel DB (1994) An introduction to simulated evolutionary optimization. Neural Netw, IEEE Trans 5(1):3–14
Füller J, Möslein KM, Hutter K, Haller JBA (2010) Evaluation games–how to make the crowd your Jury. In: Fähnrich KP, Franczyk B (eds) Lecture Notes in Informatics (LNI 175), Proceedings of the Informatik 2010: Service Science – Neue Perspektiven für die Informatik”. Leipzig 2010, pp 955–960
Gabora L (2005) Creative thought as a non Darwinian evolutionary process. J Creat Behav 39(4):262–283
Gendreau M. and Potvin, J-Y (2010), Handbook of Metaheuristics, Springer, New York
Gero JS (1996) Creativity, emergence and evolution in design. Knowl Based Syst 9(7):435–448
Gero JS, Kazakov VA (1997) Learning and re-using information in space layout planning problems using genetic engineering. Artif Intell Eng 11(3):329–334
Goldberg DE (1989) Genetic algorithms in search, optimization, and machine learning. Addison-Wesley, Reading
Hampton JA (1988) Overextension of conjunctive concepts: evidence for a unitary model of concept typicality and class inclusion. J Exp Psychol Learn Mem Cogn 14(1):12–32
Holland JH (1975) Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence. University of Michigan Press, Ann Arbor
Kijkuit B, Van Den Ende J (2007) The organizational life of an idea: integrating social network, creativity and decision-making perspectives. J Manag Stud 44(6):863–882
Kittur A, Nickerson JV, Bernstein MS, Gerber EM, Shaw AD, Zimmerman J, Lease M, Horton JJ (2013) The future of crowd work. In: Proceedings of 2013 ACM conference on Computer Supported Collaborative Work (CSCW ’13), San Antonio
Kohn NW, Paulus PB, Choi YH (2011) Building on the ideas of others: an examination of the idea combination process. J Exp Soc Psychol 47(3):554–561
Kosorukoff AL (2001) Human based genetic algorithm. In: Proceedings of systems, man, and cybernetics, 2001 IEEE international conference on, IEEE, Tucson
Kosorukoff AL, Goldberg DE (2001) Genetic algorithms for social innovation and creativity
Kyriakou H, Engelhardt S, Nickerson JV (2012) Networks of innovation in 3D printing. Paper presented at the workshop on information in networks
Lessig L (2008) Remix: Making art and commerce thrive in the hybrid economy. Penguin Pr, New York
Moscato P, Cotta C (2010) A modern introduction to memetic algorithms. In: Handbook of metaheuristics. Springer, New York, pp 141–183
Nickerson JV (2013) Crowd work and collective learning. In: Littlejohn A, Margaryan A (eds) Technology-enhanced professional learning. Routledge, New York
Nickerson JV, Sakamoto Y (2010) Crowdsourcing creativity: combining ideas in networks. In: Paper presented at the workshop on information in networks
Osborn AF (1953) Applied imagination. Scribner, New York
Osherson DN, Smith EE (1981) On the adequacy of prototype theory as a theory of concepts. Cognition 9(1):35–58
Page SE (2012) Aggregation in agent-based models of economies. Knowl Eng Rev 27(02):151–162

Perkins DN (2000) Archimedes' bathtub: the art and logic of breakthrough thinking. WW Norton, New York

Perry-Smith JE, Shalley CE (2003) The social side of creativity: a static and dynamic social network perspective. Acad Manag Rev 28:89–106

Quiroz JC, Louis SJ, Banerjee A, Dascalu SM (2009) Towards creative design using collaborative interactive genetic algorithms. In: Proceedings of evolutionary computation, 2009. CEC'09. IEEE congress on, IEEE

Reeves C (2003) Genetic algorithms In: Handbook of metaheuristics. Springer, pp 55–82

Secretan J, Beato N, D'Ambrosio DB, Rodriguez A, Campbell A, Folsom-Kovarik JT, Stanley KO (2011) Picbreeder: a case study in collaborative evolutionary exploration of design space. Evol Comput 19(3):373–403

Seneviratne O, Monroy-Hernandez A (2010) Remix culture on the web: a survey of content reuse on different user-generated content websites. Paper presented at the proceedings of WebSci10: extending the frontiers of society on-line

Spears WM (1992) Crossover or mutation. Found Genet algorithms 2:221–237

Stadler PF, Wagner GP (1997) Algebraic theory of recombination spaces. Evol Comput 5(3):241–275

Takagi H (2001) Interactive evolutionary computation: fusion of the capabilities of EC optimization and human evaluation. Proc IEEE 89(9):1275–1296

Tanaka Y, Sakamoto Y, Kusumi T (2011) Conceptual combination versus critical combination: Devising creative solutions using the sequential application of crowds. In: Proceedings of annual meeting of the cognitive science society, Boston

Tuite K, Smith AM, Studio EI (2012) Emergent remix culture in an anonymous collaborative art system. In: Proceedings of eighth artificial intelligence and interactive digital entertainment conference, Palo Alto

Welsh MB (2012) Expertise and the wisdom of crowds: whose judgments to trust and when. In: Proceedings of annual conference, Sapporo, Japan

Wisniewski EJ (1997) When concepts combine. Psychon Bull Rev 4(2):167–183

Wolpert DH, Macready WG (1997) No free lunch theorems for optimization. Evol Comput, IEEE Trans 1(1):67–82

Yu L (2011) Crowd idea generation. (Ph.D.), Stevens Institute of Technology

Yu L, Nickerson JV (2011) Cooks or cobblers? Crowd creativity through combination. In: Proceedings of the 29th CHI conference on human factors in computing systems, ACM Press, Vancouver

Yu L, Nickerson JV (2013) An internet-scale idea generation system. ACM Trans Interact Intell Syst 3(1), Article 2

Yu L, Sakamoto Y (2011) Feature selection in crowd creativity. In: Schmorrow DD, Fidopiastis CM (eds) Foundations of augmented cognition. Directing the future of adaptive systems. Springer, New York pp 383–392

Zhou A, Qu B-Y, Li H, Zhao S-Z, Suganthan PN, Zhang Q (2011) Multiobjective evolutionary algorithms: a survey of the state of the art. Swarm Evol Comput 1(1):32–49

Algorithms for Social Recommendation

Ido Guy

Introduction

Recommender Systems (Jannach et al. 2010; Resnick and Varian 1997; Ricci et al. 2011) have become highly popular in recent years. Films and series on Netflix, products on Amazon, videos on YouTube, and hotels on Trip Advisor are just a few leading examples of recent applications of recommender systems in leading web sites. Social recommender systems (SRSs) (Guy and Carmel 2011) are recommender systems that target the social media domain (Guy et al. 2010a). As social overload increases over users of social media, with millions of tweets, feed items, blogs, photos, and bookmarks created every day, and as millions of users are more active than ever before, establishing communities and forming online relationships, social media users are having a greater challenge locating the information most relevant to them, while social media sites have a greater challenge when trying to attract new users and maintain existing ones. SRSs aim to alleviate this challenge by applying various techniques that filter the information for a user to the most attractive and relevant pieces, usually on a personal basis.

Social recommender systems cover many areas of recommendations and take advantage of many types of data and metadata. Areas for SRSs span from recommending content to consume (Guy et al. 2009b, 2010c) and to produce (Geyer et al. 2008), through recommending people (Chen et al. 2009; Guy et al. 2009a, 2011b), tags (Sigurbjörnsson and Van Zwol 2008), and communities (Chen et al. 2008), to recommending news items in real-time social streams (Berkovsky et al. 2012; Guy et al. 2011a; Paek et al. 2010). Data used for recommendations takes advantage of the unique and public nature of social media, such as online relationships (Guy et al. 2009b), tags (Guy et al. 2010c; Sen et al. 2009), comments, and votes (Lerman 2006).

I. Guy (✉)
IBM Research, Haifa 31905, Israel
e-mail: ido@il.ibm.com

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_52,

In this chapter, we delve into a few prominent algorithms used for social recommender systems that serve as good examples for the uniqueness of the domain and the pending challenges. We focus on recommendation of people, recommendation of mixed content, and recommendation of news items in social streams. The rest of the chapter discusses in detail these three areas with examples taken from our previous research and then concludes with discussing common features and future directions.

Specifically, we discuss recommendation of people on social network sites—both people to connect with and strangers to get to know. We show that recommender systems play a key role in building the underlying graph of people that lies at the core of every social network site. We also show that people recommendation can contribute to one's social capital by pointing at people "worth" to know. We additionally show that the explicit network built can be effectively used to enhance recommendation of social media content. The latter can also benefit from tags, which are frequently used to annotate social media content and people. Additionally, in the real-time web era, activities are shared through real-time social streams that appear in very high pace and require even more sophisticated recommendation algorithms. We argue that the recommendation algorithms we show in this chapter can play a key role in future forms of human computation, helping to recruit and motivate the crowd, assigning the tasks, and reaching the best possible outcome.

People Recommendation

In this section, we discuss the area of people recommendation. Recommending people to people is a specific sub-domain of recommender systems and poses its own unique challenges, driven by the fact that the recommended entity is a person. We focus on two types of recommendations that are quite different from one another. The first recommends people the user already knows in order to connect in a social network site. We show that such recommendations are instrumental to building the network. The second type recommends strangers and is a more "speculative" type of recommender, which aims at increasing one's social reach and social capital. We believe that people recommendation would play a key role in future human computation paradigms, helping to recruit the right people for a task and motivating them appropriately. Recruitment can be of both people the user knows and of strangers, as is the case with people recommendation within SNSs.

The Value of Recommending People to Connect with

With the proliferation of social networks sites (SNSs) (Boyd and Ellison 2007), allowing users to connect with each other by sending and accepting invitations to and from one another, the need for effective people recommendation systems has

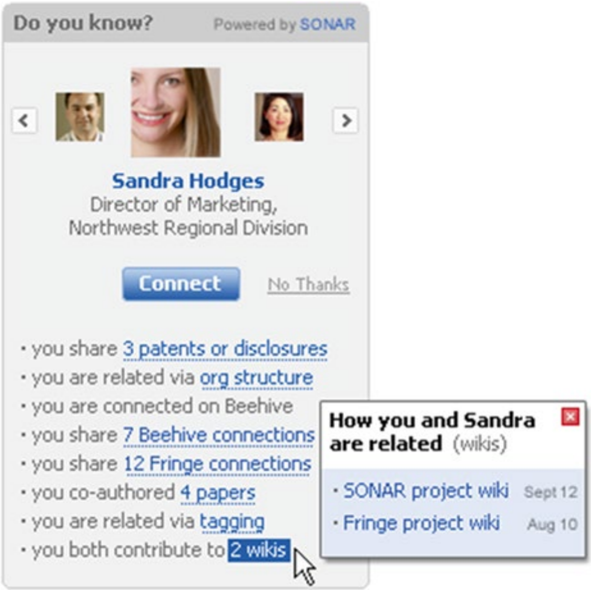


Fig. 1 The "Do You Know?" widget for people recommendation

become evident. The area of people recommendation strongly ties to social matching (Terveen and McDonald 2007), which discusses recommender systems that recommend people to people and the uniqueness of this type of recommendation with regards to privacy, reputation, trust, and interpersonal attraction. In parallel to the emergence of "people you may know" widgets on leading SNSs, such as Facebook and LinkedIn, the work on the "do you know?" (DYK) widget (Guy et al. 2009a) was the first to study the topic. The widget recommended people to connect within the enterprise, based on a rich set of implicit people-to-people relationships, derived by SONAR, a social aggregation system that collects relationships across a multitude of data sources (Guy et al. 2008).

Figure 1 illustrates the DYK widget. It enables the user to scroll though a list of recommended people one at a time. The list of people is retrieved by requesting the top 100 related people to the user as retrieved by SONAR. The recommended people are presented in descending order of relationship score. Hence, the first recommendations are those with whom the user is likely to have the strongest familiarity level, but are not connected to her yet. For each recommendation, the widget presents a picture, the person's name as a link to her profile, and a summarized list of all available evidences of the relationship to the user. These evidences are retrieved from SONAR as well. The summary would state, for example, that the recommended person and the user wrote two papers together, commented on each other's blogs three times, tagged each other in the people tagging application, share a manager, have ten mutual connections, and so forth. While hovering over each summarized evidence item, a popup appears which includes a detailed list of evidences with relevant links. In Fig. 1, the popup shows the two wikis the users share. After viewing a recommended person, the user can decide to scroll to the next or previous person,

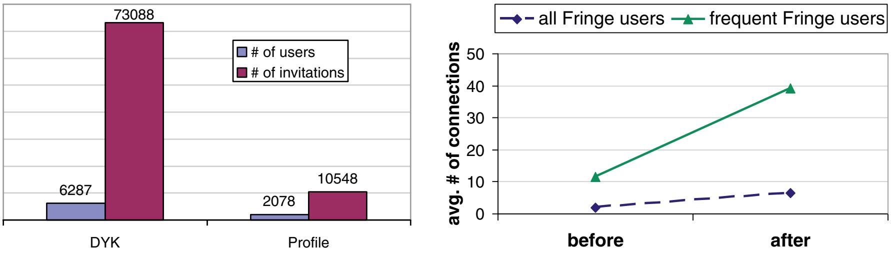


Fig. 2 Effects of the DYK widget on the Fringe Site. On the left, the number of invitations sent and the number of user sending invitations through the DYK widget, compared to the regular profile mechanism. On the right, the increase in number of connections for frequent and regular users

to invite the current person to connect, or to remove the current recommendation. The user invites the recommended person by clicking on a connect button below the name of the person. This will bring up a default invitation text which users can edit according to their wishes. The invitation is then sent out to the recommended person through email and the DYK proceeds to the next recommended person.

The DYK widget was deployed as part of the homepage of a widely used enterprise SNS, called Fringe (Farrell et al. 2007). A field study, inspecting the widget's usage over 4 months, reported a dramatic effect on the site, both in terms of number of invitations sent and in terms of number of people who send invitations (Guy et al. 2009a).

Figure 2 illustrates these results. The diagram on the left compares the usage of the DYK widget with the usage of the regular mechanism of inviting through people profiles during the 4 month period. First, it shows the number of invitations sent from the DYK widget compared to the number of invitations sent from others' profiles—73,088 invitations were sent through the DYK widget, while only 10,548 were sent through profiles. Acceptance rate was exactly 60 % both for invitations sent from DYK and for invitations sent from profiles. The identical acceptance rate indicates that while the DYK widget provoked much more invitations, their quality in terms of acceptance rate remained equal to that of the usual profile-based mechanism. In addition, while 6,287 users initiated invitations though the DYK widget, only 2,048 sent invitations through profiles. These high differences between DYK invitations and profile invitations stand in contrast to the fact that the homepage was accessed 79,108 times, while profile pages were accessed 91,964 times. The overall increase in the number of invitations in Fringe was 278 %, which resulted in an overall increase of 230 % in confirmed connections. The increase in people who sent at least one invitation was 150 %. This sharp increase took place within a period of 4 months and after the Fringe "friending" feature had been available for 15 months without recommendations.

The diagram on the right of Fig. 2 shows the substantial change in the average number of connections per user after the DYK was introduced. Frequent users, who accessed Fringe for at least 10 days during the 4-month period, had 11.6 connections on average before the DYK (stdev 22.7, median 4, max 198), and ended up with 39.1 connections on average (stdev 37.2, median 30, max 389). All Fringe users together had 2.0 on average at the beginning of the period (stdev 6.3, median 1, max 198) and 6.6 on average after the period (stdev 12.2, median 2, max 389).

Comparing People Recommendation Algorithms

In the study above, we witnessed the strong effect of people recommendation on both the number of invitations and the number of people who send invitations in an SNS. We next asked ourselves about the contribution of the aggregation algorithm to the quality of recommendations. We wanted to compare the SONAR-based algorithms with other more traditional algorithms such as friend-of-a-friend and content similarity. To this end, in a subsequent experiment (Chen et al. 2009), we examined four algorithms for people recommendation on another enterprise SNS, called Beehive. The four algorithms were:

1. Content Matching (CM)—measuring the cosine similarity between two users based the word vectors (calculated using TF-IDF) representing the content they created on Beehive. Intuitively this means u and v would be considered similar if they share many common keywords in their associated content, and even more so if only a few users share those keywords. Users similar to the recipient user u were recommended in decreasing order of similarity. We also analyzed newer and more sophisticated content similarity algorithms, including Latent Semantic Analysis (Deerwester et al. 1990) and Probabilistic Latent Semantic Analysis (Hofmann 1999). However, in a preliminary test they did not yield significantly better results.
2. Content-plus-Link (CplusL)—enhances the CM algorithm with social link information derived from the Beehive social network structure. The motivation behind this algorithm was that by disclosing a network path to a weak tie or unknown person, the recipient of the recommendation will be more likely to accept the recommendation. CplusL computes similarity the same way CM does, but boosts results by 50 % if a valid social link exists between the users (see (Chen et al. 2009) for the full definition of a valid social link). On average 77.8 % of the top ten recommendations computed with this algorithm in our experiments contained valid social link information.
3. Friend-of-Friend (FoF)—this popular algorithm, used by many social network sites, recommends people based on the number of common friends they share with the user.
4. SONAR—recommends people based on their implicit relationships mined by the SONAR aggregation system as explained above.

We experimented with these algorithms through a user survey with 230 Beehive users. Participants received 12 people recommendations and were asked to indicate for each individual whether they know this individual and whether it is a good recommendation. Results are shown in Fig. 3.

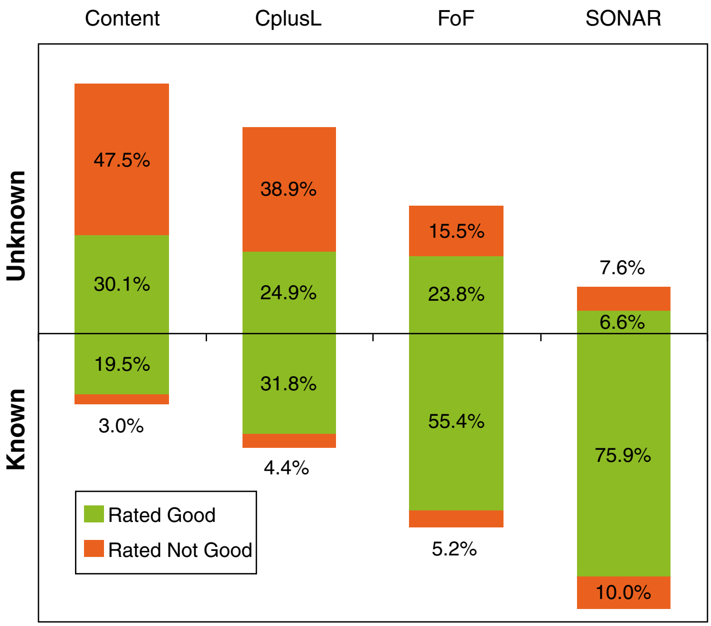


Fig. 3 Beehive recommendation rating based on four algorithms

As expected, the pure content matching algorithm recommended mostly unknown people. SONAR and FoF recommended mostly known people. On average each user already knew 85.9 % of the people recommended by SONAR, followed by the FoF algorithm with 60.6 %. In contrast, users only knew 36.2 % of the recommendations from the content-plus-link algorithm, and 22.5 % of those from the content matching algorithm ($F[3,711]=213.5,\ p<.001$). Post-hoc comparison (LSD) showed that the percentages for each algorithm were significantly different from each other ($p<.001$). These results confirm the intuition that the more relationship information an algorithm leverages, the more known people it would recommend.

Overall, our users rated 82.5 % of the SONAR recommendations as good, followed by 79.2 % for the friend-of-friend, 56.7 % for the content-plus-link and 49.6 % for the content matching algorithm ($F[3,705]=69.1,\ p<.001$). While there was no significant difference between SONAR and friend-of-friend, post-hoc comparison (LSD) showed that they have a significantly higher percentage of good recommendations than the two content-based algorithms ($p<.001$). Also, the percentage of "good" recommendations from the content-plus-link algorithm was significantly higher than basic content matching ($p<.005$). Overall, this suggests that the more known recommendations an algorithm produces, the more likely users are to consider those recommendations good.

When looking only at recommendations of known people (Fig. 3, below the center line), we can see that most of those recommendations were considered good for all algorithms (around 90 % for each algorithm). In other words, users considered recommendations of known people to be good, no matter how they were computed.

In contrast, the situation for unknown recommendations is very different in that more recommendations are considered to be not good. The number of "not good" recommendations increases from right to left, i.e. the content-based algorithm produces the highest number of recommendations not considered good. One could argue that the more strangers an algorithm recommends, the more likely users will reject or not like the recommendations.

Overall, our experimentation showed the superiority of the SONAR algorithm for recommending people to connect with in an SNS. Aggregation of data across social media sites helps achieve very accurate results in terms of who knows whom and who should connect to whom.

Stranger Recommendation

The results of the latter study also sparked our interest in recommending strangers in the enterprise. While the accuracy of such recommendations is likely to be substantially lower, the value in each good recommendation may be much higher, since a new relationship that did not exist before may be formed. We thus developed the StrangerRec system (Guy et al. 2011b), which recommends employees the user is not familiar with, but may be of interest based on common behavior on enterprise social media, such as usage of the same tags or commenting on the same blog entries.

The task of recommending unfamiliar yet interesting people is quite different from "regular" recommendation of familiar people. StrangerRS focused more on discovery and exposure to new people and less on facilitating connection within an SNS. It aimed at satisfying two rather conflicting goals: on the one hand, the recommended person should not be familiar to the user, and, on the other hand, that person should be of some interest. While accuracy of recommendations that satisfy both goals might not be high, we argue that the potential serendipity and "surprise effect" in getting a fortuitous recommendation of an interesting new person in the organization may compensate for lower accuracy (McNee et al. 2006).

To implement our stranger recommender, we used SONAR's capability to distinguish among relationship types: (1) *familiarity* (Guy et al. 2008)—people the user knows; (2) *similarity* (Guy et al. 2010b)—people with whom the user shares common interests; and (3) *interest* (Jacovi et al. 2011)—people in whom the user is interested (or interested in the user for the reverse direction—this is the only type of relationship that is asymmetric). Particularly, we used the familiarity and similarity relationships in StrangerRS. Familiarity relationships were derived based on indicators for knowing a person: either explicit indicators (being connected on an SNS or a connection through the organizational chart) or implicit indicators (co-authoring a wiki page, being member in the same project, sharing a file, etc.). Similarity relationships were derived based on common activity in social media, which serves as an

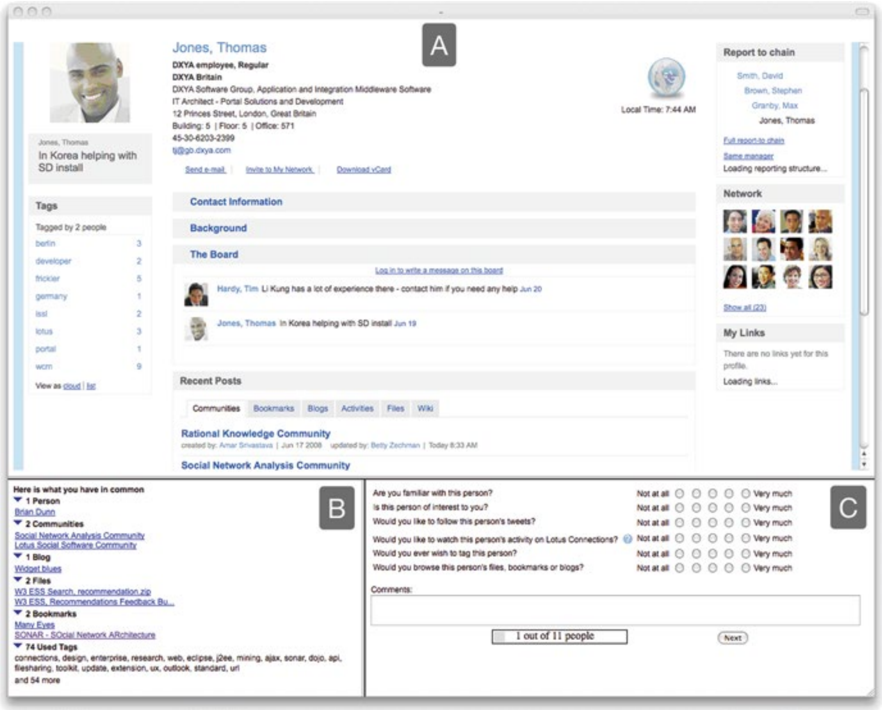


Fig. 4 User interface of the stranger recommender systems

implicit indicator for mutual interests. For example, usage of the same tag, being tagged by the same tag within the people tagging application, bookmarking the same page, membership in the same community, or commenting on the same blog, are all considered indicators of similarity relationships. To generate the recommendations, we applied a *social network composition*, i.e., a composition of two social network types. In this specific case, we subtracted the user's familiarity network (i.e., list of people s/he knows) from the similarity network (list of people s/he has common interests with) to suggest strangers who may be of interest. The rich underlying aggregation model ensured that we can derive many types of similarity relationships, while also being able to effectively filter out people the user is already familiar with.

Figure 4 demonstrates the user interface of StrangerRS. Part A shows the profile page of the recommended employee. As opposed to the DYK widget, where only few details (photo, name, job title) were shown, here the entire profile was exposed in order to provide as many hints on the recommended stranger as possible (e.g., their office location, management chain, friends, tags applied by others, or board messages). Part B shows the evidence for similar interests with the recommended individual, for instance, communities they are both member of, tags they have both used, or blogs they have both commented on. Part C shows the feedback users were asked to provide on the recommendations. Particularly, the first question (Q1) asked

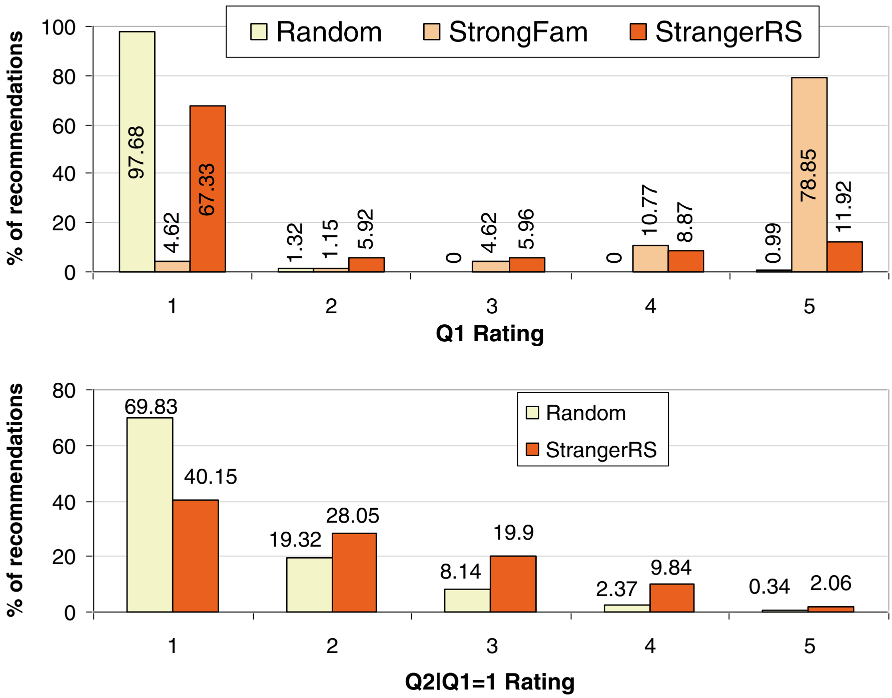


Fig. 5 Rating results for StrangerRS compared to the random and StrongFam benchmarks. The upper part shows the ratings of Q1 (familiarity with the recommended person). The *bottom* part shows the ratings of Q2 (interest in the recommended person) when Q1 = 1 (the person is a stranger)

whether the user knows the recommended person and the second (Q2) asked whether the recommended person is of interest.

Figure 5 summarizes the rating results. The upper part refers to rating of Q1. We compared StragnerRS with two benchmarks: a random person (Random) and a strongly familiar person (StrongFam). It can be seen that over two thirds (67.3 %) of StrangerRS recommendations are indeed strangers, compared to 97.7 % of the random recommendations and only 4.6 % of the strongly familiar recommendations. Hence, StrangerRS is able to recommend people who are likely to be strangers, even if not with the same likelihood as a random person.

The bottom part of Fig. 5 shows the rating results of Q2 given that Q1 was rated 1 (i.e., the person was a stranger). While 40 % of these recommendation by StrangerRS were rated as 1 (non-interesting), nearly 60 % raised some interest (compared to 30 % for random), and 30 % were rated 3 or more. Overall, even though the likelihood of StrangerRS to recommend a stranger is lower than Random, its likelihood to recommend an **interesting** stranger is higher. The latter statement is obviously true in comparison to the StrongFam. Out of a total of nine recommendations, StrangerRS was able to recommend at least one stranger rated with Q2 = 3 or higher for over two thirds of the users, and at least one stranger with Q2 = 4 or higher for over 36 % of the users. The value of such a recommendation, suggesting a stranger who is interesting, can be very high to workers.

In conclusion, while stranger recommendations have significant lower accuracy than recommendations of known people, their value lies in other aspects, such as serendipity and diversity. Practically, along time, SNSs should combine both types of recommendations. For example, friend recommendations can be suggested to new social media users who are building their initial network. Once established, stranger recommendations can help extend social circles and expand reach. Another option is to mix both friend and stranger recommendations in parallel, integrating both the higher accuracy of friend recommendations and the serendipity, or "surprise effect", of stranger recommendations. Further research needs to examine in detail how to interleave both types of people recommendations.

Recommendation of Social Media Content

Social media presents many types of content that are quite different from each other in nature. From blogs and microblogs, through social bookmarks and shared photos, videos, or files, to forums and wikis, social media users are flooded with many different content types. While various works focused on the recommendation of one specific type of social media content (Arguello et al. 2008; Lerman 2006; Seth and Zhang 2008), taking advantage of its unique characteristics to improve recommendation. This section will focus, however, on recommendation of aggregated social media content. Such recommendation not only brings the user to the most relevant items, but also implicitly suggests the types of content that are of more interest for the user.

Collaborative filtering is the most popular recommendation technique today (Goldberg et al. 1992). The most common type of collaborative filtering—user-based collaborative filtering—is based on recommending items that users with similar tastes or preferences to the target user have liked in the past. Typically, similarity between two individuals is based on the similarity of items they have liked or preferred in the past. Social media presents new opportunities for recommender systems, since social network sites, such as Facebook and LinkedIn, allow easier access to the user's list of friends. Previously to the social media era, the user's friends could be extracted only by explicitly interviewing or surveying the user or by mining sensitive contact lists from email, instant messaging, phone calls, and the like. With friend lists becoming more accessible, it was only a matter of time until studies of recommender systems would suggest and examine the use of these lists compared to the traditional use of similar users in collaborative filtering techniques. Most of the studies have shown that the list of the user's friends is likely to substantially improve collaborative filtering accuracy, when injected into it (Bonhard and Sasse 2006; Groh and Ehmig 2007; Lerman 2006; Sinha and Swearingen 2001).

In our work on recommending mixed social media content, we focused on three key research questions: (1) which network is better for recommending social media content—the user's list of familiar people or the user's list of similar people.

In theory, each of these lists can pose its own benefits for recommendation: on the one hand, the set of people the users knows best are likely to serve as the best filters for social media content; on the other hand, not all friends may have the same preferences or tastes, so the list of people who share similar interests may yield the better recommendations; (2) Which are better for recommending social media content—the set of related people to the user or the set of related tags to the user and whether a hybrid model can improve the recommendations; (3) What is the effect of explanations for both types of recommendations (people-based and tag-based) on the effectiveness of recommendation.

We experimented with IBM Connections (http://www.ibm.com/software/lotus/products/connections), a social media application suite for the enterprise that includes different types of social media applications, such as blogs, bookmarks, wikis, and files. Recommendations were made in an enterprise setting and evaluated through large-scale surveys in which hundreds of employees rated their interest in sets of 12–16 recommendations of social media items, originating from different social media applications. Effectiveness of recommendation was measured as their ability to yield more interesting items.

We foresee content recommendation playing a key role in human computation. Recommending tasks both to people and to machines who carry out a computation can be key to achieve good results. Similar mechanisms as proposed in this section can be utilized to reach effective recommendation, which are accurate, diverse, and support serendipity.

Recommendation Based on Social Relationships

Our first experiment compared recommendation based on the user's familiarity network with recommendations based on the user's similarity network. We derived both network from interaction in Connections itself. Our ability to accurately extract both types of network was proved in our previous research (Guy et al. 2010b, 2008). We distinguished between familiarity relationship, derived from data such as being friends on the enterprise SNS, tagging one another, or co-editing a wiki and the similarity network, derived from data such as common tags, common bookmarks, or common blog entries commented.

Figure 6 depicts our UI recommendation widget for providing item recommendations based on the algorithm described in the previous section. The user is presented with five items consisting of a mix of bookmarked pages, communities and blog entries. Each item has a title which is a link to the original document and a short description if available. The icon to the left of each item symbolizes its originating application—the first item in Fig. 6 is a blog entry, the second is a community, and the fourth is a bookmarked page. The user can remove an item in order to retrieve a new recommendation by clicking on the Next icon. Each recommended item includes a list of people's person names that are related to the item. Each person name provides an explanation of why the item is recommended (serving as an

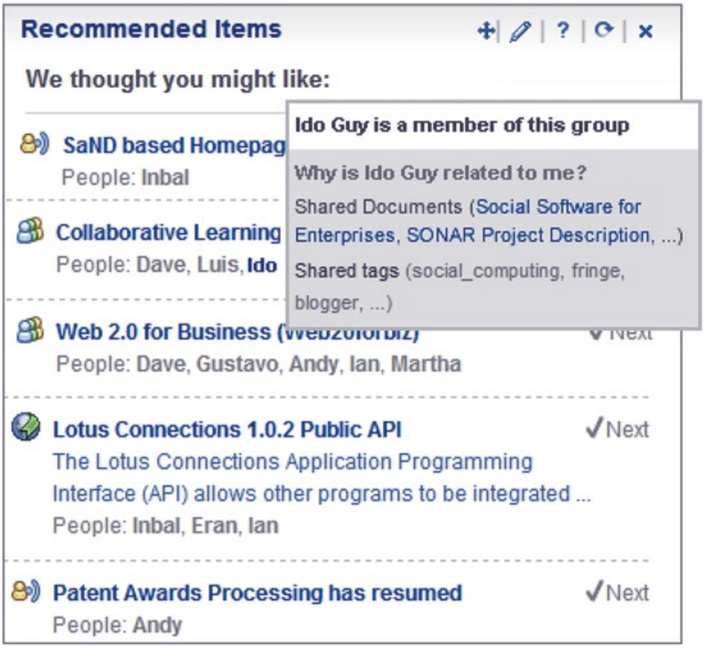


Fig. 6 Widget for recommending items based on social relationships

implicit recommender of the item). When hovering over a name, the user is presented with a popup detailing the relationships of that person to the user and to the item. In Fig. 6 the recommended items are chosen according to the similarity network of the user. The popup indicates that Ido on the one hand is a member of the recommended community and on the other hand is similar to the user as they both share a set of documents and used the same tags.

The recommender engine recommends items according to the following formula (representing the score of item i for user u):

$$RS(u,i) = e^{-\alpha t(i)} \cdot \sum_{v \in N^T(u)} S^T[u,v] \sum_{r \in R(v,i)} W(r)$$

where $t(i)$ is the number of days passed since the creation date of i; α is a decay factor (set in our experiments to 0.025); $N^T(u)$ is the set of users within u's network of type T, $T \in \{familiarity,\ similarity,\ overall\}$; $S^T[u,v]$ is the SONAR relationship score between u and v based on the network of type T; $R(v,i)$ is the set of all relationship types between user v and item i (authorship, membership, etc.); and $W(r)$ is the corresponding weight for the user-item relationship type r.

Our main evaluation of the above RS was based on a user study with 290 participants who were evenly assigned to one of three groups: familiarity, similarity, and overall (the latter is a combination of both types of networks). Each participant rated

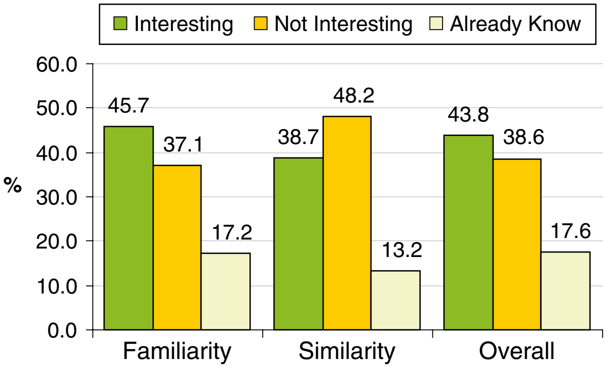


Fig. 7 Rating results for item recommender based on social relationships

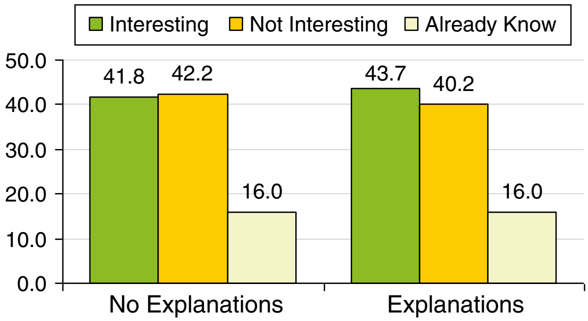


Fig. 8 Effect of explanations on interest rate in people-based recommenders

12 recommended items in two randomly-ordered phases: one phase included explanations for each recommended item, while the other did not include explanations.

Figure 7 shows the results of the comparison between the three relationship types: familiarity, similarity, and overall. Familiarity-based recommendations are significantly more interesting than similarity-based recommendations, showing that in terms of accuracy, familiarity yields better recommendations. The similarity network, however, has lower percentage of already-known items, indicating that in terms of serendipity and novelty, this network may have an advantage (at the expense of accuracy). Combining the two networks does not lead to improvement over the results of the familiarity network.

Figure 8 shows the effect of explanations. Recommendations with explanations were rated slightly higher than recommendations with no explanations. This finding was not statistically significant, but was consistent across all three network types and most noticeable for the familiarity network. It indicates that in addition to longer-term benefits (Herlocker et al. 2000), people-based explanations also instantly influence the interest in recommended items. It especially seems that

Table 1 Rating results of tags as topics of interest

%	Not interested	Interested	Highly interested
Used	16.84	38.25	44.91
Incoming	15.48	31.75	52.78
Direct	7.46	22.81	69.74
Indirect	35.38	45.38	19.23

seeing that a familiar person is related to a piece of content provides extra information that can affect the decision whether a recommendation is interesting. For example, a bookmark that was tagged by a colleague that you know and think high of may be interesting for you just because of the fact it was tagged by that colleague.

Recommendation Based on Tags

After exploring social media recommendations based on social relationships, we moved on to examine the effect of adding tag-based recommendations (Guy et al. 2010c). We first examined several types of tags as indicators of users' interests. We examined tags used by the user in different tagging systems such as the social bookmarking system and the blogging system ("used tags"), tags applied on the user by others through a people tagging application (Farrell et al. 2007) ("incoming tags"), a combination of used and incoming tags ("direct tags"), and tags put on items the user is related to by other users ("indirect tags").

In our user survey, 65 participants rated a total number of 1,037 tags of all four types. Results are summarized in Table 1. Indirect tags are rated significantly lower than all other tag types (as could be expected; such tags are noisy and should be used only in cases of data sparsity). Direct tags are rated significantly higher than all other types of tags, indicating that tags that are both used by the user and applied to her by others are the most accurate interest indicators. Interestingly, incoming tags are rated slightly higher than used tags (insignificant difference), indicating that the topics associated with the user by others are as good indicator for the user's topics of interests as the tags she used herself.

Based on these results, we built a hybrid people-tags-based recommender that suggests five types of social media items: bookmarks, blogs, communities, files, and wikis (Guy et al. 2010c). The recommender was based on a user profile that consists of both people and tags. For people, familiarity relationships were favored over similarity relationships by a factor of 3, due to the results of the people-based recommendation study (Guy et al. 2009b). For tags, we considered direct tags, due to the results of the tags user survey described above. Overall, items were recommended according to the following formula:

$$RS(u,i) = e^{-\alpha d(i)} \cdot \left[\beta \sum_{v \in N(u)} w(u,v) \cdot w(v,i) + (1-\beta) \sum_{t \in T(u)} w(u,t) \cdot w(t,i) \right]$$

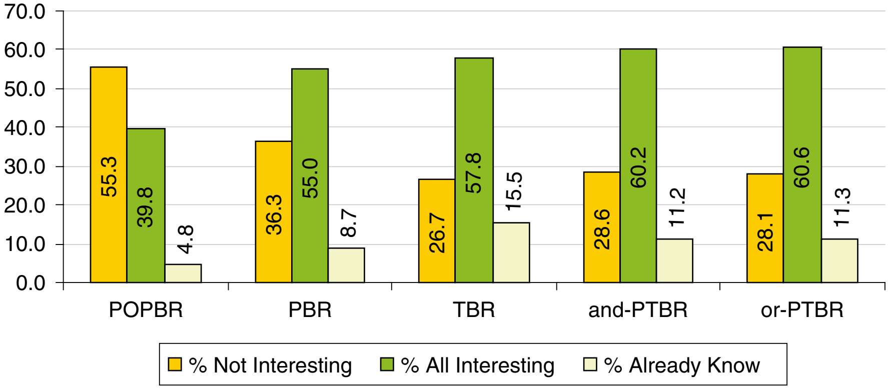


Fig. 9 Item rating results across five recommenders: a popularity-based recommender (used as a baseline), a recommender based on people only, a recommender based on tags only, and two hybrid recommenders combining both people and tags

where *d(i)* is the number of days passed since the creation date of *i*; α is a decay factor (set in our experiments to *0.025*); *β* is a parameter that controls the relative weight between people and tags and is used in our experiments to evaluate different recommenders; *w(u,v)* and *w(u,t)* are the relationship strengths of *u* to user *v* and tag *t*, as given by the user profile; *w(v,i)* and *w(t,i)* are the relationship strengths between *v* and *t*, respectively, to item *i*.

Ultimately, the recommendation score of an item, reflecting its likelihood to be recommended to the user, may increase due to the following factors: more people and/or tags within the user's profile are related to the item; stronger relationships of these people and/or tags to the user; stronger relationships of these people and/or tags to the item; and freshness of the item. We excluded items that were found to be directly related to the user. For example, we did not recommend an item the user had already commented on or had already tagged.

In our main user survey, which included 412 participants, we compared five types of recommenders: (1) POPBR—a non-personalized popularity-based recommender, as a baseline, (2) PBR—a people-based recommender, similar to the one used in the first study, considering social relationships only, (3) TBR—a tag-based recommender considering tags only, (4) and-PTBR—a hybrid recommender based on both people **and** tags, (5) or-PTBR—a hybrid recommender based on people **or** tags.

Results are displayed in Fig. 9. All types of personalized recommenders significantly outperformed the popularity-based recommender. The tag-based recommender outperformed the people-based recommender in terms of accuracy (interest ratio in recommended items). However, it posed a few shortcomings compared to the people-based recommender: higher level of expectedness, reflected in a higher percentage of already-known items; lower level of diversity, reflected in about 80 % of the recommended items being bookmarked pages; and lower effectiveness of explanations (details below). Both hybrid recommenders enjoy the benefits of both

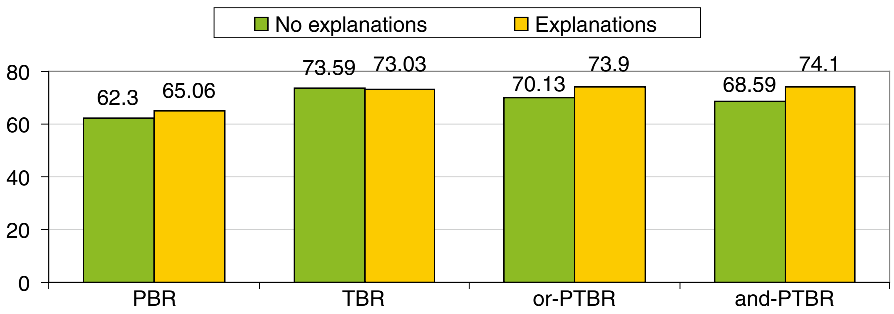


Fig. 10 Percentage of items rated interesting or very interesting with and without explanations for people-based and tag-based recommenders

"worlds"—they maintain as high accuracy as the tag-based recommender, and substantially improve on diversity and serendipity.

Figure 10 shows the effect of explanations on the different types of personalized recommenders—for each it presents the percentage of interesting items out of all recommended items. Interestingly, for the TBR, there is no effect whatsoever on the interest rate in recommended items. For the PBR, as previously found, explanations slightly improve the instant interest in recommended items. This effect is carried on to both of the hybrid recommenders, or—PTBr and and-PTBR, in which explanations also improve the accuracy. Apparently, the tag explanations do not provide any information that may lead the user to change his decision regarding her interest in the item, while for people, it maybe be the case that the explanations yield the interest in an items. Supposedly, for example, an item may not seem interesting to the user until she sees that a related person she appreciates has bookmarked the item. We note that this does not imply that "tagsplanations" (Vig et al. 2009) are not effective. They may very well have desirable long term effects in building trust with the user. Yet, they do not seem to have an instant effect on the user's interest in an item.

News Item Recommendation in Social Streams

Social streams have emerged as a means to syndicate updates about a user or a group of users within a social network site or a set of sites. The Facebook newsfeed and Twitter are two of the most popular social streams on the web today, with millions of news items generated every hour. The flood of news updates within the streams poses new challenges in terms of filtering and personalization. Bernstein et al. (2010) interviewed users of Twitter and found that they struggle to balance the promise of interesting content with the sheer volume of incoming updates. Currently, the default filtering of Twitter and Facebook is based on the list of individuals the user chooses to follow ("followees") and the list of the user's friends, respectively. However, this filtering approach is often insufficient, as some friends or followees may produce many non-interesting news items or dominate the stream, while interesting updates may also come from sources outside the circle of friends or followees. We examined the challenge of personalizing the activity stream by referring to it as a recommendation task, aimed at suggesting relevant news items to the user from the overall stream (Guy et al. 2011a). We referred to a *news item* as the basic unit of which the social stream is composed. A news item can refer to a network activity (e.g., adding a friend), an activity over an 'entity' (e.g., editing a wiki page, "liking" a file), or a status update.

We experimented with the activity stream of an enterprise social media application suite—IBM Connections (http://www.ibm.com/software/lotus/products/connections). The stream consists of status updates as well as activities across the different social media applications, which include social bookmarking, file sharing, blogging, communities, wikis, and an SNS. For example, a news item can be: *John Smith edited the wiki page Design Principles in the Cloud Computing wiki*. As mentioned, our previous work (Guy et al. 2009b; Guy et al. 2010c) has examined the recommendation of mixed entities across these applications (bookmarks, files, blog entries, communities, and wikis).

Figure 11 illustrates the user interface of the activity stream. News items are displayed in reverse-chronological order with an indication of their freshness. Each news item includes a textual description with the action, actor(s) and entity(ies) involved, and occasionally a short excerpt of the text. Also, each item includes a picture of the actor of the item and an icon indicating the originating application. Each underlined entity within the news item content is a link to its corresponding page. Overall, the stream included news items originating from eight different enterprise social media applications.

We experimented with user profiles that included three *dimensions*: people, terms, and places (resources). People included the user's set of related people. Terms included the user's related terms, as inferred using a term extraction algorithm (Carmel et al. 2012). Places (also referred to as resources) are entities for which multiple activities can occur, such as communities, wikis, blogs, or files. We extracted the user's set of related places based on those s/he already interacted with in the past.

Our evaluation was based on a user survey with 126 participants that included two phases. In the first phase, participants evaluated lists of *profile objects*, i.e., lists of people, terms, and places that make up the profile itself. In the second phase, participants evaluated news items that originated from those different profiles. For each participant, we retrieved the top ten related people, terms, and places and then produced the news items that related to them as recommendations.

Figure 12 shows the results of the second phase. Places clearly yield the highest accuracy with over three quarters of the news items rated as interesting or very interesting. People yield substantially less interesting news items, while terms produce the most noisy news items with less than 50 % rated interesting or very interesting. These results show that places are an important addition to a user's profile when it comes to news item recommendation. In case a user relates to a place, it is likely that more news from it would draw the user's interest.

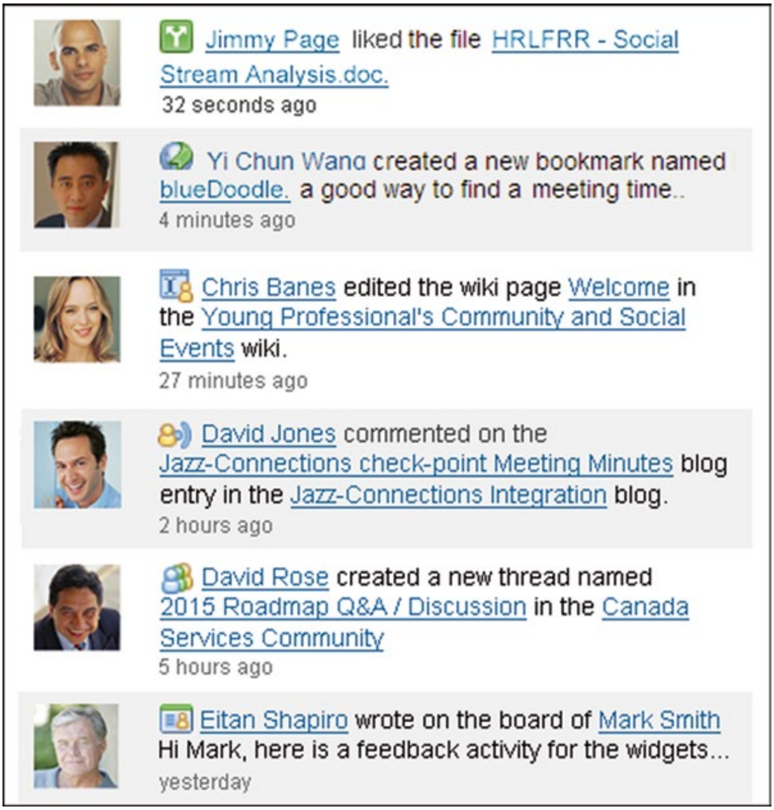


Fig. 11 User interface of the activity stream

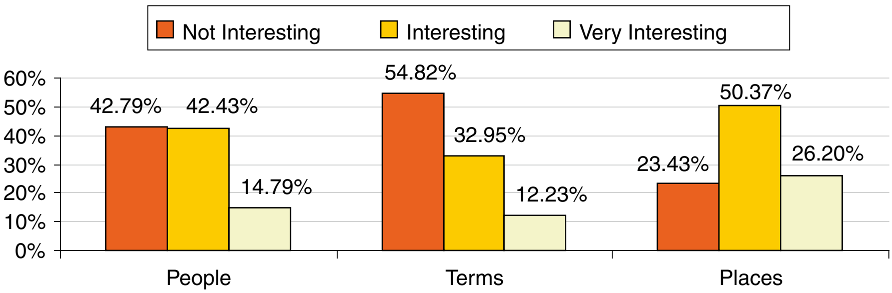


Fig. 12 News item rating results by dimension

We also considered another measurement in conjunction with accuracy that we called *throughput* (Guy et al. 2011a). Throughput measures the ability of a profile to produce news items in a given time period. There is a natural trade-off between

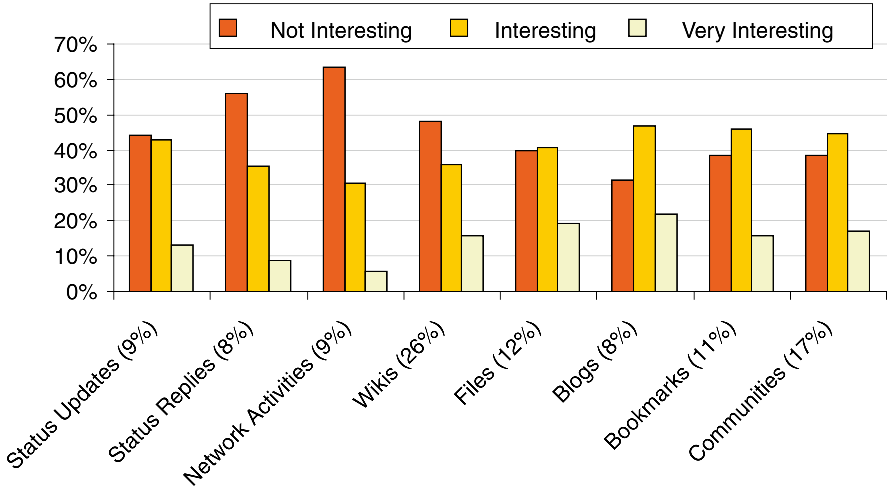


Fig. 13 News item rating by application

throughput and accuracy. Terms were able to produce many more news items than people and then places in turn. Therefore, an ideal user profile would combine people, terms, and places to achieve an optimal combination of accuracy and throughput. We left the challenge of combining all three as future work.

Figure 13 shows the rating results according to the application the items originated from. The percentage in parentheses indicates the portion of the application's items out of the entire recommended news items in our survey. It can be seen that the most accurate results are for news items that relate to entities: news items about blogs were the most interesting (68.6 % rated interesting or very interesting), followed by communities and bookmarks (61.7 % each), files (60 %), and finally wikis (51.8 %). The lower percentage of wikis can be explained by the fact that wiki is the most common type of news item, with many incremental updates on wiki edits.

Status updates also received relatively accurate results, with 55.6 % rated as interesting or very interesting (higher than wikis, but lower than all other entities). Interestingly, status replies, both on own board and on others, were considered much less interesting, with only 44 % rated interesting or very interesting. It appears that replies are perceived as an interaction between two other individuals, and are of less interest to the public. Network activities were rated as least interesting, with only 36.4 % interest. These news items also refer to interactions between two other individuals, but with even less content than the status replies.

Overall, we found that the addition of places, such as communities, wikis, and blogs to the user profile is very productive for the news item recommendation task as they yield very high accuracy—over 75 %. While producing high accuracy, places were also shown to have the lowest throughput. At the other extreme, terms produce noisy items with low accuracy and high throughput. People are the middle dimension in terms of both measurements. Hybridization of profile dimensions is likely to improve the results even further and allow more flexibility when adapting to the user's needs. Future research should examine hybridization techniques based on the characteristics of each of the profile dimensions as shown here.

The accuracy results across the various applications were quite different. Network additions and person tagging are of substantially lower interest than status updates and replies, while entity-related news have the highest interest rates. These results indicate that stream personalization methods should account for the type of originating application for a news item, as has been proposed by Berkovsky et al. (2012).

Our model was a simple one—we filtered items based on their relatedness to profile objects, i.e., if a news item relates to one or more objects in the user's profile, it would be selected, otherwise not. A more sophisticated model would apply a score per news item per user and use that score for filtering in a more flexible way. While initial works have been done in this direction (Berkovsky et al. 2012; Paek et al. 2010), we expect more advanced models to appear in the future, as the challenge of filtering social streams become more fundamental for web users.

Conclusions and Future Work

We described a rich set of algorithms and their applications in the social recommender system domain. We reviewed algorithms for recommendation of people on social network sites, including recommendation of familiar people and of strangers. We discussed recommendation of mixed social media content that is based on relationship information—both the user's familiarity and similarity network and on tags—both used tags and tags input by other individuals. And we reviewed an emerging area for SRS—recommendation of news items in social streams and the addition of "places" (resources, such as wikis, communities, and blogs) to the user profile alongside people and terms.

A common characteristic of all the recommender systems described in this work is the ability to provide explanations for recommendations. We argue that explanations should be an integral part of the recommender system algorithm. The public nature of social media allows sharing the logic behind the recommendation, such as the people and tags that yield it, in a more transparent way. Explanations were previously shown to provide longer-term user needs and build trust relationship with the user (Herlocker et al. 2000). In our work, we found other positive aspects. For example, for people recommendation, explanations lower the entry barrier for sending an invitation since they make the user feel s/he has a good "excuse" to invite that person (Guy et al. 2009a). For content recommendation, people-based explanations contribute extra information that increases the chances a recommendation would be interesting for the user (Guy et al. 2009b, 2010c). Explanations also pose a privacy challenge for social recommender systems. For example, a user who gets access to one's recommendations that are accompanied by explanations (e.g., "because you watched this video") might be exposed to sensitive data about past behavior of the user who got the recommendations. Future social recommender systems should put special care on privacy preservation.

Our recommendation techniques are generally based on a rich relationship aggregation model (Guy et al. 2008). The richness and diversity of data sources, originating from social media in the enterprise, allows providing better recommendations. Our previous work showed that the rich set of data sources substantially improves the ability to derive the user's set of familiar people (Guy et al. 2008), similar people (Guy et al. 2010b), and related topics (Guy et al. 2013). The methods we described here can be applicable for more types of social media content, such as photos, music, or video. Obviously, some generalizations were made to allow for aggregating a wide set of data sources. Future work should examine how to incorporate the preferences of the user to specific types of social media, taking advantage of the specific characteristics of application and test whether this can further improve the accuracy of recommendations.

Wes use typical social media data, such as people relationships, community membership, and tags, to produce recommendations. As said in the Introduction section, this data helps enhance traditional recommender systems techniques. In the era of big data, SRSs are exposed to an ever-growing amount of useful information that can be utilized to understand user preferences. Particularly, heavy content-based techniques should be examined as a complementary means for the data used in this work.

Recommender systems generally include three main ingredients—the data, the algorithm, and the user interface. As we saw in this chapter, the three are tightly tied together and a good algorithm is not enough to make a good recommender system, if not accompanied by a good adequate UI and a corpus of data that serves as a good ground for recommendation. In fact, it is often the case that an advanced algorithm would achieve a small improvement in recommendation accuracy, but would pose costs in other aspects, such as the ability to provide intuitive explanations. For example, when we examined content-based algorithms for people recommendation (Chen et al. 2009), more advanced techniques such as Latent Semantic Analysis (Deerwester et al. 1990) and Probabilistic Latent Semantic Analysis (Hofmann 1999) were analyzed, but since they did not yield significantly better results, we opted to apply the simpler algorithm that can provide more straightforward explanations.

In our analysis, we discussed factors other than accuracy, such as diversity and serendipity. Yet, the focus of the evaluation is primarily on accuracy and no predefined methodology is used to evaluate the combination of different evaluation measures. More emphasis on developing new evaluation methods, with focus on factors other than accuracy, should be put in future social recommendation research (McNee et al. 2006). Moreover, evaluation over time, where users get used to the system and exhaust the initial set of recommendations, while their interests keep changing, is also due for future studies. Learning from user feedback can help mitigate the expected decrease of user interest in recommendations along time.

As implied throughout this chapter, we envision social recommendation playing a central role in human computation methods. Particularly, people recommendation can help matching the right people for the right job. Explanations can help achieving higher level of engagement. Content recommendations are instrumental to task management and can help assure people (and machines) are assigned with the right activities. Finally, social streams can contribute to engagement in human computation games and help in participants' attention management.

Acknowledgements With thanks to David Carmel, Jilin Chen, Tal Daniel, Casey Dugan, Werner Geyer, Michal Jacovi, Michael Muller, Shila Ofek-Koifman, Adam Perer, Ariel Raviv, Inbal Ronen, Sigalit Ur, Erel Uziel, Eric Wilcox, Sivan Yogev, and Naama Zwerdling for jointly working on the studies described in this chapter.

References

Arguello J, Elsas J, Callan J, Carbonell J (2008) Document representation and query expansion models for blog recommendation. In: Proceedings ICWSM '08

Berkovsky S, Freyne J, Smith G (2012) Personalized network updates: increasing social interactions and contributions in social networks. In: Proceedings UMAP '12, pp 1–13

Bernstein MS, Suh B, Hong L, Chen J, Kairam S, Chi EH (2010) Eddi: interactive topic-based browsing of social status streams. In: Proceedings UIST '10, pp 303–312

Bonhard P, Sasse MA (2006) 'Knowing me, knowing you'—Using profiles and social networking to improve recommender systems. BT Technol J 24 (3): 84–98

Boyd DM, Ellison NB (2007) Social network sites: definition, history, and scholarship. J CMC 13:1

Carmel D, Uziel E, Guy I, Mass Y, Roitman H (2012) Folksonomy-based term extraction for word cloud generation. ACM TIST 3(4):60

Chen WY, Zhang D, Chang EY (2008) Combinational collaborative filtering for personalized community recommendation. In: Proceedings KDD '08, pp 115–123

Chen J, Geyer W, Dugan C, Muller M, Guy I (2009) Make new friends, but keep the old: recommending people on social networking sites. In: Proceedings CHI '09, pp 201–210

Deerwester S, Dumais S, Furnas GW, Landauer TK, Harshman R (1990) Indexing by latent semantic analysis. J Amer Soc Info Sci 41(6):391–407

Farrell S, Lau T, Nusser S, Wilcox E, Muller M (2007) Socially augmenting employee profiles with people-tagging. In: Proceedings UIST '07, pp 91–100

Geyer W, Dugan C, Millen DR, Muller M, Freyne J (2008) Recommending topics for self-descriptions in online user profiles. In: Proceedings RecSys '08, pp 59–66

Goldberg D, Nichols D, Oki BM, Terry D (1992) Using collaborative filtering to weave an information tapestry. Commun ACM 35 (12): 61–70

Groh G, Ehmig C (2007) Recommendations in taste related domains: collaborative filtering vs. social filtering. In: Proceedings GROUP '07, pp 127–136

Guy I, Carmel D (2011) Social recommender systems. In: Proceedings WWW '11, pp 283–284

Guy I, Jacovi M, Shahar E, Meshulam N, Soroka V, Farrell S (2008) Harvesting with SONAR: the value aggregating social network information. In: Proceedings CHI '08, pp 1017–1026

Guy I, Ronen I, Wilcox E (2009a) Do you know? recommending people to invite into your social network. In: Proceedings IUI '09, pp 77–86

Guy I, Zwerdling N, Carmel D, Ronen I, Uziel E, Yogev S, Ofek-Koifman S (2009b) Personalized recommendation of social software items based on social relations. In: Proceedings RecSys '09, pp 53–60

Guy I, Chen L, Zhou MX (2010a) Workshop on social recommender systems. In: Proceedings IUI '10, pp 433–434

Guy I, Jacovi M, Perer A, Ronen I, Uziel E (2010b) Same places, same things, same people? Mining user similarity on social media. In: Proceedings CSCW '10, pp 41–50

Guy I, Zwerdling N, Ronen I, Carmel D, Uziel E (2010c) Social media recommendation based on people and tags. In: Proceedings SIGIR '10, pp 194–201

Guy I, Ronen I, Raviv A (2011a) Personalized activity streams: sifting through the "river of news". In: Proceedings RecSys '11, pp 181–188

Guy I, Ur S, Ronen I, Perer A, Jacovi M (2011b) Do you want to know? recommending strangers in the enterprise. In: Proceedings CSCW '11, pp 285–294

Guy I, Avraham U, Carmel D, Ur S, Jacovi M, Ronen I (2013) Mining expertise and interests from social media. In: Proceedings WWW '13
Herlocker, JL, Konstan JA, Riedl J (2000) Explaining collaborative filtering recommendations. In: Proceedings CSCW '00, pp 241–250
Hofmann T (1999) Probabilistic latent semantic analysis. UAI'99
IBM Connections—Social Software for Business: http://www.ibm.com/software/lotus/products/connections
Jacovi M, Guy I, Ronen I, Perer A, Uziel E, Maslenko M (2011) Digital traces of interest: deriving interest relationships from social media interactions. In: Proceedings ECSCW '11, pp 21–40
Jannach D, Zanker M, Felfernig A, Friedrich G (2010) Recommender systems: an introduction. Cambridge University Press, New York
Lerman K (2006) Social networks and social information filtering on0020digg. In: Proceedings ICWSM '07
McNee SM, Riedl J, Konstan JA (2006) Being accurate is not enough: how accuracy metrics have hurt recommender systems. Proceedings CHI '06 EA, pp 1097–1101
Paek T, Gamon M, Counts S, Chickering DM, Dhesi A (2010) Predicting the importance of newsfeed posts and social network friends. In: Proceedings AAAI (Vol 10)
Resnick P, Varian HR (1997) Recommender systems. Commun ACM 40(3):56–58
Ricci F, Rokach L, Shapira B (2011) Introduction to recommender systems handbook. Springer
Sen S, Vig J, Riedl J (2009) Tagommenders: connecting users to items through tags. In: Proceedings WWW '09, pp 671–680
Seth A, Zhang J (2008) A social network based approach to personalized recommendation of participatory media content. In: Proceedings ICWSM '08
Sigurbjörnsson B, Van Zwol R (2008) Flickr tag recommendation based on collective knowledge. In: Proceedings WWW '08, pp 327–336
Sinha R, Swearingen K (2001) Comparing recommendations made by online systems and friends. In: Proceedings DELOS-NSF workshop on personalization and recommender systems in digital libraries
Terveen L, McDonald DW (2007) Social matching: a framework and research agenda. ACM Trans Comput-Hum Interact 12(3):401–434
Vig J, Sen S, Riedl J (2009) Tagsplanations: explaining recommendations using tags. In: Proceedings IUI '09, pp 47–56

Part VI
Participation

Participation

Winter Mason

Introduction

Why do people participate in human computation systems, and how can high-quality participation be encouraged? This is the key question addressed in this section. This question can be decomposed into more focused questions: What draws them to participate? What motivates them, drives them to keep working or contributing? How does this differ with respect to different types of human computation systems? How do you improve the quality of participation? Can one design a system to guide the participants to provide highest quantity and best quality output? Is there an inherent tradeoff between quantity and quality?

These questions are critical to the success of human computation. Although the topics covered in the previous sections are fundamental to human computation—the infrastructure underlying the systems, the algorithms that operate on the systems, and the specific techniques to which these are applied—without the proper motivation, none of this excellent work can be applied. You can design an amazing vehicle, but without fuel and a spark it will never go anywhere. The same is true of human computation systems.

A superficial consideration of the question might lead one to think Amazon's Mechanical Turk has solved this problem of incentivizing workers; by paying the appropriate wage, workers are motivated to complete the available tasks. However, the reality is much more complex. Prior research has shown that the way in which tasks are listed on Mechanical Turk affect which jobs are selected (Chilton et al. 2010), that the size of the incentive does not automatically lead to better results (Mason and Watts 2009), and that financial reward is actually only a (small) part of why Turkers choose to do a task (Ipeirotis 2010). In fact, the question of what motivates people has been the subject of research for nearly a century, since

W. Mason (✉)
Stevens Institute of Technology, Hoboken, NJ 07030, USA
e-mail: m@winteram.com

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_53,

Freud went "beyond the pleasure principle" (Freud 1920), and has been the focus of volumes of research in psychology, marketing, and economics. As we see in this section, properly incentivizing workers can affect the efficiency and performance with which participants do the work required for human computation, as well as the long-term commitment of people to the system.

In the first chapter, Chamberlain, Kruschwitz, and Poesio introduce three fundamental reasons users may choose to get involved in human computation projects: intrinsic interest in the project, which is the usual driver for human computation systems such as Wikipedia and citizen science projects; extrinsic motivation, such as payments for micro-tasks on Amazon's Mechanical Turk; and entertainment, which is the driver behind so-called "games with a purpose" (von Ahn 2006). It is this last motivation that the authors focus on, particularly with respect to a project known as "Phrase Detectives". The purpose of this project is to annotate text with links between anaphors (e.g., *he* or *it*) and the part of speech being referenced (e.g., *John Smith* or the table). These annotations can be used to aid automatic interpretation of text. The authors use this task as an example of the challenges one faces when trying to convert a human computation system into a game, and techniques that can be used to encourage participation in the system.

In the next chapter, Reed and colleagues describe qualitative studies of participants in the Zooniverse collection of citizen science projects. This exploration drills down into the "intrinsic interest" motivation outlined in the first chapter through a series of surveys. The authors discover three key factors for participation in these citizen science projects: social engagement, interaction with the website, and the desire to help.

An interesting perspective on human computation is to consider the group working in the system as possessing a sort of collective intelligence that is different from the intelligence of the individuals making up the group. In their chapter, Woolley and Hashmi discuss how those participating in a human computation contribute to an intelligence that is determined by the interactions between the group members, as much as it is on the characteristics of those members. Further, they describe a framework based on this research that describes tools that can be used to enhance the collective intelligence of groups. These include tools that facilitate social perceptiveness and tools that encourage equality of participation.

The tools to facilitate collective intelligence suggested by Woolley and Hashmi are offered in the context of online collaboration—that is to say, through computer mediated communication (CMC). In the following chapter, Santuzzi and colleagues discuss the importance of recognizing that factors that affect CMC may be different than those that are important in face to face collaborations. For example, when designing a human computation system, issues such as trust and leadership may be different in online CMC than they would be in offline collaborations. The authors point out that disciplinary biases in research perspectives can affect any analysis of why people are participating in human computation systems, so this chapter should be required reading for anyone planning to study the social components of computer mediated communication in human computation systems.

Finally, the section concludes with a chapter by Arpita Ghosh that demonstrates how a game-theoretic approach—one that considers the equilibrium behavior between agents trying to maximize their utility, given the constraints of the system—can be used to understand why humans behave in a certain way in different human computation systems, and how these systems can be designed to incentivize the behavior desired by the creator of the system. Ghosh illustrates the approach with three types of human computation systems: games with a purpose, such as the "Phrase Detectives" game from Chamberlain's chapter; applications that use crowd-sourced judgments; and applications that aggregate quality estimates from the crowd. She concludes the chapter and section by highlighting future directions, including research on the interactions between intrinsic and extrinsic motivations.

The diversity of perspectives and methods for understanding participation in human computation that are demonstrated in this section reflect the same diversity that is observed in research on motivation and incentives generally. It is obvious this research benefits from or may even require the convergence of these methods, from the qualitative studies described in Reed and colleagues' chapter, to the experimental work by Woolley and Hashmi, to the formal analysis described by Ghosh. Moreover, these methods may discover different reasons for participation and means of improving participation when focused on citizen science projects (Reed et al.), games (Chamberlain et al.), computer mediated communication (Santuzzi et al.) or traditional platforms like Mechanical Turk (Ghosh).

This research also suggests future directions for understanding the human equation in human computation systems. For instance, what are the tradeoffs between extrinsic and intrinsic motivations, in the short run and the long run, in microtask situations (like Mechanical Turk) versus collaborative communities (like Wikipedia)? How can you design a system to best take advantage of the typically skewed distribution of contributions in human computation systems? Can citizen science projects utilize all of the different incentives outlined in this section, or are the interest-based communities somehow different? How important is the community sense to motivation in human computation systems generally?

Although the work presented in this section represent great strides in the study of participation in human computation, there is a broad range of questions yet to be answered.

References

Chilton LB, Horton JJ, Miller RC, Azenkot S (2010) Task search in a human computation market. In: Proceedings of the ACM SIGKDD workshop on human computation, Washington, D.C., USA, pp 1–9

Freud S (1920) Jenseits des lustprinzips, vol 13. Internationaler Psychoanalytischer Verlag, Leipzig

Ipeirotis PG (2010) Analyzing the amazon mechanical turk marketplace. XRDS: Crossroads ACM Mag Stud 17(2):16–21

Mason WA, Watts DJ (2009) Financial incentives and the performance of crowds. In: Proceedings of the ACM SIGKDD workshop on human computation, Paris, France

von Ahn L (2006) Games with a Purpose. Computer, 39(6), 92–94

Methods for Engaging and Evaluating Users of Human Computation Systems

Jon Chamberlain, Udo Kruschwitz, and Massimo Poesio

Introduction

One of the most significant challenges facing some Human Computation Systems is how to encourage participation on a scale required to produce high quality data. This is most relevant to systems where non-expert volunteers perform tasks, with the system aggregating the result. Issues relating to participant psychology are applicable to any system where humans (and subsequently human error) are involved.

The willingness of Web users to collaborate in the creation of resources is clearly illustrated by Wikipedia[1]: allowing users free reign of encyclopaedic knowledge not only empowers mass participation but the resulting creation is high quality. This can be seen as a good example of the broad term **collective intelligence** where groups of individuals do things collectively that seem intelligent (Malone et al. 2009).

The utility of collective intelligence became apparent when it was proposed to take a job traditionally performed by a designated employee or agent and outsource it to an undefined large group of Internet users through an open call. This approach, called **crowdsourcing** (Howe 2008), revolutionised the way traditional tasks could be completed and made new tasks possible that were previously inconceivable due to cost or labour limitations.

One use for crowdsourcing can be as a way of getting large amounts of human work hours very cheaply as an alternative to producing a computerised solution that may be expensive or complex. However, it may also be seen as a way of utilising human processing power to solve problems that computers, as yet, cannot solve, termed **human computation** as defined by von Ahn (2006).

[1] http://www.wikipedia.org

J. Chamberlain (✉) • U. Kruschwitz • M. Poesio
University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, England
e-mail: jchamb@essex.ac.uk; udo@essex.ac.uk; poesio@essex.ac.uk

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_54, © Springer Science+Business Media New York 2013

An application of collective intelligence, crowdsourcing and human computation is to enable a large group of collaborators to work on tasks normally done by a few highly skilled (and paid) workers and to aggregate their work to produce a complex dataset that is robust and allows for ambiguity. Enabling groups of people to work on the same task over a period of time in this way is likely to lead to a collectively intelligent decision (Surowiecki 2005).

Using this method of collecting and aggregating decisions from a large, distributed group of non-expert contributors it is possible to approximate a single expert's judgements (Albakour et al. 2010; Feng et al. 2009; Snow et al. 2008).

User Motivation in Collaborative Systems

Three variations of collaboration over the Internet have been successful in recent years and are distinguished by the motivations of the participants.

1. The first variation is where the motivation for the users to participate already exists. This could be because the user is **inherently interested** in contributing, for example in the case of Wikipedia or citizen science projects such as GalaxyZoo[2] and Open Mind Commonsense[3] (now ConceptNet[4]). Users may also be intrinsically motivated because they need to accomplish a different task, for example the reCAPTCHA[5] authentication system.
2. As most tasks are neither interesting nor easy to integrate into another system, a second variation of crowdsourcing called **microworking** (or microtasking) was developed, for example Amazon Mechanical Turk.[6] Participants (sometimes called Turkers) are paid small amounts of money to complete HITs (Human Intelligence Tasks) uploaded by Requesters. The tasks can be completed very quickly, however this approach cannot be scaled up for large data collection efforts due to the cost.
3. A third approach for collecting and validating data used in human computation is to entertain the user whilst they complete the tasks, typically using games. The **games-with-a-purpose (GWAP)** approach has been used for many different types of crowdsourced data collection including text, image, video and audio annotation, biomedical applications, transcription, search results and social bookmarking (Chamberlain et al. 2013; Thaler et al. 2011).

[2] http://www.galaxyzoo.org

[3] http://openmind.media.mit.edu

[4] http://conceptnet.media.mit.edu

[5] http://www.google.com/recaptcha

[6] https://www.mturk.com

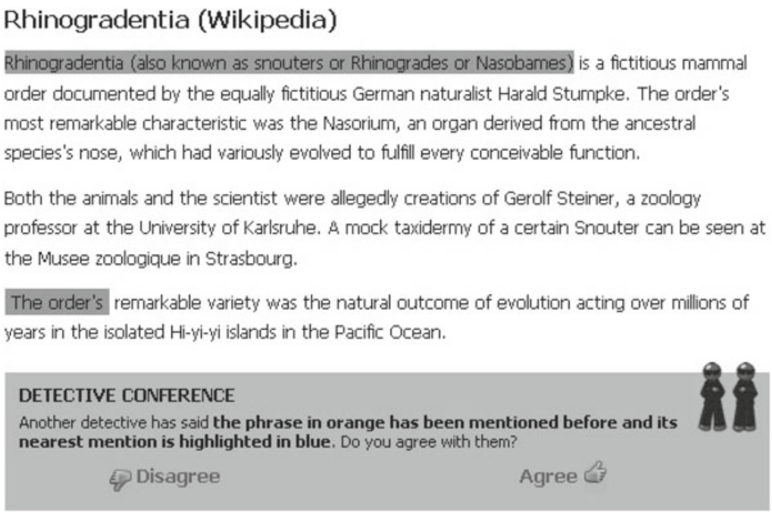


Fig. 1 Detail of a task in Phrase Detectives

There is huge potential for the general public to become engaged in Human Computation Systems and to collaborate in producing resources that would not be possible to achieve using other methods.

This chapter discusses methods that can be used to motivate and engage users. As an example, we look at how these methods were used in Phrase Detectives,[7] a Human Computation System developed by the University of Essex (England) to annotate text documents with a crowd. The conclusion summarises the benefits and limitations of using such methods in Human Computation Systems.

Phrase Detectives

Phrase Detectives (PD) is primarily a GWAP designed to collect data about English (and subsequently Italian) anaphoric co-reference (Chamberlain et al. 2008; Poesio et al. 2013).[8]

The architecture is structured around a number of tasks that use scoring, progression and a variety of other mechanisms to make the activity enjoyable (see Fig. 1).

---

[7] http://www.phrasedetectives.com

[8] Anaphoric coreference is a type of linguistic reference where one expression depends on another referential element. An example would be the relation between the entity *'Jon'* and the pronoun *'his'* in the text *'Jon rode his bike to school.'*

The aim of the project is not only to annotate large amounts of text, but also to collect a large number of judgements about each linguistic expression to preserve ambiguity that can be used to improve language processing algorithms.

A version of PD was developed for Facebook[9] in order to investigate the utility of social networking sites in collaborative annotation systems.

Methods to Engage and Evaluate Users

There have been several recent attempts to define and classify collaborative approaches in collective intelligence and distributed human computation (Quinn and Bederson 2011; Malone et al. 2009; Wang et al. 2010). We focus on four main areas:

1. Designing the Task
2. Attracting Users
3. Motivating Users
4. Evaluating Users

Designing the Task

Whilst design considerations can be somewhat generalised, it is worth noting a fundamental challenge for human computation systems. The goal here is to **collect data and reward users without directly knowing the quality of their work** (either by the system knowing the answer beforehand or by manual correction after the data is collected). Methods for motivating users without being able to provide specific feedback are discussed in detail later in the chapter.

Using an Appropriate Interface for Your Users

When designing any interface it is essential to **know your target audience**. Individual, social and socio-technical factors will all determine how successful the interface is at engaging users and what type of data will be contributed.

Wikipedia style open interfaces will invite a different type of user experience than a microworking or gaming approach and the expectations of the users need to be met in order for them to continue using the interface. Game interfaces should be graphically rich, although not at the expense of usability, and aimed at engaging a specific audience (i.e., a game aimed at children may include more cartoon or stylised imagery in brighter colours than a game aimed at adults).

[9] http://www.facebook.com

Interfaces should **provide a consistent metaphor and work flow**. For this PD used a detective metaphor, with buttons stylised with a cartoon detective character and site text written as if the player was a detective solving cases. The tasks should be integrated in such a way that task completion, user evaluation and work flow form a seamless experience.

Interfaces deployed on the Web should observe the normal guidelines regarding browser compatibility, download times, consistency of performance, spatial distance between click points, etc.[10]

Designing the Tasks

Whilst the design of the interface is important, it is the design of the task that determines how successfully the user can contribute data. The task design has an impact on the speed at which users can complete tasks, with clicking being faster than typing. For example, a design decision to use radio buttons or freetext boxes can have a significant impact on performance (Aker et al. 2012).

In PD the player is constrained to a set of predefined options to make annotations, with freetext comments allowed (although this is not the usual mode of contribution in the game). The pre-processing of text allows the game play in PD to be constrained in this way but is subject to errors in processing that also need to be fixed.

Considering Task Difficulty

The inherent difficulty of the task can provide a challenge to more experienced users and they need to be motivated to rise to the challenge of difficult tasks.

There is a clear difference in the performance of users when we consider the difficulty of tasks in GWAP (Chamberlain et al. 2009a). One way to measure this is to **use a Gold Standard** (a set of tasks that you have the answers for) or to **use inter-annotator agreement** that is created by experts or by the users themselves.

PD compared the responses from 2 experts across a range of tasks and found that they mostly agreed with each other (average 94 %). When comparing the responses produced by users of the game, the agreement would be in a similar range to expert agreement for simple tasks (average 90 %) but much lower for more difficult tasks (average 71 %) (Chamberlain et al. 2009a).

Setting Time Limits

A time limitation will elicit spontaneous answers from users, whereas no limitations gives users time to make a more considered response. The design of the task must

[10] http://www.usability.gov/guidelines

balance the increase in excitement a timed element can offer with the need to allow users time to give good quality answers.

The timing of tasks is usually required in the game format, either as a motivational feature or as a method of quality control (or both) (von Ahn and Dabbish 2008). In PD there are no timing constraints, although the time taken to perform a task is used to assess the quality of annotations. As the task in PD is text based, it was considered important to give players time to read documents at a relatively normal speed whilst completing tasks and this was confirmed by usability studies of the interface.

Measuring System Performance

System performance can be measured by the speed at which the users can process the input source (e.g. text, images) and deliver their response (e.g. a click, typing). This measure is called **throughput**, the number of labels (or annotations) per hour (von Ahn and Dabbish 2008). As well as measuring how well the task is presented in the interface, throughput is also an indication of task difficulty and cognitive load on the users.

Related to throughput is the **wait time** for tasks to be done. Most crowdsourcing systems allow data collection in parallel (i.e., many participants can work at once on the same tasks), although validation requires users to work in series (i.e., where one user works on the output of another user). Whilst the throughput gives us a maximum speed from the system, it is worth bearing in mind that the additional time spent waiting for a user to be available to work on the task may slow the data collection. Some systems deployed on Amazon Mechanical Turk pay workers a small retainer to be act as an on demand workforce (Bernstein et al. 2012).

Attracting Users

In order to attract the number of participants required to make a success of the system, it is not enough to develop an attractive interface; it is also necessary to develop effective forms of advertising. The number of websites competing for attention is huge and without some effort to raise the profile, it will never catch the attention of enough users.

Advertising

Not all advertising methods are equally successful and it is important to evaluate which works best for the task interface, delivery platform and target audience demographics. Traditional banner or pay-per-click advertising may go some way to attracting users, however in a rapidly changing landscape of Internet habits it would be worth investigating novel methods of delivery. For example, with a system that produces lots of content a dynamic and active Facebook news feed would engage more users in a social network rather than a static banner advert.

PD had a modest budget for pay-per-click advertising and considerable effort was made to promote the project in local and national press, on science websites, blogs, bookmarking websites, gaming forums, special interest email lists, conferences, tutorials and workshops.

The importance of promoting an interface should not be underestimated and **an advertising budget (both time and money) should be allocated** at an early stage.

The success of advertising methods can be analysed with user tracking tools such as Google Analytics.[11] This can be used to not only investigate the most successful venues for advertising to your audience, but also to analyse their behaviour when they come to your site. A useful figure is the bounce rate (the percentage of single-page visits, where the user leaves on the page they entered on) which shows how many casual users are being converted to users of the interface. Analysis of PD traffic data showed that Facebook pay-per-click banner adverts had a very high bounce rate (90 %), meaning that 9 out of 10 users that came from this source did not play the game. For this reason advertising budget was redirected to other sources of users.

Using Social Networks

Given the social nature of Human Computation it seems logical to deploy systems on platforms where the users are already networked. In recent years social networking has become the dominant pastime online. As much as 22 % of time online is spent on social networks like Facebook, Twitter and others. This is three times the amount of time spent emailing and seven times the amount of time spent searching the Internet.[12]

The success of social network games such as Cityville, with over 50 million active players each month, or The Sims, Farmville and Texas HoldEm Poker, with over 30 million active monthly players each, show that the potential for large scale participation is possible using social networking platforms.[13]

Social incentives can be made more effective when the interface is embedded within a social networking platform such as Facebook. In such a setting, users motivated by the desire to contribute to a communal effort may share their efforts with their friends, whereas those motivated by a competitive spirit can compete against each other. Surveys have shown that the majority of social game players start to play because of a friend recommendation.[14, 15]

---

[11] http://www.google.co.uk/analytics

[12] http://mashable.com/2010/08/02/stats-time-spent-online

[13] http://www.appdata.com

[14] http://www.infosolutionsgroup.com/2010_PopCap_Social_Gaming_Research_Results.pdf

[15] http://www.lightspeedresearch.com/press-releases/it's-game-on-for-facebook-users

Motivating Users

There are three main incentive structures that can be used to motivate users: personal; social; and financial (Chamberlain et al. 2009b). These directly relate to other classifications of motivations in previous research: Love; Glory; and Money (Malone et al. 2009). All incentives should be applied with caution as rewards have been known to decrease annotation quality (Mrozinski et al. 2008).

It is important to distinguish between **motivation to participate** (why people start doing something) and **motivation to contribute** (why they continue doing something) (Fenouillet et al. 2009). Once both conditions are satisfied we can assume that a user will continue contributing until other factors such as fatigue or distraction break the cycle. This has been called **volunteer attrition**, where a user's contribution diminishes over time (Lieberman et al. 2007).

Personal Incentives

Personal incentives are evident when simply participating is enough of a reward for the user. Generally, the most important personal incentive is that the user feels they are contributing to a worthwhile project; however personal achievement and learning can also be motivating factors.

Projects may initially attract collaborators because they are contributing to a resource from which they may directly benefit and these are usually the people that will be informed first about the research. However, in the long term, most contributors will never directly benefit from the resources being created. It is therefore essential to provide some more generic way of expressing the benefit to the user.

This was done in PD with a BBC radio interview by giving examples of natural language processing techniques used for Web searching. Although this is not a direct result of the language resources being created by the project, it is the case for efforts of the community as a whole, and this is what the general public can understand and be motivated by.

People who contribute information to Wikipedia are motivated by personal reasons such as the desire to make a particular page accurate, or the pride in one's knowledge in a certain subject matter (Yang and Lai 2010). This motivation is also behind the success of *citizen science* projects, such as the Zooniverse collection of projects (Raddick et al. 2010) (see also the chapter on citizen science participation by Reed, et al.), where the research is conducted mainly by amateur scientists and members of the public.

When users become more interested in the purpose of the project than the incentives it becomes more like a citizen science approach where users are willing to work on harder tasks, provide higher quality data and contribute more.

Social Incentives

Social incentives reward users by improving their standing amongst their peers (their fellow users and friends). By tracking the user's effort they can compete in leaderboards and see how their efforts compare to their peers. Assigning named levels for points awarded for task completion can be an effective motivator, with users often using these as targets i.e., they keep working to reach a level before stopping (von Ahn and Dabbish 2008), however results from PD do not support this (Chamberlain et al. 2012).

News feed posts are a simple way users can make social interactions from an interface that is integrated into social networks such as Facebook or Twitter. PD allows its players to make an automatically generated post to their news feed which will be seen by all of their friends.[16]

These posts include a link back to the game and has been a very important factor in recruiting more users, as well as motivating existing users by social incentives.

Financial Incentives

Financial incentives reward effort with money. Direct financial incentives reward the user for the completion of a task or for successfully competing against other users (for example, achieving a high score). The former is the main method of motivating users of microworking systems. The per-task reward however may encourage users to manipulate the system, to do minimum work for maximum reward.

Indirect financial incentives reward the user irrespective of the work they have done such as entering each completed task into a lottery where the winner is randomly selected (although doing more tasks would increase your chance of winning).

In PD and other games indirect financial incentives were sent as Amazon vouchers by email to the winners as this allows the prize to be invoiced, tracked and collected with minimum administrative effort.

Whilst financial incentives seem to go against the fundamental idea behind GWAP (i.e., that enjoyment is the motivation), it actually makes the enjoyment of potentially winning a prize part of the motivation. Prizes for high scoring players will motivate hard working or high quality players but the prize soon becomes unattainable for the majority of other players. By using a lottery style financial prize the hard working players are more likely to win, but the players who only do a little work are still motivated. Prize-based financial incentives present a risk that not enough work will be collectively done by the conclusion of the prize period, however if the users are correctly motivated it should prove much more cost-effective than pay-per-task incentives.

[16] Since the initial development of PD Facebook has changed how posts are displayed. Posts from the game now appear on the user's profile and in a news ticker.

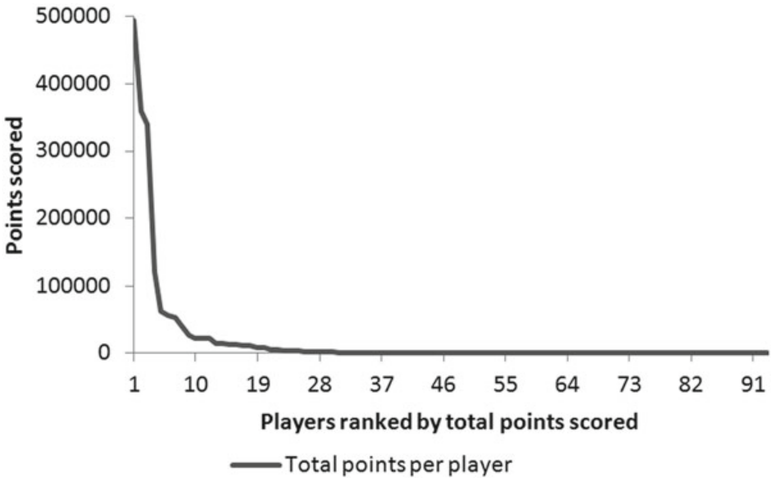


Fig. 2 Chart showing each player on the x-axis ranked by total points scored (approximately equivalent to workload) in Phrase Detectives

Whilst financial incentives are important to recruit new users, a combination of all three types of incentives is essential for the long term success of a project (Smadja 2009).

Evaluating Participation and Contribution

We can measure the success of advertising and the motivation to join the project (motivation to participate) by how many users have registered over the period of time. However, this may not be a good predictor of how much work will be done, how fast it will be completed or of what quality it will be.

Participation of users to contribute is a way to assess whether the incentives of an approach are effective. We measure motivation to contribute by the average lifetime participation.

One observation that is apparent in most crowdsourcing systems is the uneven distribution of contribution per person, often following a Zipfian power law curve—see Fig. 2 (Chamberlain et al. 2012).

An approach to improve data quality would be to focus training and incentives on the few users that are contributing significantly. However, the influence of users who only contribute a little should not be undervalued as in some systems it can be as high as 30 % of the workload (Kanefsky et al. 2001) and this is what makes the collective decision making robust. Increasing the participation from the "long tail" is key to improving the quality of the human computation.

Evaluating Users

The strategies for quality control address five main issues:

1. Training Users
2. Reducing Genuine Mistakes
3. Allowing for Genuine Ambiguity
4. Controlling Malicious Behaviour
5. Identifying Outliers

Training Users

A training stage is usually required for users to practice the task and to show that they have sufficiently understood the instructions to do a real task. The task design needs to **correlate good user performance with producing good quality data**. The level of task difficulty will drive the amount of training that a user will need and the training phase has been shown to be an important factor in determining quality and improvement in manual annotation (Dandapat et al. 2009).

Training should assume a layman's knowledge of the task and should engage the participant to increase their knowledge to become a pseudo-expert. The more they participate, the more expert they become. This graduated training makes a rating system (where the user is regularly judged against a gold standard) essential to give appropriately challenging tasks.

Most projects, at least initially, will have a core of collaborators to test and perform tasks and these are most likely to be friends or colleagues of the task designers. It can therefore be assumed that this base of people will have prior knowledge of the task background, or at least easy access to this information. These pre-trained collaborators are not the "crowd" that crowdsourcing needs if it is to operate on a large scale nor are they the "crowd" in the wisdom of the crowd.

Reducing Genuine Mistakes

Users may occasionally make a mistake and press the wrong button. Attention slips need to be identified and corrected by validation, where users can examine other users' work and evaluate it. Through validation, poor quality interpretations should be voted down and high quality interpretations should be supported (in the cases of genuine ambiguity there may be more than one). The validation process is a second stage to the data collection, that allows the task to be more varied, to make the data collection more efficient (validation is only required when there is disagreement) and to create a sense of user community and responsibility. Validation thus plays a key role as a strategy for quality control.

Unlike open collaboration in Wikipedia, it is not advisable to allow players of GWAP to go back and correct their mistakes, otherwise a player could try all possible variations of an answer and then select the one offering the highest score. In this sense the way players work together is more “collective”, where individual work is aggregated after collection, than “collaborative”, where users work more directly with each other.

Allowing for Genuine Ambiguity

The strength of Human Computation Systems is the ability to capture ambiguity in the data. Systems should not only aim to select the best, or most common, answer or annotation from users but also to preserve all inherent ambiguity, leaving it to subsequent processes to determine which interpretations are to be considered spurious and which instead reflect genuine ambiguity.

Collecting multiple judgements about linguistic expressions is a key aspect of PD. In the current configuration, eight players are asked to express their judgements on a task. If they do not agree on a single interpretation, four more players are then asked to validate each interpretation.

Validation has proven very effective at identifying poor quality interpretations. The value obtained by combining the player annotations with the validations for each interpretation tends to be zero or negative for all spurious interpretations.

Controlling Malicious Behaviour

Controlling cheating may be one of the most important factors in Human Computation System design. All crowdsourcing systems attract spammers, which can be a very serious issue (Feng et al. 2009; Mason and Watts 2009; Kazai 2011). However, in a game context we can expect spamming to be much less of an issue as the work is not conducted on a pay-per-task basis.

Nevertheless, several methods are used in PD to identify players who are cheating or who are providing poor annotations. These include checking the player’s IP address (to make sure that one player is not using multiple accounts), checking annotations against known answers (the player rating system), preventing players from resubmitting decisions (Chklovski and Gil 2005) and keeping a blacklist of players (von Ahn 2006).

A method of profiling players was also developed for PD to detect unusual behaviour. The profiling compares a player’s decisions, validations, skips, comments and response times against the average for the entire game—see Fig. 3. It is very simple to detect players who should be considered outliers using this method (this may also be due to poor task comprehension as well as malicious input) and their data can be ignored to improve the overall quality.

	System	Good player	Bad player
ANNOTATIONS			
Total Annotations:	1423078	4587	11018
Average Annotation Time:	00:00:07	00:00:07	00:00:04
Total (Ratio) DN:	955520 (0.67)	1495 (0.33)	10935 (0.99)
Total (Ratio) DO:	378256 (0.27)	2696 (0.59)	58 (0.01)
Total (Ratio) PR:	79172 (0.06)	334 (0.07)	24 (0)
Total (Ratio) NR:	13395 (0.01)	64 (0.01)	2 (0)
VALIDATIONS			
Total Validations:	608982	3848	5256
Total (Ratio) Agree:	200174 (0.33)	1186 (0.31)	8 (0)
Ave Agree Time:	00:00:09	00:00:08	00:00:18
Total (Ratio) Disagree:	408808 (0.67)	2662 (0.69)	5248 (1)
Ave Disgree Time:	00:00:08	00:00:07	00:00:02
OTHER			
Total Skips:	51616	142	26
Skip per annotation:	0.04	0.03	0
Total Comments:	26593	229	0
Comment per annotation:	0.02	0.05	0

Fig. 3 Player profiling in Phrase Detectives, showing the game totals and averages (*left*), a good player profile (*centre*) and a bad player profile (*right*) taken from real game profiles. The bad player in this case was identified by the speed of annotations and that the only responses were DN in Annotation Mode and Disagree in Validation Mode. The player later confessed to using automated form completion software

Identifying Outliers

It would be possible to ignore contributions from users who have a low rating (judged against a gold standard) however without a gold standard it is difficult to judge the performance of a user.

Variables such as annotation time could be a factor in filtering the results. An annotation in PD takes between 9 and 11 seconds and extreme variation from this may indicate that a poor quality decision has been made.

A different approach could be to identify those users who have shown to provide high quality input. A knowledge source could be created based on input from these users and ignore everything else. Related work in this area applies ideas from citation analysis to identify users of high expertise and reputation in social networks by, for example, adopting the HITS algorithm (Yeun et al. 2009) or Google's PageRank (Luo and Shinaver 2009).

Conclusion

This chapter discussed methods that can be used to engage, motivate and evaluate users of crowdsourced Human Computation Systems.

Interfaces should be attractive enough to encourage users to contribute. The design of the task itself will be determined in part by the complexity of the data being collected. By identifying the difficult or ambiguous tasks, the pre- and post-processing can be improved and the human input can be maximised to produce the highest quality resource possible given the inherent difficulty of the task. The task design should be streamlined for efficient collection of data and the throughput (annotations per hour) of the system is a good measure of this. The additional time spent waiting for a user to be available to work on the task may also slow the system.

Most users will not benefit directly from their participation, however their connection to the project and sense of contribution to science are strong motivating factors with the citizen science approach, where users are willing to work on harder tasks, provide higher quality data and contribute more. Motivational issues are less of a concern when users are intrinsically motivated to participate, as they will directly benefit from their contribution.

It is common for the majority of the workload to be done by a minority of users. Motivating the right kind of users is a complex issue and is as important as attracting large numbers of users. Controlling cheating may be one of the most important factors in crowdsourcing design and is especially problematic for a microworking approach where users are paid on a per-task basis.

The issue of data quality is an area of continuous research. The ultimate goal is to show that resources created using Human Computation Systems potentially offer higher quality and are more useful by allowing for ambiguity. By quantifying the complexity of the tasks, human participants can be challenged to solve computationally difficult problems that would be most useful to machine learning algorithms.

Acknowledgements The original Phrase Detectives game was funded as part of the EPSRC AnaWiki project, EP/F00575X/1.

References

Aker A, El-Haj M, Albakour D, Kruschwitz U (2012) Assessing crowdsourcing quality through objective tasks. In: Proceedings of LREC'12, Istanbul

Albakour M-D, Kruschwitz U, Lucas S (2010) Sentence-level attachment prediction. In: Proceedings of the 1st information retrieval facility conference. Volume 6107 of lecture notes in computer science, Vienna. Springer, pp 6–19

Bernstein MS, Karger DR, Miller RC, Brandt J (2012) Analytic methods for optimizing realtime crowdsourcing. CoRR

Chamberlain J, Poesio M, Kruschwitz U (2008) Phrase detectives: a web-based collaborative annotation game. In: Proceedings of the international conference on semantic systems (I-Semantics'08), Graz

Chamberlain J, Kruschwitz U, Poesio M (2009a) Constructing an anaphorically annotated corpus with non-experts: Assessing the quality of collaborative annotations. In: Proceedings of the 2009 workshop on the people's web meets NLP: collaboratively constructed semantic resources, Singapore

Chamberlain J, Poesio M, Kruschwitz U (2009b) A new life for a dead parrot: incentive structures in the phrase detectives game. In: Proceedings of the WWW 2009 workshop on web incentives (WEBCENTIVES'09), Madrid

Chamberlain J, Kruschwitz U, Poesio M (2012) Motivations for participation in socially networked collective intelligence systems. In: Proceedings of CI2012, Boston

Chamberlain J, Fort K, Kruschwitz U, Mathieu L, Poesio M (2013) Using games to create language resources: successes and limitations of the approach. In: ACM transactions on interactive intelligent systems, volume The People's Web Meets NLP: collaboratively constructed language resources. Springer pp 3–44

Chklovski T, Gil Y (2005) Improving the design of intelligent acquisition interfaces for collecting world knowledge from web contributors. In: Proceedings of K-CAP '05, Banff

Dandapat S, Biswas P, Choudhury M, Bali K (2009) Complex linguistic annotation – no easy way out! a case from Bangla and Hindi POS labeling tasks. In: Proceedings of the 3rd ACL linguistic annotation workshop, Singapore

Feng D, Besana S, Zajac R (2009) Acquiring high quality non-expert knowledge from on-demand workforce. In: Proceedings of the 2009 workshop on the people's web meets NLP: collaboratively constructed semantic resources, Singapore

Fenouillet F, Kaplan J, Yennek N (2009) Serious games et motivation. In: 4eme conference francophone sur les environnements informatiques pour l'apprentissage humain (EIAH'09), vol. Actes de l'Atelier "Jeux Serieux: conception et usages", Le Mans

Howe J (2008) Crowdsourcing: why the power of the crowd is driving the future of business. Crown Publishing Group, New York

Kanefsky B, Barlow N, Gulick V (2001) Can distributed volunteers accomplish massive data analysis tasks? In: Lunar and planetary science, XXXII, Houston

Kazai G (2011) In search of quality in crowdsourcing for search engine evaluation. In: Proceedings of the 33^{rd} european conference on information retrieval (ECIR'11), Dublin

Lieberman H, Smith DA, Teeters A (2007) Common consensus: a web-based game for collecting commonsense goals. In: Proceedings of IUI, Honolulu

Luo X, Shinaver J (2009) MultiRank: reputation ranking for generic semantic social networks. In: Proceedings of the WWW 2009 workshop on web incentives (WEBCENTIVES'09), Madrid

Malone T, Laubacher R, Dellarocas C (2009) Harnessing crowds: mapping the genome of collective intelligence. Research Paper No. 4732-09, Sloan School of Management, Massachusetts Institute of Technology, Cambridge

Mason W, Watts DJ (2009) Financial incentives and the "performance of crowds". In: Proceedings of the ACM SIGKDD workshop on human computation, Paris

Mrozinski J, Whittaker E, Furui S (2008) Collecting a why-question corpus for development and evaluation of an automatic QA-system. In: Proceedings of ACL-08: HLT, Columbus

Poesio M, Chamberlain J, Kruschwitz U, Robaldo L, Ducceschi L (2013) Phrase detectives: utilizing collective intelligence for internet-scale language resource creation. ACM transactions on interactive intelligent systems 3:1–44

Quinn A, Bederson B (2011) Human computation: a survey and taxonomy of a growing field. In: CHI, Vancouver

Raddick MJ, Bracey G, Gay PL, Lintott CJ, Murray P, Schawinski K, Szalay AS, Vandenberg J (2010) Galaxy zoo: exploring the motivations of citizen science volunteers. Astronomy Educ Rev 9(1):010103

Smadja F (2009) Mixing financial, social and fun incentives for social voting. World Wide Web Internet And Web Information Systems

Snow R, O'Connor B, Jurafsky D, Ng AY (2008) Cheap and fast—but is it good?: Evaluating non-expert annotations for natural language tasks. In: EMNLP '08: Proceedings of the conference on empirical methods in natural language processing, Honolulu

Surowiecki J (2005) The wisdom of crowds. Anchor, New York

Thaler S, Siorpaes K, Simperl E, Hofer C (2011) A survey on games for knowledge acquisition. Technical report STI TR 2011-05-01, Semantic Technology Institute

von Ahn L (2006) Games with a purpose. Comput 39(6):92–94

von Ahn L, Dabbish L (2008) Designing games with a purpose. Commun ACM 51(8):58–67

Wang A, Hoang CDV, Kan MY (2010) Perspectives on crowdsourcing annotations for natural language processing. Language resources and evaluation, pp 1–19

Yang H, Lai C (2010) Motivations of wikipedia content contributors. Comput Hum Behav 26:1377–1383

Yeun CA, Noll MG, Gibbins N, Meinel C, Shadbolt N (2009) On measuring expertise in collaborative tagging systems. In: Proceedings of WebSci'09, Athens

Participating in Online Citizen Science: Motivations as the Basis for User Types and Trajectories

Jason T. Reed, Ryan Cook, M. Jordan Raddick, Karen Carney, and Chris Lintott

Introduction

Citizen science relies on attracting and maintaining sufficient amounts of volunteer activity to fulfill project goals (for a review of citizen science and its relationship to human computation, see the chapter by Lintott and Reed). Within the last decade, a growing subset of citizen science projects known as *virtual citizen science* (VCS; Wiggins and Crowston 2011) have increasingly incorporated forms of computer-mediated communication such as email, blogs, and websites into project activities to reach a larger base of potential volunteers. Not only does this increase the breadth of the potential volunteer pool, but also the amount of effort contributed towards the projects' goals. For example, the eBird project increased the size and availability of its catalogue of bird watching information because it created a website for bird-watchers from around the world to record their own efforts as well as access those of other birdwatchers (Wiggins 2011). As another example, members of the public used the gaming interface of the FoldIt project to solve a protein-folding problem in a matter of weeks that puzzled professional researchers for years (Khatib et al. 2011). The Zooniverse, a site focused on citizen science, offers a variety of VCS projects to more than 800,000 registered volunteers. In addition to contributing to the science behind dozens of peer-reviewed publications, Zooniverse volunteers

J.T. Reed (✉) • R. Cook • K. Carney
Adler Planetarium, Chicago, USA
e-mail: jreed@adlerplanetarium.org; rcook@adlerplanetarium.org; kcarney@adlerplanetarium.org

M.J. Raddick
Johns Hopkins University, Baltimore, USA
e-mail: raddick@jhu.edu

C. Lintott
University of Oxford, UK
e-mail: chris@zooniverse.org

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_55,

also made serendipitous discoveries of previously unseen astronomical objects and classes of galaxies (Cardamone et al. 2009; Lintott et al. 2009).

This chapter focuses on understanding the motivations associated with participants volunteering their time and effort to VCS. We begin by describing the Zooniverse and the behavior and motivation of existing volunteers on the site. We then describe our research on the behavior and motivations of first-time Zooniverse volunteers. Finally, we end this chapter with a summary of our findings and a discussion of future research about the motivation to participate in VCS.

The Zooniverse

At present, the Zooniverse functions as a central hub for 19 active and three completed projects. Zooniverse projects contain a primary science task to which volunteers contribute their time and effort. These projects use different interfaces depending on what type of volunteer activity is required to create valid and reliable data for the research investigation (Fortson et al. 2012).

In addition to the primary science tasks, Zooniverse projects also have online forums and blog posts from project scientists. These forums were integral to the serendipitous discoveries made in the Galaxy Zoo project because they provided a platform for volunteers to post pictures of interesting objects that professional astronomers later confirmed as new astronomical phenomena like a new class of galaxy (Reed et al. 2013b). However, the scientists behind the Zooniverse wanted more integration between volunteer activity in the primary science tasks and these discussion tools to encourage volunteers to make more serendipitous discoveries and engage in more aspects of the process of scientific inquiry (e.g., analyzing data, writing papers). As a result, the "Talk" tool was added to provide a single destination to encourage conversation and collaboration among Zooniverse volunteers through activities like posting on various task and social forums, creating personal pages, and tagging collections of objects interesting to the volunteers. Beginning with the Planet Hunters project, all subsequent Zooniverse projects ask volunteers as they complete a set of data classifications whether they would like to use the Talk tools to discuss any aspects of their activity.

Volunteer Behavior in the VCS Projects of the Zooniverse

Analysis of the Zooniverse website user logs suggests that volunteers vary in their use of the primary science task and the Talk tools. Much of the activity in Zooniverse projects occurs in the primary science tasks of the project and is often done by a small proportion of the overall volunteer base. This same skew in behavior also occurs in online discussions or communication amongst project volunteers, with a few people making the majority of the statements. Furthermore, the volunteers with a history of participating in the primary research tasks of a project are also more likely to use the "Talk" tools.

These patterns of volunteer behavior also appear in similar online environments like Wikipedia that depend on volunteer contributions to projects (Ortega 2009). As such, one should not view the behavior of Zooniverse volunteers as anomalous or problematic. Instead, we suggest that different kinds and amounts of volunteer activity on a VCS project reflect different forms of meaningful contribution, especially from the volunteers' own points of view. The task then becomes how to best describe these differing reasons for volunteering. We begin by reviewing research into the possible motivations for volunteers with a history of contributing to VCS projects.

Motivation of Existing VCS Volunteers

Existing research suggests a variety of reasons to volunteer in VCS. For example, volunteers from a variety of projects that required large tasks to be divided into smaller pieces of work among multiple volunteers and their computers completed an online survey about their motivations for participating. These volunteers reported reasons for contributing such as contributing to scientific research, effectively using available computer resources, acquisition of technical knowledge, and competition or interaction with other volunteers (Holohan and Garg 2005). Volunteers in astronomy-based VCS projects participated not only for individually important reasons, but also because they wanted to fit in with, identified with, or support the efforts of relevant groups (Nov et al. 2011). A survey of both the scientists who created and volunteers who participated in a data curation VCS project suggested that volunteers participated to increase their own welfare as well as that of the project (Rotman et al. 2012).

Although these studies provide valuable insights in possible motives for volunteering in VCS, it is tricky to synthesize their findings because each study assessed volunteer motivations in different VCS projects. A more comprehensive sense of motivations may come from a set of studies about participation in the Galaxy Zoo project. The Galaxy Zoo is one of the VCS projects in the Zooniverse that asks volunteers to participate in actual scientific research activities like transforming raw data into forms suitable for analysis. Analysis of 20 interviews of Galaxy Zoo volunteers yielded 12 different volunteer motives such as contributing to scientific research, learning about galaxies, making discoveries, interacting with other people, teaching other people, looking at pleasing images, fun, helping, amazement about the vastness of the universe, interest in the Galaxy Zoo project, interest in the field of astronomy, and interest in science in general (Raddick et al. 2010). A subsequent survey of a larger sample of Galaxy Zoo volunteers corroborated these 12 motives and found that contribution to science was most important among them (Raddick et al. in press). Mankowski et al. (2011) analyzed the content of online forum posts of Galaxy Zoo and also found a similar set of motives for participation: social interaction, interest in astronomy-related topics like the space race, spiritual aspects of the project material, and a strong pleasing aesthetic appeal of the galaxy images. Because these studies all examined the same VCS project, a clearer picture of the motivations related to volunteering emerges.

Because Galaxy Zoo was its first successful project, the Zooniverse provides a natural extension of this research about motivation to participate in VCS projects. Reed et al. (2013a) surveyed 199 registered Zooniverse volunteers using a web-based survey of 54 items. The questions for the survey used items from previous research on motivation to participate in online activities as well as items created specifically for the survey. Volunteers' responses were analyzed using an exploratory factor analysis to examine any latent motivation constructs underlying the responses to individual survey items. The results of the exploratory factor analysis suggested three general categories of motivation for participating in the Zooniverse:

1. Social Engagement – awareness of and interaction with other members of the Zooniverse community;
2. Interaction with Website – sense of awareness, facility, and enjoyment from using the various features of Zooniverse projects; and
3. Helping – how participants experience positive feelings from helping or volunteering to participate in the Zooniverse projects.

Such results advance our understanding of motivation to participate in VCS because they build on the existing body of knowledge while making contributions like indicating the presence of underlying motivational dimensions. Some previous research (e.g., Holohan and Garg 2005; Raddick et al. 2010) hypothesized motives after gathering and analyzing the data, whereas other research (e.g., Nov et al. 2011; Rotman et al. 2012) judged their findings in regards to a priori models of motivation. In the same vein, we let underlying dimensions of motivation arise from the data while relying on existing theory and hypotheses to interpret any findings. As such, we are encouraged that these results are consistent with a reanalysis of a subset of items from Raddick et al. (in press) that suggested participation in Galaxy Zoo was primarily driven by engagement with the project's content, social opportunities, and scientific aspects (Reed and Raddick 2012).

Motivation of First-Time VCS Volunteers

Yet for all of the research on motivation of existing VCS volunteers, none that we are aware of has addressed their first experience with a particular VCS project. Although all volunteers must go through this critical period of initial contact with a VCS project, its transient nature makes it difficult to study volunteer reactions as they occur. As such, we conducted interviews with visitors to a Midwestern planetarium after their first use of one of the astronomy related Zooniverse projects. Interviews would allow us to capture maximal information from the expected small number of volunteers about their reactions to their first experience with a VCS project.

Reed and Carney (2011) interviewed 30 visitors to a Midwestern planetarium after they used one of three astronomy-related Zooniverse projects for about 10 min.

When asked to describe what they did on the project websites, visitors provided accurate descriptions and details of the activities and features of the Zooniverse project they used. While completing the various project tasks, the visitors commented primarily about the mechanics of the task itself and the challenge it presented. They also commented on the opportunities for personal learning, discovery, and contribution to the project afforded by even their brief time spent on the Zooniverse project. Visitors overwhelmingly enjoyed what they did, with more than 90 % of evaluative comments being positive in nature. Importantly, 75 % of the visitors actually revisited the Zooniverse project website they used during the interview sessions on their own and did more of its primary science activity. To better understand these results, we next consider them in combination with the results of our survey of existing Zooniverse volunteers.

Summary and Future Directions

Taken as a whole, our survey of existing Zooniverse volunteers and our interviews of first-time Zooniverse volunteers offer intriguing insights into and future directions for research about motivation to participate in VCS. Not surprisingly, different motives appear to be relevant at different points in a volunteer's lifespan in a VCS project (Crowston and Fagnot 2012). First-time Zooniverse volunteers commented more about their own behaviors and reactions during their interviews, whereas many of the survey items endorsed by existing Zooniverse volunteers noted the presence of and possible interactions with other Zooniverse volunteers.

Although volunteers of both types consider their personal reactions and motivations, these may be especially relevant to first-time volunteers as they familiarize themselves with what is possible in the VCS projects. It is analogous to a person learning to drive a car for the first time in that new drivers focus primarily on the what they must do to functionally operate the car; once these basic actions become more familiar, the drivers can then turn their attention to other aspects of the driving experience like their subjective sensations or the other drivers around them. This initial learning curve can also be an enjoyable experience, as evidenced by the positive reactions and return visits from the first-time Zooniverse volunteers. These differences in motivation could also be related to how different Zoonivese project websites display differing degrees of "stickiness" like amount of time volunteers spend on project activity and the number of project webpages visited, as well as differing degrees of usability (Reed et al. 2012a, b). Future research needs to explore how the design features of the websites are relevant to different motivations at different points in a volunteer's lifespan.

Conversely, there are also commonalities in the motives of first-time and existing Zooniverse volunteers. The comments from first-time Zooniverse volunteers about learning and contributing to the projects corroborate our survey of existing Zooniverse users, indicating motivations like interacting with the website and helping with the project. Given that participation in citizen science has the potential to increase volunteers' knowledge or appreciation of science (Bonney et al. 2009; Price and Lee in press; Trumbell et al. 2000; but see Brossard et al. 2005; Cronje et al. 2011), more research needs to be done on how participation in different parts of the Zooniverse project websites afford volunteers direct contact with the scientific process.

Recall that Zooniverse volunteers tend to spend the majority of their time participating in the primary science tasks of the Zooniverse websites. This was especially true of first-time Zooniverse volunteers in both their initial and return visits to the project websites. These areas of the website may offer opportunities to engage in science that are specific to a particular skill or area of study. For example, the primary science task of the Galaxy Zoo project draws on pattern recognition skills relevant to the morphology of objects like galaxies. It would be illuminating to compare use of the primary science task with how volunteers use the "Talk" tools and how these activities might allow for deeper and more varied types of science experiences (e.g., serendipitous discoveries and discussion about particular observations). Knowing what kinds of science experiences are available with different VCS activities would further refine our understanding of motives for volunteering in VCS. Future work should determine what relationships exist among volunteers' motivations for participation, their actual patterns of engagement with VCS websites, and any demonstrated changes in their understanding of and attitudes towards science.

Conclusion

VCS relies on the efforts of large numbers of volunteers from all over the world to conduct reliable and valid scientific research. The increasing number of VCS projects attest to their attractiveness to both researchers who create them and volunteers who participate. A better understanding of the variety of motives that drive the different types of volunteers to engage in different types of VCS activity can help create projects that provide maximal benefits to all parties involved. Our research on motivations for participation in the Zooniverse advance the understanding of this critical topic and should ultimately lead to a better understanding of how to truly create more authentic and engaging science experiences for the volunteers.

Acknowledgments This material is based upon work supported by the National Science Foundation under Grant No. 0917608. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

Bonney R, Cooper CB, Dickinson J, Kelling S, Phillips T, Rosenberg KV, Shirk J (2009) Citizen science: a developing tool for expanding science knowledge and scientific literacy. BioScience 59:977–984

Brossard D, Lewenstein B, Bonney R (2005) Scientific knowledge and attitude change: the impact of a citizen science project. Int J Sci Educ 27:1099–1121

Cardamone CN, Schawinski K, Sarzi M, Bamford SP, Bennert N, Urry CM, Lintott C et al (2009) Galaxy Zoo green peas: discovery of a class of compact extremely star-forming galaxies. Mon Not R Astron Soc 399:1191–1205

Cronje R, Rohlinger S, Crall A, Newman G (2011) Does participation in citizen science improve scientific literacy? A study to compare assessment methods. Appl Environ Educ Commun 10:135–145

Crowston K, Fagnot I (2012) The motivational arc of massive online collaboration, Retrieved from http://floss.syr.edu/content/motivational-arc-massive-virtual-collaboration

Fortson L, Masters K, Nichol R, Borne K, Edmondson E, Lintott C, Raddick J, Schwaniski K, Wallin J (2012) Galaxy Zoo: morphological classification in citizen science. In: Way MJ, Scargle JD, Ali K, Srivastava AN (eds) Advances in machine learning and data mining for astronomy. CRC Press, Boca Raton, pp 213–236

Holohan A, Garg A (2005) Collaboration online: the example of distributed computing. J Comput-Mediat Commun 10, Retrieved from http://jcmc.indiana.edu/vol10/issue4/holohan.html

Khatib F, DiMaio F, Foldit Contenders Group, Foldit Void Crushers Group, Cooper S, Kazmierczyk M, Gilski M, Krzywda S, Zabranska H, Pichova I, Thompson J, Popović Z, Jaskolski M, Baker D (2011) Crystal structure of a monomeric retroviral proteasesolved by protein folding game players. Nat Struct Mol Biol, Retrieved from http://homes.cs.washington.edu/~zoran/NSMBfoldit-2011.pdf

Lintott CJ, Schawinski K, Keel W, Van Arkel H, Bennert N, Edmondson E, Thomas D et al (2009) Galaxy Zoo: "Hanny's Voorwerp", a quasar light echo? Mon Not R Astron Soc 399:129–140

Mankowski TA, Slater SJ, Slater TF (2011) An interpretive study of meanings of citizen scientists make when participating in Galaxy Zoo. Contemp Issues Educ Res 4:25–42

Nov O, Arazy O, Anderson D (2011) Technology-mediated citizen science participation: a model. In: Proceedings of the AAAI international conference on weblogs and social media (ICWSM 2011), Barcelona

Ortega F (2009) Wikipedia: a quantitative analysis. (Unpublished doctoral dissertation), Universidad Rey Juan Carlos, Mostoles

Price CA, Lee H (in press) Changes in participants' scientific attitudes and epistomological beliefs during an astronomical citizen science project. J Res Sci Teach

Raddick MJ, Bracey G, Gay PL, Lintott CJ, Cardamone C, Murray P, Schawinski K, Szalazy AS, Vandenberg J (in press) Galaxy Zoo: motivations of citizen scientists. Astron Educ Rev

Raddick J, Szalay A, Vandenberg J, Bracey G, Gay P, Lintott C, Murray P, Schawinski K (2010) Galaxy Zoo: exploring the motivations of citizen science volunteers. Astron Educ Rev 9, Retrieved from http://aer.aas.org/resource/1/aerscz/v9/i1/p010103_s1

Reed J, Carney K (2011) Perceptions of the Zooniverse by first-time users. Paper presented at the annual meeting of the American educational research association, New Orleans

Reed JT, Raddick MJ (2012) Factor structure of motivations for participating in online citizen science. Talk presented at the annual meeting of the Midwestern psychological association, Chicago

Reed J, Carney K, Rodriguez W (2012a) Classifying key design features of virtual citizen science projects. Poster presented at the conference on public participation in scientific research, Portland

Reed J, Rodriguez W, Rickhoff A (2012b) A framework for defining and describing key design features of virtual citizen science projects. In: Proceedings of the 2012 iConference, Toronto, CA, pp 623–625

Reed JT, Raddick MJ, Lardner A, Carney K (2013a) An exploratory factor analysis of motivations for participating in Zooniverse, a collection of virtual citizen science projects. In: Proceedings of the 46th annual Hawaii international conference on systems sciences, Wailea, 7–10 Jan 2013

Reed JT, Smith A, Parrish M, Rickhoff A (2013b) Using contemporary collective action to understand the use of computer mediated communication in virtual citizen science. To appear in Agarwal N, Wigand RT, Lim M (eds) Online collective action: dynamics of the crowd in social media, Springer

Rotman D, Preece J, Hammock J, Procita K, Hansen D, Parr C, Lewis D, Jacobs D, Biswas A (2012) Dynamic changes in motivation in collaborative citizen science projects. In: Proceeding of the 2012 ACM conference on computer supported cooperative work, CSCW'2012, Seattle

Trumbull DJ, Bonney R, Bascom D, Cabral A (2000) Thinking scientifically during participation in a citizen-science project. Sci Educ 84:265–275

Wiggins A (2011) eBirding: technology adoption and the transformation of leisure into science, poster presented at the 2011 iConference, Seattle

Wiggins A, Crowston K (2011) From conservation to crowdsourcing: a typology of citizen science. In: Proceedings of the 44th annual Hawaii international conference on system sciences, Koloa, 4–7 Jan 2011

Cultivating Collective Intelligence in Online Groups

Anita Williams Woolley and Nada Hashmi

The incidence of humans collaborating via computer mediation is rising over time; groups collaborate online to author encyclopedias, to write software, to optimize search engines and to solve a whole range of problems from uncovering the structure of an enzyme to documenting blotches on the surface of Mars.

One main way in which combined human and computer groups manifest in the daily lives of individuals and organizations is in the form of online or computer-mediated teams. Online teams have become so common in organizations today that surveys estimate that as many as 78 % of professional workers and executives have at some point worked in online teams (Martins et al. 2004; The Economist Intelligence Unit 2009). Online teams are used in almost all industries and in a variety of areas, such as scientific innovation (Fiore 2008), software development, customer service, sales and R&D (Carmel and Agarwal 2001; Hertel et al. 2005; McDonough et al. 2001).

This chapter examines human computation through the lens of online collaboration. It begins by considering the challenges associated with assessing group performance. This leads to a discussion of collective intelligence as a measure of group effectiveness, and considers the factors that influence this measure of group performance. These factors then serve as a focal point for developing techniques that foster the emergence of greater collective intelligence in human computation systems that manifest as collaborative groups. Next, consideration is given to how these techniques can help overcome limitations of the virtual communication medium and ultimately give rise to unprecedented degrees of collaboration efficacy. Finally, new research directions are identified.

A.W. Woolley (✉)
Carnegie Mellon University, Pittsburgh, USA
e-mail: awoolley@cmu.edu

N. Hashmi
Massachusetts Institute of Technology, Boston, USA

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_56,

Historical Challenges of Performance Measurement in Online Teams

Studies have compared the performance of traditional teams and online teams with mixed and sometimes conflicting results. While some studies report greater effectiveness for online teams (i.e., Sharda et al. 1988), others found that online teams could not outperform traditional teams (McDonough et al. 2001; Warkentin et al. 1997). Still others detected no difference between the two types of teams (Burke and Aytes 1998; Burke and Chidambaram 1996; Galegher and Kraut 1994).

The disparate conclusions regarding the performance of online teams reflect the difficulties in assessing team performance more generally. Considerable work in fields such as social psychology, organizational behavior, and industrial psychology has been conducted to characterize factors that predict group performance on individual tasks. Traditionally, performance has been examined in terms of an “input-process-output” model, where researchers observe or manipulate inputs to the teams (such as individual differences, task definition, and resources), then measure process variables, and finally observe the effects on performance (Ilgen et al. 2005). Much of this research explores why groups so often appear to under-perform, given the potential of the individuals in the group (Steiner 1972; Tziner and Eden 1985). While some of the tasks that have been examined in teams are complex and multi-faceted, such as tasks performed by top management teams (e.g., Kilduff et al. 2000; Wiersema 1992) or product development teams (Katz 1984), even in these domains performance has been examined as the outcome of a particular task at a particular point in time, despite the wide array of subtasks necessary for a team’s success. Thus, conclusions about the performance of computer-mediated teams can vary as a function of the group’s task, the technology available, or both, making generalization of conclusions across studies quite difficult.

Collective Intelligence in Human Groups

Research on collective intelligence in groups was motivated initially by a desire to measure group performance in a manner that would generalize across tasks and settings (both face-to-face and online) and predict a group’s performance on future tasks. In exploring alternate ways of conceptualizing and measuring group performance, initial studies of collective intelligence in human groups built upon work in individual psychology and concepts for understanding and predicting individual performance. Psychologists have repeatedly shown that a single statistical factor—often called “general intelligence” or “g”—emerges from the correlations among individual people’s performance on a wide variety of cognitive tasks (Deary 2000; Spearman 1904). But, perhaps surprisingly, until recently none of the research on team performance has systematically examined whether a similar kind of “collective intelligence” exists for groups of people.

Recent research has sought to explore the degree to which the concept of intelligence generalizes to groups. In two studies with 699 individuals, working in 192 groups of size two to five, researchers found converging evidence of a general collective intelligence factor that predicts a group's performance on a wide variety of tasks (Woolley et al. 2010).

The groups in this study spent approximately 5 h together in the laboratory, working on a series of tasks that required a range of qualitatively different collaboration processes (McGrath 1984). Example tasks included brainstorming uses for a brick, creating a logistical plan for a shopping trip, accurately typing a large amount of text into a computer, discussing a moral reasoning problem, and answering questions from an individual intelligence test.

In a factor analysis of all the groups' scores, the first factor accounted for 43 % of the variance in performance on all the different tasks. This is consistent with the 30 %–50 % of variance typically explained by the first factor in a battery of individual cognitive tasks (Chabris 2007). In individuals, this factor is called "intelligence." For groups, this first factor is called "collective intelligence" or "*c*," and it is a measure of the general effectiveness of a group on a wide range of tasks. Mathematically, this collective intelligence factor is a weighted average of the subtask scores, with the weights calculated to maximize the correlation with all the subtasks.

In addition to the tasks used to calculate c, each group also completed a more complex "criterion task." In the first study, groups played checkers as a team against on online computer opponent. In the second study, groups completed an architectural design problem. Both of these tasks required a combination of several of the different collaboration processes measured by the individual tasks in the collective intelligence battery. As expected, *c* was a significant predictor of team performance on both these criterion tasks, and—surprisingly—the average individual intelligence of group members was not.

The researchers also investigated what characteristics of a group predicted *c*. They found that the average and maximum intelligence of individual group members was correlated with *c*, but only moderately so. So having a group of smart people is not enough, alone, to make a smart group.

But there were three other group characteristics that were also significant predictors of *c*. First, there was a significant correlation between *c* and the average social perceptiveness of group members, as measured by the "Reading the Mind in the Eyes" test (Baron-Cohen et al. 2001). This test measures people's ability to judge other's emotions from looking only at pictures of their eyes. Groups with a high average score on this test were more collectively intelligent than other groups.

Second, *c* was negatively correlated with the variance in the number of speaking turns by group members. In other words, groups where a few people dominated the conversation were less collectively intelligent than those with a more equal distribution of conversational turn-taking.

Finally, *c* was significantly correlated with the proportion of females in the group, with groups having more females being more collectively intelligent. This result, however, is largely mediated by social perceptiveness since, consistent with previous research, women in the sample scored better on this measure than men. In a regression analysis with all three variables (social sensitivity, speaking turn variance, and percent female), all had similar predictive power for *c*, though only social perceptiveness reached statistical significance.

Taken together, these results provide strong support for the existence of a general collective intelligence factor (*c*) that predicts the performance of a group on a wide range of tasks in a variety of settings, and a consistent relationship with social perceptiveness and equality of participation among group members as significant predictors.

Mechanisms of Collective Intelligence: Balancing Convergence and Divergence

The main aim of the second section of this chapter is to identify factors enabling collective intelligence in online groups. Equality of participation and social perceptiveness are two factors consistently related to collective intelligence in our studies, as described above. They are also essential to enabling the balancing of convergence and divergence in collectives more generally (Woolley and Fuchs 2011). Here we elaborate more on the role of convergent and divergent thought in collective performance, and then consider more specifically the tools and mechanisms that promote these properties in online collectives.

Balancing Convergent and Divergent Thinking

Some argue that collective intelligence emerges from the collaboration and competition of many individual entities. Research on collective intelligence has argued that central to collectively intelligent systems is the capability to engage in both convergent and divergent modes of thought, as well as to leverage the insights from reflection into course-correcting changes (Bloom 2000; Woolley and Fuchs 2011). *Convergent thinking* is thinking that proceeds toward or converges on a single answer. In contrast, *divergent thinking* moves outwards from a problem in many directions. Both convergent and divergent thinking are necessary to collective intelligence; convergence enables decisions and the possibility of moving forward, while divergence is critical for developing the wealth of insights and ideas necessary for true innovation. However, while traditional face-to-face groups tend to excel at convergent thinking, the literature on creativity suggests that divergent thinking is an area where groups often struggle (Thompson 2003).

Both convergent and divergent thinking require particular social interaction processes to occur successfully in collectives (Larson 2009; March 1991; McGrath 1984), whether those collectives are small groups (Woolley et al. 2010) or organizations. Convergence is fostered by increased quantity and intensity of interaction; the more information group members can transfer to one another, the greater the probability of arriving at a correct answer and one that all members will support (Siegel et al. 1986). Divergence requires just the opposite; groups generate the most divergent and creative sets of ideas when members work relatively independently (Brown and Paulus 2002; Thompson 2003) to enable everyone to participate more equally and fully in idea generation. Indeed, studies of innovation in organizations encourage the development of “skunk works,” as a means of protecting the pursuit of divergent modes of thought (O’Reilly and Tushman 2008) by keeping groups of individuals pursuing different ideas relatively independent of one another (Andriopoulos and Lewis 2009; Raisch et al. 2009; see also Rosenkopf and McGrath 2011). Doing so helps prevent one set of ideas from being crowded out by or subordinated to another.

Social perceptiveness and equality of participation are likely to play an instrumental role in fostering convergent and divergent thought in collectives. Social perceptiveness allows individuals to more effectively read the nonverbal signals of others, which is associated with the ability to tune’s one message in a manner that enhances consensus-building in groups (Elfenbein et al. 2007; Elfenbein 2006; Wolff et al. 2002). At the same time, equality of participation insures that all voices, including divergent voices, are heard, raising the chances that collectives will consider a broader range of perspectives (De Dreu and West 2001). Thus, collaboration tools that can be provided to collectives in online environments to foster social perception and equality of participation are likely to enhance convergent and divergent thought and, ultimately, collective intelligence.

Tools and Mechanism That Cultivate Collective Intelligence Online

So the question remains regarding how to encourage social perceptiveness and equality of participation in online collectives? Studies of technology use in online groups suggest some places to start, as conflicting findings regarding performance in online groups seem to relate to the type of task, its reliance on primarily convergent vs. divergent properties, and the ability of the technology used to foster the appropriate processes. For example, Sharda and colleagues (1988)—who observed a high level of performance in online teams—found that groups generated a greater number of ideas using email, a technology that enhances the independence of contributions and divergent thought. Burke and Chidambaram (1996)—who found no difference between online and face-to-face groups—measured decision quality resulting from the use of online discussion boards, a medium that enhances information exchange and convergence. The disparate findings have led to many studies examining ‘task-technology-fit’ where the researchers distinguish the type of technology best suited for different tasks (e.g., Goodhue and Thompson 1995). For instance, Majchrzak et al. (2000) found video conferencing better for managing conflicts while email was better for routine tasks such as analysis of data. Others have found that online groups that rely on a wide variety of different technologies are more satisfied and perform better than those that use more limited communications tools (Kayworth and Leidner 2001). These findings support the conclusion that different technologies cultivate different processes in groups, and when those processes are well-aligned with tasks demands, they help cultivate collective intelligence.

We propose here a framework for thinking about tools in terms of their role in enhancing social perceptiveness and equality in participation, which subsequently improve the quality of convergent and divergent thought in collectives.

Tools Enabling Social Perceptiveness

Traditional teams benefit from face time that enables social cues to be relayed and picked up by other members. Online collectives face the challenge of being deprived of face time and hence these social cues. This section primarily focuses on mechanisms that enhance the availability and interpretation of social cues within online teams.

1. Signaling to flag and communicate contextual information
   Participants in face-to-face groups can use a plethora of (often unconscious) nonverbal cues, such as facial expressions and body language, to facilitate communications which are not available to online groups. While we know that many tools exist to transmit intentional, spoken language in computer mediated settings, a new crop of tools is coming on the scene intended to amplify the passive signals people use to adjust the focus of their attention and effort on collaborative products. Systems can be further designed to provide some of these passive cues automatically, or else increase the ease with which group members can actively generate similar signals.

   These tools are inspired in part by research in the field of stigmergy. First introduced by Grassé (1959), stigmergy refers to the ways cooperative animals coordinate by leaving and sensing signs in a shared environment. A classic example is the pheromone trails that ants use to optimally route and distribute their foraging behavior. Similar examples occur in many other kinds of animal cognition (Bonabeau et al. 1999; Karsai and Balazsi 2002).

   The concept of stigmergy has also been developed in the fields of robot collaboration (Holland and Melhuish 1999) and human interactions (Marsh and Onof 2008; Parunak 2006). For instance, Parunak (2006) distinguishes between explicit ("marker-based") and implicit ("sematotectonic") signaling tools. Examples of explicit signaling tools are flags individual members can use to signal their current status (i.e., "busy," "available," etc.) or emoticons they might include in messages to convey their current mood or the emotional content of messages. Implicit signals can involve the automatic capture of activities which are translated into a signal for remote collaborators, to let them know when their collaborators are distracted, uncertain, etc. An example of an implicit signal would be the capture of the rate of cursor movement or numbers of additional windows open on listener's desktop to signal remote presenters regarding the dissolution of attention being paid by listeners during a live online presentation. In collaborative problem-solving, implicit signals could be captured by measuring how long someone's cursor hovers over or revises a part of a collective product, to highlight or change the color of areas where members are less certain of input provided. Such signals could help groups coordinate work by enabling remote others to see when attention is wandering (and thus move to reengage), or by attracting more workers to parts of problems involving greater uncertainty and needing additional input.

2. Status Updates on Profiles and Social Media
   With the advent of twitter, facebook and other social media, many organizations have adopted 'enterprise microblogging (EMB)', an off-shoot of the twitter model for short updates restricted to their company network, for internal communications (Zhang et al. 2010). Studies have shown that that EMB assists in (1) awareness creation, (2) task/meeting coordination and (3) idea generation & discussion (Riemer et al. 2010; Riemer and Richter 2010). In particular, Riemer et al. (2010) noted the usefulness of EMB for the gauging and sharing of opinions on current issues. Hence, such a tool could be immensely useful for capturing and amplifying the social cues that members of a computer-mediated collective might use for gauging the mood of the group.

   Furthermore, when members of online collectives update their status on various social networking media, they are able to better relay and share a social part of themselves that otherwise remains unknown to remote collaborators. Such status updates allow others to interpret otherwise ambiguous signals, such as not hearing from someone in response to a message or query, or receiving a shorter or different response than expected. By seeing their teammates outside of the 'professional environment' and interacting with them at a more social level, increased understanding develops amongst team members, and the social cues available for contextualizing other observable behavior become richer.
3. Check-ins
   Research on effective online team leadership has noted that 'check-ins' at the beginning of a virtual meeting leads to more successful online teams (Malhotra et al. 2007; Purvanova and Bono 2009). There are different types of check-ins that can facilitate teamwork. One of them is a round robin check-in where team members share either good news or ongoing progress at the beginning of a meeting. This check-in allows team members to connect and get on the same page before conducting a meeting or a task, and provides another source of contextual information for group members to use in interpreting otherwise ambiguous signals. Online collectives have the option of doing such check-ins synchronously or asynchronously; either approach can allow participants an opportunity to share good news and/or updates, information that otherwise may not be surfaced but may help the group.

Tools Enabling Equality in Participation

1. Multi-channel Chat
   A major impediment to creativity and divergent thought in groups is what is known as "production blocking" or the decline in the number and originality of

ideas produced in interacting groups compared to the same number of individuals working alone (Thompson 2003). Production blocking stems from two different sources: first, as a result of the bottleneck that occurs in group conversation when people are hindered in sharing their ideas due to limitations on "air time" when only one person can speak at a time; and, second, when each group members' ideas are influenced or inhibited as a result of hearing others' ideas. A multi-channel chat room set up to facilitate private, public chats between members can allow equality in contributions to a collective product as well as avoid production blocking, as members can generate ideas independently in an initial step and then share them subsequently. Such a tool can facilitate the Delphi technique, a variant of the nominal group technique (Delbecq 1974), an approach in which individual ideas or inputs are developed independently and then pooled to create a final product. The Delphi technique requires a facilitator or a leader trusted by the team to aggregate responses from each member individually, ensuring equal input and avoiding production blocking. Online, a multi-channel chat allows the facilitator to privately chat with and gain input from each member. The private one-to-one chat can then be aggregated and shared in the public room, consistent with the recommendations of the Delphi technique, thus helping to foster equality of member inputs to collective products.

A multi-channel chat room set up to facilitate different tasks (each 'chat room' is a 'task') allows members to work collaboratively or asynchronously as best suits the work at hand. By allowing members to divide the tasks amongst themselves and create subgroups focused on different areas of work, collectives can avoid the bottlenecks described above and integrate a broader array of inputs into the work of the collective.

2. Shared Online Documents and Social Media

   Shared online documents serve to create organizational memory as well as ensure equality of contributions to shared products. Like multi-channel chat, shared online documents enable members to be working simultaneously and to capture ideas and inputs as they occur, rather than having to wait for a turn to speak or to work on a more traditional document sequentially.

   As described above, the use of social networking media, such as facebook, google+ and twitter, in online collectives allows the members to 'see' other members and assist in developing social perceptiveness in the team. Members are able to express themselves, their thoughts, feelings as well as feedback. However, such sites can also serve as an online resource to create organizational memory. For example, using a facebook page where ideas can be documented in full view in addition to feedback to those ideas would prevent ideas being repeated and/or forgetting ideas all together. This eliminates waste, repetition and encourages new ideas to surface. Members have greater opportunity to contribute to and build those ideas, without being hampered or blocked while waiting for others to express themselves.

3. Electronic Voting Systems

   In conventional face-to-face teams, sometimes facilitators are used who are able to keep teams on track and help make resolutions by aggregating feedback from each individual. Online, one such tool that could be used for electronic facilitation is an electronic voting system. An electronic voting system allows each team member to provide feedback and determine outcomes based on the feedback of the majority of team members. This raises the probability that the 'best answer' can percolate to the top (Hertel et al. 2005). Electronic voting systems can also ensure equal participation amongst team members, and enable members to share views anonymously if they prefer to do so. This ensures no one particular team member can dominate or make decisions for the whole team.

4. Real-Time Feedback System
   Equality in conversational turn-taking can occur organically when smart teams engage all the team members or it can be facilitated using different tools. A tool to bring about awareness amongst the team members on their participation levels is the use of real-time feedback systems. These systems keep track of contributions, inputs, and level of communication from each member and display them in full view. As such, if any one particular team member begins to dominate the conversation, it becomes apparent in real-time. Conversely, if any team member is not contributing enough, that too can be flagged. While assisting equality in participation, this tool also prevents social loafing, or the tendency of individuals to put forth less effort when working in groups, another major threat to group creativity and productivity (Karau and Williams 1993).

Furthermore, a real-time feedback system allows members to gauge each other's status and thus to assist in the development of social perceptiveness as well. Members with higher social perceptiveness can focus on the members on either extreme in terms of participation and help balance the team member's contributions.

Another type of real-time feedback tool is an electronic chart that displays ranks of users by quantity of contributions. This is similar to the real-time feedback systems discussed above, but at the team level instead of individual level. This allows teams to see how they rank, in terms of their accomplishments, against other teams. Hence, teams have benchmarks to surpass or meet. This prevents underperformance and raises motivation, enabling teams to be more creative and productive (Paulus et al. 2013).

Conclusion

Collaboration with others via computer-mediation is becoming an everyday reality for more and more of us, and offers the possibility of increasing collective intelligence beyond what is possible for traditional face-to-face teams. Initial research on the performance of online teams was mixed, but more recent research on collective intelligence suggests that the work of such groups is fostered by the same qualities that foster the work of traditional face-to-face teams—namely, social perceptiveness and equality of participation. These group attributes enhance the quality of both convergent and divergent thought, and can be facilitated by a range of established as well as newer tools in online settings.

Convergent thought is critical to generating consensus and enhancing decision making, and is most directly fostered by social perceptiveness. Social perceptiveness can, in turn, be enhanced through tools that amplify what might otherwise be subtle signals, and provide group members more contextual information about others. The ability to signal status, to communicate about mood and other events that might impact members' contributions to the group can help remote members sense what might otherwise be subtle cues, and adjust to these external influences. Beyond these explicit signals, integrating automatic or implicit signals to highlight areas of greater uncertainty or disagreement (such as text that is repeatedly revised, or data that has not yet been reviewed) is an example of how human-computer environments can be designed to further enhance collaborative work.

Divergent thought is essential to creativity and innovation, and requires the opposite conditions to those necessary for convergent thought to flourish in a group. Divergence is enhanced through periods of relatively isolated brainstorming and idea generation. It is directly fostered by the same tools that also enhance greater equality of participation among group members. Many readily available online tools serve to enhance this area of team collaboration, as online environments are well-designed for asynchronous work. Newer tools, such as real-time feedback about relative member contributions to group collaboration and collective products, can help groups preserve equality of contributions and divergent thought when collaborating in real time.

Newer areas of research in this area are inspired in part by work on stigmergy; that is, the ways cooperative animals coordinate by leaving and sensing signs in a shared environment. We know that humans sense all kinds of subtle, nonverbal cues from one another when collaborating, and that the ability to sense these cues also translates to online environments to facilitate collective intelligence in online groups. It is exciting to contemplate the ways in which computer-based tools and interaction platforms can compensate for the deficits that human teams frequently experience in sensing and interpreting such signals, to raise the level of collective intelligence even beyond what is normally observed in high performing, face-to-face groups.

The challenge for those of us who want to encourage the success of human-computer collectives is to understand how these tools can be honed to better facilitate the fundamental processes for collective intelligence. Doing so can enable an even broader level of participation, a trend that stands to benefit us all.

References

Andriopoulos C, Lewis MW (2009) Exploitation-exploration tensions and organizational ambidexterity: managing paradoxes of innovation. Organ Sci 20(4):696–717. doi:10.1287/orsc.1080.0406

Baron-Cohen S, Wheelwright S, Hill J, Raste Y, Plumb I (2001) The "reading the mind in the eyes" test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. J Child Psychol Psychiatry Allied Discip 42(02):241–251

Bloom H (2000) Global brain: the evolution of mass mind from the big bang to the 21st century. Wiley, New York

Bonabeau E, Dorigo M, Theraulaz G (1999) Swarm intelligence: from natural to artificial systems. Oxford University Press, New York

Brown VR, Paulus PB (2002) Making group brainstorming more effective: recommendations from an associative memory perspective. Curr Dir Psychol Sci 11(6):208

Burke K, Aytes K (1998) A longitudinal analysis of the effects of media richness on cohesion development and process satisfaction in computer-supported workgroups. In: System sciences, proceedings of the 31st Hawaii international conference on, vol 1. pp 135–144. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=653093

Burke K, Chidambaram L (1996) Do mediated contexts differ in information richness? A comparison of collocated and dispersed meetings. In: System sciences, proceedings of the 29th Hawaii international conference on, vol 3. pp 92–101. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=493180

Carmel E, Agarwal R (2001) Tactical approaches for alleviating distance in global software development. Softw IEEE 18(2):22–29

Chabris CF (2007) Cognitive and neurobiological mechanisms of the law of general intelligence. In: Roberts MJ (ed) Psychology Press, Hove, pp 449–491

De Dreu CK, West MA (2001) Minority dissent and team innovation: the importance of participation in decision making. J Appl Psychol 86(6):1191–1201

Deary IJ (2000) Looking down on human intelligence: from psychometrics to the brain. Oxford University Press, New York

Delbecq AL (1974) The effectiveness of nominal, Delphi, and interacting group decision making processes. Acad Manage J 17(4):605–621

Elfenbein HA (2006) Team Emotional Intelligence: What it can mean and how it can impact performance. In V. Druskat, F. Sala, & G. Mount (Eds.), The link between emotional intelligence and effective performance (pp. 165–184). Mahwah, NJ: Lawrence Erlbaum

Elfenbein HA, Polzer JT, Ambady N (2007) Team emotion recognition accuracy and team performance. Res Emot Organ 3:87–119

Fiore SM (2008) Interdisciplinarity as teamwork how the science of teams can inform team science. Small Group Res 39(3):251–277

Galegher J, Kraut RE (1994) Computer-mediated communication for intellectual teamwork: an experiment in group writing. Inf Syst Res 5(2):110–138

Goodhue DL, Thompson RL (1995) Task-technology fit and individual performance. MIS Q. 19, 2 (June 1995), 213–236

Grassé P-P (1959) La reconstruction du nid et les coordinations interindividuelles chezBellicositermes natalensis etCubitermes sp. la théorie de la stigmergie: Essai d'interprétation du comportement des termites constructeurs. Insectes Sociaux 6(1):41–80

Hertel G, Geister S, Konradt U (2005) Managing virtual teams: a review of current empirical research. Hum Resour Manage Rev 15:69–95

Holland O, Melhuish C (1999) Stigmergy, self-organization, and sorting in collective robotics. Artif Life 5(2):173–202

Ilgen DR, Hollenbeck JR, Johnson M, Jundt D (2005) Teams in organizations: from input-process-output models to IMOI models. Annu Rev Psychol 56(1):517–543

Karau SJ, Williams KD (1993) Social loafing: a meta-analytic review and theoretical integration. J Personal Soc Psychol 65:681–681

Karsai I, Balazsi G (2002) Organization of work via a natural substance: regulation of nest construction in social wasps. J Theor Biol 218(4):549–565

Katz R (1984) As research teams grow older. Res Manage 7(1):29–34

Kayworth TR, Leidner DE (2001) Leadership effectiveness in global virtual teams. J Manage Inf Syst 18(3):7–40

Kilduff M, Angelmar R, Mehra A (2000) Top management-team diversity and firm performance: examining the role of cognitions. Organ Sci 11(1):21

Larson JR (2009) In search of synergy in small group performance. Psychology Press, New York

Majchrzak A, Rice RE, Malhotra A, King N, Ba S (2000) Technology adaptation: the case of a computer-supported inter-organizational virtual team. MIS Q 24(4):569–600

Malhotra A, Majchrzak A, Rosen B (2007) Leading virtual teams. Acad Manage Perspect 21(1):60–70

March JG (1991) Exploration and exploitation in organizational learning. Organ Sci 2(1):71–87

Marsh L, Onof C (2008) Stigmergic epistemology, stigmergic cognition. Cogn Syst Res 9(1):136–149

Martins LL, Gilson LL, Maynard MT (2004) Virtual teams: what do we know and where do we go from here? J Manage 30(6):805–836

McDonough EF, Kahnb KB, Barczaka G (2001) An investigation of the use of global, virtual, and colocated new product development teams. J Prod Innov Manage 18(2):110–120

McGrath JE (1984) Groups: interaction and performance. Prentice-Hall, Englewood Cliffs

O'Reilly CA, Tushman ML (2008) Ambidexterity as a dynamic capability: resolving the innovator's dilemma. Res Organ Behav 28(1):185–206

Parunak HVD (2006) A survey of environments and mechanisms for human-human stigmergy. In: Environments for multi-agent systems II. Springer, pp 163–186. Retrieved from http://link.springer.com/chapter/10.1007/11678809_10

Paulus PB, Kohn NW, Arditti LE, Korde RM (2013) Understanding the group size effect in electronic brainstorming. Small Group Res 44(3):332–352. doi:10.1177/1046496413479674

Purvanova RK, Bono JE (2009) Transformational leadership in context: face-to-face and virtual teams. Leadersh Q 20(3):343–357

Raisch S, Birkinshaw J, Probst G, Tushman ML (2009) Organizational ambidexterity: balancing exploitation and exploration for sustained performance. Organ Sci 20(4):685–695. doi:10.1287/orsc.1090.0428

Riemer K, Richter A (2010) Tweet inside: microblogging in a corporate context. In: Proceedings of the 23rd Bled eConference, pp 1–17, Bled, Slovenia

Riemer K, Richter A, Bohringer M (2010) Enterprise microblogging. Bus Inf Syst Eng 2(6):391–394

Rosenkopf L, McGrath P (2011) Advancing the conceptualization and operationalization of novelty in organizational research. Organization Science 22(5):1297–1311

Sharda R, Barr SH, McDonnell JC (1988) Decision support system effectiveness: a review and an empirical test. Manage Sci 34(2):139–159

Siegel J, Dubrovsky V, Kiesler S, McGuire TW (1986) Group processes in computer-mediated communication. Organ Behav Hum Decis Process 37(2):157–187. doi:10.1016/0749-5978(86)90050-6

Spearman C (1904) General intelligence, objectively determined and measured. Am J Psychol 15(2):201–293

Steiner I (1972) Group process and productivity. Academic, New York

The Economist Intelligence Unit (2009) Managing virtual teams: taking a more strategic approach. http://graphics.eiu.com/upload/eb/NEC_Managing_virtual_teams_WEB.pdf on October 24, 2013

Thompson L (2003) Improving the creativity of organizational work groups. Acad Manage Exec 17(1):96–109

Tziner A, Eden D (1985) Effects of crew composition on crew performance: does the whole equal the sum of its parts? J Appl Psychol 70(1):85–93

Warkentin M, Sayeed L, Hightower R (1997) Virtual teams versus face-to-face teams: an exploratory study of a web-based conference system. Decis Sci 28(4):975–996

Wiersema MF (1992) Top management team demography and corporate strategic change. Acad Manage J 35(1):91–121

Wolff SB, Pescosolido AT, Druskat VU (2002) Emotional intelligence as the basis of leadership emergence in self-managing teams. Leadersh Q 13(5):505–522

Woolley AW, Fuchs E (2011) Collective intelligence in the organization of science. Organ Sci 22(5):1359–1367

Woolley AW, Chabris CF, Pentland A, Hashmi N, Malone TW (2010) Evidence for a collective intelligence factor in the performance of human groups. Science 330:686–688

Zhang J, Qu Y, Cody J, Wu Y (2010) A case study of micro-blogging in the enterprise: use, value, and related issues. In: Proceedings of the 28th international conference on human factors in computing systems, pp 123–132. Retrieved from http://dl.acm.org/citation.cfm?id=1753346

Human Computation and Collaboration: Identifying Unique Social Processes in Virtual Contexts

Alecia M. Santuzzi, Christopher J. Budnick, and Derrick L. Cogburn

Introduction

The rapid expansions of computer-mediated communication (CMC) and information-communication technologies (ICTs) have encouraged the use of geographically distributed work-teams and other distance collaborations among human participants (Gressgard 2011). These technology-mediated collaborations may be short-term or long-term, and may vary in the extent to which synchronous work is required. For example, an investment company might develop a long-term department of financial advisors who rely primarily on technology to communicate with each other and complete work activities with clients. As a contrasting example, an emergency project team might be established and meet only once in a web-based conference to quickly solve a problem.

Paralleling the emergence of technology-mediated collaborations, social, communication, and information science researchers continue to investigate how effectively human computation tools can facilitate participation in these activities. However, conclusions from technology-mediated collaboration studies show very little consistency regarding the variables and processes that are important to successful collaborations. As noted by Woolley and Hashmi (2013, this volume), the mixed conclusions could be driven by overlooked differences in the types of tasks involved and/or the technology available to complete those tasks.

Adding to their analysis, we take the broader stance that part of the confusion in the literature may be driven by the fact that the technology-mediated aspect of collaboration in human computation systems is often overlooked or underemphasized. Human computation comprises information-processing strategies that are required

A.M. Santuzzi (✉) • C.J. Budnick
Northern Illinois University, DeKalb, Illinois, USA
e-mail: asantuzzi@niu.edu

D.L. Cogburn
American University, Washington, D.C., USA

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_57,

of all participants who are involved in a collaboration. Importantly, those information-processing strategies likely differ in face-to-face collaborations relative to technology-mediated collaborations and, thus, might yield different social processes and task completion strategies within those collaborations.

We begin with a brief examination of the variability in the definitions and measures of social process variables in the current literature on collaborations. Using the roles of trust and leadership as key examples, we consider the differences in research priorities across disciplines as a potential source of inconsistent conclusions. Important factors that contribute to the effectiveness of technology-mediated collaborations might be omitted from continued research that is limited to the perspective of one discipline. Finally, we recommend that researchers and practitioners take a step back and build new interdisciplinary theoretical models specifically designed for technology-mediated communication and collaboration with a focus on human computation, rather than expecting theoretical models informed only by their home disciplines to translate directly to technology-mediated collaboration.

Technology-Mediated Collaboration: An Interdisciplinary Topic

Currently, social, communication, and information scientists seem to agree that leadership roles and social processes are important to consider when studying and implementing technology-mediated collaborations. A content analysis of the literature across disciplines found that leadership and trust were the two most cited concepts in empirical and non-empirical work on virtual teams and organizations (Cogburn et al. 2011).

The next logical step seems to be considering the roles of leadership, trust, and other previously cited variables in the confidence in and effectiveness of technology-mediated collaborations. In order to establish conclusions about those relationships, empirical evidence must exist, be replicated across several studies, and demonstrate consensus regarding conclusions. As noted by the authors in their own work as well as authors of several attempted meta-analyses (Baltes et al. 2002; Ortiz de Guinea et al. 2005), a disproportionately low number of existing works on technology-mediated collaboration are empirical. From the 1,186 works included in their content analysis, Cogburn et al. (2011) identified only 183 that were empirical (88 % quantitative; 12 % qualitative). The three top contributors of the published empirical research were the social (40 %), information (23 %), and management (12 %) sciences. Perhaps even more concerning is the large amount of variability in the relationships among variables, and their qualifying conditions, among the empirical studies that do exist. Various factors qualified differences between virtual and face-to-face interactions in both meta-analyses, such as the collaborating partners' anonymity and the level (group or individual) at which the effects occurred. Notably, both meta-analyses also found inconsistencies in defining and measuring concepts across disciplines. To demonstrate those inconsistencies within and across disciplines, we focus on the two most cited concepts in the literature: leadership and trust.

Technology-Mediated Leadership

The leadership concept usually refers to an individual (or group) charged with supporting, motivating, and encouraging work-related outcomes. Typically, social scientists classify leaders by style (e.g., transactional, transformational, or authentic; Avolio et al. 2009). Additionally, leaders might be task, relationship, or change-oriented (Yukl et al. 2002). By either classification, the focus of leadership traditionally has been on the proximal task and the individuals completing it.

Some research might suggest that technology-mediated leadership is similar to traditional, face-to-face leadership. For instance, the consensus seems to be that both face-to-face and technology-mediated leadership involves facilitating effective team functioning, managing the socio-emotional environment, ensuring resource availability, instilling a sense of team identity, and directing personnel (Avolio et al. 2009; Bell and Kozlowski 2007; Hoch and Kozlowski 2012; Zaccaro and Bader 2002). Each of those functions is essential to successful team performance. Leaders neglecting to fulfill any of the functions likely inhibit progress towards achieving the team objective. However, the manner in which technology-mediated leaders address those functions likely differs substantially from face-to-face teams.

As such, it seems that leadership within a technology-mediated collaboration might present unique roles and responsibilities. Often these leaders must transform a collection of geographically distributed individuals into a coherent, cohesive, and integrated work-team (Bell and Kozlowski 2007). Additionally, leaders are responsible for facilitating transitions within the technology-mediated environment whenever adding, removing, or reassigning team members. Consequently, Malhotra et al. (2007) suggested that technology-mediated leaders might need to prioritize the work environment and focus on creating an atmosphere conducive to technology-mediated interactions. Confirming this, technical researchers have observed that ICT malfunctions, differences in available hardware or technology-mediated team members' abilities, and mismatches between software and/or hardware components often challenge the effectiveness of technology-mediated leaders (Bjorn and Ngwenyama 2010). The technology-mediated leader must consider and perhaps prioritize the technological environment and how it affects human computation among team members to be effective.

Moreover, research has shown that transplanting leadership models from traditional face-to-face interactions directly into technology-mediated collaborations might not work well. Hoch and Kozlowski (2012) reported worse performance by technology-mediated teams under hierarchical leadership compared to those under distributed leadership. The traditional hierarchical leadership structure of one distinct leader with many subordinates might be ineffective in technology-mediated collaborations. Recent work demonstrates that technology-mediated teams actually show a flat, transitory leadership structure with leader roles distributed among team members compared to face-to-face collaborations (Hoch and Kozlowski 2012; Muethel et al. 2012).

Taken together, the existing research suggests that technology-mediated leadership has some qualitative differences from traditional face-to-face leadership definitions. Technology-mediated collaborations may be more effective when leadership closely monitors the technological environment and allows for flat, multi-leader structures. Research and practices excluding the oversight of technological environments and flexible leadership structures from their conceptual definition of leadership likely will show findings incongruent with results based upon definitions prioritizing the technological environment and allowing for flexible leadership roles. As elaborated subsequently, differences in leadership definitions may be due to different priorities and theories across disciplines.

Technology-Mediated Trust

The definition and function of trust is agreed upon even less in the technology-mediated collaboration literature. Many conceptual definitions have been imported from social and management research on face-to-face collaborations. Likewise, past technology-mediated collaboration research seems to have assumed that traditional variables, such as trust, would be defined similarly and play a similar role as observed in face-to-face interactions. Accordingly, past research has reported that increased trust between technology-mediated team members positively affects team performance (Brahm and Kunze 2012; Chang et al. 2011; Sarker et al. 2011) and is likely the central element binding virtual team members together (Altschuller and Benbunan-Fich 2010; Clark et al. 2010; Sarker et al. 2003). Moreover, practitioners seem to accept the notion that trust building among technology-mediated team members is an important function of an effective technology-mediated leader (Malhotra et al. 2007).

However, little agreement exists on the precise definition of trust within technology-mediated collaborations. In fact, definitions vary greatly demonstrating different focal aspects. For example, Sarker et al. (2003) relied on existing definitions and briefly defined trust in technology-mediated environments as “the degree of reliance individuals have on their remotely located team-members taken collectively” (p. 37). Yet, Clark et al. (2010) asserted that trust operates on three levels: ability, benevolence, and integrity. Specifically they conceptualized trust as a continuum comprised of those three dimensions. Others have conceptualized trust as a belief in others’ intentions and behaviors, an acceptance of dependence on another, feelings of security, a shared history, or as resulting from reliable and competent performance (Brahm and Kunze 2012; Casey 2010; Clark et al. 2010; Clases et al. 2003; Greenberg et al. 2007). These definitions define trust by affective, cognitive, or interpersonal states. One strong theme among those conceptualizations is that trust represents some form of interpersonal bond among collaborators.

Some research has suggested that such a bond requires time to develop in technology-mediated environments. Researchers suggest that trust is initially low but evolves over time in technology-mediated collaborations. Such teams may reach similar trust levels as face-to-face teams; however, this process typically occurs much more slowly for technology-mediated collaborations (Crossman and Lee-Kelley 2004; Altschuller and Benbunan-Fich 2010; Walther 1996). Given enough time, trust levels in technology-mediated teams might even rise above those of traditional teams (Walther 1996).

The course and importance of trust building in a collaboration, however, depends greatly on the conceptual definition of trust. Unlike many definitions used by social scientists, Jarvenpaa et al. (1998) definition of trust places less emphasis on temporal investment in intimacy. Instead, they use the term *swift trust*. Swift trust occurs when collaborators who are unfamiliar with each other assume that trust is present from the very beginning of the collaboration. Technology-mediated teams can thus build trust through early and frequent social interactions (Jarvenpaa and Leidner 1999). When defined in that way, trust has been shown to improve collaboration effectiveness (Jarvenpaa et al. 1998). However, this version of trust might reflect collaborators' enthusiasm and commitment to complete the proximal task rather than providing an index of any degree of intimacy. Furthermore, swift trust usually is applied to ad hoc teams gathered to maximize human capital in order to achieve a short-term objective (Robert et al. 2009). As such, swift trust might be less applicable to more long-term technology-mediated collaborations.

Adding to the inconsistencies, some research has suggested that trust might not play as large a role in technology-mediated work and collaboration as compared to its importance in face-to-face interactions (Ortiz de Guinea et al. 2005). Gonzalez et al. (2003) found that only task-based perceptions of collaborators affected performance quality, whereas social relationship perceptions did not. Therefore, if concepts such as trust are examined using a general social definition, trust would appear to be unimportant to technology-mediated collaboration. If defined as task-specific, trust would be considered very important. This highlights the critical problem of using the same term (trust) to represent conceptually different social processes.

Uniting Disciplines to Identify Unique Social Processes

As technology-mediated collaborations are interesting to many disciplines of research and practice, it is not surprising that multiple areas of study have considered the topic. Research and other literature on technology-mediated collaborations have appeared in psychology, management, communications, computer science, engineering, education, and information science. We suggest that one primary reason for inconsistencies in the understanding of technology-mediated collaborations is that researchers and practitioners continue to rely on discipline-specific theoretical models rather than considering the possibility that technology-mediated collaborations are qualitatively distinct from those models' original purposes. Relying on past models can be problematic for at least two reasons: (1) variables in the original models might have different meanings within technology-mediated collaborations, and (2) research driven by past models might miss important factors that are uniquely relevant to technology-mediated collaborations.

For example, computer and information science researchers predominantly focus on infrastructure and the usability of technological tools as applied to social interactions. Social science and management researchers tend to focus on social interactions that occur within technology-mediated environments. At first, the difference seems very subtle. Each discipline seems to assume one part of the equation is constant while examining their variables of interest. For example, information scientists might assume that theories about communication effectiveness such as Media Richness Theory (Daft et al. 1987) could simply be translated to social collaborations that rely on technology media. Similarly, management researchers might assume that theories of and research on traditional social dynamics would apply to technology-mediated collaborations simply because these situations involve some form of social dynamics.

Admittedly, any discipline relying on past theoretical models benefits from a theory-driven starting point for empirical examination. However, different theories from different disciplines highlight different priorities. This reality creates a few challenges to the understanding of an interdisciplinary topic such as technology-mediated collaboration.

First of all, researchers relying on past, discipline-specific theoretical models to inform research hypotheses run a high risk of excluding important variables already confirmed by other disciplines. For example, social science research on technology-mediated interactions rarely incorporates precisely defined variables reflecting the adoption or adaptation of technology tools. Yet, *technology adaptation* is a formally defined and well-studied variable in the information science literature. It refers to the work-team processes involved in changing the way one or more CMC/ICTs are engaged to facilitate task completion (DeSanctis and Poole 1994; Thomas and Bostrom 2010). Furthermore, effective technology adaptation has been shown to predict successful team interactions (Marjchrzak et al. 2000; Malhotra et al. 2001). On the other hand, the importance of social variables such as norms and norm violations seem to be exclusive to the social and communication sciences. For example, *chronemics*—the timing of electronic messages—has been shown to have important implications for technology-mediated communications (Walther 2002). Violating the normative chronemics during a technology-mediated collaboration can lead to negative interpersonal consequences in work teams.

Secondly, enough inconsistency appears in the current technology-mediated collaboration literature for researchers to wonder whether their theoretical models, developed for one situation (e.g., human-computer interaction), are applicable to other forms of technology-mediated collaborations. For example, the inconsistencies in the social science literature should invite social scientists to consider whether technology-mediated collaboration is qualitatively different from the traditional face-to-face collaborations that provide the basis for current theoretical models. In this new technology-mediated environment, familiar factors such as leadership and trust seem to influence team members' behaviors differently than in face-to-face environments. Therefore, a discipline-specific focus might leave important variables undiscovered. Accounting for such variables likely will provide a more complete understanding of the processes inherent to virtual communications and collaborations.

Recommendations for Bridging Disciplines

In consideration of the evidence and logic above, we recommend that researchers and practitioners step back and explore the possibility that unique information and social processes might occur in technology-mediated collaborations. Traditional collaboration theories or theories about technology and human-computer interaction might not fully account for such processes. Even the information processing strategies among individuals during collaborations might differ between face-to-face and technology-mediated environments and among the types of technologies used to facilitate collaborations. Research could take a more inductive approach to identify these qualitative differences and highlight potentially new and more important variables than what past theoretical models imply.

Rather than starting from scratch, however, we recommend that researchers across disciplines begin by consolidating their work and engaging in a systematic search to identify points of disagreement. Rather than focusing only on conclusions in the literature, we strongly recommend that researchers carefully examine the definitions and operationalizations of social process and outcome variables. Examining factors and processes across disciplines, resolving conflicts, and arriving at agreement concerning how those factors and processes are defined should provide not only a better understanding of those processes, but also a more efficient avenue for future research.

Conclusion

Technology-mediated collaborations tend to be cost-effective and allow for a diverse representation of perspectives and demographics as they defy geographic boundaries and barriers. These positive qualities suggest that technology-mediated collaborations are not simply a passing trend. However, there has been little empirical testing conducted to date. Moreover, the limited empirical work that does exist suffers from inconsistencies in the conceptual definitions of key processes and their measurement. We fully support continued, rigorous, and theory-driven empirical study of technology-mediated collaborations. Yet, we wish to encourage researchers and practitioners to be sensitive to the possibility that: (1) the conceptual meaning of variables studied in their home discipline might not match the definitions used in another domain, and (2) other disciplines may have identified critically important factors involved in technology-mediated collaboration that traditional theories in the home discipline do not highlight.

We also hope to inspire researchers and practitioners to remain open to inductively identifying novel concepts and theories unique to technology-mediated collaborations. Merely applying past theoretical models designed for other activities might exclude some unique and perhaps critical variables that would guide a better understanding of how to facilitate effective technology-mediated collaborations.

For instance, the uniqueness of technology-mediated collaborations within human computation systems seems to be underemphasized in the existing empirical literature. The information processing strategies of participants in a technology-mediated collaboration likely differ from those used in a face-to-face collaboration. It is possible that machine-based contributions of human computation systems provide technology-mediated collaborations with some flexibility to adapt to the technology-mediated environment and establish workflows that would not occur in traditional face-to-face environments. An empirical focus on human computation seems to be a fruitful avenue of continued research on technology-mediated collaborations. Based on the diverse representation of disciplines in this edited volume, a focus on human computation also may serve as a theme that unites the various interested disciplines.

Acknowledgements Funding for this project was provided by a grant from the National Science Foundation (VOSS-0838492). Findings, conclusions, or recommendations are the author(s) and not those of the National Science Foundation.

References

Altschuller S, Benbunan-Fich R (2010) Trust, performance, and the communication process in ad hoc decision-making virtual teams. J Comput-Mediat Commun 16:27–47. doi:10.1111/j.1083-6101.2010.01529.x

Avolio B, Walumbwa F, Weber T (2009) Leadership: current theories, research, and future directions. Annu Rev Psychol 60:421–449. doi:10.1146/annurev.psych.60.110707.163621

Baltes B, Dickson M, Sherman M, Bauer C, LaGanke J (2002) Computer-mediated communication and group decision making: a meta-analysis. Organ Behav Hum Decis Process 87:156–179. doi:10.1006/obhd.2001.2961

Bell B, Kozlowski S (2007) A typology of virtual teams: implications for effective leadership. Gr Organ Manage 27:14–49. doi:10.1177/1059601102027001003

Bjorn P, Ngwenyama O (2010) Technology alignment: a new area in virtual team research. IEEE Trans Prof Commun 53:382–400. doi:10.1109/TPC.2009.2034926

Brahm T, Kunze F (2012) The role of trust climate in virtual teams. J Manage Psychol 27:595–614. doi:10.1108/02683941211252446

Casey V (2010) Developing trust in virtual software development teams. J Theor Appl Electron Commer Res 5:41–58. doi:10.4067/S0718-18762010000200004

Chang H, Chuang S, Chao S (2011) Determinants of cultural adaptation, communication quality, and trust in virtual teams' performance. Total Qual Manage 22:305–329. doi:10.1080/14783363.2010.532319

Clark W, Clark L, Crossley K (2010) Developing multidimensional trust without touch in virtual teams. Mark Manage J 20:177–193

Clases C, Bachmann R, Wehner T (2003) Studying trust in virtual organizations. Int Stud Manage Organ 33:7–27. Retrieved from http://www.jstor.org/stable/40397569

Cogburn D, Santuzzi A, Espinoza, F (2011) Developing and validating a socio-technical model for geographically distributed collaboration in global virtual teams. IEEE Proceedings of the Hawaii International Conference on System Sciences (HICSS), USA, pp 1–10

Crossman A, Lee-Kelley L (2004) Trust, commitment and team working: the paradox of virtual organizations. Glob Netw 4:375–390. doi:10.1111/j.1471-0374.2004.00099.x

Daft R, Lengel R, Trevino L (1987) Message equivocality, media selection, and manager performance: implications for information systems. MIS Q 11:355–366. Retrieved from http://www.jstor.org/stable/248682

DeSanctis G, Poole M (1994) Capturing the complexity in advanced technology use: adaptive structuration theory. Organ Sci 5:121–147. Retrieved from http://www.jstor.org/stable/2635011

Gonzalez MG, Burke MJ, Santuzzi AM, Bradley J (2003) The impact of group process variables on the effectiveness of distance collaboration groups. Comput Hum Behav 19:629–648

Greenberg P, Greenberg R, Antonucci Y (2007) Creating and sustaining trust in virtual teams. Bus Horiz 50:325–333. doi:10.1016/j.bushor.2007.02.005

Gressgard L (2011) Virtual team collaboration and innovation in organizations. Team Perform Manage 17:102–119. doi:10.1108/13527591111114738

Hoch J, Kozlowski S (2012) Leading virtual teams: hierarchical leadership, structural supports, and shared team leadership. J Appl Psychol, Advance online publication. doi:10.1037/a0030264

Jarvenpaa S, Leidner D (1999) Communication and trust in global virtual teams. Organ Sci 10:791–815. Retrieved from: http://links.jstor.org/sici?sici=1047-7039%28199911%2F12%2910%3A6%3C791%3ACATIGV%3E2.0.CO%3B2-6

Jarvenpaa S, Knoll K, Leidner D (1998) Is anybody out there? Antecedents of trust in global virtual teams. J Manage Inf Syst 14:29–64. Retrieved from http://www.jstor.org/stable/40398291

Malhotra A, Majchrzak A, Carman R, Lott V (2001) Radical innovation without collocation: a case study at Boeing-Rocketdyne. MIS Q 25:229–249

Malhotra A, Majchrzak A, Rosen B (2007) Leading virtual teams. Acad Manage Perspect 21:60–70. doi:10.5465/AMP.2007.24286164

Marjchrzak A, Rice R, Malhotra A, King N, Ba S (2000) Technology adaptation: the case of a computer-supported inter-organizational virtual team. MIS Q 24: 569–600. Retrieved from http://www.jstor.org/stable/3250948

Muethel M, Gehrlein S, Hoegl M (2012) Socio-demographic factors and shared leadership behaviors in dispersed teams: implications for human resource management. Hum Resour Manage 51:525–548. doi:10.1002/hrm.21488

Ortiz de Guinea A, Webster J, Staples S (2005) A meta-analysis of the virtual teams literature. Paper presented at the symposium on high performance professional teams. Industrial Relations Centre, Queen's University, Ontario

Robert L Jr, Dennis A, Hung Y (2009) Individual swift trust and knowledge-based trust in face-to-face and virtual team members. J Manage Inf Syst 26:241–279. doi:10.2753/MIS0742-1222260210

Sarker S, Valacich J, Sarker S (2003) Virtual team trust: instrument development and validation in an IS educational environment. Inf Resour Manage J 16:35–55

Sarker S, Ahuja M, Sarker S, Kirkeby S (2011) The role of communication and trust in global virtual teams: a social network perspective. J Manage Inf Syst 28:273–309. doi:10.2753/MIS0742-1222280109

Thomas D, Bostrom R (2010) Vital signs for virtual teams: an empirically developed trigger model for technology adaptation interventions. MIS Q 34: 115–142. Retrieved from http://www.jstor.org/stable/2635011

Walther J (1996) Computer-mediated communication: impersonal, interpersonal, and hyperpersonal interaction. Commun Res 23:3–43. doi:10.1177/009365096023001001

Walther JB (2002) Time effects in computer-mediated groups: past, present, and future. In: Hinds P, Kiesler S (eds) Distributed work. MIT Press, Cambridge, pp 235–257

Woolley A, Hashmi N (2013) Cultivating collective intelligence in online groups. In: Michelucci P (ed) Human computation. Springer, New York

Yukl G, Gordon A, Taber T (2002) A hierarchical taxonomy of leadership behavior: integrating a half century of behavior research. J Leadersh Organ Stud 9:15–32. doi:10.1037/t03624-000

Zaccaro S, Bader P (2002) E-leadership and the challenges of leading e-teams: minimizing the bad and maximizing the good. Organ Dyn 31:377–387. doi:10.1016/S0090-2616(02)00129-8

Game Theory and Incentives in Human Computation Systems

Arpita Ghosh

Introduction

The Web is increasingly centered around contributions by its users: human computation is growing increasingly common as a means for accomplishing a wide range of tasks, ranging from labeling and categorization of images and other content (with workers recruited on paid crowdsourcing platforms like Amazon Mechanical Turk, or in systems based on unpaid contribution such as Games with a Purpose or Citizen Science projects like GalaxyZoo), to answering questions on online Q&A forums (such as Y! Answers, Quora, or StackOverflow, to name a few), all the way to peer-grading homework assignments in online education. But while some human computation systems consistently attract high-quality contributions, other seemingly similar ones suffer from junk or low-quality contributions, and yet others fail due to too little participation. How can we design *incentives* in these systems to elicit desirable behavior from potential participants?

There are two components to the problem of incentive design for human computation: (i) Identifying the costs and benefits of potential contributors to the system (the components that help formulate a *model* of agent behavior), and (ii) deciding how to assign rewards, or benefits, as a function of contribution (analysis and design).

The first question of identifying costs and benefits relates closely to the question of *why* do people contribute—that is, what constitutes a benefit or a *reward*? The answer to this question, of course, varies depending on the particular system in question. While some systems (such as those based on the Amazon Mechanical Turk platform), offer financial incentives for participation, a vast majority of human computation is driven by social-psychological rewards from participation; such rewards include, for example, both intrinsic motivators like fun, interest, or the satisfaction of benefiting a cause,[1]

[1] Such as furthering science in a Citizen Science project

A. Ghosh(✉)
Cornell University
e-mail: arpitaghosh@cornell.edu

P. Michelucci (ed.), *Handbook of Human Computation*, 725
DOI 10.1007/978-1-4614-8806-4_58, © Springer Science+Business Media New York 2013

as well as extrinsic social rewards such as attention, reputation or status. There is now a growing literature in social psychology addressing what motivates, or constitutes a reward for, users in such systems.[2]

But even after answering the question of why people contribute, there is a second question, which relates to how rewards are *allocated*. Given that users value rewards (by definition, and irrespective of their specific nature—financial or social-psychological), and incur costs (of time and effort) associated with different actions in the system, how rewards are assigned will influence what actions users take. That is, when a system depends on self-interested agents with their own benefits[3] and costs to participation, the quality and quantity of contributions will depend on the incentives created by the reward allocation scheme being used by the system. Given the understanding from the social psychology literature on what constitutes a reward, how should the *allocation* of these rewards be designed to incentivize desirable outcomes?

The following example illustrates the point. Consider a system with a leaderboard for top contributors (say the users who have classified the most images in a Citizen Science project like GalaxyZoo, or earned the most points in a GWAP such as the ESP game); such leaderboards appear to be strong motivators for users. While there are a number of questions related to leaderboard design, consider a very basic, simplified, question—should the system display only the top contributor, or, say, the top 5 contributors? On the one hand, if only one top-contributor 'prize' is given out, it is conceivable that users will try harder to win that solitary prize, leading to higher effort than when there are five prizes, since the presence of a greater number of prizes could mean one need not do as much to win. On the other hand, one could also argue that users will be more likely to put in effort when they know there are five prizes to be had, since they have a greater chance of winning something, so that their efforts are less likely to 'go to waste', in the second case where there are more prizes. Which of these is actually the correct prediction of behavior, when all participants are facing the same question of how much effort to put in? Now suppose these prizes are not positions on a leaderboard, but rather monetary rewards that all come out of a fixed prize budget (for example, as in a crowdsourcing contest)—in this case, should the entire budget be spent on one large prize or five smaller prizes? Again, informal arguments could be made in favor of either solution; a formal game-theoretic analysis is necessary to understand how rewards should be structured to optimally incentivize effort from contributors.[4]

A formal game-theoretic approach to incentive design, very broadly, proceeds by constructing an appropriate model where users (agents) make choices over actions, which are typically associated with costs (note that the term cost does not only refer

---

[2] The motivations of contributors in human computation are, naturally, closely related to those for user-generated content; some of the literature on which is discussed in Jian and MacKie-Mason (2012).

[3] (Arising from a range of motivations including possibly other-regarding, or 'altruistic', preferences)

[4] This particular problem is addressed in a model stylized for online crowdsourcing (contests, as well as crowdsourced content as in Q&A forums), in Ghosh and McAfee (2012).

to financial costs such as an entry fee, but is also used to refer to non-monetary quantities such as the cost to time or effort). Action choices in human computation systems can consist, for example, of the following: (i) In most[5] systems, participation is a voluntary action choice (with an associated cost, e.g., of the time required to create an account or to log in to the system to participate), and mechanisms must be designed to induce adequate participation when entry is an endogenous, strategic, choice. (ii) In many systems, agents can make a choice about how much *effort* to expend on any given task, potentially influencing the quality of their output and therefore its value to the system—mechanisms must be designed so as to induce agents to expend a high level of effort (which is more 'costly' than lower effort). (iii) Finally, in some systems, agents may hold information that they can potentially strategically misreport to their benefit, such as in voting or rating—this leads to the problem of designing mechanisms that induce agents to truthfully reveal this information. (Naturally, any real system might contain a combination of these choices, as well as others unique to its function—an example of this latter kind is the choice of the order in which to output descriptive words for images in the ESP game; see section "GWAPs").

A given design for a human computation system corresponds to, or induces, some rules that specify the allocation of rewards or benefits given each set of possible actions by agents. Note that in general, an agent's reward can depend not only on her output, but also the outputs (determined by the action choices) of other agents. Given a particular system design and the corresponding rules it induces, strategic agents will choose actions that maximize their utility (difference between benefit and cost) from the system. Agents' choices of actions lead to outputs, which in turn define the benefit, or reward, that each agent receives from the system. A vector of action choices by agents, roughly speaking, constitutes an equilibrium if no agent can improve her payoff by choosing a different action.[6]

There are two aspects to a game-theoretic, or more generally, economic, approach to incentives: analysis, and design. Analyzing equilibrium behavior under the reward allocation rules of a *given* system leads to a prediction about the behavior of agents, and therefore what kind of outcomes one might expect from that system. Choosing (or altering) the rules according to which rewards are allocated to induce agent behavior that achieves some particular outcome, or family of outcomes, constitutes *design*. While a game-theoretic approach to the analysis and design of any system with strategic agents has the general structure described above, each setting or system comes with its own unique features, depending on the choices of available actions, the nature of the available rewards and differing constraints on how they can be allocated, and *observability* of agents' outputs. In the remainder of this chapter, we will illustrate applications of the game-theoretic approach outlined above to some specific human computation domains in section "Game-Theoretic Models for Human Computation Systems", and then discuss how the same kind of approach

[5] Albeit not all systems; peer-grading in online education being a prominent example

[6] A number of different *equilibrium* concepts exist to predict how strategic agents will behave under a given mechanism; see, for instance, Nisan et al. (2007).

can be applied to reward design in the context of gamification, and rewarding contributors for their overall site participation in section "Incentivizing Consistent Effort: Gamification and Game Theory". We conclude with a discussion of challenges and directions for further work in section "Challenges and Further Directions".

Game-Theoretic Models for Human Computation Systems

In this section, we will look at three instances of game-theoretic analysis and design for human computation systems to illustrate the game-theoretic approach outlined in the previous section. Of course, these are not the only examples of game-theoretic analysis in the context of human computation; we briefly mention two other domains of interest.

The DARPA red balloon challenge[7] was a highly publicized instance of human computation—in the sense of a distributed network of human sensors—that required incentivizing the rapid mobilization of a large number of participants on a social network. The challenge, run in December 2009, consisted of locating ten 8-foot high red balloons that had been moored at ten unknown locations throughout the US; the first team to correctly identify the locations of all ten balloons would receive a cash prize of $40,000. For a team to win the challenge, it was necessary not only to recruit members who would look for and report sightings of the balloons themselves, but also to incentivize recruits to further recruit team members, since increasing the number of searchers increased a team's chance of quickly locating the balloons. That is, in addition to the problem of incentivizing participation, a team also had to incentivize incentivizing further participation. The recursive incentive scheme used by the winning MIT team to split the prize money amongst its participants is described and analyzed in Pickard et al. (2011), and highlights some of the issues that arise in the context of incentives in human computation tasks on social networks where performance, albeit not available reward, scales with the number of participants.

Another interesting family of problems related to incentives in human computation (broadly defined) occurs in online knowledge sharing or question-answer forums, such as Y! Answers, StackOverflow, or Quora, where questions posed by users are answered by other users of the site. There is a growing literature addressing a range of questions related to incentives and strategic behavior on such online Q&A forums in a game-theoretic framework, including what reward structures elicit quicker answers from users (Jain et al. 2012), how to allocate attention rewards[8] amongst contributors (Ghosh and McAfee 2012), as well as regarding the implementability of outcomes (i.e., the number and qualities of answers received)

---

[7] http://archive.darpa.mil/networkchallenge/

[8] (By choosing which answers to display, and how often or prominently to display them)

by the 'best-answer' style mechanisms used by Q&A forums such as Y! Answers (Ghosh and Hummel 2012).

We now proceed with an analysis of incentives and strategic behavior in three human computation settings—we discuss Games with a Purpose in section "GWAPs", designing mechanisms for crowdsourced judgement aggregation in section "Crowdsourced Judgement Elicitation", and voting in the context of human computation in section "Aggregating Quality Estimates: Voting".

GWAPs

Games with a Purpose (GWAPs) (Ahn and Dabbish 2008) are an outstanding family of examples of successful human computation systems. GWAPs are games designed so that people who are ostensibly simply playing the game also simultaneously produce useful input to a computation or task which cannot be performed by computers alone. For example, the game Verbosity[9] matches two players, who both 'win' if the first player correctly guesses the word being described by the second player (who, of course, is forbidden from directly using the word). This gives the second player the incentive to produce good descriptions to successfully communicate the word, thereby generating word descriptions in the process. Another game TagATune[10] pairs two players, both of whom receive a sound clip and generate descriptions for their clips to decide whether they have the same clip or not—since players 'win' when they correctly determine whether or not they have the same clip, this creates incentives for both players to generate descriptive labels for their clips, leading to a useful set of labels for sound clips in the system.

The first and perhaps best known GWAP is the ESP game,[11] which cloaks the task of labeling images under the guise of a game. In the ESP game, two randomly paired players are given an image; both players are asked to generate single-word descriptions for that image. Players gain points when they agree with their partner on a descriptive word, or label, for the image (neither player can see her partner's choices until the two players have entered a common label). Since players do not know who their partner is because they are randomly paired by the game, they cannot coordinate on descriptions, and so the easiest way to agree on the output (i.e., a label for the image) is to base it on the input (i.e., the image itself). Thus the game design aligns the incentives of the players, who want to earn points, with that of the system, which is to generate descriptive labels for images.

But does it? The ESP game has been tremendously successful in terms of participation—it was played by over 200,000 people, collecting over 50 million tags (Ahn and Dabbish 2008) in approximately 4 years since its creation. This high participation makes it evident that the basic incentives were well-designed—fun

[9] http://www.gwap.com/gwap/gamesPreview/verbosity/
[10] http://www.gwap.com/gwap/gamesPreview/tagatune/
[11] http://www.gwap.com/gwap/gamesPreview/espgame/

was clearly a valid reward, and the game clearly generated adequate ‘fun’ reward to compensate for the effort involved in playing the game. But what about the *quality* of the labels generated? It has been observed, both anecdotally and in a more careful study by Weber et al. (2008), that the labels obtained for images in the ESP game tend to have a high percentage of colors, synonyms, or generic words—essentially, labels that do not necessarily contribute too much information about the image, and are perhaps not the most useful labels that could be collected by the system. As we see next, a game-theoretic model and analysis of the ESP game can help explain how the specific choices made for the rules of the game encourage the creation of such tags, and also suggests changes to the game design which might address this issue.

Consider a simple model (Jain and Parkes 2013) for the ESP game. Each player independently chooses one of two effort levels (low or high) to exert while playing the game. A player who chooses low effort samples labels from the most ‘frequent’, or common, set of words in the universe (such as colors, or generic common nouns), whereas a player choosing high effort samples labels from the entire universe of words; assume that players know the relative frequencies of each word they have sampled. Next, a player can choose in what *order* to output her sampled words (which are the labels she thinks of for the image). How do the rules of the ESP game affect what effort levels players choose, and the order in which they output words?

The ESP game design rewards players as follows. Each pair of players are matched for a set of 15 images, and try to label as many images as they can achieve agreement on in 2.5 min. For each image, both players enter a sequence of single-word descriptions and can move on to the next image as soon as they enter a common descriptive word, which then becomes the label for the image. Players receive points for each such successful labeling. Since players can see more images (thereby potentially earning more points, since points are awarded per labeled image) if they agree quickly on a descriptive word for each individual image, the 2.5 min time limit means that players would prefer to ‘match’, or agree on a label, as early as possible in their sequence of descriptive words for each image. Thus the design of the ESP game induces players to have utilities that can be described as *match-early* preferences (Jain and Parkes 2013), where each player obtains a higher utility from ‘matching’ earlier rather than later with her partner. What kind of player behavior, and correspondingly what kind of labels, can be expected from such ‘match-early’ preferences induced by the ESP game design?

Theorem 1 (Jain and Parkes 2013).

(Informal.) With match-early preferences, choosing low effort and returning labels in decreasing order of frequency (i.e., from most common to least common) is a Bayes-Nash equilibrium in the ESP game.

Further, it turns out that under reasonable restrictions on strategy choices, such undesirable equilibria, where players coordinate on common words, are the only

Nash equilibria[12] in the ESP game. This result helps explain exactly *how* the design choices, i.e., the specific rules of the ESP game, might lead to the observed outcomes of common or generic labels for images.

Now suppose rewards are instead designed so that the number of points received by a pair of players depends not just on the *number* of matches, but also on the *quality* of each match, based on the frequency of the agreed-upon label. Such a reward scheme, where a player's utility depends not on *when* the match occurs (i.e., at which point in the sequence of words output by the player), but rather on the frequency of the matched label, induces *rare-words* preferences. How does changing the reward structure to remove the 'need for speed', and so that agreeing on rare labels leads to higher rewards, affect equilibrium outcomes?

Theorem 2 (Jain and Parkes 2013).

(Informal.) With rare-words preferences, returning labels in decreasing order of frequency (i.e., common words first) is a strictly dominated [13] *strategy. Returning words in* increasing *order of frequency (i.e., least common words first) is an ex-post Nash equilibrium in the ESP game, conditional on both players choosing the same level of effort.*

That is, such a change in the reward design leads players to 'try' the rarer words in their sample first, leading to more useful labels than those obtained under the equilibrium strategy of trying more common words first under match-early preferences. This change in design alone, though, is not adequate to induce effort—high effort sampling need not be an equilibrium strategy in the ESP game even when rewards are modified to induce rare-words preferences. If, however, the distribution of words in the dictionary from which samples are drawn is Zipfian (as is the case for the English language), and if the rewards are designed so that utilities additionally obey a certain (multiplicative or additive) structure, high effort sampling followed by coordination on rare words now becomes an equilibrium in the game.

This analysis of the ESP game demonstrates both (i) how a game-theoretic model and analysis can explain and pinpoint in what way a particular design choice for the game leads to the observed outcomes of low-information labels (arising from coordination on common words), and (ii) what kind of reward redesign can lead, under what conditions, to high-effort coordination on rare words. In the next subsection, we investigate another family of human computation systems where a formal analysis of incentives can aid the design of reward mechanisms that induce desirable behavior from participants in the system.

---

[12] A Nash equilibrium is a set of *strategies*, one for each player, such that no player can benefit by deviating from her strategy given the strategy choices of other players; see, for instance, Nisan et al. (2007).

[13] A strategy is strictly dominated if there is another strategy that always leads to larger payoffs regardless of other players' choices, i.e., for all possible strategies of other players.

Crowdsourced Judgement Elicitation

An increasingly prevalent application of human computation is in the domain of using the crowd to make evaluations, or *judgements*. Suppose each of a set of objects has one of many possible properties or belongs to one of many categories, and the task is to judge, or evaluate, what property the object has or which category it belongs to—for instance, categorizing galaxies or identifying birds (as in Citizen Science projects), deciding whether some text content is abusive or an image is pornographic, or deciding whether a homework assignment is correct or incorrect, or what score it should get. When the number of objects to be evaluated is too large for a single expert and the evaluation cannot be accurately performed by a computer, a human computation-based solution is to replace the expert's opinion by an aggregate evaluation based on judgements from a 'crowd' of non-experts, typically recruited via some online platform. Crowdsourced judgement elicitation is now used in a wide range of applications including image classification, identifying adult content online, rating learners' translations on the language-learning site Duolingo, and most recently for peer grading in online education, where Massively Open Online Courses (MOOCs) with huge enrollments crowdsource the problem of evaluating homework assignments back to the students in the class.

Consider a worker, say, on Amazon Mechanical Turk who is classifying images, or a Duolingo user who has been asked to rate another user's translation into his native language. Such a worker could potentially just arbitrarily categorize the object (an image, a translation, and so on) into some category—incurring no effort cost, or alternately, she can put in effort to properly evaluate the object. *If* the system could check the accuracy of the worker's output (e.g., the correctness of her categorization), and reward based on accuracy, the worker might be incentivized to put in effort into making judgements more accurately—but the reason for using human computation, of course, is that the system does not have this information in the first place. Given that the only source of information about the ground truth—the true category for each object—is judgements from the crowd, how should the system reward agents based on the received reports?

This question is related, although not the same as, the growing literature on mechanisms for *information elicitation*, also pertinent to human computation. Broadly, that literature addresses the question of designing mechanisms that incentivize agents to *truthfully* reveal information they already happen to possess, such as their opinions about a product or service (as in the peer-prediction literature (Miller et al. 2005)), or their beliefs about the probabilities of an event as in prediction markets, a literature by now too vast to properly discuss here (Chap. 26, Nisan et al. 2007). The problem encountered in the crowdsourced judgement elicitation domain is somewhat different than the one addressed by this literature, since here agents (workers) do not already possess the information they are being asked to share—they must expend an *effort cost* to acquire that information in the first place. Of course, having acquired the information, the reward structure additionally needs to induce agents to truthfully report what they observe.

Given both formal studies (Ipeirotis et al. 2010) and anecdotal reports[14] of effort-shirking by raters under ad-hoc or output-independent reward structures in real-world systems, there is a need for mechanisms that will incentivize agents to exert effort to make useful judgements on their tasks. Suppose an agent's utility is the difference between the reward she receives, and the cost of the effort she puts in, aggregated over all the tasks she performs. A mechanism for judgement elicitation in such human computation settings should make it 'most beneficial', if not the only beneficial strategy, for agents to not just *report* their observations truthfully, but to also to expend effort to *make* the best observations they can in the first place, rather than simply making arbitrary reports. Also, it is even more important here to ensure that the payoffs from an outcome where all agents blindly and consistently report the same observation (such as declaring all content to be good) are strictly smaller than the payoffs from truthfully reporting observations of the actual input, since declaring all tasks to be of some predecided type (without even observing the input) requires no effort and therefore incurs no cost, whereas actually putting in effort to make observations about the input will incur a nonzero cost. Dasgupta and Ghosh (2013) provide a simple model for this setting of crowdsourced judgement elicitation with unobservable ground truth, where an agent's proficiency—the probability with which she correctly evaluates the underlying ground truth (i.e., the true category or property of the object)—is determined by her *strategic choice* of how much effort to put into the task. They provide a mechanism—a set of rules which determines how to allocate rewards to agents—$\mathcal{M}$, for binary information elicitation for multiple tasks when agents have such endogenous (i.e., strategically determined) proficiencies, that has the following properties.

Theorem 3 (Dasgupta and Ghosh 2013).

Exerting maximum effort into making judgements, followed by truthful reporting of observations is a Nash equilibrium in mechanism $\mathcal{M}$. Further, this is the equilibrium with maximum payoff *to all agents, even when agents have different maximum proficiencies, can use mixed strategies, and can choose a different strategy for each of their tasks.*

Informally, the main idea behind the mechanism $\mathcal{M}$ is to use the presence of *multiple* tasks and ratings to estimate a reporting statistic that identifies and penalizes *blind*, or low-effort, agreement—since the only source of information about the ground truth comes from agents' reports, it is natural to use agreement as a proxy for accuracy, and reward an agent for agreement with another agent's evaluation of the same task. However, rewarding only for agreement can lead to low-effort equilibria with high payoffs (for instance, where all agents report the same observation independent of the input and therefore always agree), which is undesirable. The mechanism $\mathcal{M}$ therefore does reward agents for agreeing with another 'reference' report on the same task, but also penalizes for *blind agreement* by subtracting out a

[14] Such as in Duolingo and peer-grading systems

statistic term, which is based on the extent of the agreement that would be 'expected anyway' given these agents' reports over all the other tasks they rate. This statistic term is designed so that agents obtain nonzero rewards *only* when they put in effort into their observations, and so that reward is increasing in effort: this yields the maximum payoff property of the full effort-truthful reporting Nash equilibrium.

This crowdsourced judgement setting thus demonstrates another instance in which game-theoretic models and mechanism design provide useful input into the incentive-centric design of a broad family of human computation systems, where—given the accounts of effort shirking by raters under ad-hoc or output-independent reward structures in real-world systems—properly incentivizing agents is key to obtaining worthwhile, or valuable, input from the humans in the system.

Aggregating Quality Estimates: Voting

We illustrate a third kind of incentive problem in human computation by examining settings where user ratings are used to compute the (absolute or relative) quality of online content, such as photographs on Flickr, reviews on Amazon or Yelp, shared articles on Reddit, and so on. Rating and ranking are natural applications for human computation—in all the examples we just mentioned, it is hard for a computer to accurately process the task at hand, which is inferring content quality or rankings (for example, how does Flickr know whether a photograph is appealing?), whereas humans can easily accomplish the task.

Where do incentives and game theory come in? In a number of such voting or rating contexts, the set of people producing ratings is not disjoint from, and often has high overlap with, the set of people producing the content or objects[15] to be rated (for example, consider a community of photographers such as on Flickr, who both post photos themselves, and rate other contributors' photos). Since having a high relative rating for one's own content is desirable (highly-ranked content receives more attention, which seems to be clearly desired by contributors), a contributor who is rating other contributions might have an incentive to strategize her votes so as to increase her relative ranking—for instance, by downvoting other highly-rated contenders. A natural question then is the following: Is it possible to design a scheme for aggregating ratings that can 'get at' the true qualities, or perhaps the true underlying ranking of objects, or identify the set of the k-best objects, when the creators of the objects being rated are also the raters?

A simple abstract model for this problem is studied in Alon et al. (2011). Suppose, for simplicity, that the set of raters is exactly the same as the set of creators of the content; abstractly, this can be modeled by a voting scenario where the set of agents

[15] Note that these objects can also be the producers themselves, rather than only the content produced, as might be the case when constructing rankings of users based on their contributions in some online community.

who vote are identical to the set of candidates being voted on.[16] Consider a directed graph over this set of n agents, where an edge from agent i to agent j is taken to mean that i 'upvotes' or supports (for example, likes the content produced by) agent j.[17] Suppose the system wants to find the k most popular agents—for example, a site might want to prominently display the k most popular contributions. Each agent is only interested in being selected in this set of k 'winners', and so may misreport its opinions, or ratings, to this end. A *mechanism* in this setting is a way to aggregate the set of votes from the n agents into a set of k selected agents. Is it possible to design a mechanism which is simultaneously *strategyproof*—i.e., where no agent can benefit by misreporting which other agents she approves (or does not approve) of, i.e., her edges—as well as *approximately optimal*, in the sense that the total number of votes on the chosen set of k agents is 'close' to (i.e., not much smaller than) the total votes for the k most popular agents? Alon et al. (2011) analyze strategic behavior in this model to first show a surprising impossibility result:

Theorem 4 (Alon et al. 2011).

For any number of agents $n \geq 2$, and any number of winners k between 1 and $n - 1$, there is no deterministic *strategyproof k-selection mechanism with a finite approximation ratio.*

However, Alon et al. (2011) constructs a *randomized* mechanism (i.e., where the choice of the set of k winners also depends on the outcome of some random coin tosses) which is both strategyproof, and selects a reasonable set of agents:

Theorem 5 (Alon et al. 2011).

For any k between 1 and $n - 1$, there is a randomized *k-selection mechanism that is both strategyproof, and has an approximation ratio*[18] *no worse than 4; this mechanism is approximately optimal as k diverges.*

Together, these results, based on a formal analysis of strategic behavior in a simple voting model, establish the tradeoffs that the designer of a human computation-based rating or ranking system should expect to find when dealing with self-interested users—while no simple (i.e., deterministic) mechanism for aggregating ratings can be both strategyproof and optimal for all inputs, there exists a more complex (randomized) mechanism that can eliminate any benefits from misreporting while also not compromising the quality of the winner set too much, especially as the size of that set diverges.

[16] An example of such a situation, outside of the context of human computation or the Internet, is the election of the pope in the papal conclave.

[17] For readers familiar with the voting literature, this setting is a special case of *approval voting* where the set of voters coincides with the set of options.

[18] That is, the set of winners obtains at least 1⁄4 as many votes as the k most popular agents

Incentivizing Consistent Effort: Gamification and Game Theory

In the previous section, we saw the role of formal game-theoretic analysis and design in three human computation contexts—specifically, we saw how rewards, or benefits, for particular tasks can be restructured to provide incentives to agents to undertake the 'right', i.e., system-desired, behaviors. In this section, we will discuss an application of game-theoretic techniques to a broader class of incentives for participation: an increasing number of human computation systems are now accompanied by corresponding *online communities*, with discussion forums, leaderboards, reputation scores, and various other features, all of which also provide rewards (typically of a social-psychological nature) to participants, albeit not for performance on a particular task. While our previous analyses looked at incentives and cost-benefit tradeoffs from a *single* action or contribution, there are also rewards that relate directly to the identity of a *contributor* typically based on her overall contribution, rather than to single actions or contributions. In this section, we will discuss very recent work on formal approaches to designing incentives that motivate *overall contribution* in human computation systems via their communities.[19]

A common theme in a growing number of online communities and social media sites relying on user contributions is *gamification*—via badges, leaderboards, and other such forms of (competition or accomplishment based) social-psychological rewards. These rewards, meant to provide an incentive for participation and effort on a given system or site, usually reflect various site-level accomplishments based on a user's cumulative 'performance' over multiple contributions. Such badges or top-contributor lists clearly appear to motivate users, who actively pursue and compete for them—for example, users on StackOverflow are observed to increase their effort levels when they get close to the contribution level required for a badge (Anderson et al. 2013), and there are entire discussion communities on the Web centered around how to break into Amazon's Top Reviewer list or how to maintain a Top Contributor badge on Yahoo! Answers, while users who have just earned entry into top contributor lists often find an increased number of negative votes from other users attempting to displace them.

Given that the rewards created by these virtual badges and leaderboards appear to be valued by users (a phenomenon that appears to be quite general, occuring across a range of online communities) and that participating and putting in the effort required to obtain them is costly, a particular way of allocating these rewards creates a corresponding set of incentives, or more formally, induces a *mechanism* in the presence of self-interested contributors. So gamification also involves reasoning about incentives in a game-theoretic sense—given that there are several different

[19] For a broad set of general guidelines on incentivizing participation and engagement in online communities, see Kraut et al. (2012).

ways to 'gamify' a site, how should these rewards for overall contribution be designed to incentivize desired levels of contribution? For instance:

1. What incentives are created by mechanisms induced by an *absolute* standard of output that must be met to earn a badge (such as a threshold number of images that must be tagged, or questions that must be answered), and what incentives are created by a *competitive*, or relative, standard, such as top-contributor badges or leaderboards? And how do these 'compare'?
2. When badges are awarded for meeting absolute standards, should multiple badges be awarded, and if yes, how should they be 'placed' relative to each other in terms of the accomplishments required to earn successively higher levels of badges?
3. Consider a very simple form of a relative standard, corresponding to handing out an (identical) 'top-contributor badge' to some set of 'best' contributors on the site. How exactly should badges for competitive standards be specified—should the site award some fixed number of top-contributor badges *independent* of the number of actual participants, such as a Top 10 Contributors list (call this mechanism $\mathcal{M}_\rho^p$), or should the number of winners be some fraction of the number of *actual* participants (mechanism $\mathcal{M}_\rho^c$)? Note that since participation in all these human computation systems is a voluntary choice, the number of actual contributors is *not fixed* apriori, but rather is determined by the choices made by self-interested users—so these two specifications are *not* equivalent.

This family of questions brings us to the frontiers of research on game theory for human computation, which we summarize below. First we address the questions about what kinds of incentives are created by absolute and relative standards mechanisms. Call the awarding of badges for achieving some absolute standard, say α, of output (such as receiving α positive ratings for one's contributions, or labeling α images correctly), an absolute standards mechanism $\mathcal{M}_\alpha$. Call the awarding of badges for belonging amongst some set of top ρ contributors to the site a relative standards mechanism $\mathcal{M}_\rho$. Easley and Ghosh (2013) investigates the existence and nature of equilibrium outcomes in these two classes of mechanisms in a simple game-theoretic model where users who value badges (presumably for social-psychological reasons), and have a cost to effort, strategically choose whether to participate and how much effort to put into the site.[20]

Easley and Ghosh (2013) find that even the existence of equilibria for relative standards mechanisms $\mathcal{M}_\rho$ depends on *how* the number of top contributor awards ρ is specified (i.e., whether there are a fixed number of top-contributor badges that will be awarded, or whether the number of badges scales as a fraction of the number of actual participants)—this is due to endogenous participation, i.e., the fact that users make a voluntary choice about whether to participate depending on the

[20] An equilibrium here consists of some level of participation and some level of effort from participants, such that no participant can benefit from either dropping out or choosing to exert a different level of effort, and no non-participant would prefer to participate.

rewards being offered. While the two versions of the relative standards mechanism behave identically for ρ lying in a certain range, the result below suggests that at least for settings that are reasonably captured by the model in Easley and Ghosh (2013), the mechanism corresponding to announcing a fixed number of top-contributor badges that is independent of the number of actual participants is a more robust mechanism than one that declares some fraction of participants to be winners, i.e., where the number of winners scales with the number of actual contestants.

Theorem 6 (Easley and Ghosh 2013).

(Informal.)

1. *For relative standards mechanisms $\mathcal{M}_\rho$, equilibria exist for all values of $\rho > 0$ if the site specifies ρ as a fraction of* potential *contributors, i.e., as a* fixed number *of winners, but not if ρ refers to a* fraction *of* actual *contributors.*
2. *For absolute standards mechanisms $\mathcal{M}_\alpha$, equilibria exist for all possible values of the standard α. However, there is a maximum standard $\alpha_{\max}$ such that the only equilibria for all standards higher than $\alpha_{\max}$ involve zero participation, leading to no contributions.*

This equilibrium analysis suggests an interesting contrast between using relative and absolute standards for rewarding overall contribution—while $\mathcal{M}_\rho^p$ elicits non-zero participation in equilibrium for every value of $\rho > 0$, $\mathcal{M}_\alpha$ can lead to zero equilibrium participation when α is too large. However, there is also a *partial* equivalence between absolute and relative standards $\mathcal{M}_\alpha$ and $\mathcal{M}_\rho^p$, of the following form. Every absolute standard $\alpha \leq \alpha_{\max}$ leads to an equilibrium outcome that is identical, in terms of induced effort and participation, to the equilibrium outcome in the relative standards mechanism with some appropriate value of $\rho \in [\rho_{\min}, 1)$, where $\rho_{\min} > 0$ is the equilibrium fraction of winners at the standard $\alpha_{\max}$—and in fact, the value of ρ that elicits the *maximum* effort from contributors occurs at a relative standard ρ that lies in this range $[\rho_{\min}, 1)$. So for a site designer who wants to optimize elicited effort, and has adequate information about the parameters of the population to choose an optimal value of the standard α or ρ, the absolute and relative standards mechanisms are equivalent. In the absence of such information, however, or with uncertainty about the population's parameters, a 'top contributor' style mechanism $\mathcal{M}_\rho^p$ based on competitive standards that always elicits non-zero equilibrium participation might be, informally speaking, more desirable than an absolute standards mechanism.

Finally, we ask a question about multiple badges—consider badges that are handed out for absolute achievements. At what levels of achievement should badges should be awarded to sustain effort on the site, and how should they be designed to steer user behavior towards different actions on the site? Anderson et al. (2013) address this question in a model where there is a multi-dimensional space representing the possible types of actions on the site. Users have a time-discounted value to earning badges and incur a cost when they choose actions from a distribution that differs from their preferred mixture of actions on the site. If users act to maximize their utility in this model of costs and benefits, how should badges be placed to align ideal user behavior with users' utility-maximizing actions? Anderson et al. (2013) finds that the effectiveness of badges in inducing desirable behavior depends significantly on their 'placement' (i.e., for what level of contribution they are awarded), with the optimal location being, roughly speaking, one that is hard to achieve and therefore motivates users for a significant length of 'time' (contributions). Also, multiple badges should be 'spread out' with roughly equal values, rather than placing them at nearby levels of contribution, suggesting that multiple smaller rewards provide more effective incentives than a small number of larger rewards at least in settings that are well-described by the model in Anderson et al. (2013).

The literature on a game-theoretic approach to overall contributor reward design is very young, and has looked at the most immediate questions under relatively simple models and reward structures. There are a number of questions still to be modeled and answered, an immediate one being the design of leaderboards. In contrast to top-contributor badges, not all 'winners' receive equal rewards in leaderboards since arguably, the reward from placing first (or say in the top 5 positions) is somewhat larger than, say, ranking 100th on the leaderboard, even in a site with a large population. A number of interesting game-theoretic questions arise, starting from the very basic question of how many positions the leaderboard should have to optimally elicit effort from contributors; this question is related to our motivating example early in this chapter, and a first step towards such questions, although in a model with perfectly observable outputs, is taken in Ghosh and McAfee (2012).

Finally, a commonly used reward structure is that of user reputations. The question of how to design—and use and update—user reputations to create the right incentives in a human computation system is one that can draw from a vast body of literature on the design on reputation systems (Chap. 27, Nisan et al. 2007), but comes with challenges unique to human computation systems that will require the development of convincing new models and schemes [21]: In addition to differences in details from the models in prior work on reputation systems (for example, in the context of electronic marketplaces such as EBay or Amazon), there are also potentially fundamental differences that might arise due to the differences in the nature of the rewards that agents seek from these systems, which are primarily financial in online marketplaces but to a large degree social-psychological (such as status or reputation within a community) in human computation systems. We briefly explore these ideas in section "Challenges and Further Directions".

Challenges and Further Directions

In the previous sections, we saw how a game-theoretic, or more broadly, an economic approach, can help with analyzing strategic behavior and incentive design in human computation systems. But there remain many challenges, unique to such online

[21] For preliminary work on social norms for reputation, see Ho et al. (2012).

contribution domains, that need to be understood before we can fully develop the game-theoretic foundations for incentives in human computation. First, of course, there are a number of immediate questions regarding theoretical modeling and analysis. In addition to questions we have already alluded to in previous sections, there is also an interesting family of problems arising from the diversity of roles that participants play in many systems (for example, contribution versus moderation in an online community). How should incentives be designed to ensure that each participant is incentivized to properly contribute to her role(s) in the system, given that different roles might require different incentives, and that these incentives could potentially interact with each other? A principled framework that helps answer this question will need to begin with new models that appropriately capture such multi-role participation as well as interactions between different sets of incentives—an issue relates, at least in spirit, to the question of what incentives are created by simultaneously using different forms of gamification on a site. A further question along these lines, arising from the voluntary nature of participation, is how to structure incentives to also induce different potential participants to *choose* their socially optimal roles in the system.

In addition to problems related to modeling and theoretical analysis, there are also a number of cross-disciplinary questions. One family of problems lies at the interface of game theory and *interaction design*. By influencing usability, and usage, the design of the user interface in a human computation system also interacts with incentives in a game-theoretic sense—after all, any game-theoretic analysis involves modeling the behavior of the agents (i.e., users) in the system, which is determined not only by its rules for reward allocation but also by its interface. As a very simple example, consider a system that rewards contributors based on the quality of their outputs, as measured by the ratings, or votes, provided by users who view these contributions. An interface design which leads to very little rating by users (for example, a hard-to-find rating button or an overly complex menu of options), or one that leads to ambiguity in the meaning of a rating (such as a thumbs-up button which is interpreted by some users to mean 'Helpful' and others to mean 'I agree') results in 'noisier' ratings than an interface which elicits meaningful votes from a large number of users. A greater degree of noise, roughly speaking, means that reward depends on effort in a more uncertain way, which in turn affects the incentives for agents to put in effort in the system. It is easy to see that even in this specific example there is much more to consider at the interface of interaction design and incentives, such as the question of *which* users are allowed to rate contributions, and whether raters are offered a more or less expressive set of ratings to choose from. Another example of the connection between interaction design and game theory can be found in the context of badges and gamification—how much information about users' behavior and performance is revealed to other users can potentially affect users' valuations of badges, and consequently their strategic choices; see Sect. 5.3 in Easley and Ghosh (2013). Generally, therefore, how users respond to a given mechanism in a strategic or game-theoretic sense, as well as the space of available mechanisms itself, can depend on the choice of interface in the interaction design phase—an ideal design paradigm would take into account both the influence of the user interface and the reward allocation rules on user behavior to provide an integrated, complete approach to the design of incentives in human computation systems.

Finally, a very important family of questions relate to properly understanding contributor motivations and rewards in a more nuanced fashion. One particularly interesting issue that is pertinent to most human computation systems is that of *mixed incentives*: unlike in most traditional economic analysis, human computation systems typically involve a *mixture* of potential contributor rewards. Systems with financial rewards for contributing, such as Amazon Mechanical Turk, mix two entirely different kinds of rewards (financial and social-psychological); even in systems without financial incentives, there are usually multiple social-psychological rewards, either intrinsic or site-created: for instance, von Ahn and Dabbish (2008) describes fun as the primary motivator in the ESP game, but there are also social-psychological rewards from leaderboards (competition) as well as from successful 'collaboration' with partners on the image labeling task.

How do people—the agents in a game-theoretic model—value these different kinds of rewards in combination, and also, how do they value them relative to each other? What happens when virtual points are used to create an economy with money-like properties (a currency for exchange of goods and services), versus using virtual points to create psychological rewards (such as status)? Second, how do social-psychological rewards, even individual ones, aggregate in terms of the perceived value to contributors? While utility from money—both in terms of value as a function of total wealth, and the change in value of wealth with time—is a relatively well-studied subject in the economics literature, very little is known or understood about how social-psychological rewards aggregate, and how they retain (or gain or lose) value over time; also, unlike financial rewards, this could be partially controlled by system design. Understanding how multiple rewards influence incentives when they occur simultaneously in a system, and how social-psychological rewards provide value—starting with understanding agent preferences from a behavioral economics perspective, and then integrating this understanding into formal game-theoretic models—is an essential component to a strong foundation for incentive design for human computation, and one of the most exciting directions for future work in this area.

References

Alon N, Fischer F, Procaccia A, Tennenholtz M (2011) Sum of us: strategyproof selection from the selectors. In: Proceedings of the 13th conference on theoretical aspects of rationality and knowledge (TARK), Groningen

Anderson A, Huttenlocher D, Kleinberg J, Leskovec J (2013) Steering user behavior with badges. In: 22nd international world wide web conference (WWW'13), Rio de Janeiro

Dasgupta A, Ghosh A (2013) Crowdsourced judgment elicitation with endogenous proficiency. In: Proceedings of the 22nd ACM international world wide web conference (WWW), Rio de Janeiro

Easley D, Ghosh, A (2013) Incentives, gamification, and game theory: an economic approach to badge design. In: Proceedings of the 14th ACM conference on electronic commerce (EC), Philadelphia, 2013

Ghosh A, Hummel P (2012) Implemeting optimal outcomes in social computing. In: Proceedings of the 21st ACM international world wide web conference (WWW), Lyon, 2012

Ghosh A, McAfee RP (2012) Crowdsourcing with endogenous entry. In: Proceedings of the 21st ACM International World Wide Web conference (WWW), Lyon, 2012

Ho C, Zhang Y, Vaughan J, Schaar MVD (2012) Towards social norm design for crowdsourcing markets. In: Proceedings of the AAAI workshop on human computation, San Francisco

Ipeirotis P, Provost F, Wang J (2010) Quality management on amazon mechanical turk. In: Proceedings of the ACM SIGKDD workshop on human computation (HCOMP), Washington, DC

Jain S, Parkes D (2013) A game-theoretic analysis of the ESP game. ACM Trans Econ Comput 1(1):3

Jain S, Chen Y, Parkes D (2012) Designing incentives for online question-and-answer forums. Games and Economic Behavior, forthcoming

Jian L, MacKie-Mason JK (2012) Incentive-centered design for user-contributed content, Oxford Handbook of the Digital Economy, edited by Martin Peitz and Joel Waldfogel, forthcoming

Kraut R, Resnick P, Kiesler S, Ren Y, Chen Y, Burke M, Kittur N, Riedl J, Konstan J (2012) Building successful online communities: evidence-based social design. MIT, Cambridge

Miller N, Resnick P, Zeckhauser R (2005) Eliciting informative feedback: the peer-prediction method. Management Science 51(9):1359–1373

Nisan N, Roughgarden T, Tardos E, Vazirani V (2007) Algorithmic game theory. Cambridge University Press, New York

Pickard G, Pan W, Rahwan I, Cebrian M, Crane R, Madan A, Pentland A (2011) Time-critical social mobilization. Science 334:509–512

von Ahn L, Dabbish L (2008) Designing games with a purpose. Commun ACM 51(8):58-67

Weber I, Robertson S, Vojnovic M (2008) Rethinking the ESP game. Technical report, Microsoft Research

Part VII
Analysis

Analysis: An Introduction

Kristina Lerman

In many systems considered in this book, computation is an emergent property of a large population of interacting individuals. The role of analysis is to uncover and validate the microscopic mechanisms that govern an individual's behavior. The products of analysis are descriptive models and theories of individual behavior, and a framework that explains the collective behavior that arises from interactions among many individuals. In addition to being descriptive, the models are often used to predict emergent collective behavior and motivate the design of future human computational algorithms and user interfaces that support them.

As social interactions have moved online, they have left behind rich traces of human behavior to be analyzed by researchers. The growing abundance of data has made it possible to study the cognitive, psychological, social and cultural mechanisms that govern individual and social behavior. The chapters in this section present a variety of analytic approaches for studying these data, as well as the insights obtained by these approaches. The techniques described in this section are diverse: they range from empirical analysis to simulations, mathematical modeling, network analysis, as well as higher-level approaches that examine socio-cultural principles of behavior and synchrony in social groups. These analytic techniques are indispensable tools of any scientist or engineer who wants to understand collective behavior or harness its computational power.

K. Lerman (✉)
USC Information Sciences Institute, Marina del Rey, CA, USA
e-mail: lerman@isi.edu

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_59,

Empirical Analysis

Empirical analysis identifies patterns in data that reveal common behavioral principles. Empirical analysis is frequently used in the exploratory investigation of data, and is the first step in constructing mathematical or computational models of a system.

The chapter on empirical analysis by Kristina Lerman argues that massive data about social interactions has enabled quantitative "in-vivo" studies of human behavior. Although such studies lack the controls of human subject experiments that are conducted by social scientists in a laboratory setting, the vast scale of the data set, plus the naturalistic setting under which it was collected, can complement laboratory studies and even provide important insights into human behavior. Specifically, the chapter illustrates how empirical analysis was used to demonstrate the manner in which psychological factors affect online interactions. Psychologists and cognitive scientists have long known that people have a limited capacity to process information, the phenomenon that we refer to as "limited attention." As shown in the chapter, this constraint has important implications for social communication and information spread, and should not be ignored by the designers of social computing systems.

One of the main challenges of working with big social data is its heterogeneity. People are extremely diverse regardless of the behavior or feature being measured. In order to identify patterns or trends in behavior, data scientists normally aggregate data across all individuals. However, "the ruses of heterogeneity" may lead to erroneous conclusions, e.g., when the average behavior does not apply to any single individual. Instead, one must separate individuals into populations that are as homogeneous as possible. As the chapter describes, this can drastically change the inferred models of behavior. For example, when measuring how people respond to repeated exposures to information, aggregating over all people suggests that after some point, exposures may inhibit response. However, once we recognize limited attention as an important factor in social media interactions and separate social media users into distinct populations based on their cognitive load, their response to repeated exposures changes dramatically. Now, likelihood of a response increases monotonically for all populations, with no evidence of inhibition.

Computational Analysis

Computational analysis uses simulations to validate individual-level microscopic models and explore their population-level outcomes. Beyond validating our understanding of human behavior, agent-based models can be used to test, in simulation, different candidate social systems before they are launched, for example, to ensure that desired behavior is achieved. The chapter on computational analysis by Weng and Menczer describes agent-based modeling, a popular framework for studying via simulations the behavior of multiple autonomous interacting agents. Each agent's behavior is described by a set of rules, the so-called microscopic model. The researcher then evolves the multi-agent system in time, starting from some initial state, by simulating interactions between agents, and monitors its collective behavior. The researcher can then adjust the microscopic rules of an individual agent's behavior until the desired system-level behavior is observed.

The computational analysis chapter describes the use of agent-based models to explain two empirical observations about the spread of information in social media: the daily variety of topics people tweet about remains relatively constant over time, and people tend to tweet about topics they tweeted about in the past. Using a model of individual agent behavior that includes finite attention and memory, Weng and Menczer are able to recreate the empirically observed macroscopic trends in the persistence of topics on Twitter. This study provides further evidence for the importance of limited attention in online social interactions. Further, they show that a combination of heterogeneous social network structure and finite agent attention is sufficient to explain the emergence of broad diversity of persistence times and popularities of various topics on Twitter.

Mathematical Analysis

Mathematical modeling provides an alternative tool for studying the collective behavior of a population of interacting agents, and it has a long tradition in ecology, epidemiology, and population dynamics. The stochastic modeling approach discussed by Tad Hogg in this section is a type of mathematical analysis that considers each agent as an automaton, i.e., a set of states and probabilistic transition rules between them—in other words as a stochastic Markov process. The Markov process then serves as a template for deriving a series of equations describing the collective behavior of a population of identical agents. The models are then be solved mathematically to obtain how the collective behavior changes in time. Note that agent-based models are related to mathematical analysis, since both start with the set of rules describing individual agent behavior; however, their routes to obtaining collective behavior are different. One advantage to mathematical analysis is that the dependence of collective behavior on some parameter can be explicitly specified, sometimes analytically, whereas in computational analysis such dependence can be discovered only after performing multiple agent-based simulations with different parameter values. It is important to remember that the mathematical model describes average collective behavior, not necessarily the behavior of any specific agent.

Despite this constraint, stochastic modeling can be used to make predictions about collective social behavior. The chapter describes an application of the stochastic modeling approach to predict the popularity of content in social media. This is an important problem in social media and crowdsourcing, since the volume of new user-generated content is continuously growing, making it critically important to provide users with tools to help them identify interesting content in a timely manner. In the chapter, Hogg links user behavior on the social news aggregator Digg to features provided by the site, such as the ability to see news stories recommended by friends or to easily see the stories Digg promoted to the front page. Hogg explains how to construct equations describing user behavior on Digg, and solve them to see how the popularity of news stories evolves in time. In addition to helping evaluate different design options for the site, stochastic modeling can be used to predict how popular different stories will become.

Network Analysis

Social computing systems are, by definition, social, and social interactions are highly structured, with communities of strongly tied individuals interacting more frequently and intensely than weakly tied individuals. Understanding how the structure of social interactions affects the dynamics of emergent behavior is crucial for designing efficient and robust social computing platforms.

The structure of social networks is highly heterogeneous, with long-tailed distribution of connectivity, assortativity, community structure, and other types of correlations between nodes. Heterogeneous network structure affects the popularity of information, as discussed in chapters by Lerman, and Weng and Menzer. The chapter in this section on network analysis by Aram Galstyan examines how network structure affects dynamics of information flow. Specifically, the chapter studies the role of community structure, a common property of social networks, in how far and how quickly information diffuses on the network. Repeating the themes of earlier chapters, Galstyan employs mathematical analysis and simulations to demonstrate that community structure profoundly changes dynamics of information flow and which nodes should be targeted to maximize information spread.

Social Synchrony

Social synergy is a much sought-after phenomenon where the total effect of social activity is much greater than the sum of individual efforts. However, as Xuan and Filkov point out in their chapter, synergy is difficult to define and measure. Instead, they examine synchrony, a concept related to synergy, as a mechanism for distributed coordination of collective behavior in social groups. Synchrony occurs in many natural systems in which cells, insects and birds coordinate their activity through local interactions, in order to improve the well-being or evolutionary fitness of all participating organisms. For example, male fire flies that synchronize their patterns of light flashing are collectively much brighter and are able to attract mates from farther away than they would individually. By analogy, synchrony via distributed synchronization in social systems is also thought to offer benefits, such as efficiency and robustness.

Xuan and Filkov combine the various types of analyses described in this section—empirical, mathematical, and network analysis—to study social synchrony. The chapter describes various mathematical models of synchrony and discusses the effect that network structure has on dynamics of synchronization. They illustrate with a case study of open source software (OSS). To ensure success of an OSS project, developers and users have to work together to write code, discover and fix bugs, etc. Xuan and Filkov present an analysis of OSS projects that quantifies the degree to which social synchrony is present in the project. Such analysis could be used to monitor the robustness and efficiency of social computing systems.

Analytical Gaming

Social behavior is a product of complex cognitive, psychosocial, and cultural processes. Individual's decisions are affected by his or her values, priorities, and reason, but also by cognitive biases, perceptions of risk, and the decisions of other individuals which are equally complex. The chapters in this section described approaches to disentangling these factors through various analytic techniques. In their chapter, Sanfilippo, Riensche and Haack, describe an alternative approach to the analysis of social systems—analytical gaming. Their approach uses gameplay to recreate real-world scenarios in which human players make decisions and use their judgements to solve problems. Gameplay allows analysts to use actual humans to calibrate the impact of "humanness" on the social behavior and scenario outcomes. These data are then used to calibrate agent-based socio-cultural models, which can be later deployed in other scenarios. This approach allows analysts to identify and quantify the psycho-social and cultural factors that contribute to the complex behavior of social computing systems.

Conclusion

The following chapters should be viewed as introductions to each type of analysis. They should serve as springboards for a motivated reader to learn more about each topic.

Social Informatics: Using Big Data to Understand Social Behavior

Kristina Lerman

Introduction

Modern communications technologies, notably email and more recently social media, have enabled people to interact on an unprecedented scale. The social networks that emerge from these interactions can amplify information (Wu et al. 2004; Gruhl and Liben-nowell 2004), mobilize massive ad-hoc teams (Pickard et al. 2011) and political movements (Lotan et al. 2011), help people discover information (Adamic and Adar 2005; Lerman 2007) and make new connections. In addition to making social networks ubiquitous, social media has given researchers access to massive quantities of data for analysis. These data sets offer a rich source of evidence for studying the structure of networks and the dynamics of individual and group behavior, and ask new questions about social communication. How far and how fast does information spread? How do people respond to new information? What are the mechanisms of information spread and how do individual's cognitive limitations affect them?

We have addressed these questions through a large scale analysis of data from two social media sites: Digg and Twitter. Despite having different functionality and user interface, both sites are used in remarkably similar ways by people to share information with others, thus enabling us to uncover principles of social behavior that generalize across platforms. The social news aggregator Digg allows users to *submit* links to news stories and *recommend* stories submitted by other users by voting for them. On Twitter, users *tweet* short text messages, that often contain links to news stories, or *retweet* messages of others. Both sites allow users to link to others whose activity (i.e., votes and tweets) they want to follow. Upon visiting Twitter, a user is presented with a list of messages most recently tweeted or retweeted by the

K. Lerman (✉)
USC Information Sciences Institute, Marina del Rey, CA, USA
e-mail: lerman@isi.edu

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_60, © Springer Science+Business Media New York 2013

followees of the user, i.e., other users whom the given user follows. Similarly, on Digg a user sees a list of news stories recently recommended by those a user follows. By recommending a story, or retweeting a message, in turn, the user acts to further spread the information contained in that story or message.

We trace the flow of information from users to their followers on these sites (using URLs as unique markers of information) and measure its properties. We find that information does not spread to as many people as predicted by a simple model that is commonly used to describe the spread of information. Our attempts to resolve this puzzle illuminates the critical role that individual's limited attention plays in social media.

Social Information Sharing

We studied social information sharing on Digg and Twitter, two popular social media sites for sharing news and other content. For our study, we tracked how items, uniquely identified by URLs, were shared by users. Details of data collection from both sites are described in Lerman et al. (2012).

Figure 1 shows the statistics of social behavior on Digg and Twitter, including the distribution of the number of followers ((a) and (d)) and activity ((b) and (e)), i.e., number of votes or retweets made by each user. While the overwhelming majority of users on both sites shared fewer than ten items (URLs) with followers, a handful of users shared thousands of items over the period of a month. Such heavy-tailed distributions are typical of social production and consumption of content, where a small but non-vanishing number of items generate uncharacteristically large amount of activity, and have been observed in voting on Essembly (Hogg and Szabo 2009), edits of Wikipedia articles (Wilkinson 2008), and music downloads (Salganik et al. 2006) and other and real-world complex networks (Clauset et al. 2009).

The total number of times the URL was shared reflects its popularity. The distribution of popularity on both sites is long-tailed (Fig. 1c, f). It appears that information in social media rarely goes "viral" (Ver Steeg et al. 2011; Goel et al. 2012). The vast majority of items fail to spread at all, reaching only a handful of users. Even the most popular items spread to at most a few thousands users, which is a tiny fraction of the follower graph. Moreover, the distribution of popularity on the two sites is strikingly different: while the distribution of popularity on Digg is well described by a log-normal (shown as the red line), with the mean of 614 votes, there is no preferred popularity for retweeted URLs on Twitter. What gives rise to the difference in distributions of popularity? Wu and Huberman (2007) proposed a phenomenological model that explained the log-normal distribution of popularity on Digg as a byproduct of competition for attention for news stories and their decaying novelty. In contrast, we find that the difference can be explained by Digg's promotion mechanism, which highlights a handful of stories on its popular front page. The test this hypothesis, we gathered statistics about more than 20 K stories submitted to Digg over the course of 1 day in July 2010. The distribution of popularity of these stories is similar to Twitter (Fig. 1c). Of these stories, about 100 were promoted to the front

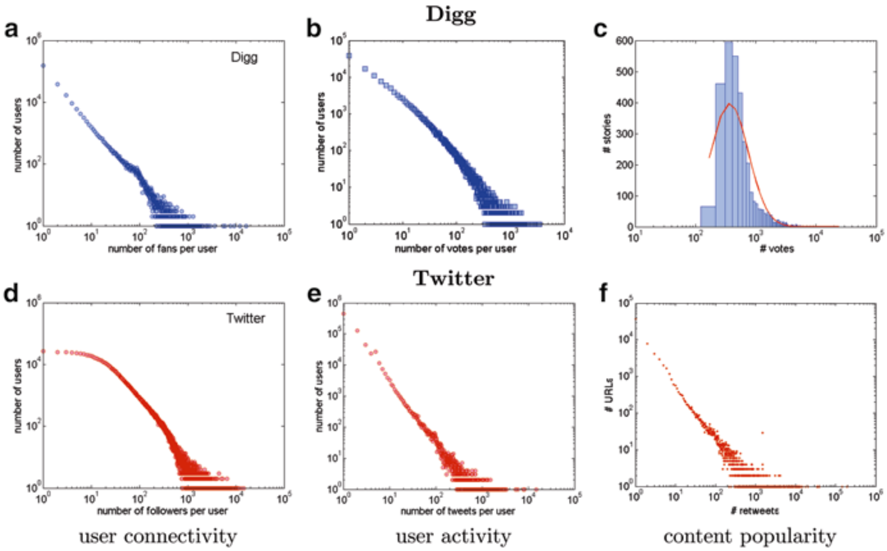


Fig. 1 Characteristics of user activity on Digg and Twitter. Distribution of the number of followers per user on the two sites, distribution of activity, which is given by the number of votes (on Digg) and retweets (on Twitter), and the distribution of popularity of content, as measured by the total votes received by news stories on Digg and the total number of times the URL was retweeted on Twitter. *Red line* in the distribution of votes received by Digg stories is log-normal fit to data

page and their popularity continued to grow. The final popularity of the promoted stories had a log-normal distribution. Therefore, we conclude that the log-normal popularity distribution is a by-product of selection by the promotion algorithm.

A Simple Model of Information Diffusion

Why does some content become popular but not other? How does information spread between people? In order to answer these questions, we need a model of social contagion that describes the microscopic dynamics of the spread of information. One of the simplest such models is the independent cascade model (ICM) (Newman 2022; Kempe et al. 2003; Gruhl and Liben-nowell 2004; Anagnostopoulos et al. 2008), which has been used to describe the spread of a disease in a population (Hethcote 2000). In this model, each exposure of a healthy person by an infected friend leads to an independent chance of the healthy person contracting the disease, and spreading it to her own followers thereby creating a cascade of infections. The likelihood that an exposure leads to an infection is set by pathogen's transmissibility, i.e., how contagious it is. When ICM is stated in the language of information spread, each exposure of a naive individual by an informed friend (e.g., via a tweet), creates an independent chance of information transmission. Therefore, the likelihood that the naive individual becomes informed should increase monotonically with the number of exposures.

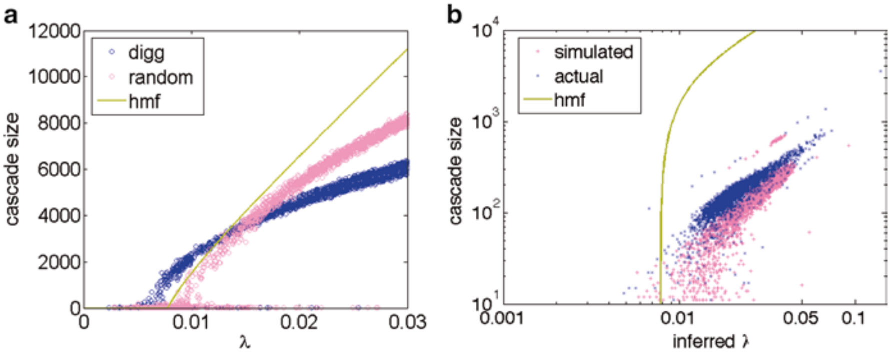


Fig. 2 Cascade size as a function of transmissibility λ. (**a**) Comparison of simulated cascades on the Digg follower graph and on the randomized graph with the same degree distribution. (**b**) Comparison of real and simulated cascades on the Digg graph that are produced using empirical exposure function. Theoretical predictions for a graph of the same size are shown by the *bold line*

Simulations of Information Diffusion

The dynamics of the independent cascade model has been well-studied. Specifically, it is known that there exists a critical value of transmissibility below which the disease does not spread, but above which it reaches a substantial fraction of the population, resulting in an epidemic (Castellano et al. 2009; Satorras and Vespignani 2001; Wang et al. 2003). Moreover, the expected size of an epidemic outbreak of a pathogen with a given transmissibility can be theoretically calculated (Moreno et al. 2002).

Our simulations of the independent cascade model on the Digg follower network confirm these expectations. Starting with random seed node, we generate a cascade as follows (see Ver Steeg et al. (2011) for details). Each time a node is infected, it will attempt to infect each follower independently with probability given by the transmissibility λ. The cascade stops when no new nodes are infected. The number of infected nodes, i.e., cascade size, is shown in Fig. 2, where each point represents a single simulated cascade with transmissibility λ. Dark gray dots represent cascades on the original Digg follower graph, while light gray dots represent cascades on a randomized version of the Digg graph with the same degree distribution. Both curves manifest a critical value of transmissibility, called the *epidemic threshold*, above which cascades spread to a significant fraction of the graph.[1] The location of the epidemic threshold is accurately predicted by the inverse of the largest eigenvalue of the adjacency matrix of the graph (Wang et al. 2003): $\lambda_c^{digg} = 0.00587$ for

[1] Note that even above epidemic threshold, cascades that start in an isolated region of the graph will die out.

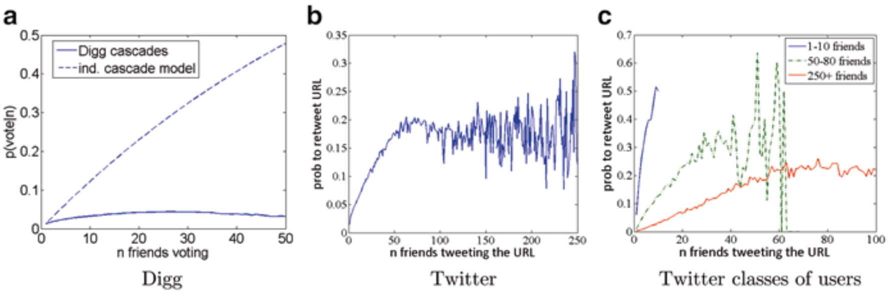


Fig. 3 Response to multiple exposures. (**a**) and (**b**) Show probability of infection given n infected friends aggregated over all users on (**a**) Digg and (**b**) Twitter. Plot (**c**) shows exposure response of Twitter users after they are separated into different classes based on their cognitive load, i.e., the number of friends they follow. *Dashed line* in (**a**) show exposure response predicted by the independent cascade model. (**a**) Digg. (**b**) Twitter. (**c**) Twitter classes of users

the original Digg graph and $\lambda_c^{rand} = 0.00928$ for the randomized graph. The size of theoretically predicted cascades is depicted by the gold line, which accurately characterizes both the threshold and growth of cascades on the randomized graph.

Figure 2 presents a puzzle. Information in social media spreads to a far smaller fraction of the population than predicted by the epidemic model. This is not because these URLs have low transmissibility: the Digg dataset, consists of URLs that have been selected for the front page. Nor does it appear to be due to network structure: while structure of the real Digg graph reduces the size of cascades in simulations compared to the randomized graph, it does not suppress it nearly enough to account for the observed sizes of actual outbreaks.

Exposure Response

A potential explanation for why information spread in social media fails to reach epidemic proportions can be found in how people respond to repeated exposures to information, that is, the probability they will rebroadcast the information via a retweet or a vote after multiple friends have tweeted about it or voted for it. According to the independent cascade model, the probability a node becomes infected, e.g., by voting for a story on Digg, increases monotonically with the number of infected neighbors n it has. This probability is given by the exposure function:

$$p_{ICM}(\text{ infected} \mid n \text{ infected friends}) = 1 - (1 - \lambda)^n.$$

To measure the exposure function on Digg and Twitter, we isolated users who had exactly n infected friends but did not become infected themselves, from users who had n infected friends before they themselves became infected. The solid lines in Fig. 3a, b show the probability of Digg and Twitter users respectively to become infected when exposed to information by n friends, averaged over all users. Exposure response on both sites is qualitatively similar. As the number of exposing friends increases, a user's probability to become infected goes up initially, but after a point additional exposure does not further increase response, and may in fact inhibit it. This behavior is similar to adoption of hashtags reported by Romero et al. (2011). In contrast, the dashed line in Fig. 3a depicts exposure response for the independent cascade model. ICM dramatically overestimates infection probability.

When we simulated information diffusion on the Digg follower graph using the empirical exposure response function measured from the data (Ver Steeg et al. 2011), the resulting cascades were dramatically smaller, as shown in Fig. 2b. In fact, the size of simulated cascades (pink dots) is similar to those of real information cascades on Digg (blue dots). It appears that failure to respond to exposures to information stops social epidemics.

Limited Attention in Information Diffusion

We still have a puzzle: why do users fail to respond to repeated exposures by friends? One potential explanation could be that users become "innoculated" to information. In other words, if a user did not find information interesting upon first exposure, she will not find it worthy of spreading upon subsequent exposures. The real explanation is both simpler and more interesting: in a nutshell, users do not see the exposures, and hence do not respond to them.

Our study of how Twitter users respond to messages from friends demonstrated that users are far more likely to retweet a recent message than an old one, and that the more friends a user follows, the less likely he or she is to retweet an older message (Hodas and Lerman 2012). We invoke the concept of limited attention (Kahneman 1973) to explain why people are less likely to retweet older messages. In order to retweet some information, a user first has to find it by wading through a stream of other messages. Reading tweets, however, requires mental effort, of which people have a limited reserve. Limited attention constrains how deeply into his or her stream the user will browse before getting tired, bored or distracted. Since both Twitter and Digg display messages in reverse chronological order, with the most recent message at the top of the screen, the user is far more likely to see recent messages than older ones that are buried deep in their stream. In addition, the more friends the user follows, the faster a message gets buried, and the less likely the user is to see it.

Limited attention alters how well-connected users, i.e., those who follow many others, respond to information. The exposure response functions shown in Fig. 3a, b have been aggregated over all users. These users form a highly heterogeneous group with a wide range of capabilities and motivations to consume and share information. By conflating together behaviors of different types of individuals, heterogeneity may in fact obscure simpler individual behavior (Vaupel and Yashin 1985). Indeed, when we separate users into more homogeneous subpopulations, a different picture emerges. Figure 3c shows the exposure response function of Twitter users who were separated into subpopulations based on their cognitive load, i.e., total amount of information in their stream. The number of messages in a user's stream is, on average, proportional to the number of friends he or she follows; therefore, we divide users into subpopulations based on the number of friends they follow. A dramatically different picture of exposure response emerges. Now, the response of users within each population increases monotonically with the number of exposures, similar to the ICM. However, unlike ICM, the response of better connected users is suppressed, due to the greater demands placed on their limited attention. The aggregated exposure response in Fig. 3b appears to saturate, because the better connected, and less responsive, users contribute to the right-hand portion of the exposure curve.

This result gives us a better picture of what is going on. Unlike the spread of a virus, which is boosted by hubs, or highly connected people, who create multiple opportunities for the virus to spread, information cascades are suppressed by such users. A cascade stops when it reaches such hubs, because they are less likely to see the message and retweet it, since there are so many other messages competing for their limited attention. Once the response of the highly connected users is encoded into a model of contagion, it leads to smaller cascades.

Discussion

Access to large data sets containing traces of social interactions has created new opportunities to study social behavior. One of the main challenges in analyzing such data is its heterogeneity. People vary greatly in their abilities and motivations, and aggregating over all individuals can sometime lead to erroneous conclusions. This effect, known as "heterogeneity's ruses" (Vaupel and Yashin 1985), was demonstrated above in how people respond to exposures to information in social media. When averaged over all users, it may appear that the more times an individual is exposed to information, the less likely he or she is to spread it. However, when we divide people into more homogeneous populations based on the number of friends they follow, exposure response changes qualitatively. Now, individual response within each population increases monotonically: the more times a user sees information, the more likely he or she is to spread it. However, users with more friends are overall less sensitive than users with few friends. The revised exposure response explains a puzzling observation with which we started this chapter: information in social media does not spread very far. It appears that decreased sensitivity to exposure of highly connected people inhibits social contagion and prevents information from spreading.

The challenge of analysis is to segment the data appropriately. In our analysis, we divided people into classes based on their cognitive load, or the volume of information in their stream. This decision was motivated by our discovery of the role that limited attention plays in the spread of information in social media. Users appear to expend finite effort or time on discovering content. Since users with many active friends have many more messages in their stream to process than users with few friends, the well connected users are less likely to discover, and spread, any specific message. For other problems, other segmentations of data may be desirable. As the amount of social data increases, finer segmentations of data into more homogeneous populations will be statistically feasible, leading to finer-grained models of human behavior.

References

Adamic LA, Adar E (2005) How to search a social network. Soc Netw 27(3):187–203

Anagnostopoulos A, Kumar R, Mahdian M (2008) Influence and correlation in social networks. In: Proceedings of the 14th ACM SIGKDD international conference on knowledge discovery and data mining, Las Vegas. ACM, New York, pp 7–15

Castellano C, Fortunato S, Loreto V (2009) Statistical physics of social dynamics. Rev Mod Phys 81(2):591–646

Clauset A, Shalizi CR, Newman MEJ (2009) Power-law distributions in empirical data. SIAM Rev 51(4):661+

Goel S, Watts DJ, Goldstein DG (2012) The structure of online diffusion networks. In: Proceedings of the 13th ACM conference on electronic commerce (EC 2012), Valencia

Gruhl D, Liben-nowell D (2004) Information diffusion through blogspace. In: Proceedings of the international world wide web conference (WWW), Geneva, pp 491–501

Hethcote HW (2000) The mathematics of infectious diseases. SIAM Rev 42(4):599–653

Hodas N, Lerman K (2012) How limited visibility and divided attention constrain social contagion. In: ASE/IEEE international conference on social computing, Amsterdam

Hogg T, Szabo G (2009) Diversity of user activity and content quality in online communities. In: Proceedings of international conference on weblogs and social media (ICWSM), San Jose

Kahneman D (1973) Attention and effort. Prentice Hall, Englewood Cliffs

Kempe D, Kleinberg J, Tardos É (2003) Maximizing the spread of influence through a social network. In: KDD '03: proceedings of 9th international conference on knowledge discovery and data mining, Washington DC, pp 137–146

Lerman K (2007) Social information processing in social news aggregation. IEEE Intern Comput Spl Issue Soc Search 11(6):16–28

Lerman K, Ghosh R, Surachawala T (2012) Social contagion: an empirical study of information spread on digg and twitter follower graphs. arXiv:1202.3162

Lotan G, Graeff E, Ananny M, Gaffney D, Pearce I, Boyd D (2011) The revolutions were tweeted: information flows during the 2011 Tunisian and Egyptian revolutions. Int J Commun 5:1375–1405

Moreno Y, Pastor-Satorras R, Vespignani A (2002) Epidemic outbreaks in complex heterogeneous networks. Eur Phys J B Condens Matter Complex Syst 26(4):521–529

Newman MEJ (2022) Spread of epidemic disease on networks. Phys Rev E 66(1):016128+

Pickard G, Pan W, Rahwan I, Cebrian M, Crane R, Madan A, Pentland A (2011) Time-critical social mobilization. Science 334(6055):509–512

Romero DM, Meeder B, Kleinberg J (2011) Differences in the mechanics of information DiffusionAcross topics: idioms, political hashtags, and complexcontagion on twitter. In: Proceedings of world wide web conference, Lyon

Salganik MJ, Dodds PS, Watts DJ (2006) Experimental study of inequality and unpredictability in an artificial cultural market. Science 311:854

Satorras RP, Vespignani A (2001) Epidemic spreading in scale-free networks. Phys Rev Lett 86(14):3200–3203

Ver Steeg G, Ghosh R, Lerman K (2011) What stops social epidemics? In: Proceedings of 5th international conference on weblogs and social media, Barcelona

Vaupel JW, Yashin AI (1985) Heterogeneity's ruses: some surprising effects of selection on population dynamics. Am Stat 39(3):176–185

Wang Y, Chakrabarti D, Wang C, Faloutsos C (2003) Epidemic spreading in real networks: an eigenvalue viewpoint. In: IEEE symposium on reliable distributed systems, Florence 0:25+

Wilkinson DM (2008) Strong regularities in online peer production. In: EC'08: Proceedings of 9th conference on electronic commerce, Chicago. ACM, New York, pp 302–309

Wu F, Huberman BA (2007) Novelty and collective attention. Proc Natl Acad Sci 104(45):17599–17601

Wu F, Huberman B, Adamic L, Tyler J (2004) Information flow in social groups. Phys A 337(1): 327–335

Computational Analysis of Collective Behaviors via Agent-Based Modeling

Lilian Weng and Filippo Menczer

Introduction

Agent-based modeling (ABM) is a class of computational analysis tools that is widely used for simulating system dynamics when the system consists of multiple autonomous and interacting individual components—named *agents*. Each agent follows its own decision-making processes according to a set of rules and contextual information from history, other agents, and possibly other environmental settings. The sets of behavioral rules can be identical for all agents (*homogeneous*) or different from agent to agent (*heterogeneous*). For example, in prisoner dilemma games, every agent follows the same strategy to negotiate; in an ecosystem, some agents play the roles of producers while others are consumers.

Eric Bonabeau summarized the key characteristics of ABMs from the perspectives of *capturing emergent phenomena, natural descriptions of systems*, and *flexibility* (Bonabeau 2002). The key point of ABM is to describe a system by setting up the behavioral strategies of its constituent agents. ABMs are often applied to validate individual-level configurations by comparing the patterns that result at the system level with empirical data. Model predictions are typically obtained by computational simulations, in which the outcomes of interactions between agents are repetitively calculated. This approach makes it possible to make predictions that reach beyond those derived by pure mathematical methods, when the model cannot be solved analytically.

L. Weng (✉) • F. Menczer
Center for Complex Networks and Systems Research, School of Informatics and Computing, Indiana University Bloomington, Bloomington, IN, USA
e-mail: weng@umail.iu.edu; lilian.wengweng@gmail.com; fil@indiana.edu; fmenczer@gmail.com

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_61,

Agent-based modeling has now obtained a central role in the study of natural systems. A large body of literature has been developing in the past few years about the internal characteristics of agents, their activities, connectivity, and multi-agent features (Castellano et al. 2009). Biological, ecological, human collaborative systems, and society can be naturally translated into an agent-based framework. ABM techniques are therefore employed in domains that include biology, ecology, cognitive science, epidemiology, and the social sciences. Let us consider a few domains to demonstrate the usage of ABMs in practice.

Social dynamics. The Axelrod model (Axelrod 1997) investigated cultural dynamics by modeling individuals as nodes (agents) in networks in which whether a person interacts with another depends on the similarity between their statuses. Global convergence and the persistence of diversity are two important ingredients explored in this setting. Holme and Newman (2006) proposed a model in which each agent is associated with an opinion. At each time step, agents either change their opinions to match neighbors, or re-wire links toward agents with similar opinions. The model can capture the dual process of social influence via opinion changes and selection via re-wiring of connections. Agent-based models have also been applied to the study of the birth and decline of scientific disciplines (Sun et al. 2013). The evolution of disciplines is guided by social interactions among agents representing scientists. Disciplines emerge from splitting and merging of social communities in a collaboration network. This model is capable of reproducing various empirical observations about the relationships between disciplines, scholars, and publications. Many more models of social dynamics are reviewed in the literature (Castellano et al. 2009).

Network Evolution. In the above examples, agents are connected in a network structure. This is often the case in ABMs, especially in the context of social systems. The edges in the network represent social relationships between pairs of individuals. Some models focus specifically on the local rules that regulate the growth and evolution of the network and lead to its observed global topology. Models have explored many different strategies of how an agent creates connections (Erdös and Rényi 1960; Watts and Strogatz 1998; Barabási and Albert 1999). For example, the phenomenon of linking to well-connected nodes (e.g., people or Web pages) is described by preferential attachment mechanisms (Barabási and Albert 1999). Other ingredients considered in network evolution models include homophily (McPherson et al. 2001) and triadic closure (Granovetter 1973; Shi et al. 2007; Leskovec et al. 2008).

Diffusion. Information and innovation spread on networks, and we can observe the cascades that ensue as agents are infected. The diffusion process is affected by the actions of agents and the underlying network structure. Watts studies the cascade sizes and vulnerability of the system to global cascades using a simple spreading process on random networks (Watts 2002). Pastor-Satorras and Vespignani simulated classical epidemic models on scale-free networks (Barabási and Albert 1999), revealing that infections always survive no matter how small the spreading rate (Pastor-Satorras and Vespignani 2001). Goetz et al. (2009) proposed an agent-based model of blog dynamics, where each agent is associated with mechanisms capturing both the topology and temporal features.

Case Study: Meme Competition for Limited Attention

Here we present a case study using agent-based modeling to understand how limited human attention may constrain competition among ideas (Weng et al. 2012).

The advent of social media (Lazer et al. 2009; Vespignani 2009) has lowered the cost of information production and broadcasting, boosting the potential reach of each idea or *meme* (Dawkins 1989). However, the abundance of information to which we are exposed through online social networks and other socio-technical systems is exceeding our capacity to consume it, increasing the competition among ideas for our finite attention. As a result, the dynamic of information is driven more than ever before by the economy of attention (Simon 1971). In this context one of the most challenging problems is the study of the competition dynamics of ideas, information, knowledge, and rumors (Crane and Sornette 2008; Lerman and Ghosh 2010; Wu and Huberman 2007; Moussaid et al. 2009). Studying limited user attention is motivated by the cognitive limit on the number of stable social relationships that one can sustain (Dunbar 1998; Gonçalves et al. 2011). However, it is hard to disentangle the effects of limited attention from many concurrent factors, such as the underlying network structure (Watts 2002), the activity of users (Asur et al. 2011), homophily (McPherson et al. 2001), and the intrinsic quality of the information (Bakshy et al. 2011).

We can think of the collective actions of many individuals, as they decide which information to propagate through the network, as a computational system that produces a concrete output, namely, a small number of very popular memes while the majority of memes go mostly unnoticed. An agent-based model allows to test micro-level hypotheses about this type of human computation: can certain individual behaviors be responsible for the patterns observed at the collective level?

The design of agent operation strategies in our model in inspired by empirical observations about individual behaviors, outlined in section “Empirical Observations”. We then describe an agent-based toy model of meme diffusion and compare its predictions with the empirical data in section “Model Description”. Finally, in section “Simulation” we show that the social network structure and our finite attention are both key ingredients of the diffusion process, as their removal from the model leads to results inconsistent with the empirical observations.

Empirical Observations

We investigate this problem using a sample of data from *Twitter*, a micro-blogging platform that allows millions of people to broadcast short messages through social connections. Users post short messages (“tweets”), subscribe to (“follow”) people to receive their tweets, and forward (“retweet”) selected posts to their followers. Posts may contain special topic labels (“hashtags”), which we use to identify *memes* operationally. This provides us with a quantitative framework to study the competition for attention in the wild.

Here we outline two empirical findings that motivate both our question and the main assumptions behind our model (Weng et al. 2012). First, the attention of a user is independent from the overall diversity of information discussed in a given period. A user's daily breadth of attention (measured through Shannon entropy) remains roughly constant and bound irrespective of system diversity, which varies greatly day by day. Second, users are more likely to retweet memes about which they posted in the past, suggesting that user memory is an important component for modeling information diffusion.

Model Description

We propose an agent-based model to simulate the retweeting behavior of Twitter users, which explicitly incorporates the above observations about *memory-based user interests* and *limited attention*.

Our basic model assumes a frozen friends/followers network of agents. An agent maintains two time-ordered lists of memes: a *screen* that stores received memes and a *memory* that records posted memes, capturing endogenous interests. Users pay attention to memes in these lists only. At each time step, an agent is randomly selected with uniform probability to transmit a few memes to neighboring agents. The selected agent can generate a new meme or forward some memes from the list and store the posted memes in memory. Neighbors in turn pay attention to a newly received meme by placing it at the top of their screens.

To model limited user attention, both screen and memory have a finite capacity so that memes only survive for a finite amount of time. Figure 1 illustrates details of the model.

Simulation

To evaluate the simulation outcomes of the model, we measured several regularities in the empirical data. The meme lifetime and popularity display long-tailed distributions, meaning that a few memes gain huge popularity.

Our aim is to determine a minimal set of individual-level assumptions necessary to interpret these collective patterns. To evaluate the role of the competition among memes for limited user attention, we simulated variations of the model with *stronger* or *weaker* competition by tuning the length t_w of the time window in which posts are retained in an agent's screen or memory. A shorter time window leads to less attention and thus strong competition, while a longer time window allows for weaker competition. As shown in simulation results, stronger competition fails to reproduce the large observed number of long-lived memes (Fig. 2a), while weaker competition cannot generate extremely popular memes (Fig. 2b).

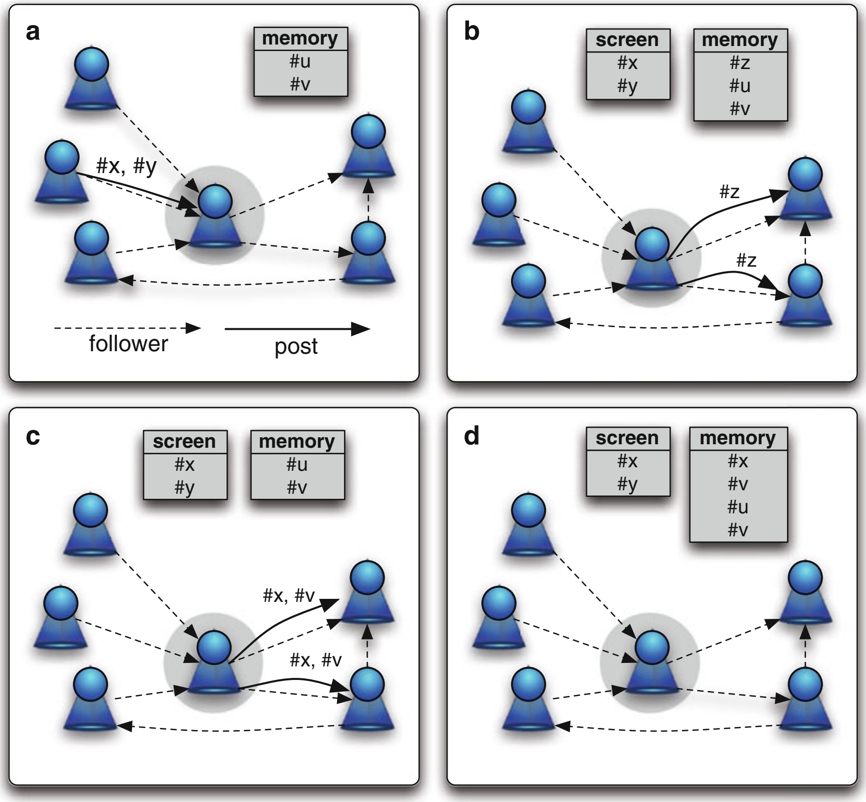


Fig. 1 Illustration of the meme diffusion model (Weng et al. 2012). Tables show the memory and screen of the center user. (**a**) The memory contains memes `#u` and `#v`, that the center user has produced in the past. Memes `#x` and `#y` are received along follower links. (**b**) The received memes, `#x` and `#y`, appear on the screen. With probability p_n, the center user who is selected at this time step posts a new meme `#z` to his followers. (**c**) Otherwise, with probability $1-p_n$, the user scans the screen. Each meme h in the screen catches the user's attention with probability p_r; with probability p_m a random meme from memory is triggered, or h is retweeted with probability $1-p_m$. In the illustrated case, `#x` is retweeted and `#v` is triggered by `#y`. (**d**) All memes posted by the user are stored in memory. Parameters p_n, p_r and p_m are estimated from the empirical data

To gauge the role of underlying network structure in shaping the diffusion process, we simulated the model on both the real social network and a random network. The model is able to reproduce the main empirical features on the social network, while the observed heterogeneity is largely reduced on a random network (Weng et al. 2012). The structure of the network is thus another key ingredient of the system dynamics.

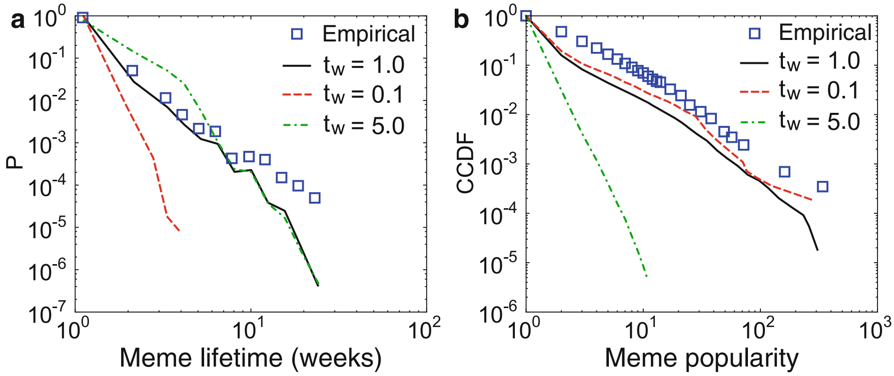


Fig. 2 Evaluation of model by comparison of simulations with empirical distributions of meme lifetime (*left*) and meme popularity (*right*). We simulate the model on the real social network with different levels of competition; posts are removed from screen and memory after t_w time units. We compare the standard model ($t_w = 1$) against versions with less and more competition ($t_w = 5$ and $t_w = 0.1$, respectively)

Conclusion

In this chapter we discussed agent-based modeling as a powerful computational tool to analyze system dynamics, and in particular to study global patterns by simulating individual behaviors and interactions between agents (Bonabeau 2002).

As an example, we presented a case study where agent-based modeling was used to shed light on how memes compete for limited attention on social networks (Weng et al. 2012). The computational approach allows us to demonstrate that, surprisingly, a combination of social network structure and competition for finite user attention is a sufficient condition for the emergence of broad diversity in meme popularity and lifetime, without having to assume exogenous factors.

In general, agent-based models can be very useful in identifying minimal hypotheses consistent with collective patterns generated by human computation systems.

References

Asur S, Huberman BA, Szabo G, Wang C (2011) Trends in social media: persistence and decay. In: Proceedings of international AAAI conference on weblogs and social media, Menlo Park

Axelrod R (1997) The complexity of cooperation: agent-based models of competition and collaboration. Princeton University Press, Princeton

Bakshy E, Mason WA, Hofman JM, Watts DJ (2011) Everyone's an influencer: quantifying influence on twitter. In: Proceedings of ACM international conference on web search and data mining, Hong Kong

Barabási A-L, Albert R (1999) Emergence of scaling in random graphs. Science 286:509–511

Bonabeau E (2002) Agent-based modeling: methods and techniques for simulating human systems. Proc Natl Acad Sci 99(Suppl 3):7280–7287

Castellano C, Fortunato S, Loreto V (2009) Statistical physics of social dynamics. Rev Mod Phys 81(2):591

Crane R, Sornette D (2008) Robust dynamic classes revealed by measuring the response function of a social system. Proc Natl Acad Sci 105(41):15649–15653

Dawkins R (1989) The selfish gene. Oxford University Press, Oxford

Dunbar RIM (1998) The social brain hypothesis. Evol Anthr 6:178–190

Erdös P, Rényi A (1960) On the evolution of random graphs. Magyar Tud. Akad. Mat. Kutató Int. Közl 5:17–61

Goetz M, Leskovec J, McGlohon M, Faloutsos C (2009) Modeling blog dynamics. In: Proceedings of international AAAI conference on weblogs and social media, San Jose

Gonçalves B, Perra N, Vespignani A (2011) Validation of dunbar's number in twitter conversations. PLOS One 6:e22656

Granovetter M (1973) The strength of weak ties. Am J Sociol 78:1360–1380

Holme P, Newman MEJ (2006) Nonequilibrium phase transition in the coevolution of networks and opinions. Phys Rev E 74(5):056108

Lazer D, Pentland A, Adamic L, Aral S, Barabási AL, Brewer D, Christakis N, Contractor N, Fowler J, Gutmann M, Jebara T, King G, Macy M, Roy D, Alstyne MV (2009) Computational social science. Science, 323(5915):721–723

Lerman K, Ghosh R (2010) Information contagion: an empirical study of the spread of news on digg and twitter social networks. In: Proceedings of international AAAI conference on weblogs and social media, Washington DC

Leskovec J, Backstrom L, Kumar R, Tomkins A (2008) Microscopic evolution of social networks. In: Proceedings of SIGKDD international ACM conference on knowledge discovery and data mining, Las Vegas, pp 462–470

McPherson M, Lovin L, Cook J (2001) Birds of a feather: homophily in social networks. Annu Rev Sociol 27(1):415–444

Moussaid M, Helbing D, Theraulaz G (2009) An individual-based model of collective attention. In: Proceedings of European conference on complex systems, Warwick

Pastor-Satorras R, Vespignani A (2001) Epidemic spreading in scale-free networks. Phys Rev Lett 86:3200–3203

Simon H (1971) Designing organizations for an information-rich world. In: Greenberger M (eds) Computers, communication, and the public interest. Johns Hopkins, Baltimore, pp 37–52

Shi X, Adamic LA, Strauss MJ (2007) Networks of strong ties. Physica A 378:3347

Sun X, Kaur J, Milojevic S, Flammini A, Menczer F (2013) Social dynamics of science. Sci Rep 3(1069)

Vespignani A (2009) Predicting the behavior of techno-social systems. Science 325(5939):425–428

Watts DJ (2002) A simple model of global cascades on random networks. Proc Natl Acad Sci 99(9):5766–5771

Watts DJ, Strogatz SH (1998) Collective dynamics of 'small-world' networks. Nature 393: 440–442

Weng L, Flammini A, Vespignani A, Menczer F (2012) Competition among memes in a world with limited attention. Sci Rep 2:335

Wu F, Huberman BA (2007) Novelty and collective attention. Proc Natl Acad Sci 104(45):17599–17601

Yang L, Sun T, Mei Q (2012) We know what @you #tag: does the dual role affect hashtag adoption? In: Proceedings of international ACM world wide web conference, Lyon

Stochastic Modeling of Social Behavior on Digg

Tad Hogg

Introduction

Stochastic models are a general approach to analyzing and predicting the behavior of large interacting systems when details of individual components of the system are either unavailable or of secondary interest compared to overall or aggregate behaviors. Stochastic models treat these details approximately as probabilistic influences on the system behavior. This chapter describes this approach for modeling social behavior of users and content on social media web sites.

Social media sites such as Twitter, Digg, Flickr, Delicious, and YouTube, allow people to post or find interesting content, talk about content they find interesting, and interact with friends and like-minded people. The success of such sites depends on how well they enable people to achieve these goals. This includes the quality of the user experience (Rashid et al. 2006) and the extent to which the collective action of the users focuses attention on interesting content.

A key question is how the performance of social media relates to web site designs, particularly the choices of content directed to users' attention. One approach to this question uses machine learning and data mining. In this approach, statistical regression-based methods classify large data sets according to features in the data. Such methods can identify correlations among sets of features or behaviors, which are then used to predict outcomes in new cases. However, these approaches are limited in their ability to identify causal mechanisms. Another approach uses experiments, especially with multiple randomly-selected groups (Salganik et al. 2006). Comparing observed performances of different designs is a powerful approach in identifying causal relationships. Unfortunately, such experiments are seldom practical in the social media domain due to the need to recruit large numbers of users, the

T. Hogg (✉)
Institute for Molecular Manufacturing, Palo Alto, CA, USA
e-mail: tad@stanfordalumni.org

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_62,

challenge of creating realistic social scenarios for those users and the difficulty of running multiple long-term experiments.

In contrast to these approaches, stochastic models can identify mechanisms relating the design of social media sites to their collective behavior (Lerman 2007; Hogg and Szabo 2009; Iribarren and Moro 2009; Castellano et al. 2009; Hogg and Lerman 2012). In general, a stochastic model consists of a set of states and probabilistic transitions among these states. For application to social media, these states and transitions define a set of key features of the site's design, its users and the content available to those users on the site. By comparing predictions of the models to observed user behavior, such mechanistic models can indicate ways to improve social media services by identifying key mechanisms leading to successful outcomes. The rest of this chapter describes this modeling approach and applies it to Digg, a news aggregator web site.

Stochastic Models of Social Dynamics

Descriptions of social media typically focus on aggregate behavior of the large numbers of users described by *average* quantities. These include average rate at which users contribute and rate content, and the rate they form links to other users. Stochastic models of social media are similar to such models used in demographics, epidemiology (Ellner and Guckenheimer 2006) and macroeconomics, where the focus is not to reproduce the results of a single observation, but rather to describe the typical behaviors and relations among aggregate quantities, such as vaccination policy and fraction of infected population or interest rates and employment.

Stochastic models represent an individual entity, whether a user or contributed content, as a stochastic process with a few states. This abstraction captures much of the complexity seen on social media sites by viewing individual's actions as inducing probabilistic transitions between states. For simplicity, we focus on processes that obey the Markov property, namely, a user whose future state depends only on her present state and the input she receives from the site and other users. A Markov process is succinctly captured by a *state diagram* showing the possible states of the user and conditions for transition between those states. This approach is similar to compartmental models in biology (Ellner and Guckenheimer 2006). For instance, in epidemiology such models track the progress of a disease as shifting individuals between states, or compartments, such as susceptible and infected.

A key requirement for designing stochastic models is to ensure the state captures enough of the variation in individual behavior to give a useful description of aggregate system properties. A sufficient condition for the usefulness of focusing on average behavior is that variations around the average are relatively small. In many stochastic models, variations are indeed small due to many independent interactions among the components and the short tails of the distributions of these component behaviors. Ensuring the models are useful is particularly challenging when individual activity follows a long-tail distribution (Newman 2003), such as seen in some epidemics (Lloyd-Smith et al. 2005) and commonly found in social media

(Wilkinson 2008). In these cases, typical behaviors differ significantly from the average and we have no guarantee that the averaged approximation is adequate (Sornette 2004). Instead we must test its accuracy for particular aggregate behaviors by comparing model predictions with observations of actual behavior. As described below, including user link information as part of the state accounts for enough of this variation to provide reasonable accuracy. In particular, including this information significantly improves predictions compared to direct extrapolation of voting rates without accounting for the properties of the web site user interface. More elaborate versions of the stochastic approach give improved approximations when variations are not small, particularly due to correlated interactions (Opper and Saad 2001) or large individual heterogeneity (Moreno et al. 2002).

In summary, the general stochastic modeling framework requires only specifying the aggregate states of interest and how individual user behaviors create transitions among these states. The modeling approach is best suited to cases where the users' decisions are mainly determined by a few characteristics of the user and the information they have about the system. This is a reasonable approximation for social media sites that provide relatively few ways for users to find content and learn about other users, i.e., only via a small number of options provided by the web site's graphical interface. These system states and transitions lead to equations describing the average rate of transitions among the states. Solutions to these equations then estimate how aggregate behavior varies in time and depends on the characteristics of the users involved.

A Model of Digg

With over six million registered users, the social news aggregator Digg was an early and popular crowd-sourced news portal on the Web. Digg allows users to submit and rate news stories by voting on, or 'digging', them. Each day has over 16,000 new submissions. Every day Digg promotes about a hundred stories to the front page. The choice of stories to promote accounts for the reaction of users to newly submitted stories. Thus, Digg's front page is emergent, created by the collective decision of its many users.

A newly submitted story goes on the *upcoming* stories list, where it remains for a period of time, typically 24 hours, or until it is promoted to the front page, whichever comes first. The default view shows newly submitted stories as a chronologically ordered list, with the most recently submitted story at the top of the list, 15 stories to a page. To see older stories, a user must navigate to page 2, 3, etc. of the upcoming stories list. Promoted stories (Digg calls them 'popular') are also displayed as a chronologically ordered list on the *front pages*, 15 stories to a page, with the most recently promoted story at the top of the list. To see older promoted stories, user must navigate to page 2, 3, etc. of the front page. Users vote for the stories they like.

Digg allows users to designate friends and track their activities. The friend relationship in Digg is asymmetric, as illustrated in Fig. 1. For example, in the network

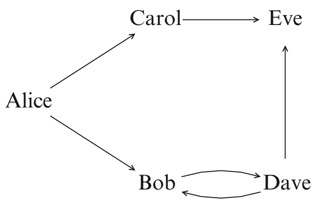


Fig. 1 A friends network. The *arrows* indicate direction of information flow about a user's activities. For instance, Eve is a *fan* of both Carol and Dave, who are Eve's designated *friends*. Bob and Dave are mutual friends and hence fans of each other

shown in the figure, Bob and Carol follow Alice, i.e., have designated Alice as a friend. We say Bob and Carol are Alice's *fans*. In the figure, Bob and Dave follow each other, i.e., are each other's fan. In the context of crowd-sourced news, typically a user will choose to follow another person who posts content the user found interesting. In effect, each user can follow a personal set of news "editors" whose judgement the user finds appealing.

A newly submitted story is visible in the upcoming stories list, as well as to the submitter's fans through the friends interface. With each vote, a story becomes visible to the voter's fans through the friends interface, which shows the newly submitted stories that user's friends voted for. For example, suppose Alice submits a story. Her fans, Bob and Carol, can then see the story in their friends interface. We call such users *submitter's fans* with respect to that story. The remaining users do not immediately see the story in their friends interface: they are *non-fans* of the submitter. As users vote for the story, their fans gain the ability to see the story via the friends interface. Such users who are not also fans of the submitter are called *other fans*. Thus, if Carol votes for the story, her fan, Eve, now has the story in her friends interface so switches from the non-fan to the other fan category.

Stories submitted or voted on by a user's friends are highlighted in the friend's interface, so those stories are more readily seen during that user's next visit to Digg than other stories. Thus fans are more likely to see, and hence have an opportunity to vote on, stories than non-fan users who visit Digg. We account for this design of the Digg user interface in the choice of states for the stochastic model. Specifically, Fig. 2 shows the state diagram for user behaviors with respect to a single story.

A story's submitter provides its first vote, so the submitter is in the *vote* state. Other users start in either the *submitter's fans* or *non-fans* state. A user visiting Digg may view the story, either via the friends interface (if a fan of the submitter or prior voter) or via another Digg page such as upcoming or front page stories. If so, the user moves to the *view* state. Users in the *view* state may decide to vote on the story, moving to the *vote* state. In that case, any other user who has not voted for the story and is a fan of the new voter but not a fan of any prior voter, moves from the *non-fans* to *other fans* state. The stochastic model specifies transition probabilities for users moving between these states, based on how users navigate web sites and the diversity of how stories appeal to the user community (Hogg and Lerman 2012).

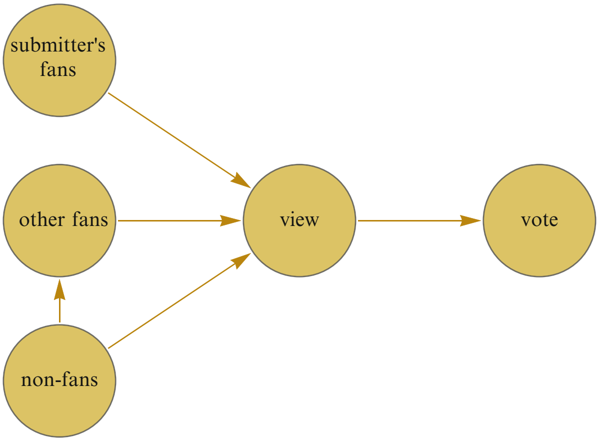


Fig. 2 State diagram of user behavior for a single story. The probability a user visiting Digg views the story depends on the user's position in the network of followers for the submitter and prior voters. Although not shown explicitly in this diagram, the transition from viewing a story to voting on it also depends on this property of the users

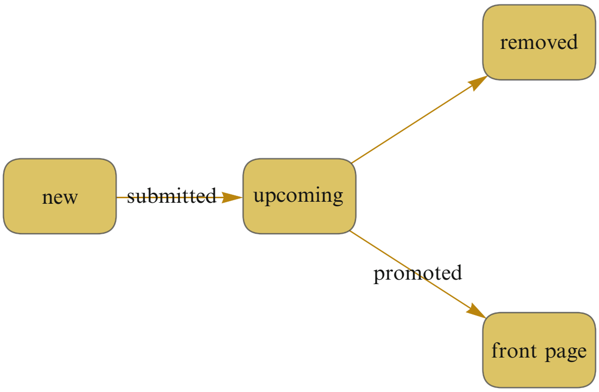


Fig. 3 State diagram for a story. A new story discovered by a Digg user is submitted to the upcoming list. If the story is popular enough with other users, it is promoted to the front page. A story not promoted after sufficient time (usually within a day) is removed

As users vote, the story transitions through its own set of states, shown in Fig. 3. In addition to these states, showing its position in Digg, the story accumulates votes from users. The number of votes determines whether the story is promoted and how it is shown to users.

We evaluated this model using voting activity and snapshots of the social network from June 2006 to June 2009. We used this data to estimate parameters describing the transition rates for the state diagrams. With these parameter values, we can solve the stochastic model to produce predicted behavior. We focus on how stories accumulate votes, which is a key property of crowd-sourced news.

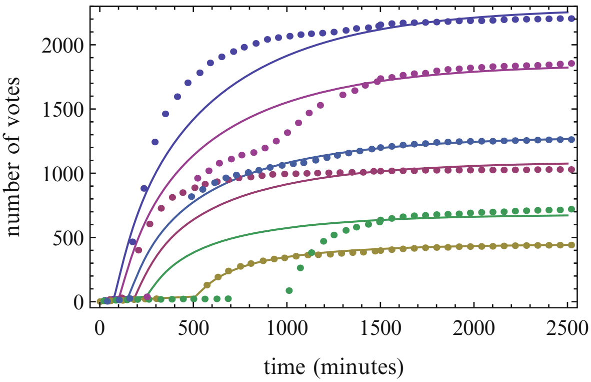


Fig. 4 Evolution of the number of votes received by six stories compared with model solution

Figure 4 shows the behavior of six stories that were promoted to the front page and the corresponding solution from the model. In the model, a story has two characteristics: the number of fans of the story's submitter and the probability a user seeing the story will vote for it (the story's *interestingness* or *quality*), which depends on whether the user is a fan of the submitter or one of the other prior voters. Overall there is qualitative agreement between the data and the model, indicating that the features of the Digg user interface we considered can explain the patterns of collective voting. Specifically, the model reproduces three generic behaviors of Digg stories: (1) slow initial growth in votes of upcoming stories; (2) more interesting stories are promoted to the front page (inflection point in the curve) faster and receive more votes than less interesting stories; (3) however, as first described in Lerman (2007), better connected users are more successful in getting less interesting stories promoted to the front page than poorly-connected users.

These observations highlight a benefit of the stochastic approach: identifying simple models of user behavior that are sufficient to produce the aggregate properties of interest. Solutions to the model indicate how collective outcomes depend on the user population: in this case, not only the appeal of a story to the user community, as intended for crowd-sourcing, but also the disproportionate effect of highly-connected users. Thus while the stochastic model primarily describes typical story behavior, we see it gives a reasonable match to the actual vote history of individual stories. Nevertheless, there are some cases where individual stories differ considerably from the model, particularly where an early voter happens to have an exceptionally large number of fans, thereby increasing the story's visibility to other users far more than the average value.

Predicting Story Popularity

By separating the impact of story quality and social influence on the popularity of stories on Digg, a stochastic model of social dynamics supports two novel applications: (1) estimating story quality for the user community from the evolution of its

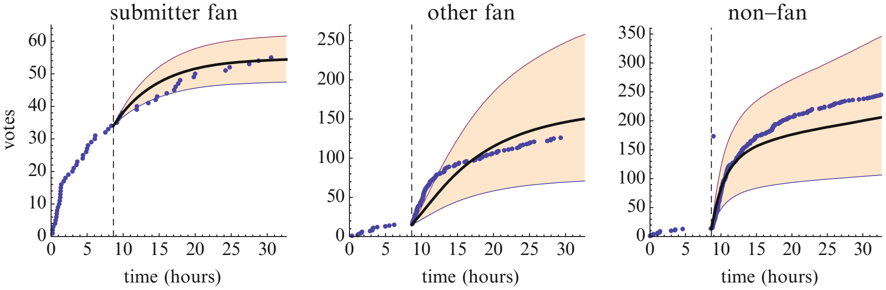


Fig. 5 Predictions compared to actual votes (*dots*) for each type of user for one story. The figure shows predictions made at promotion (*black line*) and the growth in the 95 % confidence interval of the prediction up to 24 h after promotion. The *dashed vertical line* shows the story's promotion time

observed popularity, and (2) predicting its eventual popularity based on users' early reactions to the story. For the first application, by estimating story quality from the evolution of its popularity, the model identifies the distribution of quality among the stories (Hogg and Lerman 2012). This distribution is not directly apparent from the data itself which confounds effects of changing visibility of the stories with their appeal to the users.

Predicting popularity in social media from intrinsic properties of newly submitted content is difficult (Salganik et al. 2006). However, users' early reactions provide some measure of predictability (Hogg and Szabo 2009; Kaltenbrunner et al. 2007; Lerman and Galstyan 2008; Szabo and Huberman 2010). To predict how popular a story will become, we use the early votes, including those cast before the story is promoted, to estimate how interesting it is to the user community. With this estimate, the model determines, on average, the story's subsequent evolution. These predictions are for expected values and cannot account for the large variation due, for example, to a subsequent vote by a highly connected user which leads to a much larger number of users seeing the story and, consequently, to a larger number of votes than expected.

We can improve predictions from early votes by using the observed distributions of quality (Hogg and Lerman 2012) as a prior probability for story quality. Using Bayes theorem to combine this prior distribution with the observations of early votes for a story improves predictions, especially during the period shortly after a story's submission when it has not yet received many votes. Figure 5 shows predictions made for one story based on the votes it received at the time it was promoted. The model's predictions for subsequent votes approximately reproduce the observed votes for this story for each class of user considered in this model.

Examining predictions on other stories (Hogg and Lerman 2012) shows that errors generally decrease when predictions are made later, as would be expected since later predictions are based on reactions from more users. Of more interest is the difference among the type of votes, particularly for votes from other fans. Early votes are mainly from submitter's fans and non-fans, so the ability to predict differences in behavior for those groups based on early votes could be useful in quickly distinguishing stories likely to be of broad or niche interest to the user community.

Overall, the model reasonably predicts votes from submitter's fans and non-fans, but is much less accurate for votes from other fans. One reason for this difference is the relatively small number of other fan votes while a story is upcoming. Specifically, the number of other fans starts at zero. Only a vote by a non-fan can increase the number of other fans, and upcoming stories have low visibility to non-fan voters. Even after a number of users become other fans due to prior votes for the story by other users, it takes some time for those other fans to return to Digg. Thus there are relatively few early other fan votes, leading to poor estimates for the story's appeal to those users. Moreover, the relatively small number of other fans means a single early voter with many fans can significantly change the number of other fans away from its average value used in the model. These factors lead to the relatively large errors in predicting the other fan votes. As a direction for future work, this observation suggests predictions would benefit from including measurements of the social network of the voters to determine the actual number of other fans at the time of prediction rather than using an estimate based on the model.

For crowd-sourcing web sites, predicting whether a story will attract a large number of votes is often more important than predicting the precise number of votes it will receive. Such predictions form the basis of using crowd sourcing to select a subset of submitted content to highlight (Lerman and Galstyan 2008). As an example of this application, we can use the model to predict whether a story will receive more than the median number of votes of each type of user based on votes received up to various times. Depending on when the prediction is made, classification error for submitter fans and non-fans is around 10 %. However, for other fans, the error is around 40 %, again reflecting the challenge of estimating story interestingness from the small number of votes by such users (Hogg and Lerman 2012).

As a more sophisticated application of prediction, a web site could use prediction of likely popularity as part of decisions of what content to highlight to users and combine with sponsored content such as ads. In this case, it would be useful not only to have the prediction but also an indication of how well early user behavior allows for accurate predictions, i.e., an estimate of the confidence of the model's prediction for individual stories. The stochastic approach can provide such estimates. For a given set of model parameter values, prediction variability comes from differences in estimated quality of the story to various types of users (Hogg and Lerman 2012). If the values are poorly determined, predictions will be unreliable. We estimate the values from early votes using maximum likelihood estimation, i.e., finding the values most likely to produce the observed early reaction to the story according to the model. Examining how sharply peaked the likelihood is around this maximum indicates how tightly the model constrains the estimated values. In particular, this gives not only the most likely values for the story's appeal, but also confidence intervals for those values for each type of user. Solving the model using the most likely values gives the prediction, while solving the model using the extent of the confidence intervals provides an estimated range for the prediction accuracy.

As one example, Fig. 5 shows how confidence intervals grow with time subsequent to the time the prediction was made. In this case, the predictions are made when a story is promoted, based on the votes the story has at that time. Confidence intervals grow as the prediction is projected further into the future and quantify the decreasing reliability of predictions over longer time intervals. Subsequently, as the story receives more votes the predictions can be updated based on that new information. Since they are based on more information about the story, predictions made at later times generally have smaller confidence intervals. Thus the confidence intervals, which are computed from the vote information available at the time of prediction, indicate how well the model can predict votes over various time intervals. However, the confidence intervals do not account for all the sources of error. For instance, in some cases the error is considerably larger than the confidence interval, and only about a third of the actual votes 24 h after promotion are within the confidence intervals, whereas we would expect about 95 % of the stories to be within the intervals (Hogg and Lerman 2012). Additional variation could be due, for instance, to votes by exceptionally well-connected users that significantly increase the story's visibility compared to the average value assumed with the model. Large prediction errors can also arise from poor estimates of the sizes of the groups of each type of user at the time of prediction, which is particularly an issue for the other fans, as discussed above.

Conclusion

The stochastic framework provides a simple modeling approach to incorporate details of user behaviors based on information available on the web site. Creating such models is a straight-forward translation of the web site's user interface into a set of states for users and content, as illustrated for Digg. Solving the model provides insights into how aggregate behavior arises from the interaction between user behaviors and web site design. In particular, user models can help distinguish aggregate behaviors that arise from intrinsic properties of the stories (e.g., their interestingness to the user population) from behavior due to the information the web site provides, such as ratings of other users and how stories are placed in the site. In addition to explaining empirically observed phenomena (e.g., it is easier for submitters with more fans to get a story promoted to the front page, even when the story is less interesting), stochastic models also have predictive power.

Social media has transformed the Web into a participatory medium and potentially a powerful new computational platform. As people interact online, their collective activity and the structure of the Web itself are becoming increasingly complex and dynamic. The stochastic framework described here models emergent behaviors in social media. This framework represents individual dynamic entities as stochastic processes and allows the modeler to relate aggregate behaviors to these descriptions. As illustrated for the social news site Digg, these models indicate how people respond to design choices for the web site's user interface and allow predicting how the user community will react to contributed content based on the behavior of the first users to see the new content.

References

Castellano C, Fortunato S, Loreto V (2009) Statistical physics of social dynamics. Rev Mod Phys 81(2):591–646

Ellner SP, Guckenheimer J (2006) Dynamic models in biology. Princeton University Press, Princeton

Hogg T, Lerman K (2012) Social dynamics of Digg. EPJ Data Sci 1:5

Hogg T, Szabo G (2009) Diversity of user activity and content quality in online communities. In: Proceedings of the third international conference on weblogs and social media (ICWSM2009), San Jose. AAAI, pp 58–65

Iribarren JL, Moro E (2009) Impact of human activity patterns on the dynamics of information diffusion. Phys Rev Lett 103:038702

Kaltenbrunner A, Gomez V, Lopez V (2007) Description and prediction of slashdot activity. In: Proceedings of 5th Latin American web congress (LA-WEB 2007), Santiago

Lerman K (2007) Social information processing in social news aggregation. IEEE Intern Comput Spl Issue Soc Search 11(6):16–28

Lerman K (2007) Social networks and social information filtering on Digg. In: Proceedings of international conference on weblogs and social media (ICWSM-07), Boulder

Lerman K, Galstyan A (2008) Analysis of social voting patterns on Digg. In: Proceedings 1st ACM SIGCOMM workshop on online social networks, Seattle

Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM (2005) Superspreading and the effect of individual variation on disease emergence. Nature 438:355–359

Moreno Y, Pastor-Satorras R, Vespignani A (2002) Epidemic outbreaks in complex heterogeneous networks. Eur Phys J B Condens Matter Complex Syst 26(4):521–529

Newman MEJ (2003) The structure and function of complex networks. SIAM Rev 45(2): 167–256

Opper M, Saad D (eds) (2001) Advanced mean field methods: theory and practice. MIT, Cambridge

Rashid AM, Ling K, Tassone RD, Resnick P, Kraut R, Riedl J (2006) Motivating participation by displaying the value of contribution. In: Proceedings of the ACM conference on human-factors in computing systems (CHI 2006), Atlanta. ACM, New York, pp 955–958

Salganik MJ, Dodds PS, Watts DJ (2006) Experimental study of inequality and unpredictability in an artificial cultural market. Science 311:854–856

Sornette D (2004) Critical phenomena in natural sciences: chaos, fractals, selforganization and disorder: concepts and tools, 2nd edn. Springer, Berlin

Szabo G, Huberman BA (2010) Predicting the popularity of online content. Commun ACM 53(8):80–88

Wilkinson DM (2008) Strong regularities in online peer production. In: Proceedings of the 2008 ACM conference on E-Commerce, Chicago, pp 302–309

Activation Cascades in Structured Populations

Aram Galstyan

Introduction

One of the most interesting properties of real networks is modularity, i.e., the tendency of nodes to partition themselves into *communities* (Girvan and Newman 2002; Newman 2006). Loosely speaking, a community is a group of nodes for which the density of links within a group is higher than across the groups. Those communities might represent groups of individuals with shared interests in online social networks, topic-specific research communities in co-authorship networks, and so on.

Much recent research has focused on methods for detecting and analyzing community structure in networks (for a recent review of existing approaches see Fortunato (2010) and references therein). However, the dynamical properties of modular and correlated networks have started to attract attention only recently (Arenas et al. 2006; Galstyan and Cohen 2007; Gleeson 2008; Melnik et al. 2012; Payne et al. 2009).

Understanding the impact of group structure on network dynamics is important for social computing applications. Consider, for instance, word-of-mouth (or viral) marketing of a new product. If different consumer groups have different rating criteria for the product, or different reaction to marketing strategies, then one needs to model how influence propagates within and across communities to predict whether the product will be a hit, or confined to a small subset of consumers. Similarly, understanding how a political message propagates within and across partisan constituencies could be very important for designing effective political campaigns.

Here we report our analysis of a simple dynamical process in networks with community structure. We consider a threshold-based dynamical process on networks (Watts 2004) where the nodes can be in two states, *passive* or *active*. The actual

A. Galstyan (✉)
USC Information Sciences Institute, Marina del Rey, CA, USA
e-mail: galstyan@isi.edu

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_63, © Springer Science+Business Media New York 2013

meaning of those states is application-dependent (e.g., in viral marketing activation might correspond to purchasing a product). Starting from initial configuration with only a handful of nodes in the active state, we consider a discrete-time dynamics where at each time step, a passive node becomes active if the number of his active neighbors exceeds some predefined threshold. This process is iterated until none of the nodes change his state.

We study the dynamical properties of the above model for networks composed of two loosely coupled communities. Our main observation is that if the initially active nodes (*seeds*) are contained in one of the communities, then under certain conditions the cascading process has a two-tiered structure, that is, the peaks of the activation dynamics in each community are well separated in time. Furthermore, depending on the link density between and across the groups, and the fraction of seed nodes, the activation might either die out, spread to one of the groups, or spread to both groups. In particular, for a given network, there is a *critical* fraction of the seed nodes, so that below this critical threshold the activation process is contained, while above the threshold the activation spreads throughout the network. This critical behavior has implications for problems such as influence maximization, where one intends to select initial target nodes so that the size of the resulting cascade is maximal. In particular, we demonstrate that simple target selection strategies that neglect the network community structure can yield overly sub-optimal results.

The rest of the paper is organized as follows: In the next section we formally introduce the cascade model and present its mean-field analysis for networks with structural heterogeneity-random graphs consisting of two loosely coupled subgraphs (communities). We then elaborate on the implications of the analysis on the influence-maximization problem, and present experiments on synthetically generated networks. We conclude the paper by discussing our main results in the context of the existing literature and pointing out open research questions.

Mean Field Analysis of the Activation Dynamics

Cascade Model

There are a number of approaches for modeling activation cascades on networks (see Borge-Holthoefer et al. (2013) for a recent survey). In this paper we use the Linear Threshold Model (Granovetter 1978) (LTM), which, starting from a set of initially active nodes, propagates the activation through a threshold-based mechanism. Let $\mathcal{N}_i$ be the set of active neighbors of node i. Then the node i is activated whenever

$$\sum_{j \in \mathcal{N}_i} w_{ij} \geq \theta_i \tag{1}$$

Here w_{ij} is the normalized weight of the link between the nodes i and j, $\sum_j w_{ij} = 1$, and θ_i is the activation threshold for the node i. Usually, θ-s are assumed to be random variables from some distribution, reflecting the uncertainty about individuals.

To simplify the analysis, here we use a modified version of the linear threshold model, where the threshold condition is applied not to the fraction of active neighbors, but their number. We stress, however, that our main results are valid for the fractional threshold model as well, provided that it demonstrates a phase-transition behavior.[1]

Let us associate a binary state variable with each node, $s_i \in \{0, 1\}$, where the states 0 and 1 correspond to passive and active states, respectively. Then the dynamics of the process is characterized by the following set of equations:

$$s_i(t+1) = \Theta\left[\sum_j W_{ij} s_j(t) - h_i\right] \quad (2)$$

where $\Theta(x)$ is the step function,[2] h_i is the activation threshold for the ith node, and W is the adjacency matrix of the network: For the sake of simplicity, we consider the case of an unweighted graph, so that the entries in the adjacency matrix are either 0 or 1. Equation 2 is iterated until steady state is achieved, that is, none of the nodes changes its state upon further iteration.

We have previously developed a mean-field theory of activation dynamics on modular graphs (Galstyan and Cohen 2007) in the case when the thresholds were the same for all the nodes, $h_i = H$. Here we generalize the framework to the case when nodes have different activation thresholds, drawn from a specified distribution P_h.

Activation Cascades in Single-Community Networks

Let us first focus on a single-community network, and consider a graph composed of N nodes, where each of the $N(N-1)/2$ edges is present with probability p. In the limit of large N, the resulting degree distribution of nodes in this network is the Poisson distribution with a mean $z = pN$.

Let $\rho_h(t)$ be the fraction of active nodes with activation threshold h at time t. Initially, it equals to the fraction of nodes that have been targeted, $\rho_h(t=0) = \rho_{h,0}$. We assume that probability for a node to be selected as a seed is independent of its activation threshold, so that $\rho_{h,0} = \rho_0$. The total fraction of active nodes is $\rho(t) = \sum_h \rho_h(t) P_h$. Further, let $P(k;t)$ be the probability that a randomly chosen node is connected with exactly k active nodes at time t. It is easy to see that at time $t = 0$, k is given by Poisson distribution with a mean $p N_0 \equiv z \rho_0$.[3] To study the dynamics of the process, we need to estimate these distributions for later times. To do so, here we use

[1] Furthermore, we would like to argue that the modified model with integer threshold also seems more plausible from the social-choice standpoint. Indeed, it is hard to imagine that, when trying to make a decision based on our friends' recommendations, we normalize the number of recommendations by the total number of our friends.

[2] $\mathbf{\Theta}(x) = 1$ if $x \geq 0$, and $\mathbf{\Theta}(x) = 0$ otherwise.

[3] Strictly speaking, $P(k;t)$ is given by a binomial distribution $B(N_0, p)$. However, in the limit of large network sizes considered here, we approximate the binomial distribution by the Poisson distribution as it simplifies the analysis.

the so called *annealed approximation*, which has been used to study the dynamical properties of random boolean networks (Derrida and Pomeau 1986; Derrida and Stauffer 1986; Rohlf and Bornholdt 2002). Within the annealed approximation, one averages over the disorder by "rewiring" the network after each iteration. Since during the rewiring process all edges are equally likely, it is easy to see that $P(k;t)$ is still given by a Poisson distribution: However, the mean now depends on the fraction of active nodes $\rho(t) = \sum_h \rho_h(t) P_h$:

$$P(k;t) = e^{z\rho(t)} \frac{[z\rho(t)]^k}{k!} \tag{3}$$

Consider all the nodes with an activation threshold h. On the first step of the cascading process, the fraction of active nodes among those is given by $\sum_{k \geq h} P(k; t = 0)$. In later iterations, the fraction of active nodes can be calculated as follows. There are $N_h(1-\rho_h(t))$ passive nodes at time t, and each one of these nodes will probability $\sum_{k \geq h} P(k;t)$. Also, due to the rewiring, some of the $N_h(\rho_h(t)-\rho_{h,0})$ active nodes will switch to passive state with the rate $\sum_{k<h} P(k;t)$. We note that *the initially targeted nodes are not allowed to de-activate*. Combining these together, and using the normalization condition $\sum_{k=0}^{\infty} P(k;t) = 1$, we obtain the following set of equations

$$\rho_h(t+1) = 1 - (1-\rho_{h,0}) Q(h; z\rho(t)) \tag{4}$$

where $Q(h,x) = \sum_{k<h} e^{-x} x^k / k!$ is the regularized gamma function.

To get the total fraction of activated nodes, we multiply Eq. 10 by P_h and sum over h, which yields

$$\rho(t+1) = 1 - (1-\rho_0) \sum_{h=0}^{\infty} P_h Q(h; z\rho(t)) \tag{5}$$

Equation 5 describes the dynamics of the cascading process in the network. For a fixed connectivity z, the dynamics depends on the fraction of initially targeted nodes, ρ_0, as well as on the threshold distribution function P_h. Let us elaborate on the latter dependence in more details. First of all, we assume that $P_0 = 0$, i.e., there are no nodes that activate spontaneously, aside from the initially targeted nodes. Furthermore, simple inspection shows that the dynamical properties of the model depend on the fraction of nodes with threshold $h = 1$, P_1. We call these nodes *vulnerable* since they will activate whenever one of their neighbors is active. Clearly, if the fraction of the vulnerable nodes is sufficiently large, a single node might trigger a global cascade throughout the network. Without going into much mathematical details, we simply observe that such a global cascade will happen whenever the vulnerable nodes form a giant connected component, which, for the random Erdos–Renyi graphs translates into $P_1 z \sim 1$. In this paper we focus on the case when P_1 is either zero, or sufficiently small, $P_1 \ll 1/z$, so that for a network of size N, the number of nodes required to cause a global cascade must be of order $O(N)$.

For the latter case, the analysis of Eq. 5 yields the following observation: For a given connectivity z, there is a critical fraction ρ_c such that for $\rho_0 < \rho_c$ the activation

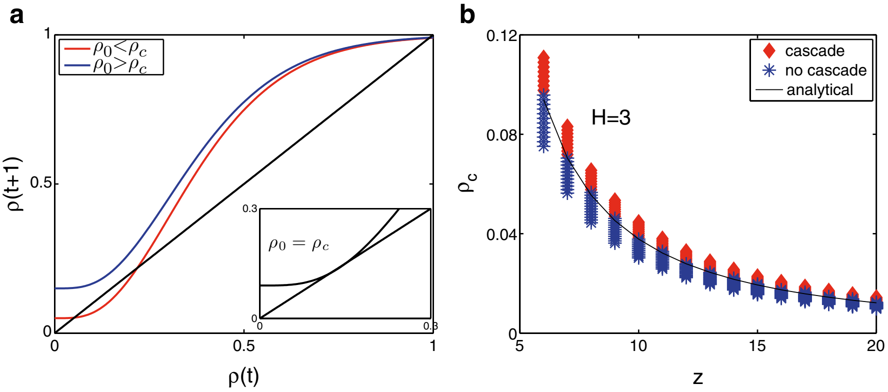


Fig. 1 (**a**) Graphical representation of Eq. 5 for below-critical (*red*) and above-critical (*blue*) values of ρ_0. The *inset* shows the equation (in the vicinity of the solution) for the critical value $\rho = \rho_c$. (**b**) The critical connectivity plotted against the fraction of seed nodes for the threshold parameter $H = 3$. The *solid line* shows the phase boundary obtained analytically

process is localized, while for $\rho_0 > \rho_c$ activation spreads to all the nodes in the network. This is schematically illustrated in Fig. 1a, where we plot $\rho(t+1)$ against $\rho(t)$. Note that the intersections characterize the steady state of the dynamics, or in other words, the fraction of activated nodes at the end of the cascading process. Note, that there is always one intersection around $\rho(t+1) = \rho(t) \approx 1$. For smaller ρ_0, however, there is another stable fixed point. One can calculate the critical density by requiring that the left hand side of Eq. 5 be tangential to the right hand side, as indicated in the inset of Fig. 1a. This yields the following expression for the critical density:

$$\rho_c = 1 - \left[z e^{-x_0} \sum_{k=0}^{\infty} P_{k+1} \frac{x_0^{k-1}}{(k-1)!} \right]^{-1} \tag{6}$$

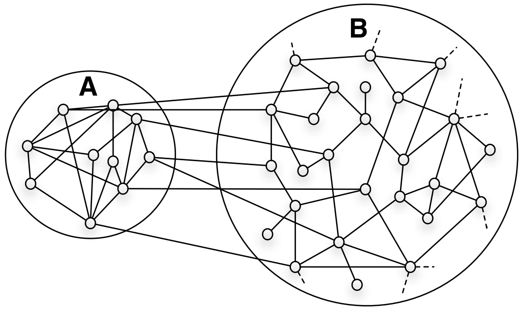
where x_0 satisfies the following equation:

$$1 - \frac{x_0}{z} = \frac{\sum_{k=0}^{\infty} P_{k+1} \frac{x_0^{k-1}}{(k-1)!}}{\sum_{k=0}^{\infty} (1 - D_k) \frac{x_0^{k-1}}{(k-1)!}} \tag{7}$$

Here $D_k = \sum_i \leq_k P_i$ is the cumulative distribution function for the activation thresholds.

In Fig. 1b we compare the analytical prediction with simulation results for the case when all the activation thresholds are set to $h = 3$. The simulations were done for a graph with 5×10^4 nodes, and for 50 random trials. Each pair (z, ρ_c) was considered to be above the critical line if a global cascade was observed in the majority of trials for that parameters. One can see that the agreement of analytical prediction and the simulation results are excellent.

Fig. 2 Schematic illustration of a bi-community network



Activation Cascades in Bi-community Networks

Now let us focus on heterogeneous networks where not all the links have the same probability. In particular, here we focus on networks that are composed of a relatively small, tight community that is connected with a larger population of nodes, as schematically depicted in Fig. 2. We call the nodes in the first and the second community as A and B, respectively. Note, that the group B itself might be comprised of a larger number of sub-communities. This is the case for the networks that we use in our experiments. From the analysis perspective, however, we assume that the links are homogeneously distributed within each community. In other words, we assume that each community is represented by a random Erdos–Renyi graph of N_a and N_b nodes, respectively, and the interaction between two communities are introduced by linking each of the N_aN_b with a uniform probability. Such a network is fully characterized by within-group connectivities z_a, z_b, and the across the group connectivities $z_{ab} \equiv (N_b/N_a)z_{ba}$.

For the sake of simplicity, let us assume that the cascading dynamics in group A is not affected by the nodes in group B. This is a reasonable assumption as long as there are not that many active B-nodes, which is usually the case at the beginning of the cascading process. Thus, the activation dynamics of A nodes is still governed by the Eq. 5. For the B nodes, the activation dynamics is given by a similar equation, with the only difference that it is affected by active A nodes:

$$\rho_b(t+1) = 1 - (1-\rho_{b,0})\sum_{h=0}^{\infty} P_h Q(h; z_b\rho_b(t) + z_{ba}\rho_a(t)) \tag{8}$$

The steady state fraction of active B nodes satisfies the following equation:

$$\rho_b^s = 1 - (1-\rho_{b,0})\sum_{h=0}^{\infty} P_h Q(h; z_b\rho_b^s + z_{ba}\rho_a^s) \tag{9}$$

where $\rho_a{}^s$ is the steady state fraction of active A nodes. Thus, the presence of the active A nodes facilitates the activation of B nodes, and the effect depends on the across the group connectivity z_{ba}. Specifically, if z_{ba} is very small, then the activation dynamics in group B can be described as in the previous section. Namely, there

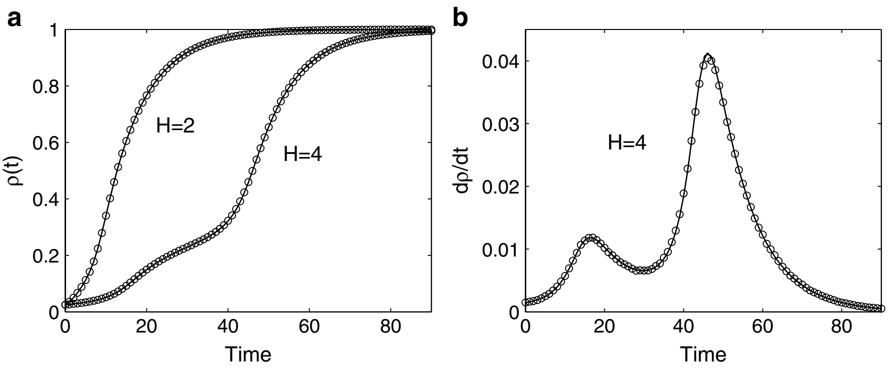


Fig. 3 Analytical (*solid lines*) and simulation (*circles*) results for the activation dynamics. The *upper panel* shows the fraction of active nodes vs. time for threshold parameter $H = 2$ and $H = 4$. The *lower panel* shows the activation rate $d\rho/dt$ vs. time for $H = 4$

is a threshold fraction of seed nodes so that above the threshold all the B will be eventually activated. However, even below the threshold, there is a possibility of a global cascade in group B if the across the group connectivity z_{ba} is sufficiently large. Indeed, our analysis has shown (Galstyan and Cohen 2007) that for a fixed within-group connectivity z_b, there is a critical across the group connectivity $z_{ba}^{\ c}$ so that for $z_{ba} > z_{ba}^{\ c}$ the activation will propagate from group A to group B and cause a global cascade.

Now let us look at the transient dynamics of the activation cascade; see Galstyan et al. (2009) for more details. In the continuous time limit, the dynamics can be written as

$$\tau \frac{d\rho_{a,b}}{dt} = 1 - \rho_{a,b} - (1 - \rho_{a,b}^0) \sum_{h=0}^{\infty} P_h Q[h; z_b \rho_b(t) + z_{ba} \rho_a(t)] \tag{10}$$

Let $\rho(t) = \alpha\rho_a(t) + (1-\alpha)\rho_b(t)$, $\alpha = N_a/(N_a + N_b)$, be the fraction of active nodes in the whole network. In Fig. 3 we compare the solutions obtained from Eq. 10 with the results of simulations on randomly generated graphs for the same network parameters but two different values of the threshold parameter. The parameters of the network are $N_a = 5{,}000$, $N_b = 15{,}000$, $z_{aa} = z_{bb} = 15$, $z_{ab} = 4$. The fraction of seed nodes is $\rho_a^0 = 0.1$, and $\tau^{-1} = 0.1$. The simulations are averaged over 100 random realizations.

We note that the agreement between the analytical prediction and results of the simulations is quite good. The network settles to the same steady state for both values of the threshold parameter H: that is, all of the nodes are activated at the end of the cascading process. However, the transient dynamics depend on the threshold parameter H. For $H = 2$, activation spreads very quickly through both communities and after a short interval all of the nodes are activate. For $H = 4$, on the other hand, the fraction of active nodes seems to saturate, then, in later iterations, $\rho(t)$ increases rapidly and eventually all the nodes become active. In Fig. 3b we plot the rate of activation process $d\rho/dt$ vs. time for $H = 4$. Apparently, the peak rates of activation in the two communities are separated in time. We call this phenomenon *two-tiered dynamics*. We would like to note that previously such a multi-peak structure has been observed in Gupta et al. (1989), where the authors studied the impact of different mixing patterns on the spread of sexually transmitted infection.

Influence Maximization

We now focus on influence maximization in modular networks. From the algorithmic standpoint, the influence maximization problem can be stated as follows (Domingos and Richardson 2001; Kempe et al. 2003): Given a social network, an influence model, and a set of nodes S, let $\sigma(S)$ be the expected number of nodes that will be activated by the end of the cascading process. Then, for a given *budget* M, the influence maximization problem is concerned with finding the set S of size M that maximizes the return $\sigma(S)$. While this problem is known to be NP hard for the many influence models, several approximate methods have been developed. An important result established in Kempe et al. (2003) states that for a class of models that obey the so called *diminishing returns* property, a simple hill-climbing algorithm, which works by greedily selecting the next best candidate node, yields a solution which is guaranteed to be within $\sim 63\,\%$ of the optimal. This result was further extended to more general models (Kempe et al. 2005; Mossel and Roch 2007).

It is quite safe to assume that the diminishing returns property is satisfied in saturated, or near-saturated, niche markets. However, those models might fail to capture the dynamics of emerging markets, where the condition of the sub-modular growth can be violated. Indeed, many economical and social phenomenon are better described in terms of critical phase transitions, where a huge growth is observed only after some threshold conditions are met. Here we are interested in this latter case. As we demonstrate below, in such critical systems, the structural properties of networks can play a significant role in the cascading dynamics. Consequently, selection strategies that discard the community structure might result in sub-optimal solution to the influence maximization problem. The intuition is as follows: since the critical number of nodes necessary to cause a cascade for a given connectivity grows linearly with the network size, then it might be beneficial to target the smaller group first and cause an activation cascade in that group. Afterwards, the activation will propagate through the larger network, provided that the density of links between the groups is sufficiently strong.

To validate this observation, we performed experiments on synthetic random graphs as well as real-world citation networks, using both integer and fractional versions of the linear threshold model.[4] We examined several different targeting strategies. The results presented below are for the *random selection* (RS), and

[4] Due to space restrictions, here we report our findings only for the integer threshold model on synthetic graphs.

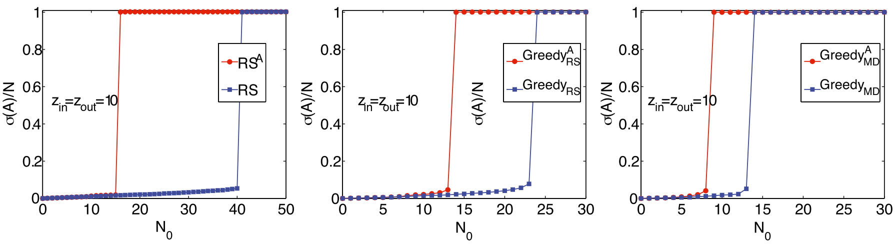


Fig. 4 Results for the integer-threshold LTM

greedy selection with two different tie-breaking mechanisms in case there are more than one candidates for selection: A random tie-break, where one of the candidates is chosen randomly, and a *maximum degree* tie-break, where the candidate with the maximum number of links is selected. We denote the corresponding algorithms as $G_{R\,S}$ and $G_{M\,D}$. Furthermore, we complemented each of those strategies by another strategy, which work exactly the same way, but now the candidate nodes are selected only from the community A. The corresponding strategies will be differentiated by a superscript A: RS^A, $G_{RS}{}^A$, and G^A_{MD}.

We constructed synthetic networks using a generative model known as *stochastic block model* (Holland et al. 1983). Namely, we assume that the network is composed of m groups, with N_m nodes in each. Each pair of nodes within the same group are linked with probability $p_{i\,n}$, while the pairs across the groups are linked with probability $p_{o\,u\,t}$. Thus, the corresponding connectivities within and across the groups are $z_{i\,n} = p_{i\,n}N_m$ and $z_{out} = p_{out}(N - N_m)$, respectively. In the experiments below we used $m = 10$, and $N_m = 100$, so that the total network size is $N = 1{,}000$. We assume that one of those ten groups constitute the group A, while the remaining nine communities form the group B.

In Fig. 4 we plot the fraction of activated nodes against the number of targeted nodes for the integer-threshold model, and for different selection strategies. The connectivities are set to $z_{in} = z_{out} = 10$. The integer thresholds were chosen randomly and uniformly from the interval $[2, 10]$. One can see that the selection strategies that explicitly target nodes from the smaller community are generally much more efficient, compared to the targeting from the general population of nose. Namely, for small and large values of N_0, both methods have a similar performance. However, there is a window $[N_1{}^c, N_2{}^c]$, within which the selection of A nodes is clearly superior. Recalling the analysis from the previous section, it is clear that $N_1{}^c$ corresponds to the critical threshold for which the activation spreads throughout group A, and the spills into the rest of the network. If one targets nodes from the general population, on the other hand, this critical effect does not come into play until later, when larger number of nodes, $N_0{}^2$, have been selected. The difference $N_2{}^c - N_1{}^c$ depends on the particular selection strategy (e.g., greedy, random selection, etc.), as well as the size of the network. For instance, for random selection strategies, the difference can be estimated as $\rho_c(N_b - N_a)$, where ρ_c is the critical fraction of seed nodes required to cause a global cascade (see section "Activation Cascades in Single-Community Networks").

Discussion

We have examined linear threshold model of activation cascades in structured heterogeneous networks. We demonstrated that for models with critical behavior, the structural properties of the network, and specifically, its community structure, can have a strong impact on the cascading process. For two-community networks, we demonstrated that by targeting nodes from the smaller community, one can achieve a cascade with fewer number of seed nodes. This effect is especially significant if the sizes of two communities are vastly different.

We note that the networks considered here mimic scenarios where innovations are introduced through a small community of early adopters. In this respect, our work is related to the organizational viscosity model of Krackhardt (1997) and McGrath and Krackhardt (2003) that describes the diffusion of ideas in an organization. In their approach, the organization is modeled as a number of interacting sub-units, with closer social ties within each unit. When the organization has a more or less homogenous structure, then a newly introduced idea cannot survive unless it is initially adopted by a large number of individuals. However, if the network describing the interaction of sub-units meets certain structural conditions, then the idea might take over the whole population even starting from a small number of initial adopters.

While the analysis shown here was for Erdos–Renyi networks, a similar behavior is observed also for communities with power-law degree distribution; see Galstyan et al. (2009). One important implication of the heavy tail is that it might affect networks dynamical properties, and, in some cases, suppress critical behavior. Finally, we note that the binary-state, single-stage model considered here might be too naive to capture certain dynamical processes on real-world networks. A number of authors have started examining multi-stage models that allow for more fine-grained notion of influence (Bruyn and Lilien 2008; Melnik et al. 2013). Another important extension is enabling nodes with more elaborate temporal dynamics, where the activity patterns can be sustained and reinforced over time (Piedrahita et al. 2013). Understanding the impact of network modularity on more elaborate dynamical models is an interesting future problem.

References

Arenas A, Diiaz-Guilera A, Peerez-Vicente CJ (2006) Synchronization reveals topological scales in complex networks. Phys Rev Lett 96:114102

Borge-Holthoefer J, Baos RA, Gonzlez-Bailn S, Moreno Y (2013) Cascading behaviour in complex socio-technical networks. J Complex Networks 1(1):3–24

Bruyn AD, Lilien GL (2008) A multi-stage model of word-of-mouth influence through viral marketing. Int J Res Mark 25(3):151–163

Derrida B, Pomeau Y (1986) Random networks of automata: a simple annealed approximation. Europhys Lett (EPL) 1(2):45–49

Derrida B, Stauffer D (1986) Phase transitions in two-dimensional kauffman cellular automata. Europhys Lett (EPL) 2(10):739–745
Domingos P, Richardson M (2001) Mining the network value of customers. In: KDD'01: proceedings of the 7th ACM SIGKDD international conference on knowledge discovery and data mining, San Francisco. ACM, New York, pp 57–66
Fortunato S (2010) Community detection in graphs. Phys Rep 486(35):75–174
Galstyan A, Cohen P (2007) Cascading dynamics in modular networks. Phys Rev E 75:036109
Galstyan A, Musoyan V, Cohen P (2009) Maximizing influence propagation in networks with community structure. Phys Rev E 79:056102
Girvan M, Newman ME (2002) Community structure in social and biological networks. Proc Natl Acad Sci USA 99(12):7821–7826
Gleeson JP (2008) Cascades on correlated and modular random networks. Phys Rev E 77:046117
Granovetter M (1978) Threshold models of collective behavior. Am J Sociol 83:1420–1443
Gupta S, Anderson RM, May RM (1989) Networks of sexual contacts: implications for the pattern of spread of HIV. AIDS 3:807–817
Holland PW, Laskey KB, Leinhardt S (1983) Stochastic blockmodels: first steps. Soc Netw 5(2):109–137
Kempe D, Kleinberg J, Tardos E (2003) Maximizing the spread of influence through a social network. In: KDD'03: proceedings of the 9th ACM SIGKDD international conference on knowledge discovery and data mining, New York, NY, USA, ACM, pp 137–146
Kempe D, Kleinberg J, Tardos E (2005) Influential nodes in a diffusion model for social networks. In: Proceedings of 32nd international colloquium on automata, languages and programming (ICALP), Springer-Verlag, Berlin, Heidelberg
Krackhardt D (1997) Organizational viscosity and the diffusion of controversial innovations. J Math Sociol 22(2):177–199
McGrath C, Krackhardt D (2003) Network conditions for organizational change. J Appl Behav Sci 39(3):324–336
Melnik S, Porter MA, Mucha PJ, Gleeson JP (2012) Dynamics on modular networks with heterogeneous correlations. CoRR abs/1207.1809
Melnik S, Ward JA, Gleeson JP, Porter MA (2013) Multi-stage complex contagions. Chaos Interdiscip J Nonlinear Sci 23(1):013124
Mossel E, Roch S (2007) On the submodularity of influence in social networks. In: STOC'07: proceedings of the 39th annual ACM symposium on theory of computing, San Diego. ACM, New York, pp 128–134
Newman MEJ (2006) Modularity and community structure in networks. Proc Natl Acad Sci USA 103(23):8577–8582
Payne JL, Dodds PS, Eppstein MJ (2009) Information cascades on degree-correlated random networks. Phys Rev E 80:026125
Piedrahita P, Borge-Holthoefer J, Moreno Y, Arenas A (2013) Modeling self-sustained activity cascades in socio-technical networks. arXiv:1305.4299
Rohlf T, Bornholdt S (2002) Criticality in random threshold networks: annealed approximation and beyond. Physica A 310(1):245–259
Watts D (2004) A simple model of global cascades on random networks. Proc Natl Acad Sci USA 99:5766

Synchrony in Social Groups and Its Benefits

Qi Xuan and Vladimir Filkov

Introduction

Self-organized synchrony is a group behavior which commonly occurs in nature. For example, groups of insects (Sullivan 1981), birds (Emlen 1952), and fish (Shaw 1978) can coordinate their moves and speeds with their neighbors so that they can all move together, behavior called swarming, flocking, schooling, and herding, for different kinds of species. Other examples of such behavior include fireflies that flash in unison (Mirollo and Strogatz 1990), pacemaker cells in the heart (Kuramoto and Yamagishi 1990), neural activities in cognitive processing (Fries 2005), etc. Synchrony is also a staple in social settings: choir singing (Müller and Lindenberger 2011), synchronization of applause in concert goers (Neda et al. 2000), and the formation of public opinion (Haken 2004) are easily recognizable examples. Another example is the collaboration in decentralized communities, e.g. among developers in Open Source Software (OSS) projects (Pinzger and Gall 2010; Xuan et al. 2012; Posnett et al. 2013). Yet other examples include the herd behavior among stock market traders (Scharfstein and Stein 1990; Chiang and Zheng 2010), the collective attention and emotion waves in online communities (Lehmann et al. 2012; Schweitzer and Garcia 2010), and language mimic (Gonzales et al. 2010).

It may be surprising that such synchronized behavior arises spontaneously without overall coordination and centralized authority. In fact, in all those groups synchronization emerges spontaneously, driven by simple decisions made by individuals in

Q. Xuan (✉)
University of California, Davis, CA 95616-8562, USA

Zhejiang University of Technology, Hangzhou, 310023, China
e-mail: qxuan@ucdavis.edu

V. Filkov
University of California, Davis, CA 95616-8562, USA
e-mail: filkov@cs.ucdavis.edu

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_64,

the group, based on limited sensory input of the behavior of their immediate neighbors. Staying synchronized with others takes effort, and thus comes at some cost to the individual. Thus, there are benefits to being synchronized, ranging from higher attractiveness to mates (fireflies) and evading predators (school of fish), to expressing forceful appreciation (concert goers).

Understanding the emergent behavior of complex systems which lack centralized governance would greatly enhance our understanding and interaction with the world around us. Recently, computer scientists have much benefited from observing self-organized biological systems and simulating their distributed rules in order to solve computational problems efficiently. E.g., a number of artificial intelligent algorithms (Navlakha and Bar-Joseph 2011; Anthony and Bartlett 2009; Dorigo and Blum 2005; Poli et al. 2007) were proposed to solve computational tasks of non-trivial difficulty (Vellido et al. 1999; Singh et al. 2009; Merkle et al. 2002; Aghdam et al. 2009; Gaing 2003). Meanwhile, these natural rules were also adopted to design distributed control schemes (Blaabjerg et al. 2006; Yu et al. 2012; Yan and Chen 2013) for groups of artifacts in order to deal with complex tasks, e.g., formation of spacecrafts (Beard and Hadaegh 2001) and robotic drumming (Crick et al. 2006). Effective study of synchrony in nature and society requires the use of quantitative analysis methods and data sets exemplifying such behavior.

Here we review work on social synchrony, a phenomenon arising when a group of people perform similar actions in a short period of time, actions which, over time, lead to the accomplishment of tasks of significant complexity (Choudhury et al. 2009). Although not all naturally occurring social synchrony is well understood, a significant corpus of work on these questions has amassed. A typical property of social synchrony is that individuals can obtain some information of others' behavior, followed by a simple modification of one's own behavior. Repeating this behavior leads to the emergence of the self-organized collective. This leads to several important questions that we and others have asked:

1. Synchrony is easy to describe and observe, but how can synchrony be measured and modeled in social groups?
2. If social ties among individuals and their behavior are in correlation, then what is the role of the social network structure on their synchronization?
3. Why do individuals synchronize their activities with each other, i.e., what is the benefit of synchronization? Does it lead to synergy?

This review chapter is structured around the above questions, and thus will elaborate on the quantitative aspects of social synchrony modeling, including specific metrics and models, the impact of social structure on the ability to synchronize, and the possible benefits of synchronization for the individual and community. Formal mathematical descriptions are used in the following sections for completeness; the chapter can be followed and understood sans the mathematical formalism.

Where appropriate, we will also summarize our results on the subject. Our own research work has recently centered on understanding self-organization in those social networks formed to achieve specific tasks, which we call task-oriented networks. To that end, we have focused significant attention on Open Source Software communities.

An established avenue for creating social capital, and reachly rewarding for the volunteer participants, OSS are examples of projects where people work in the absence of a coordinating hierarchy, to create snippets of code which when put together become complex artifacts of useful software. Some popular OSSs are Apache web server, Linux operating system, and the Mozzila web browser, but thousands of others exist. The software developers in OSS can be thought of as collaborating remotely on programming tasks, code integration, documentation writing, bug fixing, etc., while coordinating their work via electronic communication or by sharing examples. At the end of the chapter we present a case study on synchronization of software developers' activities in the Apache web server project.

Metrics and Models for Social Synchrony

Information exchange is necessary to achieve synchrony. A social network describes the links through which pairs of individuals exchange information. The following model is often used to describe the dynamics in social networks (Yan and Chen 2013; Park et al. 2006; Arenas et al. 2008; Gómez-Gardeñes et al. 2007):

$$\dot{x}_i(t) = F(x_i(t)) + \delta \sum_{j \in \pi_i} G(x_i(t), x_j(t)), i = 1, 2, \ldots, N \tag{1}$$

where $x_i(t)$ and π_i are the state and the neighbor set of individual v_i, δ is the coupling strength, $F(\cdot)$ is the individual dynamics, and $G(\cdot)$ is the coupling function through which different individuals interact with each other. The group of individuals $v_1, v_2, \ldots, v_N$ are considered mathematically synchronized if and only if

$$\sum_{i,j=1}^{N} x_i(t) - x_j(t) \to 0, \tag{2}$$

as $t \to \infty$ (Yu et al. 2012; Lü and Chen 2005; Li et al. 2007). However, Eq. (2) cannot be directly used to measure the synchrony in real systems because of their finite life span, and also because individuals may not take actions at the exactly same time, i.e., there might be short delays between their actions. To overcome these limitations, recently, several quantitative methods were proposed to measure social synchrony more realistically.

Sun et al. (2011) modeled synchrony in a group of cows, of two different behavioral states, eating or lying down. If we denote by $\tau_i(k)$ the kth time at which cow v_i switches to certain state (switching action), then the synchrony of this state between cows v_i and v_j is measured by

$$\Delta_{ij} = \frac{1}{K} \sum_{k=1}^{K} | \tau_i(k) - \tau_j(k) |, \tag{3}$$

where it is assumed that the two cows have the same number of switching actions. A smaller value of Δ_{ij} indicates more synchrony of the two individuals. Then, for

N cows, the group synchrony is measured by averaging over all pairwise synchronies:

$$\Delta = \langle \Delta_{ij} \rangle = \frac{1}{N^2} \sum_{i,j=1}^{N} | \Delta_{ij} |. \tag{4}$$

An alternative metric for synchrony is to directly count how often all individuals have the same state (Fæevik et al. 2008).

In many real cases different individuals may be active at very different rates in any given time interval, and each action may last for only a very short period of time, or even be discrete, i.e., the activities may form a Zero-measure set on the time axis. For example, in Open Source Software projects, different software developers have different rhythms of submitting changes to the software, and only the times when they submitted the changes are recorded. To address this situation, we have proposed the following more general metric for synchrony.

1. *Identify activity bursts.* From the time-series of activities for each individual, identify activity bursts based on a one-dimensional clustering method, i.e., first inter-activity time intervals larger than a predefined time window θ are obtained, then the activities between two consecutive large intervals are grouped as an "active burst", with occurrence time equal to the average of the times of the first and the last activities in this burst.
2. *Smooth bursts.* Let Γ_i be the set of all occurrence times of activity bursts of individual v_i. The smoothing function of the activity bursts is constructed by using Gaussian kernels (Moon 2001), as follows:

$$\varphi_i(t) = \frac{1}{|\Gamma_i|} \sum_{\xi \in \Gamma_i} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\xi)^2}{2\sigma^2}}. \tag{5}$$

3. *Calculate synchrony through correlation.* For each pair of individuals v_i and v_j, their centralized curves are obtained by subtracting the corresponding average value in the time interval $[T_L, T_U]$, where T_L and T_U are the minimum and maximum elements in the set $\Gamma_i \cup \Gamma_j$, respectively. Their synchrony is calculated by the Pearson correlation coefficient (Chatterjee and Price 1991) between the two centralized curves. Similarly, the group synchrony is calculated by averaging over all pairwise synchronies.

The metrics above calculate synchronies but don't tell us if those values are significantly different than those that would result from chance synchronization. To calculate the significance of the results we need a random or null model of behavior for all possible activities. One null model example is the uniform model, and another is a class of models that results from randomly permuting the labels on the events in the time series (bootstrapping) (Xuan et al. 2012). Using such models, the data can be randomized many times, each resulting in a population, and then pairwise synchronies can be computed for the individuals in each population. This procedure will yield a distribution with which the statistical significance of the real case can be assessed, using tests such as the t-test or the Wilcoxon-Mann-Whitney test.

The Impact of the Network Architecture on Synchrony

Based on the mathematical model represented by Eq. (1), we can see that synchrony may depend on the underlying network structure. As a result, it is of much scientific interest to characterize the kinds of networks which can facilitate synchrony.

In many theoretical works (Park et al. 2006; Arenas et al. 2008; Barahona and Pecora 2002; Hong et al. 2004; Motter et al 2005; Lerman and Ghosh 2012), it is simply assumed that

$$G(x_i(t), x_j(t)) = H(x_j(t)) - H(x_i(t)), \tag{6}$$

where $H(\cdot)$ is called the output function. Equation (6) is intuitive by considering that each individual is cooperative and hopes to be in an activity state close to those of its neighbors. By substituting Eq. (6) into Eq. (1), we have

$$\dot{x}_i(t) = F(x_i(t)) - \delta \sum_{j=1}^{N} L_{ij} H(x_j(t)), i = 1, 2, \ldots, N, \tag{7}$$

where L is the Laplacian matrix with its element $L_{ij} = -1$ if v_i and v_j are neighbors in the network, $L_{ii} = k_i$ if v_i has degree k_i, and $L_{ij} = 0$ otherwise. If the network is connected, i.e., there is a path between each pair of nodes, the Laplacian matrix has the eigenvalues satisfying $0 = \lambda_1 < \lambda_2 \leq \lambda_3 \leq \ldots \leq \lambda_N$.

Nishikawa et al. (2003) found that the network's ability to synchronize is determined by λ_N/λ_2: the smaller that ratio, the less difficult it is to synchronize the dynamics of the nodes, and vice versa. Then, the question is which kind of networks have relatively small ratio of λ_N over λ_2. Several studies (Barahona and Pecora 2002; Donetti et al. 2003; Xuan et al. 2009) proved that the ratio is mainly determined by two factors: *small world* property and *homogeneity*. That is, a group of individuals are more likely to synchronize with each other when they are close to each other, i.e., have short average distance, and meanwhile have similar social status, i.e., have similar degrees. Thus, it is easy to infer that the fully connected network has the maximum synchronization ability since it has the minimum average distance and all the nodes have exactly the same degree. In fact, it can be proved that, in a fully connected network of N nodes, λ_2 and λ_N have the same value N, so that the ratio λ_N/λ_2 is equal to 1, which is the minimum over all connected networks (Chen et al. 2012). However, in most real cases, an individual cannot establish and keep the social ties with all others in a social system, especially when the system is large. Therefore, it is of much interest to identify the optimal network structures for synchrony under the condition that the average degree is fixed and much smaller than the network size. Donetti et al. (2003) proposed a method to minimize the eigenvalue ratio by a rewiring process, and they found that the optimal networks have extremely homogeneous structure, i.e., very small variance in degree, node distance, betweenness, and loop distributions (Costa et al. 2007), properties similar to those of *Cage* graphs (http://mathworld.wolfram.com/CageGraph.html) studied by many mathematicians. We obtained the same result by adopting another method (Xuan et al. 2009), where the average shortest path length rather than the ratio is minimized by a rewiring process under the condition that all nodes have exactly the same degree. In fact, we found that the average shortest path length and the ratio λ_N/λ_2 are linearly correlated in the optimization process. Since such optimization algorithms are always very time-consuming, we also proposed a growth model to obtain sub-optimal structure of large-scale networks in this work.

Most real-world networks have heterogeneous and modular structure (Barabási and Albert 1999; Girvan and Newman 2002; Ravasz et al. 2002; Xuan et al. 2006). When looking inside, it was found that hub nodes and the links connecting different modules play key roles in the synchronization process (Choudhury et al. 2009; Park et al. 2006; Wang and Chen 2002). For example, theoretical analysis (Wang and Chen 2002) proved that the network of individuals are more likely to be synchronized when those highly connected individuals are selected as leaders (they are not influenced by others), i.e., smaller number of leaders are needed, as compared to the random case, while empirical studies of the popular social site *Digg* (Choudhury et al. 2009) also indicate that large-scale social synchronies are more likely to arise if initialized by individuals with larger numbers of connections. Recent studies of synchrony on modular networks can also provide some useful insights. In fact, synchrony always occurs within each module at group level because the nodes in each module are always highly connected, almost like a fully connected subnetwork. However, the steady states of different modules may be independent from each other, i.e., the global synchrony cannot be achieved at system level, unless there are enough between-module links including some random and long-range links among these modules (Park et al. 2006; Oh et al. 2005; Zhou et al. 2007). These findings indicate that the links connecting different modules are important for the systemic behaviors.

The Kuramoto model (Arenas et al. 2008; Acebrón et al. 2005) may be the most well-known model to study the synchronization on networks. In this model, $F(\theta_i) \equiv \omega_i$ and $G(\theta_i, \theta_j) \equiv \sin(\theta_j - \theta_i)$, where ω_i is the natural frequency of node v_i, θ_i rather than x_i is adopted as the state of a node in order to keep these symbols the same as those in the related references, and the time t is omitted for simplicity. Then, we have the following collective dynamics:

$$\dot{\theta}_i = \omega_i + \delta \sum_{j \in \pi_i} \sin(\theta_j - \theta_i), i = 1, 2, \ldots, N. \tag{8}$$

The synchronization here means that a group of individuals with different natural frequencies may oscillate with the same mean frequency when their coupling strength exceeds some critical point determined by the network structure. Note that this model can be theoretical analyzed, and Arenas et al. (2008) have provided a detailed review for this kind of study, which will not be extendedly discussed here. In fact, the Kuramoto model on networks has a simple linear form:

$$\dot{\theta}_i = \omega_i + \delta \sum_{j=1}^{N} L_{ij} \theta_j(t), i = 1, 2, \ldots, N. \tag{9}$$

Recently, Lerman and Ghosh (2012) proposed a more general linear model by replacing the Laplacian matrix L in Eq. (9) by $R \equiv \alpha\, I - A$ in order to describe non-conservative social and biological processes more appropriately. The synchronization process partly depends on network structure, as a result, it can also be used to identify the network structure (Arenas et al. 2006; Boccaletti et al. 2007; Li et al. 2008; Fortunato 2010), e.g., detect communities. Interestingly, Lerman and Ghosh (2012) found that the identified network structure may be different by using different kinds of interactions in the synchronization scheme, which suggests that such methods for identifying local structures in complex networks must be used with great care.

Benefits of Social Synchrony: Toward Synergy

One of the reasons that a group of individuals prefer to take similar actions in certain time is that they want to deal with complex tasks more efficiently, in other words, they see synchrony as a way for the group to gain more than what each individual puts in. Thus, they aim to achieve *s*ynergy, defined as the creation of a whole that is greater than the sum of its parts (French et al. 2008). There are a dozen of such examples in nature. Ants are more likely to follow others in the same colony in order to perform better when they search and carry food as a group (Deneubourg et al. 1983; Dorigo et al. 1996). Male fireflies synchronize their flashing rhythm in order attract more females in a wide-range area (Otte 1980; Lewis and Cratsley 2008). A larger flocking of birds can help them detect approaching predators with a higher probability (Siegfried and Underhill 1975), meanwhile, formation flight can also reduce the flying cost on aerodynamics aspect (Hummel 1983), which can explain why groups of birds always present special shapes when they migrate over a long distance.

There are also many reasons for humans to synchronize our actions with others: Macrae et al. (2008) found that the synchrony of movements during social exchanges may facilitate the person perception process, e.g., the memory for an interaction partner's characters can be enhanced during this process. Hove and Risen designed experiments to show that interpersonal synchrony increases affiliation with a group (Hove and Risen 2009), similar to the effect of mimicry (Lakin et al. 2003), which may provide evidence for the hypothesis that such phenomena may play important role in social cohesion (Freeman 2000). While recently Paladino et al. (2010) suggested that synchrony may also have a magic to blurs self-other boundaries. All of these psychological findings indicate that social synchrony is selected evolutionarily, which may help a group of people increase their cooperative ability to better solve complex social tasks, as validated by Gonzales et al. (2010), Valdesolo et al. (2010), and Wiltermuth and Heath (2009) in their empirical studies. Moreover, Woolley et al. (2010) suggested that such cooperative ability can be characterized as a general collective intelligence factor, i.e., they found that the group performances on different tasks are significantly positively correlated, while the average and maximum performances of individual group members are not, and this factor can further be used to predict the group performance on other tasks. More on *collective intelligence* can be found in Woolley and Hashmi's chapter of this book. Having chosen a metric and model of synchrony as described in the above sections, synergy can be studied as an outcome, by modeling it in terms of the observed synchronizations in the groups or in the whole system. Great attention must be paid to following good modeling habits to avoid colinearities and other statistical obstacles.

A Case Study of Synchrony in Open Source Software Systems

Open Source Software systems provide a good platform to analytically study social synchrony and synergy among people. In OSS, groups of volunteer software developers create a software artifact by sharing programming experiences, finding bugs, or committing to files directly. OSS resemble ecological systems (Posnett et al. 2013) in that in addition to the actual developers, they attract thousands of users and other contributors looking to gain knowledge. These human resources, in turn, make the software grow faster and become better by providing feedbacks and joining the ranks of developers occasionally. Pavlic and Pratt, in another chapter of this book, compare eusocial insect behavior with human behavior conceptually in the context of OSS on a variety of dimensions.

Here, we look at projects from the Apache Software Foundation, and show how to validate whether developers prefer to work together or not, i.e., we show how to measure social synchrony and demonstrate that it is prevalent in these projects. We selected the six projects *Ant*, *Axis2_java*, *Cxf*, *Derby*, *Lucene*, and *Openejb* because they contain most developers so that we can get most meaningful statistical results. The data, gathered on March 24th, 2012, contains both the commit-code-to-file (commits) activities and the communication activities (emails) among developers. For each commit in a project, we have gathered the developer ID, file ID, the submitting time in seconds, and the numbers of added and deleted lines of code in each file. For each email communication activity, we have the sender ID, receiver ID, and the sending time in seconds.

Based on this data we calculated group synchrony. First, we filtered the data by selecting the files committed to by at least ten developers, and considered each month from the first to the last commit time as a time window. For each file f_i, out of a total of M across all six projects, we counted the number of developers, denoted by $n_i(t)$, that committed to this file in each time window t. Let X_i the total number of months in the time interval and $Y_i = \max_t n_i(t)$. Then, for each f_i, we obtained an $X_i \times Y_i$ binary count matrix A_i, with its elements $A_i(t, n_i(t)) = 1$ and the others equal to zero.

Note that the count matrix A_i shows that developers worked together in the same month on the same file, which, however, may be largely dependent on their own working rhythms, i.e., Y_i will be very large if the developers worked on the file frequently and will be very small otherwise. Therefore, to establish a baseline, we need to create simulated count matrices for comparison. To do that, we randomized the data as follows. If developer v_j committed to the file in h_{ij} months, we randomly permuted these h_{ij} active months among the total Y_i months. We repeated that process 100 times and got 100 binary matrices, denoted by $B_i^l, l = 1, 2, \ldots, 100$, for these random cases. Note that the real and simulated matrices may have different sizes, in

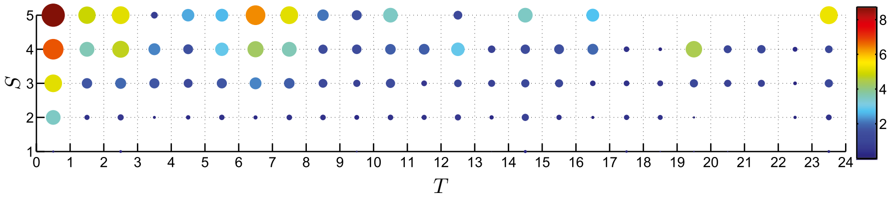


Fig. 1 The visualization of the significance matrix C. Here, S is the group size and T is a month in the first 2 years for each file since it was created. The elements with $a_{ij}<5$, $b_{ij}<0.1$, or $c_{ij}\leq 0$ are not shown. The point size is proportional to the value of the corresponding element in matrix C

which case we then expand the smaller matrices by filling them with zeros, so that all these matrices have the exactly same size. When considering all M files together, we also expand smaller matrices by the same method, and still denote them by A_i and B_i^l, $i = 1,2,\ldots,M, l = 1,2,\ldots,100$. Then, we can calculate the real and simulated matrix counts by:

$$A = \sum_{i=1}^{M} A_i, \quad B = \frac{1}{100}\sum_{i=1}^{M}\sum_{l=1}^{100} B_i^l \tag{10}$$

respectively. Based on these two matrices, we can get a significance matrix C with each element calculated by $c_{ij} = (a_{ij}-b_{ij})/b_{ij}$, which shows how significantly differently than chance the developers prefer to work together as a group at a certain scale. Here, only the elements satisfying $a_{ij}\geq 5$ and $b_{ij}\geq 0.1$ are considered. The significance matrix C for the first 2 years of the lives of the files is visualized in Fig. 1, where we can see that developers indeed prefer to work together as a group at larger scale, and the absence of most points when $S = 1$ indicates that they seldom work alone.

Conclusions

In this chapter, we have described social synchrony, and reviewed proposed metrics and models for it. We also discussed its possible benefits in social groups, especially how it leads to synergy among participants. We applied those methods to the analysis of distributed software development as a case study. In our analysis, we successfully discovered group synchrony of code developers when they commit to files, demonstrating the utility of this technique.

Future work involves extending this technique to identify synchrony patterns in OSS systems, based on which more realistic synchrony models for code developers can be created. These methods can also be used to analyze other social communities, where people cooperate with each other to finish complex tasks, e.g., online knowledge communities like Wikipedia, or question and answer communities such as Stack Overflow, where people share knowledge by shaping answers for technical problems together.

Acknowledgements We gratefully acknowledge support from the Air Force Office of Scientific Research, award FA955-11-1-0246. QX acknowledges support from the National Natural Science Foundation of China (Grants No. 61004097 and No. 612732122).

References

Acebrón JA, Bonilla LL, Vicente CJP, Ritort F, Spigler R (2005) The Kuramoto model: a simple paradigm for synchronization phenomena. Rev Mod Phys 77(1):137–185

Aghdam MH, Ghasem-Aghaee N, Basiri ME (2009) Text feature selection using ant colony optimization. Expert Syst Appl 36(3):6843–6853

Anthony M, Bartlett PL (2009) Neural network learning: theoretical foundations. Cambridge University Press, Cambridge

Arenas A, Díaz-Guilera A, Pérez-Vicente CJ (2006) Synchronization reveals topological scales in complex networks. Phys Rev Lett 96(11):114102

Arenas A, Díaz-Guilera A, Kurths J, Moreno Y, Zhou C (2008) Synchronization in complex networks. Phys Rep 469(3):93–153

Barabási A-L, Albert R (1999) Emergence of scaling in random networks. Science 286(5439):509–512

Barahona M, Pecora LM (2002) Synchronization in small-world systems. Phys Rev Lett 89(5):054101

Beard RW, Hadaegh FY (2001) A Coordination architecture for spacecraft formation control. IEEE Trans Control Syst Technol 9(6):777–790

Blaabjerg F, Teodorescu RE, Liserre M, Timbus AV (2006) Overview of control and grid synchronization for distributed power generation systems. IEEE Trans Ind Electron 53(5):1398–1409

Boccaletti S, Ivanchenko M, Latora V, Pluchino A, Rapisarda A (2007) Detecting complex network modularity by dynamical clustering. Phys Rev E 75(4):045102(R)

Chatterjee S, Price B (1991) Regression analysis by example. Wiley, New York

Chen J, Lu J-A, Zhan C, Chen G (2012) Laplacian spectra and synchronization processes on complex networks. In: Thai MT, Pardalos PM (eds) Handbook of optimization in complex networks: theory and applications. Springer, Heidelberg, pp 81–113

Chiang TC, Zheng D (2010) An empirical analysis of herd behavior in global stock markets. J Bank Financ 34(8):1911–1921

Choudhury MD, Sundaram H, John A, Seligmann DD (2009) Social synchrony predicting mimicry of user actions in online social media. In: The proceedings of the 2009 international conference on computational science and engineering, Vancouver, Canada, pp 151–158

Costa LF, Rodrigues FA, Travieso G, Boas PRV (2007) Characterization of complex networks: a survey of measurements. Adv Phys 56(1):167–242

Crick C, Munz M, Scassellati B (2006) Synchronization in social tasks: robotic drumming. In: The proceedings of the 15th IEEE international symposium on robot and human interactive communication, Hatfiled, UK, pp 97–102

Deneubourg JL, Pasteels JM, Verhaeghe JC (1983) Probabilistic behaviour in ants: a strategy of errors? J Theor Biol 105(2):259–271

Donetti L, Hurtado PI, Muñoz MA (2003) Entangled networks, synchronization, and optimal network topology. Phys Rev Lett 95(18):188701

Dorigo M, Blum C (2005) Ant colony optimization theory: a survey. Theor Comput Sci 344:243–278

Dorigo M, Maniezzo V, Colorni A (1996) Ant system: optimization by a colony of cooperating agents. IEEE Trans Syst Man Cybern Part B: Cybern 26(1):29–41

Emlen JJT (1952) Flocking behavior in birds. The Auk 69(2):160–170

Fæevik G, Tjentland K, Løvik S, Andersen IL, Bøe KE (2008) Resting pattern and social behaviour of dairy calves housed in pens with different sized lying areas. Appl Anim Behav Sci 114:54–64

Fortunato S (2010) Community detection in graphs. Phys Rep 486(3–5):75–174
Freeman W (2000) A neurobiological role of music in social bonding. In: Wallin NL, Merker B, Brown S (eds) The origins of music. MIT, Cambridge, pp 411–424
French R, Schermerhorn JR, Rayner C, Rees G, Rumbles S, Hunt JG, Osborn RN (2008) Organizational behaviour. Wiley, New York
Fries P (2005) A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. TRENDS Cogn Sci 9(10):474–480
Gaing Z-L (2003) Particle swarm optimization to solving the economic dispatch considering the generator constraints. IEEE Trans Power Syst 18(3):1187–1195
Girvan M, Newman MEJ (2002) Community structure in social and biological networks. Proc Natl Acad Sci USA 99(12):7821–7826
Gómez-Gardeñes J, Moreno Y, Arenas A (2007) Paths to synchronization on complex networks. Phys Rev Lett 98(3):034101
Gonzales AL, Hancock JT, Pennebaker JW (2010) Language style matching as a predictor of social dynamics in small groups. Commun Res 37(1):3–19
Haken H (2004) Synergetics: introduction and advanced topics. Springer, Heidelberg
Hong H, Kim BJ, Choi MY, Park H (2004) Factors that predict better synchronizability on complex networks. Phys Rev E 69(6):067105
Hove MJ, Risen JL (2009) It's all in the timing: interpersonal synchrony increases affiliation. Soc Cogn 27:949–960
http://mathworld.wolfram.com/CageGraph.html
Hummel D (1983) Aerodynamic aspects of formation flight in birds. J Theor Biol 104(3):321–347
Kuramoto T, Yamagishi H (1990) Physiological anatomy, burst formation, and burst frequency of the cardiac ganglion of crustaceans. Physiol Zool 63(1):102–116
Lakin JL, Jefferis VE, Cheng CM, Chartrand TL (2003) The chameleon effect as social glue: evidence for the evolutionary significance of nonconscious mimicry. J Nonverbal Behav 27:145–162
Lehmann J, Gonçalves B, Ramasco JJ, Cattuto C (2012) Dynamical classes of collective attention in Twitter. In: The proceedings of the 2012 international world wide web conference committee, Lyon, pp 251–260
Lerman K, Ghosh R (2012) Network structure, topology, and dynamics in generalized models of synchronization. Phys Rev E 86(2):026108
Lewis SM, Cratsley CK (2008) Flash signal evolution, mate choice, and predation in fireflies. Annu Rev Entomol 53:293–321
Li C, Sun W, Kurths J (2007) Synchronization between two coupled complex networks. Phys Rev E 76(4):046204
Li D, Leyva I, Almendral JA, Sendiña-Nadal I, Buldú JM, Havlin S, Boccaletti S (2008) Synchronization interfaces and overlapping communities in complex networks. Phys Rev Lett 101(16):168701
Lü J, Chen G (2005) A time-varying complex dynamical network model and its controlled synchronization criteria. IEEE Trans Autom Control 50(6):841–846
Macrae CN, Duffy OK, Miles LK, Lawrence J (2008) A case of hand waving: action synchrony and person perception. Cognition 109:152–156
Merkle D, Middendorf M, Schmeck H (2002) Ant colony optimization for resource-constrained project scheduling. IEEE Trans Evol Comput 6(4):333–346
Mirollo RE, Strogatz SH (1990) Synchronization of pulse-coupled biological oscillators. SIAM J Appl Math 50(6):1645–1662
Moon BS (2001) A gaussian smoothing algorithm to generate trend curves. Korean J Comput Appl Math 8(3):507–518
Motter AE, Zhou C, Kurths J (2005) Network synchronization, diffusion, and the paradox of heterogeneity. Phys Rev E 71(1):016116
Müller V, Lindenberger U (2011) Cardiac and respiratory patterns synchronize between persons during choir singing. PLoS One 6(9):e24893
Navlakha S, Bar-Joseph Z (2011) Algorithms in nature: the convergence of systems biology and computational thinking. Mol Syst Biol. doi:10.1038/msb.2011.78

Neda Z et al (2000) Self-organizing processes: the sound of many hands clapping. Nature 403:849–850
Nishikawa T, Motter AE, Lai Y-C, Hoppensteadt FC (2003) Heterogeneity in oscillator networks: are smaller worlds easier to synchronize? Phys Rev Lett 91(1):014101
Otte D (1980) On theories of flash synchronization in fireflies. Am Nat 116(4):587–590
Oh E, Rho K, Hong H, Kahng B (2005) Modular synchronization in complex networks. Phys Rev E 72(4):047101
Paladino MP, Mazzurega M, Pavani F, Schubert TW (2010) Synchronous multisensory stimulation blurs self-other boundaries. Psychol Sci 21:1202–1207
Park K, Lai Y-C, Gupte S (2006) Synchronization in complex networks with a modular structure. Chaos 16(1):015105
Pinzger M, Gall HC (2010) Dynamic analysis of communication and collaboration in OSS projects. In: Collaborative software engineering. Springer, Heidelberg, pp 265–284
Poli R, Kennedy J, Blackwell T (2007) Particle swarm optimization. Swarm Intell 1(1):33–57
Posnett D, D'Souza R, Devanbu P, Filkov V (2013) Dual ecological measures of focus for software development. In: The proceedings of the 2013 international conference of software engineering, San Francisco
Ravasz E, Somera AL, Mongru DA, Oltvai ZN, Barabási A-L (2002) Hierarchical organization of modularity in metabolic networks. Science 297(5586):1551–1555
Scharfstein DS, Stein JC (1990) Herd behavior and investment. Am Econ Rev 80(3):465–479
Schweitzer F, Garcia D (2010) An agent-based model of collective emotions in online communities. Eur Phys J B 77(4):533–545
Shaw E (1978) Schooling fishes: the school, a truly egalitarian form of organization in which all members of the group are alike in influence, offers substantial benefits to its participants. Am Sci 66(2):166–175
Siegfried WR, Underhill LG (1975) Flocking as an anti-predator strategy in doves. Anim Behav 23:504–508
Singh KP, Basant A, Malik A, Jain G (2009) Artificial neural network modeling of the river water quality-a case study. Ecol Model 220(6):888–895
Sullivan RT (1981) Insect swarming and mating. Fla Entomol 64(1):44–65
Sun J, Bollt EM, Porter MA, Dawkins MS (2011) A mathematical model for the dynamics and synchronization of cows. Phys D: Nonlinear Phenom 240(19):1497–1509
Valdesolo P, Ouyang J, DeSteno D (2010) The rhythm of joint action: synchrony promotes cooperative ability. J Exp Soc Psychol 46(4):693–695
Vellido A, Lisboa PJG, Vaughan J (1999) Neural networks in business: a survey of applications. Expert Syst Appl 17:51–70
Wang XF, Chen G (2002) Pinning control of scale-free dynamical networks. Phys A: Stat Mech Appl 310(3–4):521–531
Wiltermuth SS, Heath C (2009) Synchrony and cooperation. Psychol Sci 20:1–5
Woolley AW, Chabris CF, Pentland A, Hashmi N, Malone TW (2010) Evidence for a collective intelligence factor in the performance of human groups. Science 330(6004):686–688
Xuan Q, Li Y, Wu T-J (2006) Growth model for complex networks with hierarchical and modular structures. Phys Rev E 73(3):036105
Xuan Q, Li Y, Wu T-J (2009) Optimal symmetric networks in terms of minimizing average shortest path length and their sub-optimal growth model. Phys A: Stat Mech Appl 388(7):1257–1267
Xuan Q, Gharehyazie M, Devanbu P, Filkov V (2012) Measuring the effect of social communications on individual working rhythms: a case study of open source software. In: The proceedings of 2012 ASE/IEEE international conference on social informatics, Washington, DC
Yan F, Chen G (2013) Distributed consensus and coordination control of networked multi-agent systems. In: Kocarev L (ed) Consensus and synchronization in complex networks. Springer, Heidelberg, pp 51–68
Yu W, DeLellis P, Chen G, Bernardo MD, Kurths J (2012) Distributed adaptive control of synchronization in complex networks. IEEE Trans Autom Control 57(8):2153–2158
Zhou T, Zhao M, Chen G, Yan G, Wang B-H (2007) Phase synchronization on scale-free networks with community structure. Phys Lett A 368(6):431–434

Psychosocial and Cultural Modeling in Human Computation Systems: A *Gamification* Approach

Antonio Sanfilippo, Roderick Riensche, Jereme Haack, and Scott Butner

Introduction

Human decision making is hard to capture in computation systems because it is strongly driven by insight and creativity and influenced by cognitive biases. Qualities such as the ability to focus on what is perceived to be most important and the capacity to make quick decisions by insight and intuition make human judgment uniquely effective (Gigerenzer 2007; Gladwell 2005). However, the same qualities can also be responsible for fallacious reasoning when judgment is affected by memory limitations (Miller 1956), lack of knowledge/expertise (Klein 1998), and biased judgment due to factors such as increased confidence in extreme judgments and highly correlated observables (Kahneman and Tversky 1973), positive framing (Tversky and Kahneman 1981), "groupthink" (Janis 1972; Surowiecki 2004), and premature commitment to a single expected outcome (Heuer 1999). The ability to understand weaknesses and leverage strengths in the human decision process is crucial in designing human computation systems which effectively benefit from and complement human intelligence.

In this chapter, we review some of the challenges and opportunities regarding the integration of human decision making into human computation systems and discuss ways in which challenges can be met to avail ourselves of the opportunities afforded. We discuss a *gamification* approach in which gameplay is applied to real-world problems to develop social intelligence and support analysis and decision-making through a concerted reasoning effort that interleaves human and machine intelligence. We describe a systematic methodology for integrating modeling algorithms within a serious gaming environment in which role-playing by human agents provides updates to model nodes and the ensuing model outcomes in turn influence the behavior of the human players. The approach implements a strong functional partnership between

A. Sanfilippo (✉) • R. Riensche • J. Haack • S. Butner
Pacific Northwest National Laboratory, Richland, WA, USA
e-mail: Antonio.sanfilippo@pnnl.gov

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_65, © Springer Science+Business Media New York 2013

human players and computer models that leverages modularity and independence across participating agents and components to facilitate the connection between model and game structures. We illustrate an embodiment of this approach with reference to the characterization of transactions in illicit nuclear trafficking.

Decision Making and Risk Perception

Risk perception plays a major role in regulating human decision-making. For example, experiments performed to see how people evaluated probabilities (Kahneman and Tversky 1973, 1974; Tversky and Kahneman 1981) demonstrate that people are risk-averse with respect to gains, but risk-seeking about losses. A certain outcome is preferred over a gamble with a higher expected utility which presents the possibility of total loss, while the hope for the chance of losing nothing is preferred over a sure but smaller loss. Understanding how cognitive and cultural biases impact risk perception is therefore crucial in designing a strategy for integrating the human decision making process in human computation systems.

Psychometric approaches to risk perception (Kahneman and Tversky 1974; Starr 1969; Slovic 1987) have made significant strides in identifying cognitive factors responsible for influencing the individual perceptions of risk. Social and cultural approaches (Douglas and Wildavsky 1982) have broadened the scope of risk perception research by focusing on how the perception of risk reflects an individual's commitment to competing cultural and political views. Several promising interdisciplinary efforts are underway to integrate the psychometric and sociocultural perspectives into a unified cultural and cognitive framework of risk perception (Kahan et al. 2010; Kasperson et al. 2003).

Kahneman's and Tversky's groundbreaking experimental and theoretical work on the psychology of decision-making (Kahneman and Tversky 1973, 1974; Tversky and Kahneman 1981) paved the way to the identification of patterns of deviation in human judgment that occur under risk. Through empirical observation, Kahneman and Tversky (1974) propose that people rely on simple heuristics when exerting judgment under uncertainty. These heuristics are crucial in streamlining human decision-making so as to achieve an ideal balance between judgment effectiveness and use of cognitive-processing and information resources. However, they also lead to cognitive biases. For example, Tversky and Kahneman (1981) showed through a series of experiments that different ways of framing the same risk information can have diametrically opposite responses. In one of these experiments, subjects were asked to choose health intervention options to combat a disease outbreak expected to kill 600 people. The first choice was between program A, which would save 200 people, and program B, which would either save all the people with a 1/3 probability or no people with a 2/3 probability. Most subjects preferred the guarantee that 200 people be saved (A) rather than risking everyone dying (B). However, when asked to choose between program B and program C, in which 400 people would die, most subjects chose program B, even though the expected outcomes of programs A and C are identical in terms of casualties. The overtly expressed certain death of 400 people is less acceptable than the two-in-three chance that all would die.

Risk perception research within the psychometric paradigm has increasingly emphasized the role of affect and emotion on risk perception. The impact of fear/dread, outrage, familiarity and uncertainty/lack of control were demonstrated early on to be important determinants of risk perception (Slovic 1987). Recent work has focused on capturing the emotion components into an affect heuristic, according to which positive and negative affect is modulated by information about benefits and risks (Slovic et al. 2005).

Risk perception is also regulated by social and cultural identity factors. As individuals, we typically form judgments within a social context. Consequently, our assessment of risk is filtered through concerns about safety, power, justice and legitimacy that are germane to the social enclave with which we identify. Our perception of risk thus reflects our individual commitment to specific cultural values, as opposed to alternative ones. Following this line of reasoning, the cultural theory of risk (Douglas and Wildavsky 1982) explains variance in the perception of risk in terms of social and cultural values to which allegiance grants taking higher risks. The polemic surrounding the human-papillomavirus (HPV) vaccination is a good example of how critical one's own commitment to specific cultural values determines the willingness to accept a higher or lower risk. HPV is responsible for 70 % of cervical cancers, 80 % of anal cancers, 60 % of vaginal cancers, and 40 % of vulvar cancers (De Vuyst et al. 2009). Since 2006, when the U.S. Food and Drug Administration approved the first preventive HPV vaccine, political dispute has hindered a plan to vaccinate US girls against HPV, amid claims that the vaccine causes harmful side effects and promotes unsafe sex among teens (Kahan 2010a). Interestingly, experimental evidence (Kahan et al. 2010) shows that when the arguments pro and against HPV vaccination are conveyed in such a manner as to reduce "biased assimilation" (i.e., the propensity to credit and dismiss information so as to confirm one's own prior beliefs), opinion polarization diminishes. People react more open-mindedly towards achieving scientific consensus instead of forming risk perceptions that reflect their commitments to controversial views of ethics and morality.

While psychometric and sociocultural approaches have emphasized diverse facets of human behavior that shape how people perceive risk, both sides have long recognized that an integration of the two perspectives is highly desirable. This intellectual advancement has led to the establishment of a new approach to risk perception, known as *cultural cognition of risk*, as an interdisciplinary endeavor that draws from several social science disciplines including psychology, anthropology, political science, sociology, and communications (Kahan et al. 2010; Slovic 2006). According to cultural cognition, people form perceptions of risk which conform with the behavior they and their peers find honorable and socially beneficial. For example, people who subscribe to individualistic values are inclined to value commerce and industry and accept or doubt environmental risks ensuing from such activities, while people who subscribe to egalitarian and communitarian values tend to regard commerce and industry as sources of inequality and are more critical of environmental risks (Kahan 2010b; Kahan et al. 2006). Experimental studies have provided strong support for this hypothesis and established a new paradigm for the study of risk perception based on attitudinal measurements (Kahan 2010a). These measurements are now starting to be used to support the creation of models and simulations of risk perception (Burns and Slovic 2007). Human computation systems can benefit from the integration of these models to manage human judgment biases due to the psychosocial amplification of risk.

Serious Gaming as a Psychosocial and Cultural Aware Human Computation System

Analytical or serious gaming provides a unique opportunity to address cognitive and cultural biases in human decision-making through the use of role-playing and gameplay. These game mechanics leverage people's natural desires for competition, achievement, status, self-expression, altruism, and closure to engage people in collaborative problem solving. For example, game logics can be used as a control mechanism to compare, contrast and measure (e.g. via scoring) different problem solving strategies. Such a control mechanism is usually be represented as a set of rules implemented by a human game master or a computer model (or a combination of the two) that regulate outcomes during gameplay. Using resources allocated to each role, and the game logic and activities, human players can update model parameters and engender new model outcomes which in turn influence the behavior of the human players in the game. The approach ensures a strong functional partnership between human players and computer models that regulate or/and predict role-play behavior, while maintaining a high degree of independence and greatly facilitating the connection between model artifacts (e.g. computational agents), human players, and game structures. The outcome of this approach is a collaborative decision-making process which exploits cognitive and cultural awareness to engage human creativity and reduce the impact of biases on human judgment.

Background

Due to their great potential as an aid to understanding complex issues, role-playing games (RPG) are currently being widely tested for learning and training purposes, and to a lesser extent for analysis and decision making. In RPGs, players endeavour to enact the roles of fictional characters within a narrative, either through literal acting or through a process of structured decision-making in which the players' actions succeed or fail according to a formal system of rules and guidelines (Sanfilippo et al. 2010; Cover 2010). RPGs can be played live as tabletop, live-action, or computer games. Tabletop RPGs are conducted through discussion, while in live action role-playing games players physically perform their characters' actions (Tychsen 2006). Both tabletop and live-action RPGS rely on a game master to administer and the rules and setting of the game and referee its outcomes. Computer RPGs exist both as

Table 1 Criteria that promote a system's ability to bridge across human judgment and machine inference (Adapted from Sanfilippo et al. 2010)

Social interaction	Games that allow multiple human players to interact (negotiate, compromise, etc.) contribute more than those that do not bridge the gap between human and machine reasoning, as human judgment is often performed as a collective activity
Adaptability & flexibility	Degree of role restriction, e.g. are the players' roles solely determined by the roles in the model? Can new roles be defined in the game by grouping model parameters at will to match the player's wishes?
	Can the players specify new model rules and parameters?
	Can the game outputs be modified to fit decision making requirements?

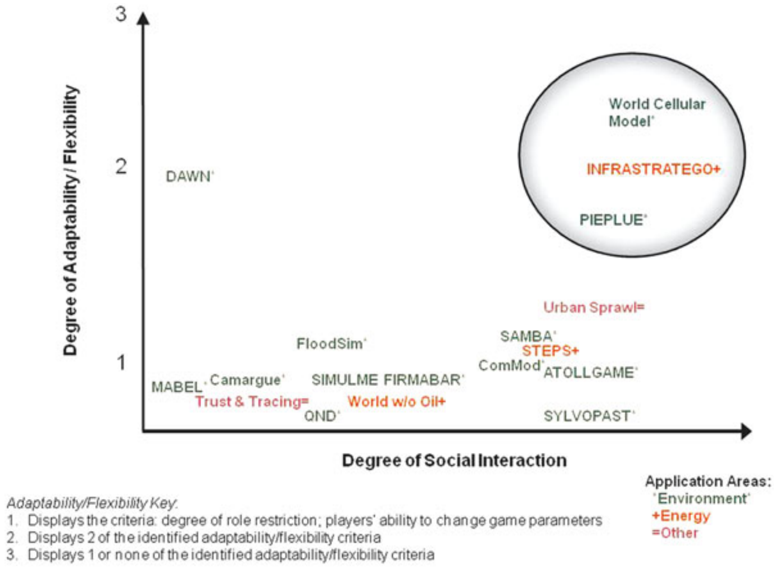


Fig. 1 Comparison of selected game systems (Adapted from Sanfilippo 2010)

multi-player games, such as massively multiplayer online role-playing games (MMORPGs), and single-player games. MMORPGs can be regarded as an implementation of tabletop or live-action RPGS, in which discussion and live performance is computationally mediated and the function of the game master is partially or fully automated.

In a recent literature survey, (Sanfilippo et al. 2010) found that approximately 32 % of 684 RPG systems examined for the period 2000–2008 address analysis and decision support. Only a small proportion of this subset focuses on how to bridge across human judgment and machine inference, with reference to criteria such as the ability of the game to facilitate interactions across human players and accommodate a player's request for a resource or outcome not represented in the game (Table 1).

Figure 1 below illustrates how a selection of the systems surveyed rank according to the criteria described in Table 1. As shown, only a few systems (those enclosed in the circle) appear to have a high degree of social interaction and at the same time exhibit some degree of adaptability/flexibility. For example, in the *World Cellular Model* (Valkering et al. 2007) multiple players interact with the goal to "survive" in a sustainable world. Players can modify some aspects of game rules through a weighted vote, include new events and drivers, and enact a game scoping stage based on previous game outcomes; however, the players' roles seem to be rigidly determined by the roles in the model. The *Infrastratego* game (Kuit et al. 2005) is highly interactive including up to 40–50 participants and players can negotiate the introduction of new rules with the game master/controller. The *Pieplue* game (Barretau and Abrami 2008) enables participatory decision making; rules are set up in advance, but new parameters can be introduced into the game.

Analytical Gaming

Analytical Gaming (AG) (Sanfilippo et al. 2010; Riensche et al. 2009, 2000) provides an environment in which analysts and decision makers can engage in interactive role-play to critique each other's ideas and action plans in order to achieve preparedness in real-world situations. AG facilitates creation and execution of games analogous to traditional tabletop simulation exercises. One application of the AG approach is to generate virtual evidence, by recording the behaviors of players, for calibrating model parameters in the absence or sparseness of real-world evidence. AG may also be configured as a collaborative and interactive interface to computer models. In constructing such environments, (Sanfilippo et al. 2010; Riensche et al. 2009, 2000) set a number of goals, including:

- Define interfaces that allow inclusion of computerized data sources (e.g., models/simulations, historic datasets) in an interactive environment.
- Define interfaces that allow display of an environmental state (informed by the aforementioned data sources) to players in ways that are naturally intuitive and realistic.
- Define interfaces by which players may interact with the environment (and by extension, other players and underlying models).
- Construct software architectures to implement these interfaces in such a way that the architectural "building blocks" are reusable across multiple distinct games.
- Leverage the use of common software architecture across multiple games to collect data regarding player actions and environmental/model states during game play, which can be used to reconstruct a history of game play(s) and to analyze the context of player actions and interactions.

As described in (Riensche et al. 2009, 2000), the abstract architecture that implements such goals includes the following notions:

- *Domain models*–the applicable computational models that we can use to drive changes in the game environment.

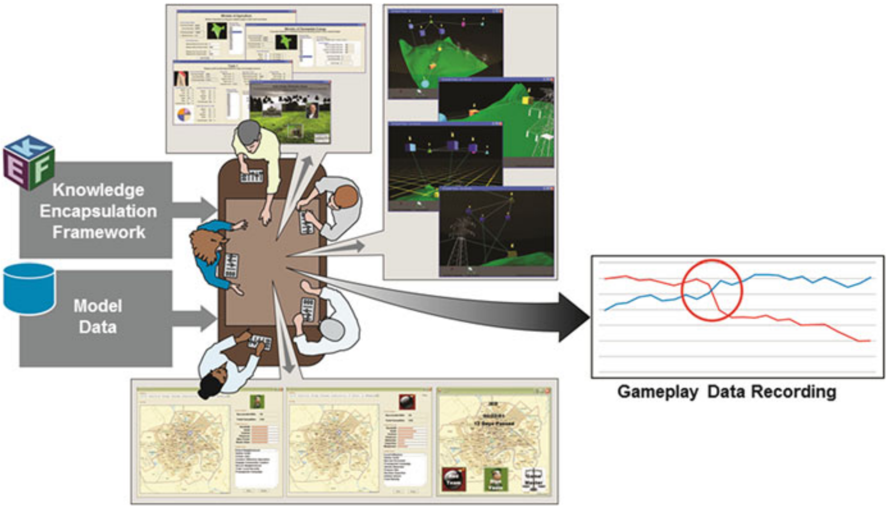


Fig. 2 Analytical gaming

- *Roles*–the roles we will ask players to assume. Identification of roles includes determining how and why a player in a particular role would be involved in the scenario represented by a game, what that player's objectives would be, and the means by which the player may influence the environment.
- *Game parameters*–the underlying data parameters that describe the state of the environment.
- *Game elements*–the user interface devices by which information is exchanged between the environment and users.
- *Handles*—the game elements which users may directly manipulate.

Figure 2 provides a graphic representation of the analytical gaming concept and its components.

Application: Illicit Nuclear Trafficking

Illicit nuclear trafficking networks are a serious security threat. These networks can directly lead to nuclear proliferation, as state or non-state actors attempt to identify and acquire nuclear weapons-related expertise, technologies, components, and materials. The ability to characterize and anticipate the key nodes, transit routes, and exchange mechanisms associated with these networks is essential to influence, disrupt, interdict or destroy the function of the networks and their processes. One of the major challenges in addressing these requirements is the lack of reliable data that can be used to develop and evaluate computational models. For example, the

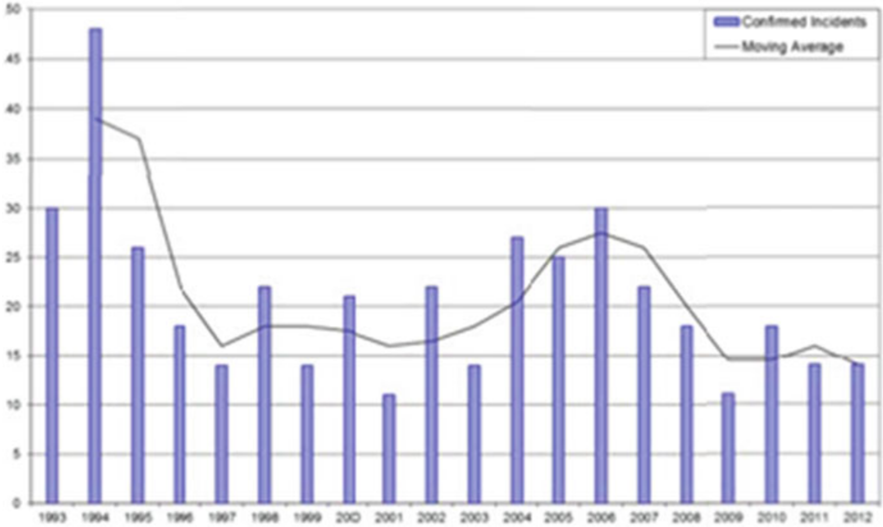


Fig. 3 Incidents reported to the ITDB involving unauthorized possession and related criminal activities, 1993–2012 (Adapted from the IAEA Incident and Trafficking Database, 2013 Fact Sheet, http://www-ns.iaea.org/downloads/security/itdb-fact-sheet.pdf)

total number of known incidents of illegal possession and movement of nuclear material or radioactive sources and attempts to sell, purchase or otherwise use such material for illegal purposes, for the period 1993–2012 is just a few hundred (Fig. 3). Consequently, the use of machine intelligence to infer and forecast patterns of illicit nuclear trafficking from historical data has limited reach. Instead of purely deductive models, we need generative models of illicit nuclear trafficking.

Sanfilippo et al. (2011) describe a prototype analytical game that provides an environment where human and machine intelligence can be jointly harnessed to meet the requirements and challenges of developing generative models of illicit nuclear trafficking (henceforth "INT game"). The INT game focuses primarily on human behavioral dynamics, in particular communications, deception, deal-making, and influencing. The game was developed using a simplified framework, where a subset of the real life contexts in which illicit trafficking occurs is selected. This methodology is akin to the practice in biological research to recreate in vitro components of an organism that have been isolated from their usual biological surroundings in order to permit a more detailed and convenient analysis than can be done with the whole organism.

Initially the game is developed as a tabletop exercise to identify and articulate game elements and their behavior. Once the structure of the game has reached maturity, it is implemented as a computer-based game. In the INT game, one player is given the objective of obtaining a set of commodities required to achieve nuclear weapon readiness (e.g. acquire uranium ore, computer and fissile core fabrication capabilities, nuclear reactor equipment, and weapon delivery systems), while some other players seek to prevent the achievement of this goal, and still other players

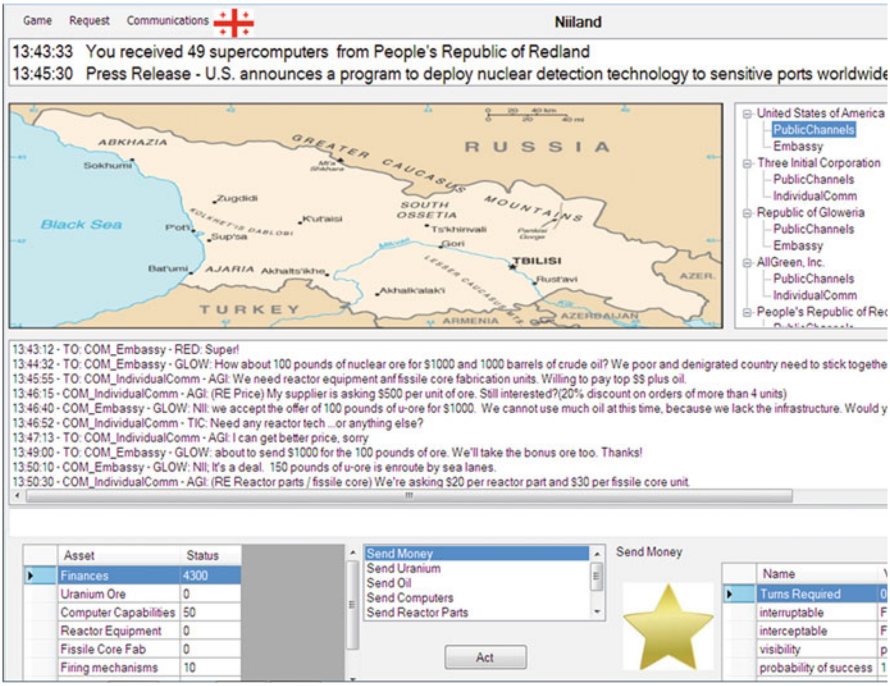


Fig. 4 A player's interface to the game (Adapted from Sanfilippo et al. 2011)

who may have nuclear commodities are motivated by other objectives such as profit within or outside the bounds of legality. Players are not told in advance who the would-be proliferator is.

Game roles include a mix of *Countries* and *Companies*. All players managed resources (*Commodities*) that were divided into three categories based on their role in the nuclear weaponization model: General Use (i.e., unrelated to nuclear weapons), Dual Use (e.g., items that could serve purposes in both nuclear energy and nuclear weapons production, and Focused Use (items that are only useful in production of nuclear weapons). Potential player actions included primarily sending of communications, attempts to intercept communications of other players, and initiating transfers of money and *Commodities*. All communications and actions were moderated by a Game Master, with whom the players could also negotiate addition of ad hoc actions.

Figure 4 provides a view of player's application screen half-way through playing a game session. The player's aim is to acquire assets (lower left in Fig. 4) which would enable the construction of nuclear weapons. In carrying out this aim, the player communicates with the other players through instant messaging to

- Use finances and other resources available to the player (e.g. crude oil) to acquire nuclear material and capabilities, and
- Cover his/her real intents to escape interception by controlling actors (e.g. the US).

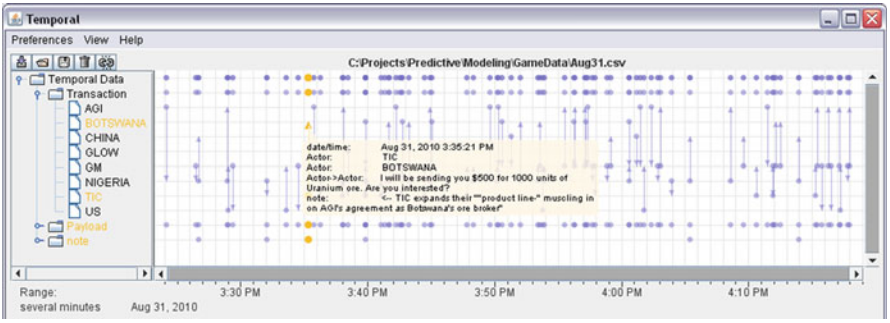


Fig. 5 Mining game results

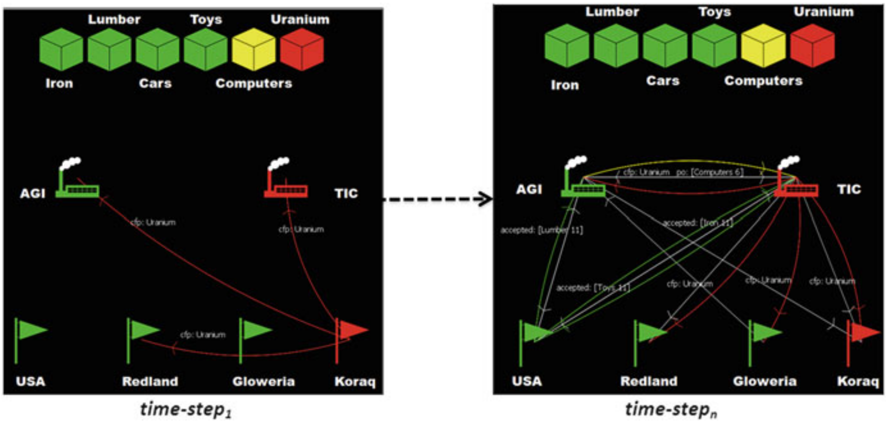


Fig. 6 Time series simulation output of an illicit nuclear trafficking agent-based model

Each player has specific roles in the game and each plays to his/her own advantages in making alliances tactically or strategically, as needed. Occasionally, in a random fashion, intelligence is leaked by the game master that reveals covert/deceptive operations. Players can negotiate with the game master authorization to perform new activities and be rewarded with new assets in the event the new activity is successfully carried out.

Every time the INT game is played, the game results are stored and analyzed to characterize the behavior of players (Fig. 5). The analysis of players' behavior is then used to calibrate an agent-based model of how the exchange of goods and know-how may play out through time series simulations with reference to developments of ongoing behaviors and the emergence of new behaviors, as shown in Fig. 6. Agent-based modeling (ABM) provides an ideal way of capturing the evolution of networking structure emerging from proliferation activities and knock-on effect from the behaviors of specific actors involved, such as the observables from the role-play activity described above.

Once the illicit nuclear trafficking ABM is sufficiently calibrated, it is integrated with the game. Model parameters, roles and activities are matched with roles, assets and activities in the game so players' behavior is both regulated by and perturbs the model's simulations as discussed in (Sanfilippo et al. 2009–2011). The emerging approach is still in its experimental stage and if successfully implemented can be instrumental in enabling analysts and policymakers to plan strategic action in influencing, disrupting, interdicting or destroying the function of illicit nuclear networks and their processes, and can be integrated with a radiation detection approach to address medium and short medium analysis and intervention objectives.

Conclusions

The integration of psychosocial and cultural processes that affect human judgment is crucial in designing human computation systems which effectively leverage and complement human intelligence. In this chapter, we have argued that gamification helps achieve such an integration. The goal of gamification is to apply gameplay to real-world problems in order to develop social intelligence through a concerted reasoning effort that exposes judgment biases and promotes creativity by interleaving human and machine intelligence. The gameplay data which results from this endeavor provide content that can be used to train and calibrate behavioral models. This is a significant achievement, especially in those domains where using historical data has limited value, either because there is not enough data available, as in the illicit nuclear trafficking problem discussed in this chapter, or because the operational context changes so rapidly, as in the cybersecurity domain. The models trained on the data generated through gameplay can be linked back to the game to increase the complexity and or level of automation of the game. This process can be repeated iteratively to develop human computation systems capable of making more complex and powerful inferences.

Partly due to its novelty, there is no shortage of challenges and opportunities for this novel endeavor. Ubiquitous access to the Internet, mobile telephony and technologies such as digital photography and digital video have enabled social media application platforms such as Facebook, YouTube, and Twitter that are altering the nature of human social interaction. The fast increasing pace of online social interaction introduces new opportunities to articulate a gamification approach to human computation systems that integrates psychosocial and cultural factors that influence human judgment. However, online behavior tends to differ from non-virtual behavior in ways that we still do not fully comprehend. Moreover, despite the great progress in understanding how humans make decision under risk, the integration of psychometric, emotive and cultural factors that impinge on risk perception is still largely unexplored. Another important question is the evaluation of game-based human computation systems. The level of human engagement elicited by these systems is certainly an important metric that can be assessed with relative ease. However, other performance metrics such as the reliability and effectiveness of the analysis and decision-making outcomes these systems generate may be harder to measure. A resolution of these challenges and the ensuing ability to reap the related benefits will largely determine the success human of computation systems based on gamification techniques.

References

Barretau O, Abrami G (2008) Variable time scales, agent-based models, and role-playing games: the PIEPLUE river management game. Simul Gaming 38(3):364–381

Burns WJ, Slovic P (2007) The diffusion of fear: modeling community response to a terrorist strike. J Def Model Simul Appl Methodol Technol 4:298–317

Cover JG (2010) The creation of narrative in tabletop role-playing games. McFarland & Company, Jefferson, NC, p 6

De Vuyst H, Clifford GM, Nascimento MC, Madeleine MM, Franceschi S (2009) Prevalence and type distribution of human papillomavirus in carcinoma and intraepithelial neoplasia of the vulva, vagina and anus: a meta-analysis. Int J Cancer 124(7):1626–1636

Douglas M, Wildavsky AB (1982) Risk and culture: an essay on the selection of technical and environmental dangers. University of California Press, Berkeley

Gigerenzer G (2007) Gut feelings: the intelligence of the unconscious. Penguin Books, New York

Gladwell M (2005) Blink: the power of thinking without thinking. Little, Brown and Company, Boston

Heuer RJ Jr (1999) Psychology of intelligence analysis. Center for the Study of Intelligence, Central Intelligence Agency, Washington, DC

Janis I (1972) Victims of groupthink: a psychological study of foreign-policy decisions and fiascoes. Houghton, Mifflin, Boston

Kahan D (2010a) Fixing the communications failure. Nature 463(7279):296–297

Kahan D (2010b) Fixing the communications failure. Nature 463:296–297

Kahan D, Slovic P, Braman D, Gastil J (2006) Fear of democracy: a cultural critique of sunstein on risk. Harv Law Rev 119:1071–1109

Kahan D, Braman D, Cohen G, Gastil J, Slovic P (2010) Who fears the HPV vaccine, who doesn't, and why? An experimental study of the mechanisms of cultural cognition. Law Hum Behav 34(6):501–516

Kahneman D, Tversky A (1973) On the psychology of prediction. Psychol Rev 80:237–251

Kahneman D, Tversky A (1974) Judgment under uncertainty: heuristics and biases. Science 185(4157):1124–1131

Kasperson J, Kasperson RE, Pidgeon N, Slovic P (2003) The social amplification of risk. Cambridge University Press, Cambridge, UK

Klein G (1998) Sources of power. MIT Press, Cambridge

Kuit M, Mayer IS, De Jong M (2005) The INFRASTRATEGO game: an evaluation of strategic behavior and regulatory regimes in a liberalizing electricity market. Simul Gaming 36(1):58–74

Miller G (1956) The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychol Rev 63:81–89

Riensche RM, Martucci LM, Scholtz J, Whiting MA (2000) Application and evaluation of analytic gaming. In: 2009 international conference on computational science and engineering, vol 4. Vancouver, pp 1169–1173, 29–31 Aug 2009

Riensche RM, Paulson PR, Danielson GR, Unwin SD, Butner RS, Miller SM, Franklin L, Zuljevic N (2009) Serious gaming for predictive analytics. In: AAAI spring symposium on technosocial predictive analytics, AAAI Press, Menlo Park

Sanfilippo AP, Cowell AJ, Malone EL, Riensche RM, Thomas JJ, Unwin SD, Whitney PD, Wong PC (2009) Technosocial predictive analytics in support of naturalistic decision making. In: Proceedings of the 9th bi-annual international conference on naturalistic decision making, Covent Garden, London, pp 23–26, June 2009

Sanfilippo AP, Riensche RM, Unwin SD, Amaya JP (2010) Bridging the gap between human judgment and automated reasoning in predictive analytics. In: International probabilistic safety assessment & management conference, Seattle

Sanfilippo et al (2011) Technosocial predictive analytics for illicit nuclear trafficking. In: Proceeding SBP'11 proceedings of the 4th international conference on social computing, behavioral-cultural modeling and prediction, Springer-Verlag, Berlin/Heidelberg, pp 374–381

Slovic P (1987) The perception of risk. Science 236(4799):280–285

Slovic P (2006) The perception of risk. Routledge, Washington, DC

Slovic P, Peters E, Finucane ML, Macgregor DG (2005) Affect, risk, and decision making. Health Psychol 24(4 Suppl):S35–S40

Starr C (1969) Social benefits versus technological risks. Science 165(389):1232–1238

Surowiecki J (2004) The wisdom of crowds. Anchor, New York

Tversky A, Kahneman D (1981) The framing of decisions and the psychology of choice. Science 211:453–458

Tychsen A (2006) Role playing games—comparative analysis across two media platforms. In: Proceedings of the 3rd Australasian conference on interactive entertainment, Perth, Australia, December 4–6, 2006, pp 75–82

Tychsen A, Hitchens M, Brolund T, Kavakli M (2006) Live action role-playing games: control, communication, storytelling, and MMORPG similarities. Games Cult 1(3):252–275 (Sage Publications)

Valkering P, Offermans A, Tàbara D, Wallman P (2007) New tools for integrated sustainability assessment: the world cellular water game. In: Conference on the human dimensions of global environmental change. Earth system governance: theories and strategies for sustainability, Amsterdam, pp 24–26, May 2007

Part VIII
Policy and Security

Introduction to Security and Policy Section

Dan Thomsen

The rewards and rules of human collaboration systems shape the behavior of the human participants, often leading to behaviors the human computation system designers never envisioned. Most software designers make the mistake of assuming that people will follow the intent of the rules they set up in the program. But, as the rising wave of cyber-crime shows, people do what they can get away with. People will do anything they can to achieve rewards, and sometimes the reward means breaking the system for the joy of figuring out how to solve a puzzle.

Most systems suffer from concentrating on the functionality of the system and the developers add security only when they realize someone has broken the rules. Developers assume that people will behave in "cyber-space" the same way the behave in the real world. Unfortunately the anonymity, and the cognitive leverage that computers give people result in a very different behavior.

Many human computation systems have altruistic goals, and will have a core assumption that people are participating to do "good". Even, if they recognize the existence of malicious users the developer's core assumption of "good" players tamps down their ability to see just how devious a malicious player can be. For example, will they seed problems to have the other human solvers help them in support tasks for a malicious goal? How will people mix in things from the real world into the human computation environment, like greed, jealousy and network packet lengths that the developer simply has not thought of?

People have postulated that reward mechanism in human computation games must be carefully constructed to reward the exact behavior desired. If the rules are poorly designed people playing within the rules may perform legal actions that result in different outcomes from what the developers intended. Now consider at the meta-rules; what the human computation system actually enforces at a low level. What parts of the system are easy to break, spoof, or compromise? Now can you predict the final outcome of a collaboration effort?

D. Thomsen (✉)
SIFT, LLC, USA
e-mail: dthomsen@sift.net

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_66,

Growing up in rural Minnesota, there was a saying, "Locks are for honest people". Meaning the people that locks stop from entering your house probably did not plan on taking your stuff anyway. We locked the door mainly to keep neighbors from leaving apples, or other produce we would have to eat or can. The one time our home was burglarized the thieves pushed the locked front door out of its frame to get into the house. Which was unfortunate since we left the back door open in case a neighbor really had to get rid of a bushel of apples. For human computation, this means two things, first you have to have set of clearly defined mechanisms to get the human behavior you want for the honest people. This first set of obvious security mechanisms guides the honest people to the desired behavior. Second, you need hard security mechanisms to enforce that behavior, detailed auditing for when the mechanisms fail, and policy to make it clear to malicious users the penalty for malicious behavior.

There is a complex dance between security and policy. Building secure software costs money and developers must always trade security for features to reduce development time, which leaves gaps in the security enforcement model. Often policy can cover these gaps cheaper than developing the necessary software. For example, developers can easily implement password authentication, but passwords do not really identify the person, only that the person knows the secret. A complete security system includes a policy, to inform the users not to share their passwords with anyone. This allows the developers to trade development costs for a weaker form of authentication. Human computation environments must make these same types of trade-offs; implementing mechanisms to encourage desired behavior, and policy to define behavior when the mechanisms are insufficient.

In this section we have five papers covering both security and policy. Felstiner writes about labor laws standards and human computation. Computation environments could be written that enforce fair labor laws for different countries, but think of the expense for simply understanding all those different laws, let alone implementing them in software. Instead developing clear policies about what labor laws mean in a global, anonymous job market can help ensure fairness with less expense. James Caverlee also discusses exploitation in human computation systems. System developers may assume that contributors are volunteering their time, but money and rewards attract middlemen who might be packaging laborers from depressed economics for a cut of their wages, making for a very different experience for the contributors than the developers imaged.

Tom Deutsch then looks at how our neurobiology plays a role in our digital interactions, specifically as it relates to privacy, and how human computations systems need to consider those dynamics to meet their own goals. Elena Ferrari and Marco Viviani look the evolution of privacy from offline to online communities and the impact for online collaboration systems. Finally in my chapter, I look at the risks involved in human computation systems. What are the assets worth protecting? How can those assets be degraded or lost? Since you never get enough budget to secure all your assets, you need to make sure the most valuable assets get protection.

The history of computer science has shown that when developers explore a new area of using computers, developers concentrate on the exciting new features and not security. Often security does not get integrated until someone suffers. With this new area of human computation we have another chance to get security right before people accept poor security practices as the norm. We can still prevent catastrophic failures in human computation by designing in the correct balance of security mechanism and policy to aid finding new solutions.

Labor Standards

Alek Felstiner

Introduction: Human Computation and Work

Human computation, as broadly conceived in this handbook, encompasses many activities we might call "labor" or "work." Where an individual makes a conscious decision to perform a task in exchange for money or other compensation, a more traditional employment relationship may arise and, consequently, various labor laws may apply. For example, online work platforms such as Amazon's Mechanical Turk distribute computational tasks to a large pool of workers who can choose to accept and perform those tasks at the advertised pay rate. Despite the novelty of the platform, the workspace, and the payment structure, this is still recognizable as "work."

Not all forms of human computation follow such a clear model for the exchange of labor. Sometimes a group of people will collaborate online to sequence a gene, examine a galaxy, or build a software model, without any expectation of compensation. They are contributing labor in some form, but we might hesitate to call it "employment." Similarly, online gamers perform human computation within the gaming context, either as participants in a "game with a purpose," or simply to enrich their own experience in the virtual environment. They gain some satisfaction, and their labor (if it can be called that) contributes to a greater whole, but their participation somehow seems to fall outside our notion of "work."

The work/non-work distinction matters only insofar as it may help to shape the discussion of what labor standards ought to apply in human computation projects. As used here, "labor standards" refers first and foremost to the legal obligations that attach to an employment relationship. Many countries regulate wages, hours, benefits, and other aspects of work. How those laws might pertain to human computational work, generally speaking, occupies the bulk of this chapter. The variety of different employment laws, across borders and in different jurisdictions, precludes

A. Felstiner (✉)
Law Clerk, United States District Court for the District of Columbia, Washington, DC, USA
e-mail: alek.felstiner@gmail.com

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_67,

drawing firm conclusions about any particular human computation project under any particular law. Instead, this chapter aims to sketch out the features of human computation and employment law that are likely to influence legal authorities and policy makers.

To that end, the discussion below assumes that "labor standards" are primarily relevant to situations in which people consciously decide to exchange their labor for some form of compensation. The people in question act not as collaborators, volunteers, or consumers, but as "workers," in some sense. Of course the categories often overlap especially in online environments such as games or shopping sites where work, recreation and consumption weave together.

One way to separate "work" from "non-work" is to look at the nature of the bargain participants make. For example, on online labor (sometimes called "crowd labor") platforms, firms (or "requesters") offer to trade some amount of compensation for the performance of a task. The product of the labor and its ultimate benefit go to the requester, while the worker receives cash, credit, or some form of virtual currency (discussed in more depth below). This fits the traditional definition of "work," and we can see how labor standards might apply to this bargain.

Online collaborations and games with a purpose, by contrast, are difficult to characterize as "work." They may involve an exchange of labor, but the bargain functions differently. Participants act out of self-interest, altruistic volunteerism, or some combination of the two, but no one anticipates any formal compensation. What participants get instead is personal satisfaction, or future benefits in the form of better software, medicine, government, etc. Labor laws almost always exempt volunteer activities of this kind, so NASA need not worry that it owes minimum wage and overtime to the image-processing volunteers who mapped Mars.[1]

It becomes difficult to draw these distinctions in online environments where work blends with recreation and consumption, as in some social games and shopping websites. Certain games allow gamers to perform small tasks on crowd labor platforms such as Mechanical Turk, in exchange for virtual currency.[2] Those arrangements remain "work," though the participant arrived at the task in pursuit of recreation and performed it to prolong or enhance the recreation. However, an online shopper who performs a survey in exchange for free shipping cannot be described as "working" without overstretching the boundaries of the definition. To remain within a safe definition of work, this chapter assumes that any "labor standards" will apply only to bargains involving a conscious decision to trade labor for compensation, with the benefit clearly conferred upon the recipient of the labor's product.

---

[1] Szpir M (2002) Clickworkers on Mars. Am Sci 90(3):226.

[2] See Galante J (2010, June 17) Crowdflower's virtual pay for digital purchases. Bloomberg Businessweek Magazine. http://www.businessweek.com/magazine/content/10_26/b4184041335224.htm. Accessed 31 May 2013; Kit Eaton (2009) Gambit lets you be a mechanical Turk for social game credits. Fast Company. http://www.fastcompany.com/blog/kit-eaton/techno-mix/gambit-lets-you-be-mechanical-turk-social-game-credits. Accessed 31 May 2013. See also Felstiner A (2012) Regulating in-game work. J Internet Law 16(2):3.

The first section examines the regulatory obstacles and unsettled legal questions that arise when work is distributed in small chunks to a global network of casual laborers. It discusses which employment laws may apply, to whom they apply, and how they might operate in light of alternative compensation models steadily gaining ground in online economies.

The second section, "Voluntary Measures," recognizes that for many the term "labor standards" means more than the minimum obligations an employer must meet to stay on the right side of the law. We have an expectation that when people contribute labor, in exchange for compensation or not, the terms and outcome should meet some threshold level of fairness—particularly with respect to those in subservient positions, with inferior bargaining power, little control, and no prospect of extracting profit. This section offers several suggestions for voluntary measures that online work platforms can take to address emerging ethical issues. Whether covered by labor laws or not, the companies that control online work platforms can act to protect their users' reputations, privacy, and dignitary interests.

The final section identifies and briefly describes some of the opportunities created by online distributed work. Too often regulatory debates come to resemble a battle between meddling authoritarians and their laissez-faire opponents. Assuming that employment law applies to these forms of work, or should apply, this last section emphasizes aspects of the work that legal authorities contemplating regulation may see fit to preserve and encourage.

Obstacles and Open Questions

Jurisdiction and Choice of Law

Though this article discusses labor law in broad terms, systems of labor regulation vary widely from country to country. Given the globality of distributed human computation, the first step of any inquiry may be to determine which legal body has or should have the authority to administer justice. This is called "jurisdiction." The authority with jurisdiction will also need to know which labor laws to apply. Usually the answer is whatever labor laws are in effect in that jurisdiction, but where the employment relationship crosses jurisdictional boundaries in some respect, the authority may face a choice of law.

With different laws in different places, jurisdiction as a threshold question can dramatically affect the rights and obligations of the parties as well as the outcome of any dispute they seek to resolve. In some cases it is quite simple: where an employment relationship exists entirely within a jurisdiction, the authority and applicable laws would readily present themselves. In the United States, for example, if an employer is headquartered and operates within a jurisdiction, the worksite is there, and the employee lives there, the local authority in that jurisdiction decides any employment dispute and determines what laws will govern. Even where an employer is headquartered outside the jurisdiction, or operates in multiple jurisdictions, the deciding authority may change but the location of the worksite will usually determine the applicable laws. Labor laws developed with this (now sometimes outdated) model of employment in mind.

As employment relationships stretch across borders, disputes become more complex. The European Union has adopted multinational labor standards, which establish baselines in certain regulatory fields such as health and safety.[3] The EU has also issued some guidance on choice of law, aimed at tackling problems that arise when employees perform work outside their normal location.[4] For many years the International Labor Organization has also set broad standards, without much resultant uniformity, even among the member states.

Given the potential for confusion, it is no surprise that the parties often settle jurisdiction and choice of law matters through contract. They may agree to resolve any disputes in a particular forum, according to particular laws. or to submit those disputes to a neutral arbitrator. Such contractual clauses avoid jurisdictional issues, provided that both parties embrace the proposed mechanism. In practice, the party with less power—invariably the employee—is effectively compelled to acquiesce in order to obtain the job, if he or she even knows about the clause at all. Once a dispute arises, the employee may discover that litigating in the chosen forum is prohibitively distant or costly, and also that laws in that forum favor the employer.

Who Is Covered?

Applying any labor standard requires at minimum two parties that the law will recognize as an employer and an employee. "Employer" and "employee" tend to function as terms of art, not as common-sense descriptors. Their purpose is to precisely carve out, from the mass of actors engaged in "labor," just those workers and bosses the law in question seeks to regulate.

Such a seemingly simple issue can quickly grow quite fraught, due in no small part to vagueness in legal definitions.[5] Distinguishing true "employees" from self-employed contractors, who sell their "services" rather than their labor, often proves tricky. In the UK, for example, the law understands "genuinely self-employed" to mean people who run their own businesses and take responsibility for success or failure, have several customers at once, control the details of their work, can hire

---

[3] See Council Directive 89/391/EEC, June 12, 1989 ("On the introduction of measures to encourage improvements in the safety and health of workers at work"); Council Directive 2003/88/EC, Nov. 4, 2003 ("Working Time").

[4] See Council Directive 96/71/EC, Dec. 16, 1996 ("Posted Workers").

[5] For example, U.S. minimum wage law unhelpfully defines "employee" as "any individual employed by an employer," and defines "employ" as "to suffer or permit to work." 29 U.S.C. §§ 203(e)(1), 203(g). Under U.K. law, employees are workers who have an express or implied contract of employment.

others to replace or assist, and provide their own equipment.[6] In the US, courts and administrative agencies use a panoply of factors, some rather intangible—such as the degree to which the worker exercises independent judgment or serves as an "integral" part of the principal's business.[7] And the factors sometimes change depending on the law at issue.

Identifying the proper employer (or employers) is normally a matter of looking at the sign above the door, or the company name on the paychecks. However, convoluted sub-contracting models, particularly in agriculture and business services, have muddied that water as well. These work arrangements often involve a series of middlemen, some of which exercise a not insignificant amount of control over the work, worksite, and compensation. Those that ultimately receive the fruits of the labor are able to disengage completely from the process.

Legal tests for "employer" and "employee" developed with a traditional employment relationship in mind. Employees had a fixed or at least identifiable worksite, a single employer, and a relatively permanent economic connection to that employer. This labor model has already begun to erode, in various ways, but the migration of work onto online platforms transforms the employment relationship beyond anything the law's original authors could have envisioned. Crowd labor changes the usual cardinality—one employer with many employees—to a many-to-many relationship. The notion of a fixed worksite evaporates in the face of globally networked crowds and proliferating mobile technology. And microtask labor can shrink the duration of an employment relationship into a single transaction, one in a stream of such small exchanges, lasting minutes or even seconds and followed immediately by another.[8] The authors of those original laws never anticipated this. To the extent the original employment law rules survive at all, legal authorities will need to adapt them, or replace them as they become obsolete.

For example, we can fairly easily imagine how to administer minimum wage laws where a single employer has recruited a group of workers, in different locations and with different shifts, to screen photos or perform sentiment analysis. The only thing that would distinguish this form of human computation from any other traditional employment relationship is the human computation aspect. By contrast, if those workers are more distributed, and anonymous, if they have control over what tasks they accept, if they work for multiple (or nested) requesters, if their supervision is replaced with engineered redundancy—in short, if their labor is disintegrated—we still need to know whether the law applies.

---

[6] See HM Revenue and Customs (2013) Work out if you're employed or self-employed. http://www.hmrc.gov.uk/working/intro/empstatus.htm. Accessed 31 May 2013.

[7] U.S. Dept. of Labor, Wage and Hour Standards Division (2009) Fact Sheet #13: Employment relationship under the Fair Labor Standards Act. http://www.dol.gov/whd/regs/compliance/whdfs13.pdf. Accessed 31 May 2013.

[8] Pontin J (2007, Mar. 25) Artificial intelligence, with help from the humans. New York Times, p. 35; Felstiner A (2011) Working the crowd: employment and labor law in the Crowdsourcing industry. Berkeley J Employment Labor Law 32:143.

As with jurisdiction, there are no easy answers to that question. Authorities may try to simplify by relaxing existing legal standards. For example, the question of who exerts "control" over the relationship dominates the employee/contractor distinction in US and UK employment law, but perhaps that factor matters less in a distributed work environment. Authorities might also introduce some flexibility into the "employer" definition to account for the convoluted relationships between the multiple online entities that compensate and exert their influence over the "crowd" as it works.[9]

Legal authorities will have to balance expedience, fairness, and inclusion. If they draw the employer/employee definitions in full deference to the painless and uncomplicated application of the law, many workers in the grey areas will effectively lose protection. On the other hand, if the definitions become so inclusive as to disappear, even more than they already have, no one can apply the law or enter a labor market with any certainty. An ideal balance would adapt and replace in service of the original regulation's purpose(s)—with a steady eye toward the protections the original law aimed to provide and the attendant coercions it deemed necessary.

Compensation

Many jurisdictions impose restrictions on compensation and hours, including wage floors, overtime, sick pay, and parental leave. All these require an initial determination of the covered employee's work time and compensation rate. Where employees receive monetary compensation at an hourly rate, the legal questions are fairly simple, even if the employees work unusual schedules in locations spread across the globe. Piece-rate compensation traditionally occurred in textiles and agricultural labor, but now also predominates in online crowd labor. The piece-rate system complicates the math, but leaves the legal questions unchanged. For example, employees performing piecework in the US are still entitled to the hourly minimum wage and overtime, calculated using their total work hours and compensation during the workweek.[10]

Knottier legal problems arise in virtual compensation—that is, compensation in representational forms of currency or other virtual assets. Ten or fifteen years ago the law would perhaps have treated virtual compensation as a hypothetical question, but the recent surge in social networking and online games has made it an undeniable reality. Virtual currency has become a multi-billion dollar industry.[11] Demand for virtual assets continues to grow, and those who cannot afford or choose not to

[9] Felstiner A (2011) at 174–76; Felstiner A (2012) at 11–12.

[10] See U.S. Dept. of Labor (2009) Employment law guide: minimum wage and overtime pay. http://www.dol.gov/compliance/guide/minwage.htm. Accessed 31 May 2013; 29 C.F.R. § 778.111.

[11] Eldon E (2011, Dec. 7) US Virtual Goods Market To Hit $2.9 Billion In 2012, With Facebook games maturing, mobile booming. Techcrunch. http://techcrunch.com/2011/12/07/us-virtual-goods-market-to-hit-2-9-billion-in-2012-with-facebook-games-maturing-mobile-booming/. Accessed 31 May 2013.

buy the currency outright have found bustling labor markets in which to earn it. Many will earn virtual currency to enable some other online pursuit, such as gaming or online shopping. Some, online gamers perform microtasks in exchange for virtual currency, which they can then use to buy virtual goods or execute trades in their game's virtual economy. Crowdflower, a leading microwork vendor, estimates that it makes half of its payments to crowd workers in virtual currency.[12]

The first question is whether an employer can legally pay its employees in virtual currency. Laws generally require employers to pay employees in cash or its equivalent.[13] In the US, federal law actually prohibits employers from compensating employees using scrip, coupons, credits, or similar devices.[14] Virtual currency would seem to fall into one of these categories. Imaginary gold coins or space credits are not cash, or the equivalent of cash. Even where a recipient of virtual assets can immediately redeem those assets for "real money," perhaps on an informal exchange or grey market, the potential liquidity of the virtual asset does not make it equivalent to cash. Thus, virtual compensation likely runs afoul of any law requiring cash payment, and would probably cause further violations in jurisdictions (such as the US) that prohibit payment in scrip.

In the future, virtual currency may permeate real world economies such that it becomes functionally integrated with real currency, at which point legal authorities might see fit to relax the "cash or equivalent" standard. However, even leaving aside specific anti-scrip laws, a question still remains as to whether virtual currency as it exists now can actually qualify as compensation. The issuers of virtual currency have a strong interest in keeping it captive, controlled by the issuer and used exclusively within the system in ways that stimulate the virtual economy and contribute directly or indirectly to the issuer's profits. The issuer has no immediate reason to deal in real currency if gamers and shoppers will happily seek and accept a captive and proprietary virtual version.

One common practice is for the issuer to designate virtual currency as a form of "license," rather than a form of property.[15] In other words, the issuer gives the recipient a right to use the virtual asset, but does not relinquish any claim of ownership over the value that might be derived. By implication, or sometimes by explicit agreement, the issuer retains the right to revoke that license at any time. The issuer may also reserve the right to void the currency, change its value, or alter its permissible uses. Take for example the terms of service that accompany Zynga's popular social networking game, Farmville:

> You understand that while at times you may "earn" "buy" or "purchase" (a) virtual currency, including but not limited to virtual coins, cash, tokens, or points, all for use in the Service; or (b) virtual in-game items (together with virtual currency, "Virtual Items"); these

[12] Mahajan N (2010, Nov. 5) CrowdFlower gets gamers to do real work for virtual pay. Mission Local. http://missionlocal.org/2010/11/crowdflower/. Accessed 31 May 2013.

[13] See 29 C.F.R. § 531.27;

[14] 29 CFR 531.34.

[15] Felstiner A (2012), p. 15–16.

> real world terms are only being used as shorthand. You do not in fact "own" the Virtual Items and the amounts of any Virtual Item do not refer to any credit balance of real currency or its equivalent. Rather, you may purchase a limited license to use the Service, including software programs that occasionally manifest themselves as these items. The purchase and sale of the limited license referred to in these Terms of Service is a completed transaction upon receipt of your direct payment or redemption of a Zynga game card or a third party virtual currency like Facebook Credits. Any "virtual currency" balance shown in your Account does not constitute a real-world balance or reflect any stored value, but instead constitutes a measurement of the extent of your license….
>
> …Zynga reserves the right to stop offering and/or supporting the Service or a particular game or part of the Service at any time either permanently or temporarily, at which point your license to use the Service or a part thereof will be automatically terminated or suspended. In such event, Zynga shall not be required to provide refunds, benefits or other compensation to users in connection with such discontinued elements of the Service.[16]

Has someone actually been "paid" when he or she performs work and receives in return a limited license to use a virtual asset, revocable at the payer's will? Not by any current legal or common-sense definition of a "wage." Assuming the worker qualifies as an employee, and the payer as an employer, no amount of "licensing"-type language can erase the employer's obligation to pay for the work in a form of currency that counts. This is because workers generally cannot waive or sign away their right to minimum wage. For virtual compensation to satisfy existing wage and hour laws, employers will need to relinquish some of their control over the currency, at least with respect to that portion earned as wages. Issuers could still retain control over virtual currency gamers acquire through other means, such as by direct purchase, game play, or acceptance of non-work promotional offers (surveys, subscriptions, etc.). However, the issuer would need the ability to distinguish virtual currency earned through work from virtual currency otherwise acquired.

Let us assume for the sake of argument that virtual compensation will satisfy the law. Two further questions arise: (1) how can an employer determine work hours?, and (2) how do we properly value virtual currency? Emerging technology has made the first question fairly simple to answer. Though employers may argue that online work platforms make it impossible to monitor when an employee is actually working, existing systems have made tracking of activity within a virtual workspace as easy—or easier—than supervising employees on the proverbial factory floor. Employers can monitor keystrokes and cursor activity, and can automatically log employees out during inactive periods. Developing technologies are likely to refine and automate remote monitoring, further antiquating the notion of a foreman walking around with a stopwatch and a clipboard. Employers can also set certain quotas, and though an employer must pay an employee for time worked even if the employee fails to meet a quota, employers have no obligation to continue to employ or hire anyone who falls short. To a certain extent, the risk of fraud is a trade-off for the advantages that come with a remote, 24-hour workforce. And as discussed above, even that risk may prove illusory.

[16]Zynga (2012) Terms of service. http://company.zynga.com/ about/legal/terms-of-service. Accessed 31 May 2013.

Valuing virtual currency is not so simple. Wage and hour laws usually take the form of a minimum rate of pay, and virtual currency comes in all names, denominations, and orders of magnitude. Knowing that an employee received (or even that an employer explicitly guaranteed) a minimum "150 gold pieces" per hour of work, for example, tells us nothing about the employer's compliance with wage and hour laws. We need a fair and uniform method of valuing virtual currency—and not just to apply employment laws. The high trade volume of virtual assets engenders difficulty in many areas of law, from tax to tort.[17]

Where a game developer, social network, or crowd work vendor also sells the currency directly, there is already an exchange rate to apply. Authorities would have to implement it carefully given the currency control issues discussed above, and the conflicts of interest inherent when an employer controls the value of the currency with which it pays its workers' wages. But a formal exchange rate would at least offer authorities something to work with. Currency exchanges, whether formal or informal, would give authorities another reference point for valuation. It is also possible that some individual issuer or group of issuers might develop a universal virtual currency, as a way to reduce cross-platform friction and avoid individualized regulatory compliance costs. But the current incentives to keep virtual currencies proprietary make such a scenario unlikely in the near future.

Ensuring that workers receive their legal wage will require of legal authorities some brave estimation and perhaps a few shaky assumptions. Virtual currencies are so versatile and ubiquitous that the law will have to deal with them one way or another.[18]

Voluntary Measures

As human computation is still in its relative infancy, it is appropriate to ask not only what the legal authorities *may* do to enforce existing labor laws, but also what the putative and prospective employers *should* do in the general interest of fairness and decency. These are not all ethical obligations, precisely. But they implicate the ethical concerns inherent in any employment relationship, as well as the particular issues that arise in human computation. In fact, the recommendations below apply to the non-work forms of human computation as much as to compensated work. They also concentrate on aspects of the relationship that implicate labor specifically, and thus issues of privacy, intellectual property, torts, or criminal offenses do not appear though such issues certainly exist in human computation.

[17] See Camp B (2007) The play's the thing: a theory of taxing virtual worlds. Hastings Law J 59:1; Lederman L (2007) Stranger than fiction: taxing virtual worlds. N.Y.U. Law Rev 82:1620; Seto T (2009) When is a game only a game?: taxing virtual worlds. U. Cincinnati Law Rev 77:1027.

[18] In fact, looking at the wider landscape of virtual economies, it may prove easier to manage legal challenges associated with virtual currency than to confront the legal ramifications of other virtual assets and transactions. After all, modern currency is by nature notional and representative, making for a thin barrier between its virtual and "real" forms

Reputation-Building and Portability

The distributed and disintegrated character of online commerce and online labor exchange has amplified the crucial ways we measure trust and reputation on online platforms.[19] Most stakeholders rely in some way on reputation information when deciding with whom to do business, whether that business is e-commerce or online work. Online work platforms tend to feature worker reputation systems, but workers usually cannot carry their reputations from platform to platform, and the structure of the reputation system may create coercive penalties.

In the human computation arena, especially on distributed work platforms, everyone has incentive to make worker reputations buildable and fair. Those performing the work want their experience and expertise recognized, being otherwise anonymous and indistinguishable by virtue of the distributed work model. Favorable reputations allow workers to beat out other applicants and qualify for more specialized tasks. Meanwhile, those requesting the work have a corresponding interest in being able to identify experienced and qualified workers, because their other methods of doing so—pre-training every worker, or assigning work and assessing its quality at completion—require investments with no guaranteed useable return. Some requesters concerned with quality control will just give up trying to identify "good" workers and build in sufficient redundancy to allow for quality drops, but this method increases waste. Having trustworthy reputation ratings would go a long way in combatting such inefficiencies. And finally, the companies that build the work platforms have an interest in making sure such reputation systems exist and function reliably because a reliable reputation system makes the work platform more attractive to requesters and workers alike.

There is no reason to limit reputation systems to workers. Similar incentives exist with respect to requester reputations. Where deception and exploitation are prevalent, or where labor shortages occur, all fair-dealing stakeholders benefit from a system that would allow workers to consider the reputation of their potential employer before accepting a task. On Amazon's Mechanical Turk, for example, requesters can reject any work deemed unsatisfactory, even if the work actually meets the specifications. Crafty requesters may also use misleading descriptions to lure workers into accepting a task, at which point workers may feel compelled to complete the task in order to avoid the reputation damage that results from abandoning it. A requester reputation system might allow workers to avoid or even weed out unscrupulous requesters. In fact, one such user-generated ratings system for requesters on Amazon Mechanical Turk has existed since 2009.[20]

Reputation systems should allow portability as well. Though reputation has become a dominant force on certain platforms, it tends not to carry from one

---

[19] See Zittrain J (2008) Ubiquitous human computing 1–2 (Univ. of Oxford Legal Research Paper Series, Paper No. 32, 2008). http://ssrn.com/abstract=1140445. Accessed 31 May 2013.

[20] Irani L, Silberman M (2013) Turkopticon: interrupting worker invisibility in Amazon mechanical Turk. http://www.ics.uci.edu/~lirani/Irani-Silberman-Turkopticon-camready.pdf. Accessed 31 May 2013.

platform to another.[21] Making reputations portable is perhaps a slightly harder sell for certain stakeholders. Those who build the platforms have no obligation to embrace cross-platform reputation systems, and may reject portability mechanisms to preserve the competitive edge they earned by investing in a proprietary system. However, online workers may soon expect to see a reputation system in place, and will have put considerable efforts and resources into cultivating their own reputation on other platforms. Why would they join any platform that forces them to rebuild their reputations from scratch? And why would requesters choose a platform with limited single-platform reputation information over one that supplied a worker's entire cross-platform history? At that point, the platform operators' desire to attract workers and requesters may overcome any incentive to defend a proprietary reputation system.

Transparency

Some employment laws require employers to maintain employment records, and in certain cases make those records available to employees. However, those laws generally apply with respect to employees only, and the required recordkeeping may cover only wages and hours worked. A more comprehensive transparency policy would benefit employees and non-employees, and should encompass not just payment information, but also assignment descriptions, instructions, communications, and any other data related to the work. This would allow employees to keep track of their work for personal and tax purposes, substantiate their claims during disputes, and track their relationships with particular requesters over time. A transparent platform would provide employees not just a dashboard snapshot of their current and past work, but instead a kind of virtual desk and file cabinet.

Disclosures

The question of how much workers should or need to know about the work they perform is hardly unique to human computation. Almost every "real world" industry has succumbed in various ways to subcontracting, with its attendant opacities. These differ in no material way from the opacity created by disintegrating a large process into bite-sized pieces for human computation. In fact, one could argue that the absence of layering and the potential ease of lateral communication among workers actually increase the likelihood that workers will understand the nature and

[21] See Zittrain J (2008) Ubiquitous human computation. Oxford legal studies research paper no. 32, 6; Kumar S, Koster P (2009) Portable reputation: proving ownership of reputations across portals. Paper presented at the 2009 European context of awareness and trust (EuroCAT 2009), 3rd Workshop on combining context with trust, security, and privacy.

consequences of what they do. Nevertheless, at present the people performing work on crowd labor platforms have little expectation of disclosure. This section proposes that, under certain circumstances, the engineers or initiators of a computation project are obligated by the nature of the bargain to disclose the project's purpose(s).

First, if requesters are paying for the work, one might argue that they have also purchased the right to keep close to their proverbial chests any matters outside the scope of the arrangement. The workers have bargained to exchange labor for compensation, and an employer's fulfillment of that bargain does not necessarily include satisfying the workers' curiosity. In Anglo-American contract law, "consideration" is the legal term for what one party promises to another party in exchange for performance of the contractual obligations. The employer's consideration is generally limited to compensation paid to the employee for the work. Any argument for disclosure in addition to compensation would have to rely on vague notions of the worker having also earned an extra-contractual right to knowledge. Yet common decency would seem to require at least that prospective employers not lie, affirmatively or by omission, about what they plan to do with the product of the work. Beyond that, the onus seems to rest on both parties to determine how much knowledge they need and are willing to give in order to feel comfortable executing the bargain.

By contrast, where volunteers perform the computation, the requesters have a heightened obligation to disclose the nature of the project. As discussed above, volunteer labor falls outside the category of "work" precisely because the nature of the bargain involves a clear sense, on the part of the performers, of what they are getting, and to what they are contributing. Participants in a "game with a purpose" should understand or at least have access to that purpose. And those who sign up to scan satellite imagery or analyze online comments should know what their contributions may enable and with whom they may be shared. Otherwise the bargain is far from what it seems, and verges on fraud. This heightened obligation also applies where the engineer of a computation project is also in business performing the same work, and is essentially using volunteers to replace paid labor (dubious ethicality aside, such an arrangement could actually prove illegal). Finally, the disclosure imperative applies even (or perhaps especially) in situations where participants have no idea they are participating at all, such as the tasks users may perform to access a website. Volunteers deserve to fully comprehend the role they agree to play.

Other Dignitary Interests

Loosely defined, a worker's "dignitary interest" means his or her interest in receiving respect, preserving a sense of self, and remaining free from distress, humiliation, and degradation. Employers and requesters have no obligation to consider the dignitary interests of the people who perform human computation on their behalf, but doing so would make online computation platforms more appealing and hospitable.

Promoting dignitary interests can and should take a variety of forms. Where possible, platform designers should maintain the privacy of communications made through the platform. They should also provide forums for discussion and collaboration, not just to streamline work but to foster community. If disputes arise, workers should have some procedurally fair method to mediate and resolve them. Firms in the human computation industry should endeavor to treat and refer to the performers of that computation as people, with agency, and not as scalable units of commoditized labor. These policies should remain in effect even in internal communications and marketing, as such language and worldview tend to self-propagate.

Finally, though it may contradict the notion of the fungible workforce—a stream of anonymous, interchangeable workers that can be turned on and off like a faucet—designers of online distributed work platforms should attempt to involve workers in governance. They should solicit and respond to opinions, perhaps appoint ombudsmen or advisory committees, and even cede certain areas of decision-making to the collective. Or, at least, those in charge should remain open to the possibility of democratic developments and willing to embrace changes as they emerge. That means not taking retaliatory action to shut down dialogue or dissent, and not necessarily using inherent authority over the "walled garden" to promote a vision of the workforce that best suits the business model. In the long run, this kind of flexibility may lead to a more loyal, engaged community of participants. In the short term, it serves to acknowledge and reward workers' own investment, and recognize that although humans may replace computational processes, they are not computers.

Opportunities

If and when authorities make the legal interventions described above, they need not throw the baby out with the bathwater. It is possible to regulate human computation without destroying it. More important, there are aspects of human computation, unique opportunities it presents, that legal authorities should keep in mind and seek to promote.

First, distributed online work creates low-friction, low-cost avenues for transnational organizing and solidarity. Workers from different countries and circumstances perform labor on the same platforms, connecting directly with the same requesters and competing in the same labor market. This usually drives down wages from the requester's perspective, and many stakeholders count on exactly that outcome, but it also allows for unprecedented coordination. These platforms, at least in theory, eliminate many of the social and institutional barriers that would otherwise prevent workers from organizing up and down on a subcontracted supply chain. Any legal regime that fussily parcels off globally distributed work according to outdated jurisdictional boundaries risks destroying those budding solidarities.

Second, the flatness of distributed work exists in part because the barriers to entry are so low: broadband and a rudimentary laptop will suffice, with perhaps minimal remote or onsite training. We should celebrate the ways that distributed work allows marginalized workers with few opportunities to participate directly in a global and, relatively speaking, lucrative labor market. Non-profits have already begun using crowd labor to combat poverty.[22] This aspect of distributed work is precious, and deserves consideration in the discussion of whether and how to regulate online labor. For example, establishing basic wage protections will inevitably involve some oversight, and resultant bureaucracy, but authorities should take care not to impose such onerous compliance costs that the middlemen this work model recently banished have a new opening to re-enter the supply chain. There is no good reason to rebuild those institutional and infrastructural barriers.

Third, distributed work offers unique insights into how labor markets function. The more legible these online labor markets are, the more researchers, policymakers, investors, and workers themselves can glean from the mountains of data produced. Some scholars have already embraced work platforms such as Amazon's Mechanical Turk because of its low cost, large sample size, and flexibility.[23] This is not to suggest that online labor markets provide a perfect or even suitable analogue for real world markets in any particular situation, but given the likely growth in online labor of all kinds, the knowledge we gain from emerging work platforms could prove crucial in developing the next generation of online work. For example, understanding what motivates workers and makes transactions more efficient could help non-profits and state agencies encourage participation in areas where poverty, war, or climate have eliminated other sources of income.

Finally, one hopes that authorities will take into account the insight that regulating distributed online work could provide into the regulation of more established industries. We have an opportunity to rethink not just how employment regulations apply to online work, but how they apply generally to the modern economies that only vaguely resemble industrial economies of the past. As we face new legal questions, we also have a chance to re-interrogate the principles and assumptions that undergirded labor law in the now-archaic days of the traditional employment relationship. These new work platforms are but an extreme example of the ways our workplaces have changed, and legal recognition of those changes, across the board, is well past due.

---

[22] See, e.g., Samasource (2013) http://www.samasource.org. Accessed 31 May 2013.

[23] See Horton J, Rand D, Zeckhauser R (2011) The online laboratory: conducting experiments in a real world labor market. Experimental Econ 14(3):399–425.

Exploitation in Human Computation Systems

James Caverlee

Introduction

Much of this handbook has been devoted to the positive potential and possibilities unlocked by human computation systems. From specialized systems like Ushahidi (for crisis mapping), Foldit (for protein folding) and Duolingo (for foreign language learning and translation) to general-purpose crowdsourcing platforms like Amazon Mechanical Turk and Crowdflower—these systems have shown the effectiveness of intelligently organizing large numbers of people, and suggest a rich future for next generation human computation systems.

In this chapter, we turn our sights to the negative aspects of these systems. How may participants be exploited by human computation systems? How can these systems be used as a means to exploit other populations? What are the existing types of exploits and what types of exploits does the future hold? Beyond characterizing the threat horizon, we also consider efforts toward detecting exploits in human computation systems. And what are steps that can be taken toward mitigating the risk as these systems continue to mature?

Exploitation Within a Human Computation System

In this section, we present opportunities for exploitation within a human computation system. We consider exploits that target workers (who actually perform jobs), exploits that target requesters (who solicit jobs), and finally exploits that target the system as a whole. This taxonomy is intended as an initial organization of some of the exploits facing human computation systems, and should not be considered comprehensive.

J. Caverlee (✉)
Department of Computer Science and Engineering, Texas A&M University,
College Station, TX, USA
e-mail: caverlee@cse.tamu.edu

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_68,

Exploits Targeting Workers

Human computation systems rely on the support of *workers* who are tasked with supporting the overall efforts of the system.

Misrepresentation of the task. In many human computation systems, the overall effort is subdivided into smaller chunks that may be handled by individual workers. For example, a system to automatically recognize animals captured on video may provide each individual worker with access to only a few key frames from a handful of videos. As a result, the worker may have only incomplete knowledge of the ultimate goal of the task. In many settings, this incomplete knowledge is uncontroversial.

However, this compartmentalization of task knowledge may lead to workers agreeing to participate in human computation systems where the overall effort is contrary to the worker's moral, ethical, or religious grounding. For example, Jonathan Zittrain characterized this exploit as such:

> You might synthesize a new chemical that winds up being used as a poison or in a bomb. Iran's leaders could ask Turkers to cross-reference the faces of the nation's 72 million citizens with those of photographed demonstrators. Based on Mechanical Turk's current rates, Repression 2.0 would cost a mere $17,000 per protester. (Zittrain 2009)

Another example of abusing workers morals via task misrepresentation would be saying you are tracking elephant movement supposedly for conservation, but the data is used by poachers. In this way, workers may become cogs in a machine that works counter to their own interests.

Exposure to unwanted risks. Even for tasks that are agreeable to a worker, a worker in a human computation system may be exposed to risks that go beyond their reasonable expectations. In one direction, a worker may be exposed to disturbing content (say, via an image labeling task). In a separate direction, a worker may encounter misinformation spread through an otherwise legitimate task. For example, a worker may be asked to label blog posts as containing evidence of propaganda or not; through the labeling process, the worker may encounter deliberately placed misinformation designed to change the worker's perceptions of a particular candidate or political issue (e.g., climate change). Such a risk is similar to "push polling" in traditional political campaign surveys whereby a polling question is deliberately constructed to persuade (or even mislead) a respondent.

In addition to the cognitive risks of exposure, workers and their computing systems may also be subject to spam, malware, and phishing (Jagatic et al. 2007) attacks that have shown a remarkable ability to migrate to emerging systems. From email to Web to social media, and eventually to human computation systems, malicious users have shown great ability to target new populations.

Privacy leakage. Workers in a human computation system may also subject themselves to potential loss of privacy. A recent study has found that Amazon Mechanical Turk—designed to be an anonymous system—leaks private information of workers by using a single unique identifier for all Amazon accounts (Lease et al. 2013). In this way, a worker's anonymous Mechanical Turk account can be linked to the same worker's Amazon profile page, which could reveal personally identifying information. Beyond the direct negative consequences of privacy leakage (e.g., loss of user anonymity, targeted attacks on individuals de-identified), a worker's willingness to participate in human computation systems may be limited if there are perceived risks of privacy leakage.

Unsatisfactory compensation. The final exploit has been widely recognized as a potential threat in the increasingly globalized virtual workforce enabled by human computation systems (Ross et al. 2010). By drawing on workers from low income countries, there is the potential for exploitation of disadvantaged workers.

Exploits Targeting Requesters

On the other hand, there are threats to the requesters in human computation systems (or to the overall operators of the system).

Competitive disruption. In the 2009 DARPA Red Balloon Challenge, the winning MIT team reported that some participants deliberately falsified balloon sightings, whether to disrupt the overall functioning of the overall requester goal (find all of the balloons) or to disrupt the balloon sightings of competitor teams (Tang et al. 2011). In this way, groups of workers within a system or a competitor system itself may negatively impact the functioning of a target system by delaying task completion time, by degrading the quality of work being done (say, through deliberately inserting misinformation), and by adding uncertainty to the overall reliability of the system.

Poor quality work. One of the key concerns for requesters using existing systems like Amazon Mechanical Turk's crowdsourcing marketplace is the quality of work provided by workers. It is possible that the quality of work provided by workers is of lower quality than advertised by the human computation system: e.g., workers, regardless of incentivization scheme, may choose to complete as many tasks as possible while exerting little effort. For example, in a task that is answered using multichoice options, the worker might randomly select answers, or in case of tasks that require answering verbosely (review of a product, comparison between two products) the worker might use generic answers or answers off of Internet to complete the task quickly. In addition to this, another reason for poor quality of work on a human computation system could be because of the "one size fits all" expectation that requesters have of the system. The requester might observe a mismatch in worker skills between what the system can provide and what they are expecting. For example, a human computation system might mostly have English speaking workers, but a requesters task might need knowledge of Chinese that the system might not be able to satisfy. Existing systems (like Amazon Mechanical Turk) do include capabilities to track worker performance across tasks, to filter participants by native language, and other "checks and balances" to overcome some of these quality issues. However, as human computation systems increase in variety and capabilities, maintaining quality work will be a fundamental challenge.

Privacy leakage. As in the case of workers on a human computational system, a requester's privacy may be leaked. This can be either due to the design of the computation system itself, for example, a work requester's Mechanical Turk account being linked to his Amazon profile page, or it could be because of the information that the requester inadvertently added to the task like his email, company he works for, and so on. A requester's privacy leakage could result in reduced quality in work. For example on Amazon's Mechanical Turk, the requester has the right to pay or not pay a worker for the task completed based on the worker's quality of work. A worker who knows requester's details could answer the task in a biased way (praise requester's company, prefer requester's approach while two items are compared) so as to impress the requester without actually doing the task correctly. Also in case the requester rejects worker's work then a worker who knows requester's contact details can get in touch with him requesting the details or even threaten him.

Exploits Targeting the System Itself

Finally, human computation systems themselves may come under threat by external parties interested in degrading the quality of online information and threatening the usefulness of these systems. Traditional denial of service, spam, and other targeted attacks can be modified to disrupt the reliability, quality, and timeliness of human computation systems.

Exploits Targeting External Populations

In this section, we consider opportunities for malicious users to leverage human computation systems to target external (outside of the system) populations. We couple this treatment with a study of the prevalence of one type of exploit (crowdturfing), and consider additional exploits.

Crowdturfing

One growing threat is the emergence of "crowdturfing" (crowdsourcing + astroturfing), whereby masses of cheaply paid shills can be organized to spread malicious URLs in social media, form artificial grassroots campaigns ("astroturf"), and manipulate search engines. One example is the development of sites like SubvertAndProfit (www.subvertandprofit.com), which claims to have access to "25,000 users who earn money by viewing, voting, fanning, rating, or posting assigned tasks" across social media sites. These campaigns are being launched from commercial crowdsourcing sites, potentially leading to the commoditization of large-scale turfing campaigns. In a recent study of the two largest Chinese crowd-sourcing sites Zhubajie and Sandaha, Wang et al. (2012) found that ~90 % of all tasks were for crowdturfing.

Evidence of Crowdturfing

To illustrate the impact of crowdturfing, we report here a brief study of 505 campaigns collected from 3 popular Western crowdsourcing sites that host clear examples of crowdturfing campaigns: Microworkers.com, ShortTask.com, and Rapid-workers.com during a span of 2 months in 2012. Almost all campaigns in these sites are crowdturfing campaigns, and these sites are active in terms of number of new campaigns. Note that even though Amazon Mechanical Turk is one of the most popular crowdsourcing sites, we excluded it in our study because it has only a small number of crowdturfing campaigns and its terms of service officially prohibits the posting of crowdturfing campaigns. For the 505 sampled campaigns, each has multiple tasks, totaling 63,042 tasks. Based on a manual assignment, we found five major crowdturfing campaign types:

Social Media Manipulation (56 %). The most popular type of campaign targets social media. Example campaigns request workers to spread a meme through social media sites such as Twitter, click the "like" button of a specific Facebook profile/ product page, bookmark a webpage on Stumbleupon, answer a question with a link on Yahoo! Answers, write a review for a product at Amazon.com, or write an article on a personal blog.

Sign Up (26 %). Requesters ask workers to sign up on a website for several reasons, for example to increase the user pool, to harvest user information like name and email, and to promote advertisements.

Search Engine Spamming (7 %). For this type of campaign, workers are asked to search for a certain keyword on a search engine, and then click the specified link (which is affiliated with the campaign's requester), toward increasing the rank of the page.

Vote Stuffing (4 %). Requesters ask workers to cast votes. In one example, the requester asked workers to vote for "Tommy Marsh and Bad Dog" to get the best blue band award in the Ventura County Music Awards (which the band ended up winning!).

Miscellany (7 %). Finally, a number of campaigns engaged in some other activity: for example, some requested workers to download, install, and rate a particular software package; others requested workers to participate in a survey or join an online game.

Other Example Exploits

Propaganda. Crowdturfing can be leveraged for spreading misinformation and propaganda. For example, it has been recently reported that Vietnamese propaganda officials deployed 1,000 propagandists to engage in online discussions and post comments supporting the Communist Party's policies (BBC 2013). Similarly, the Chinese "Internet Water Army" can be hired to post positive comments for the government or commercial products, as well as disparage rivals (Wired 2010). Mass organized crowdturfers are also targeting popular services like iTunes (Gizmodo 2012) and attracting the attention of US intelligence operations (Guardian 2011).

Coordinated attacks. By exploiting collaboration to solve problems, newly engineered human computation systems could create novel ways of perpetrating crimes, acts of war, and other attacks. As illustration of this potential, in February 2013 a criminal syndicate infiltrated a credit card processing company, raised the withdrawal limits of ATM cards, and then distributed these ATM cards to dozens of participants around the world to simultaneously withdraw $45 million. Now imagine a similar attack coordinated via a human computation system whereby thousands of participants collaborate in a similar fashion. Beyond criminal activity, coordinated crowdsourced attacks could be used to decrypt passwords or launch cyber attacks on the computer systems of a country's adversaries. Perhaps even more troubling, a coordinated attack by a large group could mask their malicious behavior by acting collectively so that their influence on the system cannot be traced to a single aberrant individual.

Crowdsourced click manipulation. We have observed crowd workers leveraging human-powered crowdsourcing platforms to intentionally manipulate click patterns of URLs spread through social media to create conditions of artificial collective attention, in effect to create the illusion of collective attention toward increasing the population exposed to a malicious URL (say, by pushing the message containing such a URL into the day's trending topics on a system like Twitter) (Lee et al. 2013a).

Location-based deception. The rise of global-scale location sharing services (like Foursquare, and services supporting fine-grained location sharing like Instagram) allow users to connect in the physical world by revealing their footprints (typically via a "check-in" containing the user's current location that is shared through a social media service), leading to a host of positive opportunities. But these services can be misused to manipulate collective attention. In discussions with the Austin (Texas) Police Department, we have identified the threat of intentional deception through the creation of fake "check-ins" around protests so that police response may be re-directed to the wrong location.

Notice that these threats may have far reaching consequences, if successfully carried out. For example, during the recent Hurricane Sandy, several episodes of misinformation have led to confusion, errors, and slowed down humanitarian actions in affected zones, causing FEMA to formally address the issue (FEMA 2012; Meier 2012). For example, social media users posted fake storm images and spread misinformation that FEMA had run out of bottled water. Given the magnitude of the storms, FEMA has acknowledged the great role of social media as an effective means to quickly gain collective attention, but identified misinformation as a real threat to human lives.

Methods to Detect and Mitigate Exploits

Detecting exploits in human computation system is quite important, and the corresponding detection technique varies based on the type of exploit.

Reputation Systems. Many e-marketplaces and online communities use reputation systems to assess the quality of their members, including eBay, Amazon, and Digg, and reputation-based trust systems have received considerable research attention, e.g., Marti and Garcia-Molina (2006) and Resnick et al. (2000). These approaches aggregate community knowledge for evaluating the trustworthiness of participants. The benefits of reputation-based trust from a user's perspective include the ability to rate neighbors, a mechanism to reach out to the rest of the community. Along these lines, the recently proposed Turkopticon (Irani and Silberman 2013) is one such reputation system designed for human computation systems, in which workers on Amazon Mechanical Turk can rate interactions with requesters.

Policy-Based Approaches. Separately, exploits in human computation systems may be dealt with by rule-based or policy-oriented approaches. For obvious exploit like "workers getting poorly paid", it is intuitive just to compare the estimated average payment per hour for a task to the legal minimum wage rate, to decide whether workers are being exploited for lack appreciation of their efforts. In a more systematic manner, there has been some recent work on monitoring the quality of workers and their outputs. For example, Venetis and Garcia-Molina (2012) described two quality control mechanisms. The first mechanism repeats each task multiple times and combines the results from multiple users. The second mechanism defines a score for each worker and eliminates the work from users with low scores. Xia et al. (2012) provided a real-time quality control strategy for workers who evaluate the relevance of search engine results based on the combination of a qualification test of the workers (i.e., a question for which the requester already knows the answer) and the time spent on the actual task. The results are promising and these strategies facilitate reducing the number of bad workers.

Machine Learning. For more complicated exploits that target external populations, machine learning techniques can be applied. In one direction, the artifacts of a crowd powered targeting of an external population can be analyzed to develop machine learning models of the activities of the users who engaged in this activity. For example, by analyzing the social media artifacts of astroturf campaigns, researchers have developed methods to automatically detect crowd powered campaigns and the users engaged in these campaigns (Gao et al. 2010).

In a separate direction, since many current crowd turfing approaches target social media, researchers have proposed a framework for linking tasks (and their workers) on crowdsourcing sites to social media, by monitoring the activities of social media participants (Lee et al. 2013b). In this way, we can track the activities of crowdturfers in social media where their behavior, social network topology, and other cues may leak information about the underlying crowdturfing ecosystem. Based on this framework, researchers have identified the hidden information propagation structure connecting these workers in Twitter, which can reveal the implicit power structure of crowdturfers identified on crowdsourcing. Specifically, three classes of crowdturfers have been identified—professional workers, casual workers, and middlemen; based on statistical user models these users can be automatically differentiated from regular social media users.

Crowd-Based Mitigation. Finally, the crowd itself may be mobilized to mitigate exploits. How can a crowd be organized to police itself? How can a crowd detect the exploits within its own system and mitigate the impacts of exploits powered by other systems? In one direction, a crowd-powered monitoring system (akin to the Turkopticon) could be extended so that sub communities within the system validate the tasks within the system, towards reducing the opportunity of exploits to gain sufficient traction. Similarly, crowds could be deployed to monitor external communities (as on social media) for evidence of exploits; such a crowd-powered system could alert external communities of exploits and even roll-back negative actions (e.g., undoing Wikipedia vandalism). Of course such a crowd-checking-crowd system raises questions of "who watches the watchmen?" which we leave as an open and enduring question.

Summary

This chapter has presented a characterization of exploits that may target participants within human computation systems, as well as exploits that may target other populations. As crowd-powered systems continue to become more complex and of greater variety, we would expect a commensurate maturation of the exploit vectors, and (hopefully) of the technical and policy-oriented countermeasures to mitigating their impact.

References

BBC (2013) Vietnam admits deploying bloggers to support government. http://www.bbc.co.uk/news/world-asia-20982985, Jan 2013

FEMA (2012) Hurricane sandy: rumor control. http://www.fema.gov/hurricane-sandy-rumor-control

Gao H, Hu J, Wilson C, Li Z, Chen Y, Zhao BY (2010) Detecting and characterizing social spam campaigns. In: IMC, Melbourne

Gizmodo (2012) How a fake erotic fiction ebook hit the top 5 of itunes. http://gizmodo.com/5933169/how-a-fake-crowdsourced-erotic-ebook-hit-the-top-5-of-itunes, Aug 2012

Guardian T (2011) Revealed: US spy operation that manipulates social media. http://www.guardian.co.uk/technology/2011/mar/17/us-spy-operation-social-networks, Mar 2011

Irani L, Silberman MS (2013) Turkopticon: interrupting worker invisibility in amazon mechanical turk. In: ACM SIGCHI conference on human factors in computing systems, Paris

Jagatic TN, Johnson NA, Jakobsson M, Menczer F (2007) Social phishing. Commun ACM 50(10):94–100

Lease et al. M (2013) Mechanical turk is not anonymous

Lee K, Kamath K, Caverlee J (2013a) Combating threats to collective attention in social media: an evaluation. In: ICWSM, Cambridge

Lee K, Tamilarasan P, Caverlee J (2013b) Crowdturfers, campaigns, and social media: tracking and revealing crowdsourced manipulation of social media. In: 7th international AAAI conference on weblogs and social media (ICWSM), Cambridge

Marti S, Garcia-Molina H (2006) Taxonomy of trust: categorizing p2p reputation systems. Comput Netw Int J Comput Telecommun Netw 50(4):472–484

Meier P (2012) What was novel about social media use during hurricane sandy? http://irevolution.net/2012/10/31/hurricane-sandy/

Resnick P, Kuwabara K, Zeckhauser R, Friedman E (2000) Reputation systems. Commun ACM 43(12):45–48

Ross J, Irani L, Silberman MS, Zaldivar A, Tomlinson B (2010) Who are the crowdworkers? Shifting demographics in mechanical turk. In: CHI'10 extended abstracts on human factors in computing systems, CHI EA'10, New York, pp 2863–2872. ACM

Tang JC, Cebrian M, Giacobe NA, Kim H-W, Kim T, Wickert DB (2011) Reflecting on the darpa red balloon challenge. Commun ACM 54(4):78–85

Venetis P, Garcia-Molina H (2012) Quality control for comparison microtasks. In: CrowdKDD 2012, Beijing

Wang G, Wilson C, Zhao X, Zhu Y, Mohanlal M, Zheng H, Zhao BY (2012) Serf and turf: crowdturfing for fun and profit. In: WWW, Lyon

Wired (2010) The chinese online 'water army'. http://www.wired.com/beyond_the_beyond/2010/06/the-chinese-online-water-army/, June 2010

Xia T, Zhang C, Xie J, Li T (2012) Real-time quality control for crowdsourcing relevance evaluation. In: Network infrastructure and digital content (IC-NIDC), Beijing

Zittrain J (2009) Work the new digital sweatshops. Newsweek

Big Data, Dopamine and Privacy by Design

Thomas W. Deutsch

Introduction: Privacy and Mom's Cupcakes

Human computation obviously requires the engagement of the people doing the computation, and people's willingness to participate has a lot to do with the value they derive from the engagement and the "cost" of participating. Most of us would associate the "cost" of participating being the time spent but there is a second, and often under appreciated expenditure, that of the person's privacy. How privacy is handled varies substantially across the spectrum of human computation, and the people doing the computation may not be in full control of the privacy decisions they are making. There are several reasons for this, but we are going to focus our attention on an underexplored considerations based on how we, as people, actually function and make decisions. To explore the dynamics here, like all good things in life, we are going to start by talking about cupcakes.

Imagine no one told you that Mom's homemade buttercream cupcakes were bad for you—you'd eat them to the point of exploding. OK, maybe that's just me but you get the idea. This is why we label our food's nutritional content since in theory an informed consumer is a healthy consumer, or at least one making good long-term choices. In practice we know that doesn't work so well—the obesity rates in the USA as confirmation of that. So why do people engage in the irrational behavior of eating both unhealthy food and unhealthy amounts of it? It turns out that we do that for the same reasons that many consumers struggle with the notions of privacy in an increasingly virtual world. That struggle has significant implications for the future of human computation and the problems that we are collectively trying to solve. To put a fine point on it—people's willingness to work on group problems and serve as part of a human sensor network will long-term depend on their ability to trust how their engagement and inputs to the project are handled.

T.W. Deutsch (✉)
IBM Information Management, San Jose, CA, USA
e-mail: tdeutsch@us.ibm.com

P. Michelucci (ed.), *Handbook of Human Computation*,

DOI 10.1007/978-1-4614-8806-4_69,

To help understand what is happening here and how it relates to big data, digital engagement and the need for privacy in human computation, let's turn back to Mom's buttercream cupcakes for a minute. Now, to be sure they taste good (especially the vanilla cake with strawberry frosting ones), but at some level we know we probably should not be eating too many of them, or eating them too frequently. So why do we do often over indulge in too large position sizes or too often make bad nutritional choices? Well in some cases we just can't help it because of how our brains work. As it turns out we are wired to be susceptible to responding to certain food types and components in a way that in an age of surplus becomes counterproductive.[1] Fat, sugar and salt all trigger physiological reactions quite separate from our purely rational experience of eating the cupcake. Simply stated, the same physiology and brain wiring that has allowed us to survive to this point is not especially well-equipped to handle our new circumstances of surplus.[2] It is starting to become evident that the dynamics that contribute to our poor food choices have parallels in our ability to self-regulate in our digital engagements.

Challenges in How We Make Decisions: Temporal Discounting and Neurobiology

To help explore these challenges, there are two important concepts that need to be introduced. The first is the idea of temporal discounting; the second is the idea that our neurobiology and neurochemistry are in play here without our being aware of it. Temporal discounting refers to our brains tendency to discount further away events from near term ones thereby making even smaller (if fleeting) rewards now appear more valuable than larger rewards in the future.[3] Here again the digital implications are harder to grasp as at some level nearly everyone understands that eating too much can cause unpleasant issues but our long term cost of trading off privacy are not as immediately apparent as, say, an upset stomach. As one of our Editors pointed out "even eating three cup-cakes has a near term impact on how you feel as well as poorly understood long-term implications. In the digital world, we almost never have immediate consequences for poor decisions—they are always long term "costs".

Strategies in the physical world for dealing with temporal discounting—such as walking around the neighborhood before going into McDonalds to give your brain time to better weigh the true "cost" of that chocolate milkshake you are craving—don't always work so well in the digital world. It can be a bit challenging to go walk around the neighborhood to 'cool off' before using a mapping service on a mobile device if you are turning to the mapping service since you don't know your way

[1] http://www.ncbi.nlm.nih.gov/pubmed/3135745

[2] http://www.ncbi.nlm.nih.gov/pubmed/3300488

[3] http://www.ncbi.nlm.nih.gov/pmc/articles/pmc1382186/

around the neighborhood. The immediacy and intimacy of our digital engagements, which clearly contributing to the usefulness and likability of the engagements, does pose challenges in self-regulation and decision making.

The Neurobiology and neurochemistry factors are play here are much more complex than the temporal discounting challenge outlined above. Neurobiology and neurochemistry deal with how the biology and chemistry of our brains impact our behaviors,[4, 5] To illustrate some of the considerations at play here, we are going to discuss our neurobiology. Before we go any further, I hasten to point out I am going to use some of that system functioning, such as the role of dopamine, as a stand in for a much more complex set of neurobiology and neurochemistry considerations. Mapping out all of that complexity, especially given that it is a fast-evolving area of scientific inquiry, is out of the scope of this article (as well as not being my field of expertise, so please take what follows with a grain of salt.).

As it turns out, how we make decisions is under much less of our conscious control than we realize. In some cases, we make decisions before we even become aware we are making a decision, or as Soon et al. summarizes "a network of high-level control areas can begin to shape an upcoming decision long before it enters awareness."[6] Our neurobiology works to shape decisions without our being consciously aware of it, and it does this very-very quickly. This happens through a complex set of interactions, but to single out as an example one component of this we are influenced by the amount of dopamine in our systems. Dopamine is an organic chemical that serves as a neurotransmitter that our bodies synthesis in response to simulation and it plays an important role in how we respond to situations. More specifically, "….midbrain dopamine systems are involved in processing reward information and learning approach behavior."[7]

Going back to the challenge of Mom's cupcakes can help illustrate some of the neurobiology at play here. Just the thought of eating the cupcakes can trigger neurochemical reactions that make us want to eat them that much more.[8] The actual act of eating one invokes neurobiological feedback loops that encourage us to eat yet more. Yet as challenging as the cupcakes are, in some ways the digital challenge is even harder to manage. First, unlike mom's cupcakes, there is no natural satiation mechanism whereby (after, say, five or six cupcakes, maybe less if you aren't me) you actually get full. Our saturation point of experience is far higher in digital engagements than cupcakes. Even more challenging is that Mom's cupcakes don't get more and more appealing as you eat more of them as they are a fixed experience. That is to say, the cupcakes don't change their behavior to be even more appealing and thus trigger another round of reinforcement. Highly intimate and

---

[4] http://www.merriam-webster.com/dictionary/neurobiology

[5] http://www.merriam-webster.com/dictionary/neurochemistry

[6] http://www.rifters.com/real/articles/NatureNeuroScience_Soon_et_al.pdf

[7] http://jn.physiology.org/content/80/1/1.abstract?ijkey=9149a8c097da470088e9a355b467daa4a58ebd5c&keytype2=tf_ipsecsha

[8] http://www.ncbi.nlm.nih.gov/pubmed/15987666

personalized applications, increasingly enabled by big data technologies, however, do exactly that. More engagement leads to even better ability to micro-segment your likes and preferences, and that will change your experience to become even more intimate and pleasing. Whereas Mom's cupcakes don't whisper in your ear that some chocolate milk would be especially tasty right now, the more engaging digital experiences can do the equivalent of that. More digital engagement leads to more insight about you, which in turn leads an even more engaging or socially invocative experience. The more engaging experience is, especially if it is social attachment in nature, the more to appears triggers neurochemical reinforcement[9] that, whatever privacy you have surrendered for that experience was worth it, potentially without the consumer every being aware they had made a decision to do just that before consciously doing it.

Social Interactions and Potential Impact on Privacy Choices

The idea of moderation is especially challenging in the digital realm. It is well understood that we are responsive to the dopamine reward we get from social interactions and it appears this is true when the social engagement happens to be a digital one rather than a physical one.[10] Digital interactions are even more intense when we process them as "intimate," since they are more engaging, and the more engaging the greater the involvement of dopamine.[11] So when we are faced with a choice of an impersonal or socially intimate experience, we choose an intimate one, at least partially since we get the reward of the chemical "hit". All of this may happen far more rapidly than we have time to understand the ramifications of our choices, since "dopamine concentrations are now known to fluctuate on a phasic timescale (subseconds to seconds)."[12] As noted above, the notion of choice and rational handling of the privacy issues comes into clear question when our neurochemistry is moving faster than consumers ability to make temporal tradeoffs (which, as noted above, is an iffy proposition anyway).

In the natural world that intimacy, however, can't scale beyond a relatively low number of connections at a given time due to time/space/personal network limitations. Technology can overcome those natural limitations so we're engaging at a volume and pace never possible before. This similar to overeating in a new age of surplus creates a volume of reinforcing neurobiological events we struggle to effectively manage. That, in turn, sets up a feedback loop wherein surrendering privacy increases the likelihood of a reward, which in turn rewards the surrender of privacy and so on. So while at some level, many people can intellectualize that surrendering

[9] http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2889690/

[10] http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2889690/

[11] http://www.dnalc.org/view/2385-The-Neurobiology-of-Love.html

[12] http://www.clinchem.org/content/49/10/1763.full

our privacy should require some contemplative thought on the balance of what we surrender versus what is gained and the long term implications of such a trade-off, it simply may be very difficult to overcome the brain chemistry that says, "this feels good, more please." Long term considerations of privacy may not stand much of a chance at that point compared to the neurobiology at work especially give the temporal discounting issues. Privacy is hard to value, and unlike food that comes with at least some basic nutritional information, there is no third-party reference point that people involved in human computing efforts can turn to for help in making a privacy related decision.

The Role of Commercial Models to Which We Have Become Accustomed

If that were not enough of a challenge, there is another one which has proven to be a quite well-established consumer behavior: customers like free stuff. Free email, mapping, social sites, free music, free hosting, free just about everything. And of course none of that is truly free. As was seminally expressed in the following blurb "If you are not paying for it, you're not the customer; you're the product being sold."[13] Many of the activities that trigger the neurobiology that proves challenging have the double "incentive" of being supported by monetizing the consumer as the product. Now of course there is nothing wrong with that, but when combined with the other triggers we've discussed makes a powerful experiences that, short of handing out free beer (or wine based on your personal preferences, which of course we'll likely know) is about perfectly designed to functionally disincentivize privacy considerations. It is worth noting here that there is a tendency to still frame the "do I or don't I surrender privacy" in purely rational terms when that may not be how we actually make the decision. Our emotional reactions to engagement are powerful, and often influence our perception that we are making purely rational choices when we are not.[14]

Just as the food industry has learned to develop foods engineered to take advantage of our neurobiology (think sweet and salty mix in ice cream so have many of the most utilized Internet sites. To be clear, they are responding to consumer preferences. Consumers want easier, consumers want more relevant, consumers want to be better entertained. That, of course, is the challenge. The benefits from using well designed (from an engagement point of view anyway) sites are immediate, the privacy tradeoffs not immediately apparent and almost always involves stopping the behavior we enjoy to read Terms and Conditions of site usage that is not, shall we say, quite so engaging and thus not as rewarding. The neurobiology of this *is* different than in our physical lives. As a good friend of mine said "No one gets a dopamine hit from having the grocery store track their purchases through a loyalty card".

[13] http://www.metafilter.com/95152/Userdriven-discontent#3256046

[14] http://metablog.borntothink.com/wp-content/uploads/2011/07/1994-Damasio-Descartes-Error.pdf

While human computing is not confined to social sites and advertising funded sites/applications, one could argue that a majority of human computing comes as a byproduct of those activities. Problematically, it is those sites that often have the least transparent privacy policies and are designed to present an experience that invokes a neurobiological response that overrides a deliberate or methodical privacy-oriented decision making.

Examples such as http://www.patientslikeme.com/ where there is the potential for advancing understanding through shared information processing (in form of shared experiences) depend upon deeply personal information quite possible that is to be de- anonymized. The question of a person engaging in human computation can consider privacy and how their information will be shared when dealing in a social, experience related to their (or loved one's) health is debatable.

The Shortcomings of Anonymization

Anonymization has been presented as a way around this but as it as it turns out, anonymization is not very anonymizing in the age of big data. Anonymization—the basic tenant of decoupling the data from the common unique identifiers of phone number, user ID, email, or name doesn't hold up to a world enabled by big data technologies. Big data technologies offer both an expanded range of data gathering as well as increased processing power to dig into the data in more depth. One need not always dig that far however, as Ohm warned us about in 2009[15] and de Montjoye et al. recently reminded us of how "anonymized" data sets can still allow for very precise identifications of people.[16] This presents a substantial privacy challenge as our most common approach to building applications assumes that anonymity can be counted on to protect privacy, and many of the most commonly used application would simply seem to function properly if all the data that could be used for undoing anonymization were removed. It is also unclear if commercial entities could track down all the potentially de-anonymizing data in their anonymized data sets. As was recently observed:

> Removing forgotten information from all aggregated or derived forms may present a significant technical challenge. On the other hand, not removing such information from aggregated forms is risky, because it may be possible to infer the forgotten raw information by correlating different aggregated forms.[17]

It is also worthwhile to keep in mind here that the data being collected is often critical to the services being provided, as well as generating positive results from the human computing effort. As David Myers summarized in an especially well written observation: "Mobile operators' datasets help keep their networks running.

[15] http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1450006

[16] http://www.nature.com/srep/2013/130325/srep01376/full/srep01376.html

[17] http://www.enisa.europa.eu/activities/identity-and-trust/library/deliverables/the-right-to-be-forgotten

Location-based services don't work without location. We even hope big data capabilities will help us fight diseases and socio-economics problems. And, most importantly, despite the fact that most people in the U.S. and European Union insist they want better data privacy, we see time and again that this desire doesn't translate into action—people still give up their data without much consideration." Rock, meet hard place. Not surprisingly this debate has surfaced recently surrounding sugary drinks where our biological challenges to moderation and resulting personal and societal costs all require difficult trade offs.[18]

As with the great soda debate of 2013 in NYC[19], it seems pretty clear at this point there are not any easy answers here in the digital space. Users appear unlikely to spontaneously demand privacy baring some traumatic mass event, and the commercial models based on data collection have become firmly and widely embedded. To reference Myers again "we are not going to stop all this data collection, so we need to develop workable guidelines for protecting people."[20] It is unlikely that we are going to quickly evolve to a point where our neurobiology is not an issue to be considered in our online engagements, yet doing nothing does not appear to be an option. Voluntary solutions like Do Not Track,[21] which is both a technology and policy approach to giving users more control over their privacy, remain works in progress with uneven implementations.[22] Do Not Track has spawned related ideas on dealing with the issues outlined above, including the notion of Privacy By Design.

Privacy by Design Principals

Privacy By Design, an initiative by the Information and Privacy Commissioner of Ontario, Canada,[23] lays out seven key tenants that are designed to introduce some privacy protection by default. Without getting into the role of free will in all of this, Privacy By Design attempts to help protect us from ourselves by codifying an approach to the systems we interact with. The key tenants of Privacy By Design are[24];

1. ***Proactive*** not Reactive; ***Preventative*** not Remedial
   The *Privacy by Design* (PbD) approach is characterized by proactive rather than reactive measures. It anticipates and prevents privacy invasive events *before* they happen. PbD does not wait for privacy risks to materialize, nor does it offer remedies for resolving privacy infractions once they have occurred—it aims to *prevent* them from occurring. In short, *Privacy by Design* comes before-the-fact, not after.

[18] http://www.nytimes.com/2012/06/06/opinion/evolutions-sweet-tooth.html?_r=1&

[19] http://www.nytimes.com/2012/09/14/nyregion/health-board-approves-bloombergs-soda-ban.html?_r=0

[20] http://gigaom.com/2013/03/25/why-the-collision-of-big-data-and-privacy-will-require-a-new-realpolitik/

[21] https://www.eff.org/issues/do-not-track

[22] https://www.eff.org/issues/do-not-track

[23] http://www.privacybydesign.ca/index.php/about-pbd/

[24] http://privacybydesign.ca/about/principles

2. Privacy as the ***Default Setting***
   We can all be certain of one thing—the default rules! *Privacy by Design* seeks to deliver the maximum degree of privacy by ensuring that personal data are automatically protected in any given IT system or business practice. If an individual does nothing, their privacy still remains intact. No action is required on the part of the individual to protect their privacy—it is built into the system, *by default*.
3. Privacy ***Embedded*** into Design
   Privacy by Design is embedded into the design and architecture of IT systems and business practices. It is not bolted on as an add-on, after the fact. The result is that privacy becomes an essential component of the core functionality being delivered. Privacy is integral to the system, without diminishing functionality.
4. Full Functionality—***Positive-Sum***, not Zero-Sum
   Privacy by Design seeks to accommodate all legitimate interests and objectives in a positive-sum "win-win" manner, not through a dated, zero-sum approach, where unnecessary trade-offs are made. *Privacy by Design* avoids the pretense of false dichotomies, such as privacy vs. security, demonstrating that it is possible to have both.
5. **End-to-End Security—***Full Lifecycle Protection*
   Privacy by Design, having been embedded into the system prior to the first element of information being collected, extends securely throughout the entire lifecycle of the data involved—strong security measures are essential to privacy, from start to finish. This ensures that all data are securely retained, and then securely destroyed at the end of the process, in a timely fashion. Thus, *Privacy by Design* ensures cradle to grave, secure lifecycle management of information, end-to-end.
6. ***Visibility*** and ***Transparency***—Keep it ***Open***
   Privacy by Design seeks to assure all stakeholders that whatever the business practice or technology involved, it is in fact, operating according to the stated promises and objectives, subject to independent verification. Its component parts and operations remain visible and transparent, to users and providers alike. Remember, trust but verify.
7. ***Respect*** for User Privacy—Keep it ***User-Centric***
   Above all, *Privacy by Design* requires architects and operators to keep the interests of the individual uppermost by offering such measures as strong privacy defaults, appropriate notice, and empowering user-friendly options. Keep it user-centric.

None of this is easy or problem free to implement. As noted earlier many existing applications pre-date these notions and may not be able to function if tracking/tracing data were removed. It is unclear that separating the notions of privacy and security is commercially practical given how many applications have been designed. There is also the non-trivial issue of the potential need for a shift from the user's data being monetized to pay for the digital service if privacy is fully protected. The costs of not implementing, however, could be higher. If the potential of human computing is blunted by concerns of privacy, who knows what we as a society we will forgo. It would seem that a reasonable next step is an honest conversation and full disclosure of how a human computing participant's information and activities will be utilized. In a free-market, people can vote with their time and there should be no shortage of human computing projects that both have worthy goals and manage to protect the participant's privacy. I hope this was a useful discussion, and I don't know about you but I'm craving a cupcake at this point.

Privacy in Social Collaboration

Elena Ferrari and Marco Viviani

Introduction

With the expression *social collaboration* we refer to the processes of helping multiple people to interact and share information in order to achieve common goals. Nowadays, collaboration and social dissemination of information are facilitated by the Internet and *Social Network Services* (SNS). The reliance of social collaboration on SNS might seem surprising given the differences between their group-centric and individual-centric views. In particular, social collaboration services focus on group activities, identifying groups and collaboration spaces in which messages are explicitly directed at the group and the group activity feed is seen the same way by everyone. In contrast, social networking services generally focus on single personalized activities, sharing messages in a more-or-less undirected way and receiving messages from many sources into a single personalized activity feed.

Despite these differences, in current digital society a convergence between mass communication and personal communication is leading to social and community uses of online social network services. This is because the present use of social media has grown enormously, moving from a niche phenomenon to mass adoption (Gross and Acquisti 2005). For these reasons, it emerges how social interactions on the online world must not be considered as separated entities with respect to collaborating communities in the real (offline) world. In this scenario, it often comes to light how current social network services architectures do not allow to treat and analyze communities and their privacy issues in the online world as really happens in the offline world. This is due, in particular, to the fact that the online world often does

E. Ferrari (✉) • M. Viviani
Dipartimento di Scienze Teoriche e Applicate (Department of Theoretical and Applied Sciences), Università degli Studi dell'Insubria (University of Insubria), Via Mazzini 5, Varese, Italia
e-mail: elena.ferrari@uninsubria.it; marco.viviani@uninsubria.it

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_70,

not have the same boundaries and does not follow the same social norms which are more clear and common in the offline world. This disparity may exist because social norms are connected to particular situations involving users (AA.VV SPION 2011).

In the offline world, more or less clear barriers exist among situations and contexts. In this scenario, privacy is signaled by physical characteristics: e.g., low lighting, enclosed spaces, and relative isolation from others. People who want to conduct a private conversation can recognize the privacy levels of an offline space based on physical properties (Dwyer and Hiltz 2008).

In the online world, missing these clear boundaries and well defined social norms, and due to the fact that users have more control on how their identity is displayed (since each user can decide which information provide to the world), often the context is not clear and it is free to be filled. Online, privacy levels are not signaled by the inherent properties of the online social space in any clear way, except for the common assumption that nothing is private. In this scenario, some argue that privacy in online communities should be a system level requirement, rather than a group of access settings for each member: privacy should apply to an online space and not be a collection of settings attached to each individual member (Dwyer and Hiltz 2008).

Aim and Organization of the Chapter

In this chapter we address the problems connected to privacy issues in social collaboration, in particular with respect to social network services when used for social and community-based purposes.

The chapter is organized as follows: in section "Social Communities" we describe the concept of social community in the online and in the offline worlds and the relationships between them; in section "Social Networking" we investigate the community-based use of social networks also providing a brief history of their evolution; section "Privacy in Social Networks Services" arises and discusses privacy issues in social network services especially when context issues emerge from online social behavior of users, and describes some concrete privacy concerns in current SNS. Finally, section "Conclusions and Further Research" concludes the chapter.

Social Communities

The Meaning of Community

First attempts to define the concept of *community* dates back to nineteenth century, with the studies of the theorists Tönnies (Tönnies and Loomis 1957), Toqueville, Durkheim. These theorists follow the desire for a pre-modern society; in this scenario, a community can be described as a private and intimate place that stands for the basic needs of individuals (e.g., warmth, shelter, nurture, etc.), while *society* is seen as a more rational and purposeful (Kivisto 2003).

In current literature, the *existence* and the utility of the concept of community is debated. In particular, the *focus* of community varies from domain to domain: it is a cultural construct or social context for sociologists; in psychology the individual members of a community are emphasized; anthropologists concentrate on interaction among the members of a community. With such wide-ranging and diverse interpretations, the concept of community is definitely an ambiguous and abstract concept.

According to post-modernists, it is only a diluted concept unsuitable to describe current society. Bauman (2001), for example, sees community as an extension of the concept of identity.

Other authors have another vision and think that the concept of community has still its meaning. Turner, for example, sees community as an opposition to *structure*, an expression of the *social nature* of society (Delanty 2003). He calls *liminality* the expression of such a community. Liminal moments refer to events of life not subjected to instrumental rationality, and create a powerful bonding between members of society. In this vision, one obtains a feeling of belonging and relating to others when not being subjected to rules, laws, norms, etc. Then, in interacting with others, members of the community reveals the community itself (AA.VV SPION 2011).

In his hermeneutic approach on community, Cohen defines it in terms of particular kinds of awareness of reality; and as such community is a "symbolization of boundaries by which the community differentiates itself from others" (Delanty 2003).

Lyon (1986) reviews a plethora of definitions of community, noting that the vast majority enumerates three common qualities: shared place, distinctive social interaction and common ties. These three qualities are not independent, but mutually reinforcing instead. They are distinguishable theoretically, and do capture critical facets of what community is characterized for, as Nisbet (1976) observes.

Based on Lyon and other researchers' work, Carroll (2011) proposes a conceptual model of communities, comprising of collective identity, community engagement, and network of social ties.

According to Zhang et al. (2011), these three elements emphasis different underpinnings of communities: social identities as psychological foundation, social engagement as behavioral manifestation, and network of social ties as structural depiction of communities.

Following these positive perspective on communities, it is possible to divide the concept of community in two categories: *community of interest*, and *community of place*. Community of place refers to a geographical fixed community. A community of interest is based upon a common interest between members. It may be that both communities overlap each other. This teaches us that a community does not need to be anchored in a particular location, but can also exist in the *virtual*.

From Offline to Online Communities

As emerges from previous section, definitions of a community are diverse and, at times, vague. Despite this, the concept of community is frequently adopted in the digital era to describe social practices in cyberspace. In fact, individuals can share their common interests by gathering virtually in the online communities associated with social bookmarking sites, blogs and forums, regardless of their physical location. The absence of a spatial environment has not only complicated how a community have to be defined of the Web, but has also raised issues as to how communities in online environments are to be operationalized for detection and investigation (Zhang and Jacob 2012).

For these reasons, a main question we have to address when dealing with the concept of *online community* is if an online community is a reality or a virtuality with respect to classical 'offline' communities.

Thick and Thin Communities

According to Giddens (1990), *virtuality* is a product of modernity that constantly 'displaces' individuals from the places and everyday life with which they were familiar: individuals are re-located in different contexts, in which "familiarity and estrangement are recombined". Similarly, Rheingold (2000)—to our knowledge the first author having introduced the concept of *virtual community*—describes this concept connected to the Internet as an alternative reality, with capacities to transform society (Delanty 2003). When referring to virtual communities, he only considers non-existing offline communities, exclusively rooted in cyberspace. This means that, for him, virtual communities are 'communities on the Net': they do not have their counterpart in everyday life. Even further, the downfall of communities can be compensated by a virtual one (Delanty 2003). In this vision, if virtuality is the opposite of reality, it follows that a virtual community on the Web cannot be regarded as the same as—or even similar to—a traditional offline community. According to Zhang and Jacob (2012), because the online environment can only provide the illusion of reality and because a virtual community exists online, it is not part of the real world and thus cannot be understood or even discussed as a real world community might be.

However, a different and interactionist perspective about virtuality and reality is provided by Castells (1996), who includes the concept of virtuality as a part of the real world. New communities like virtual ones are built out of networks of social actors (individuals, families or social groups) (Delanty 2003). In our global network society, spatial communities are replaced by spaceless ones in the virtual space constituted by the Web. Castells affirms that "localities become disembodied from their cultural, historical, geographical meaning, and reintegrated into functional networks, or into image collages, inducing a space of flows that substitutes for the space of places". Social relations are not changed by the global network society itself; rather, by the individualism inherent in society.

To sum up, in both authors' visions, communities can be defined as personalized communities embodied in networks and centered on the individual. But where Rheingolds refers to virtual communities as *thick*, Castells would definitely speak of *thin* communities. With 'thin' we refer to a virtual reality that is an addition to the offline reality, whereas 'thick' can be seen as an equivalent of the offline reality.

Thick communities are often composed of *strong ties*: frequent contact between people who personally know each other. *Weak ties* are often related with thin communities: they are online ties between persons socially and physically distant, not bound into work structures or circle of friends.

Social Capital

The concepts described above, and their interactions, bring forward another important concept related to communities: the *social capital*. In sociology, Putnam and Bourdieu are probably the most prominent authors on this topic. Putnam defines social capital on a community level as those "features of social organisation such as networks, norms and social trust that facilitate coordination and cooperation for mutual benefit" (Baum and Ziersch 2003). Bourdieu stresses more the individual aspects in his definition of social capital, seen as "the aggregate of the actual or potential resources which are linked to possession of a durable network of more or less institutionalized relationships of mutual acquaintance and recognition" (Baum and Ziersch 2003).

Broadly speaking, it consists on the expected collective or economic benefits derived from the preferential treatment and cooperation between individuals and groups. In Web 2.0 it is fostered by the possibility to maintain long-term contact with people via weak as well strong ties. Which of these ties contribute more to social capital, it is a debated topic.

Granovetters in his paper on the strength of weak ties (Granovetter 1983) states that weak ties are more important in some situations, as for looking a job. He affirms that weak ties are more likely related to sparse networks. Hence, users that are loosely connected within virtual communities can access remote regions and obtain new and non-redundant information. In contrast, dense networks (dominated by strong ties) facilitate frequent, reciprocal and supportive contact. So, whether or not virtual communities can be labelled as thick or thin, both seem to be important for different reasons.

In his revised copy on virtual communities (Rheingold 2000), Rheingold states the following: "A social network with a mixture of strong ties, familial ties, lifelong friend ties, marital ties, business partner ties, is important for people to obtain the fundamentals of identity, affection, emotional and material support. But without a network of more superficial relationships, life would be harder and less fun in many ways. Weaker ties multiply peoples social capital, useful knowledge, ability to get things done". Following this 'optimistic' vision, weak ties in virtual communities enable users to engage and interact with a variety of other users that do not necessarily share the same interests and environments, expanding users' horizon. At the same time, virtuality offer the possibility to bring offline contacts to online environments enlarging communication possibilities.

Another author, Calhoun, also assign importance to these mediated relationships, although in his more 'pessimistic' view, we should not exaggerate these forms (Delanty 2003). Offline communities are supplemented by virtual ones, rather than substituted. Calhoun has a rather negative view on the capacity of virtual communities to enhance participation, due to compartmentalization in communities: “we are aware of others but not in discourse with them” (Calhoun 1998). This leads to categorization of individuals. This view anticipates the principle of the filter bubble, introduced by Pariser (2011) and described in section “Filter Bubble”. The filter bubble is an effect of the Internet when tailored to the personal identity of the individual, isolating him/her from other perspectives.

Social Networking

Social Network Services as Online Communities

Summarizing the definition of community as either a place or a metaphor for place in terms of shape, structure, context and experience, the application of *Social Network Analysis* (SNA) (Scott 2012) offers an efficient and productive approach for the detection and investigation of communities as complex social phenomena in social network services. Using the diagnostic tools of SNA, it is possible to capture the structure and function of communities and to provide a relatively objective interpretation of these ‘subjective’ phenomena.

As introduced before, social network services are now entangled in society and not floating around in a vacuum. In fact, in current dynamic digital society, a convergence between mass communication and personal communication is occurring. This convergence has been defined by Castells (2009) as *mass self-communication*. According to Pierson and Heyman (2011), “on the one hand mass communication because social computing tools can potentially reach a global Internet audience. On the other hand self-communication because the message production is self-generated, the potential receiver(s) definition is self-directed and the message or content retrieval is self-selected”. Hence, new social network services and tools for acting ‘socially’ can be seen as an important fraction of mass self-communication. According to boyd and Marwick (2011), social networks can serve multiple ‘public’ purposes: “they can play a civic function, serving to gather people in a democracy. But they can also play a social role, enabling people to make sense of the world around them and understand their relationships to society”.

Formally, it is still boyd that, with Ellison in boyd and Ellison (2007), defines a social network service as “a web-based service that allow individuals to construct a public or semi-public profile within a bounded system, articulate a list of other users with whom they share a connection, and view and traverse their list of connections and those made by others within the system. The nature and nomenclature of these connections may vary from site to site”.

From this definition, we can observe that, in general, a social network service is characterized by the following properties:

- It is an online service, platform, or site that focuses on facilitating the building of social networks or social relations among people;

- People can share interests, activities, backgrounds, or real-life connections;
- Each user has a virtual representation, often a profile, plus his/her social links, and a variety of additional services.

Due to these characteristics, and to mass self-communication, it is possible to recognize several online community services aspects in social network services, leading to an overlapping between the two kinds of services.

Evolution in Social Network Services

Many early online services, including Usenet (Hauben and Hauben 1997), ARPANET, LISTSERV, and bulletin board services (BBS) made efforts to support social networks via computer-mediated communication. At that time, also online services such as America Online, Prodigy, CompuServe, ChatNet, included yet some prototypical features of Social Networking Services. Early social networking services on the World Wide Web began in the form of generalized online communities such as Theglobe.com (1995), Geocities (1994) and Tripod.com (1995). These early online communities were essentially focused on bringing people together to interact with each other through chat rooms, and encouraged users to share personal information and ideas via personal web pages by providing easy-to-use publishing tools and free or inexpensive web space. Other online communities (e.g., Classmates.com) followed a different approach: they simply allowed people to link each other via email addresses. In the late 1990s, thanks to the introduction of the concept of 'user profile' as a central feature of social networking services, users started to have the possibility to compile lists of 'friends' and search for other users with similar interests. By the end of the 1990s, new social networking methods were developed, and many sites began to develop more advanced features for users to find and manage friends (Livermore and Setzekorn 2009). SixDegrees.com in 1997, followed by Makeoutclub in 2000, Hub Culture and Friendster in 2002 represented the first 'new generation' social networking services, and soon became part of the Internet mainstream. Friendster was followed by MySpace and LinkedIn. Attesting to the rapid increase in social networking sites' popularity, by 2005, it was reported that MySpace was getting more page views than Google.[1]

Facebook, launched in 2004, is currently the largest SNS (Hampton et al. 2011). According to socialbakers.com, one of the biggest Facebook statistics portals in the world, at the time of writing the total amount of users is closing in to one billion users.[2] Hence, more or less 1 out of 7 people in the world have a Facebook-account.

Not only Facebook, but also Twitter has acquired a large market share nowadays. Even if it is difficult to determine the precise amount of users on Twitter, the number of 'tweets per day' (TPD) give an indication of the usage of this medium. The average TPD in March 2010 was 50 million according to Twitter statistics. The average

[1] http://www.businessweek.com/stories/2005-07-18/news-corp-dot-s-place-in-myspace

[2] http://www.socialbakers.com/countries/continents/

TPD in February 2011 was 140 million. In 2012, with over 200 million active users, over 400 millions tweets daily have been generated, handling over 1.6 billion search queries per day.[3] This huge increase of tweets gives a strong indication on what can be defined as the hype of today. It is uncertain whether this trend will sustain itself over time. In fact, not so long ago other SNS like Friendster and Myspace, were considered as the revelation of twenty-first century. In May 2011 however, Friendster repositioned itself from a social network service to a social gaming site. Likewise the number of MySpace users has declined immensely.[4] Like Friendster, Massive Media (the company behind Netlog), acquired by Meetic, has moved its scope to dating.[5] All these 'old' social network services failed to compete with Facebook and Twitter. Maybe the future will bring the same destiny for Facebook and Twitter, maybe not.

Google+, the new social network service of Google, is a new competitor on the SNS market. With 25 million users in 2011, Google+ has been the fastest website to reach that audience size[6] and it is nowadays the second largest social networking site in the world, having surpassed Twitter in January 2013.[7] As of December 2012, it has a total of 500 million registered users, of whom 235 million are active in a given month.

Regardless the success of specific social networks, their evolutions suggest us that social relationship layers on the Internet are here to stay and continue to gain ground (AA.VV SPION 2011).

Studies on Social Network Services

Different studies on real social network services have revealed a clear connection between offline and online communities. In particular, from a large survey study, Wellman et al. (2001) argued that, besides decreasing social capital in communities, online activities can also increase and supplement social capital (described in section "Social Capital") in different cases. In fact, SNS applications provide an infrastructure for social participation in online and offline communities that facilitates user contribution, communication, and even collaboration.

When conducting research on MySpace, danah boyd and Ellison (2007) found that teenagers are motivated to go on SNS because their offline friends are there too. Parks (2011) when studying MySpace, stated that "offline and online communities are linked in ways that we are only beginning to understand." Moreover, "…it may be more accurate to say that virtual communities are often simply the online extension of geographically situated offline communities."

[3] http://www.techvibes.com/blog/twitter-users-tweet-400-million-times-2012-12-17

[4] Statistics summary for myspace.com

[5] http://pulse2.com/2012/12/23/meetic-acquires-massive-media-for-25-million/

[6] http://in.reuters.com/article/2011/08/03/idINIndia-58589020110803

[7] http://www.forbes.com/sites/anthonykosner/2013/01/26/watch-out-facebook-with-google-at-2-and-youtube-at-3-google-inc-could-catch-up/

Lampe et al. (2006), in their 2006 study on the use of Facebook, found that it is used primarily for maintaining previous, offline relationships. Again, in 2011, according to Pew research center only a small fraction of Facebook friends, are people we have never met offline: 89 % of the friends we have on Facebook, we have met more than once offline (Hampton et al. 2011). Confirming these results, an empirical study by Ellison et al. (2011) shows that getting in touch over Facebook with completely unknown people does not influence users' social capital, though getting in touch with latent or weak ties, for social information-seeking activities, has a direct impact on social capital.

In the same way, Cha et al. 2009, studying information propagation in Flickr, showed that social links are a primary way users find and share information in social media (instead of using other features such as search and hot lists).

Similarly to these works, the focus of most research lies on the individual as a user when it comes to investigating online behavior on SNS, not the community referring to an individual embedded in a particular context. Aim of this chapter is therefore to focus on this particular aspect, connected in particular with privacy issues, as emerges from following sections.

Privacy in Social Networks Services

With respect to other Web applications, social network services present new challenges concerning privacy issues. SNS are built on interaction, they are typically open systems, and have certain semantic characteristics. Each privacy-related declaration has effects beyond the interaction between one individual data subject and one data collector, effects that may concern a number of members of a community who may or may not be users of the same system (Preibusch et al. 2007).

The Context Issue

As introduced in previous sections, the architecture of SNS does not allow sensing the community in the same way an offline world does, due in particular to the absence of a clear definition of the *situation*, as a way for users to act individually and as a community. In fact, in both scenarios (online and offline), only when the condition of a clear situation is satisfied can adequate behaviors be made possible. With adequate behavior, we mean behavior that takes into account all different aspects that (can) influence behavior in a certain *context*. In the online world, a lot of self-representative information is not put into context and this influences the performance of adequate behaviors, also regarding privacy concerns.

According to Hewitt and Shulman (2010): "A definition of the situation is an organization of perception in which people assemble objects, meanings, and others, and act toward them in a coherent, organized way. A definition of the situation, in other words, organizes meanings in such a way that people can act individually and jointly". A clear definition of the situation/context is exactly what is absent on SNS. There are many aspects an individual has to take into account, if it wants to perform adequate behaviors. In an offline world more or less clear barriers between contexts exist. Most of the time we know who is present in a situation, what conduct we ought to expect from others, what role we should perform, and where the situation is located. When mass self-communication enters the picture, this more or less clear context disappears (AA.VV SPION 2011).

Social and Instrumental Privacy

When the definition of the situation is not clear, performances on SNS become difficult in relation to privacy on mainly two levels: *social privacy* and *instrumental privacy*.

The former can be defined, according to Raynes-Goldie (2010) as "the control of information flow about how and when their personal information is shared with other people". It usually deals with *disclosure*.

The latter refers to the access by governments and corporations to users data, usually via *data mining* techniques (boyd and Hargittai 2010). Instrumental privacy in online environments deals with the problem of not awareness of people about what happens with their personal information, i.e., who and why they are gathered and the possibility for users to do something about it. In this scenario, individuals often lack every ability to act in a meaningful way (Solove 2001).

Disclosure and data mining in social network services are two macro areas including several privacy issues. Concerning the former area, main topics are self-disclosure (Krasnova et al. 2009), context collapse (boyd and Ellison 2007) or context collision (Raynes-Goldie 2010), and forced disclosure (Gross and Acquisti 2005). Concerning the latter area, both emergent and well known topics are represented by filter bubble (Pariser 2011) and link prediction (Lü and Zhou 2011). All these issues in both areas refer to major gaps in the architecture of SNS. These makes it hard for users to interact, represent themselves and create communities and on top of that bear in mind their social and instrumental privacy.

Disclosure

In general, information disclosure enables an attacker to gain valuable information about a user (or a system). In social network services, disclosure is often concomitant with the social network service use itself. In fact, according to the already cited definition of SNS provided by boyd and Ellison (2007), social network services allow the creation of "public or semi-public profiles within a bounded system", they foster the articulation of lists of personal connections within the system, and they allow the transversal of these connection lists within the system. This way, with respect to the general problem of information disclosure, it has become more evident that in SNS the problem of privacy is not bounded by the perimeters of individuals but also by the privacy needs of their social networks and of the communities they belong to. When information is disclosed on SNS (voluntarily or involuntarily), personal data can be utilized not only for the primary purposes for which they were collected. They can be utilized for secondary (from the perspective of the user) purposes that are covered in the SNS's terms of use and in that sense accepted by users (e.g., targeted marketing), but they can also be utilized for other illegal or unwanted purposes, both from the point of view of the user or the members of the community the user belongs to (indirectly affected by user's information disclosure).

For these reasons, particular attention must be provided in managing 'private' and 'public' data, according to the common classification of confidentiality levels. Preibusch et al. (2007) provide two further levels for classifying data confidentiality, taking into account specific 'group' and 'community' aspects of social network services:

- *Private data*: disclosed to the SNS operator for its internal purposes only, its disclosure needs explicit consent;
- *Group data*: disclosed to the SNS operator and accessed by other users of the same SNS that are also in the same group as the user; data disclosure is limited to the group;
- *Community data*: disclosed to the SNS operator and available to all registered and logged-in users of the SNS; the data is not accessible for anonymous SNS visitors;
- *Public data*: disclosed to the SNS operator and made accessible for all SNS visitors, including anonymous visitors.

Even if the concrete details and the application (and even the interpretation) of these confidentiality levels to data depends on the SNSs implementation, their correct definition and use could help in addressing the privacy issues described in the following sections.

Self-Disclosure

Prior research has considered a range of motivations for self-disclosure in social network services. According to the works of Goffman (1959), Donath and boyd (2004) and boyd and Heer (2006), users employ a social network service as a performance of identity. Strategically presenting themselves, through the constructed profiles, users' challenge is to increase their diverse networks of social ties. Similarly, Lampe et al. (2006) note that motivations for use and disclosure within a social network service are a function of offline outcomes such as relational formation and deepening. Works by Bumgarner (2007) and Joinson (2008) illustrate the social motive of social network service use and consequent personal data disclosure: the participants' desire to connect and learn about one another. Without significant personal sharing in these sites, these motives of use would not be addressed. For this reason, recent research points out that SNSs seem to require self-disclosure by default (Joinson et al. 2011; Nguyen et al. 2012).

The earliest studies on concrete social network services, provided empirical evidence of the remarkable disclosure practices within the sites. Work by Acquisti and Gross (2006) found that students in the Carnegie Mellon University Facebook network extensively shared sensitive information such as political views and sexual orientation in Facebook, and that information shared in Facebook was generally self-reported as valid. Other studies conducted at the time in different university networks, including Stutzman (2006) and Lampe et al. (2006), further evidenced the high degree of personal disclosure within social network services. Large scale studies such as Thelwall (2008) and James and Webb (2008) provided evidence of similar disclosure phenomena in Myspace, once the leading social network service. These findings were corroborated by a national probability study conducted by Lenhart and Madden (2007).

Despite this, it seems nowadays that users are becoming more and more aware of (at least some) privacy risks connected to social networking. In their study concerning the relationship between perceived privacy and comfort with self-disclosure, Frye and Dornisch (2010) analyzed the behavior of 214 US participants. They reported that participants tended to feel more comfortable disclosing information when they perceived the communication tools as offering a higher level of privacy. Concerning Facebook, its transition to a global social network service and the changes to the interface and to site policies have altered the level of trust individuals have in Facebook itself, which was often described as the more trusted social network service (in particular when compared with Myspace (Dwyer et al. 2007). To combat the increases in privacy and decreased disclosure to a wide audience in the platform, Facebook has consistently changed the nature of sharing certain items in the platform, and the default sharing settings for new accounts.

The increased awareness of users concerning privacy issues, and the consequent better use of privacy settings provided by online social network services, may help in addressing self-disclosure issues and take part in the management of context collapse.

Context Collapse

Context collapse refers to the challenge of managing disclosure across multiple social contexts in a social network service (Marwick and boyd 2011). Also known as context collision (Raynes-Goldie 2010), it represents a problem for social privacy. It refers to the blurring of contexts in an online environment, whereas in an offline environment more or less strict barriers can be distinguished. Combined elements of mass media and personal communication makes difficult to acquire a proper self-presentation to multiple audiences for people.

On the one hand, there is the idea that this problem cannot be solved, because disclosure networks is so large that according to some authors, the concept of privacy is ‘a zombie’[8] and ‘illusory’ (Hoadley et al. 2010). As stated in AA.VV SPION (2011), practice does not afford ongoing social surveillance of an entire network, but rather alters of particular situational interest. Indeed, the potential for large-scale surveillance exists, but does not occur in practice due to segmentation, non-participation and socio-technical affordance.

On the other hand, it has been showed that users on social network services seem to have the ability for balancing personal and public information. For example they avoid certain topics maintaining, at the same time, authenticity (boyd 2008). Other strategies employed by users to manage multiple contexts in social network services, have been illustrated in the work of Lampinen et al. (2009, 2011). This range of strategies includes self-censorship, and withdrawal of content, creating more inclusive group identities, and sharing different types of content in different spaces. In addition to these behavioral and mental strategies for context and privacy management, individuals also turn towards the application of privacy settings within the site. Numerous studies documented both increased use of privacy within Facebook by students (boyd and Hargittai 2010; Vitak 2013) and the contextual application of privacy settings in relation to perceived harms (Stutzman and Kramer-Duffield 2010), even if not always privacy settings match users’ expectations (Liu et al. 2011; Special and Li-Barber 2012).

Forced Disclosure

A problem, related to context collapse, is the phenomenon of *forced disclosure*. It follows the same principle of mandatory disclosure in the field of network security, where mandatory disclosure of vulnerabilities is considered a possible solution because it provide incentives for software firms to make the software code more secure and to quickly fix vulnerabilities that are identified (Choi et al. 2010). Similarly, in social network services, forced disclosure refers to the ongoing process of clarifying private information through private information (according to Rosen (2001)). This is necessary because a lot of self-representative information on social network services is not put into context; for this reason, the only way to clarify this is to augment the amount of disclosed (even private) information on these sites. According to AA.VV SPION (2011), when private information is disclosed, the only way of clarifying this is by giving more private information, in particular in situations presenting multiple context collisions (e.g., when a person breaks up his relationship with someone and changes his status from ‘in a relationship’ to ‘single’ only a couple of people will know exactly what happened. The majority of people will not).

The concept of ‘reciprocal self-disclosure’ (Sprecher et al. 2013) can also be considered a sort of ‘de facto’ forced disclosure. This kind of disclosure is ‘forced’ in the sense that, as it has been proved, participants who disclose reciprocally reports greater liking, closeness, perceived similarity, and enjoyment of the interaction after the first interaction than participants who disclose non-reciprocally.

[8] http://technosociology.org/?p=35

Data Mining

By analyzing the *big data*, i.e., the digital breadcrumbs of human activities sensed as a by-product of the ICT systems that we use, we have today the opportunity to observe and measure how our society intimately works. These data describe the daily human activities: e.g., automated payment systems record the tracks of our purchases, search engines record the logs of our queries for finding information on the web, social networking services record our connections to friends, colleagues and collaborators, wireless networks and mobile devices record the traces of our movements and our communications.

These social data are at the heart of the idea of a knowledge society, where decisions can be taken on the basis of knowledge in these data. Social data analysis can help us understand complex social phenomena, such as mobility, relationships and social connections, economic trends, spread of epidemics, opinion diffusion, sustainability, and so on.

The opportunities of discovering knowledge from social data increase with the risk of privacy violation: during knowledge discovery, the risk is the uncontrolled intrusion into the personal data of the data subjects, namely, of the (possibly unaware) people whose data are being collected, analyzed and mined. Privacy intrusion jeopardizes trust: if not adequately countered, they can undermine the idea of a fair and democratic knowledge society.

Filter Bubble

A *filter bubble* is a result state in which a website algorithm selectively guesses what information a user would like to see based on information about the user (such as location, past click behavior and search history) and, as a result, users become separated from information that disagrees with their viewpoints, effectively isolating them in their own cultural or ideological bubbles. Prime examples are Google's personalized search results and Facebook's personalized news stream.

The term was coined by internet activist Eli Pariser as "that personal ecosystem of information that's been catered by these algorithms" (Pariser 2011); according to Pariser, users get less exposure to conflicting viewpoints and are isolated intellectually in their own informational bubble. For Pariser, the detrimental effects of filter bubbles include harm to the general society in the sense that it has the possibility of "undermining civic discourse" and making people more vulnerable to "propaganda and manipulation". This constitutes a concrete problem in particular for social network service users and the possibility for them to act as a community: according to Miconi (2013) being a bubble built upon individual tastes and preferences, it does not allow any kind of sharing: in short, everybody is 'alone' in the bubble, condemned to find his own way to knowledge. Again, the bubble it is invisible, and, unlike traditional media, it does not reveal its bias and selectiveness. For the same reason, whether users like it or not, they can not choose to enter the bubble: participants are not allowed to actively select the filter.

In addition to this problem, filter bubble presents the same privacy issues connected to algorithms collecting information concerning users: once a user has been observed, profiled and recognized on subsequent visit, according to Parsier the risk posed in the filter bubble are not undone with a simple ‘privacy settings adjustment’.

Link Prediction

Link prediction is a sub-field of social network analysis. Link prediction is concerned with the problem of predicting the (future) existence of links among nodes in a social network (Liben-Nowell and Kleinberg 2003). Link prediction is the only sub-field of SNA which has focus on links between objects rather than objects themselves. This makes link prediction interesting and different from traditional data mining areas which focus on objects.

Link prediction can lead to privacy concerns when the predicted link is between users who consider this link to be private. In this case, a sensitive link disclosure occurs. In social network data, for example, the friendship relationships of a person and the public preferences of the friends such as political affiliation, may lead to infer the personal preferences of the person in question as well. Therefore, studying how to prevent sensitive link disclosure while providing accurate link recommendations is an important problem.

To solve it, different strategies have been proposed in literature. Concerning the node data, they are usually anonymized with ‘classical’ *k*-anonymity (Samarati 2001) techniques, or more recent and refined *l*-diversity (Machanavajjhala et al. 2007) and *t*-closeness (Li et al. 2007) techniques.

For the edge data, different anonymization strategies have been proposed. In Zheleva and Getoor (2008), five possible anonymization approaches are described. They range from one which removes the least amount of information to a very restrictive one, which removes the greatest amount of relational data. Bhagat et al. (2010), provide methods to anonymize a dynamic network when new nodes and edges are added to the published network exploiting link prediction algorithms to model the evolution. Using this predicted graph to perform group-based anonymization, the loss in privacy caused by new edges can be eliminated almost entirely. In Xue et al. (2012), authors theoretically establish that any kind of structural identification attack can be prevented using random edge perturbation techniques. This is confirmed also in Díaz and Ralescu (2012).

Privacy Settings

According to previous sections, many and different are the ways leading to attempts to instrumental and social privacy of users. This is often facilitated, in current social network services, by the way privacy settings are either implemented or used.

Let us take into consideration Facebook, nowadays the most popular and widespread social network service. At the present moment, Facebook allows users to manage the privacy settings of uploaded content (photos, videos, statuses, links and notes) using five different granularities: Only Me, Specific People, Friends Only, Friends of Friends, and Everyone. Specific People allows users to explicitly choose friends (or pre-created friend lists, discussed below) to share content with. The default or ‘recommended’ privacy setting for many pieces of content is Everyone, meaning users share their content with all one billion Facebook users if they decline to modify their privacy settings. Facebook allows users to re-use Specific People privacy settings via friend lists. Users create a friend list, add a subset of their friends to it, name it, and can then select the list as a basis for privacy control. Friend lists are private to the user who creates them, unless the user explicitly chooses to display them as part of his profile. The granularity of privacy settings varies according to content type. Photos are grouped into albums, and privacy settings are specified on an album granularity (i.e., all photos in an album must have the same privacy setting). For the remaining content types, users can specify different privacy settings for each piece of content.

As introduced along the chapter, users awareness and use of these settings have changed over time. For example, from early empirical studies, Facebook users in the United States had inconsistent behavior with respect to privacy concerns, demonstrating excessive sharing of personal data and rare changes to default privacy settings (Gross and Acquisti 2005), even users who claimed to be concerned about privacy (Acquisti and Gross 2006). Still in 2006–2008 a low percentage of Facebook profiles in US were restricted to ‘friends only’ (Lampe et al. 2008). The situation was slightly different in U.K., where in 2008 the majority of the respondents (57.5 %) reported having changed the default privacy settings (Joinson 2008).

Now that more recent studies suggest that users are becoming more privacy concerned and more likely to change their privacy settings (boyd and Hargittai 2010), some problems still remain. In fact, according to Liu et al. (2011) and Madden (2012), users are not completely satisfied about social networks way to protect their privacy. The complexity of privacy settings varies greatly across different social network services. In all, according to Madden (2012), 48 % of social networks users still report some level of difficulty in managing the privacy controls on their profile. Few users (2 %) describe their experiences as ‘very difficult’, while 16 % say they are ‘somewhat difficult’. In particular, social networks users who are college graduates are significantly more likely than those with lower levels of education to say that they experience some difficulty in managing the privacy controls on their profiles. In addition to this, according to Liu et al. (2011), 36 % of the Facebook content still remains shared with the default privacy settings and, overall, privacy settings match users’ expectations only 37 % of the time, and when incorrect, almost always expose content to more users than expected.

Contextual and Demographics Privacy Concerns

The development of Facebook in 2004 as a university network represented yet a meaningful privacy boundary between students from family, employers, and municipal law enforcement. With Facebook's growth in popularity, users have to deal with the presence of multiple contextual networks in the site. As a result, the known audience and the expected audience in social network services do not always overlap (boyd and Heer 2006; Lampe et al. 2008; Stutzman and Kramer-Duffield 2010). This can be intended to mean that within a system with hundreds of articulated connections, disclosures are intended for a subset of the audience. In most cases, one does not expect their disclosure to range beyond a certain subset of alters. The implication of this finding is often in collision with discourses that argue that disclosure in a socio-technical system is intended to be public (AA.VV SPION 2011).

Demographics seem to affect privacy attitudes and behaviors of social network service users. In general men had less privacy concerns than their female counterparts, and thus tended to disclose more personal information such as telephone numbers and physical addresses on SNSs (Fogel and Nehmad 2009; Madden 2012). Female users and users who have more Facebook friends are more likely to have friends-only profiles (Stutzman and Kramer-Duffield 2010). In addition to this, individual characteristics such as Internet skill, frequency, and type of Facebook use are correlated with making modifications to privacy settings (boyd and Hargittai 2010; Madden 2012). Users display more concern about sharing with their weak-tie friends than with outsiders or companies. Stutzman and Kramer-Duffield suggest that users adopt friends-only profiles mainly to deal with unintended disclosure to their weak ties rather than outsiders (Stutzman and Kramer-Duffield 2010). Raynes-Goldie found that users cared more about protecting information from members of various social circles, rather than protecting their information from companies (Raynes-Goldie 2010).

Conclusions and Further Research

As emerges from the literature review, privacy management in social network services is receiving growing attention, in particular when connected to context. In fact, privacy risks emerge above all when individuals are forced to manage their disclosures between different situations and spheres of life, across different communities representing for example the professional and personal spheres, or even communities within an 'augmented reality'. That is, a reality we experience that superimposes a layer of virtual data on top of our actual 'sensate' reality. This mix of virtuality and reality adds useful contextual information, that could be used to better protect users' privacy. At the same time, this poses serious data inference problems.

For all these reasons, in last years, the architecture of SNS has been subjected to constant renovation. In order to help users in managing their privacy settings, 'privacy wizards' or recommendation tools have been proposed, based on the observation that real users conceive their privacy preferences based on an implicit social network structure (Fang et al. 2010). Relationship-Based Access Control (ReBAC) techniques follow the same paradigm: ReBAC is characterized by the explicit tracking of interpersonal relationships between users, and the expression of access control policies in terms of these relationships, capturing the contextual nature of relationships themselves (Fong 2011) and trust Carminati et al. (2012) among users. Taking into account concrete social network services, designers have attempted to address the problem of group context management through the inclusion of technical features enabling the grouping of contacts. The Facebook ‘Friends List’ feature allows users to aggregate friends according to individually-defined criteria, and then selectively disclose to these lists. Unfortunately, Facebook’s system is still considered in some way too complicated and/or insufficient to provide privacy at the group level (boyd and Hargittai 2010; Liu et al. 2011; Madden 2012). Google+, the Google social network, aims to “bring the nuance and richness of real-life sharing to software”. Google+ has defined ‘circles’ of life where individuals can place their contacts, and share accordingly (Kairam et al. 2012). Thanks to this intuitive feature, Google+ puts effort in making the group management process more simple. In spite of this, the Google+ ‘real name’ policy and the difficulty to enforce privacy concerns over data associated with multiple users, lead influential critics to challenge the privacy gains of Google+ (Hu et al. 2011; AA.VV SPION 2011).

All these efforts are not still sufficient in our opinion. In fact, as also boyd suggests in boyd and Marwick (2011) on the topic of privacy, the solution to this puzzle will not be to restrict data collection or to enhance individual control over specific items of data, but “to think long and hard about what happens as the data flows across networks and as the data is networked together”. In fact, in the current Social Web vision of the Net, different (contextual) graphs often unifies multiple data flows and social networks, and consequently personal information they provide (Berlingerio et al. 2011).

For these reasons, it is necessary to put more emphasis on the interconnections between offline and online world in achieving privacy, and on the concept of context both intra and inter social network services. When the architecture of SNS will be improved in a way to better take into account these issues, numerous problems connected to identity protection and privacy will be probably solved.

References

AA.VV. SPION – D2.1 – State of the Art. In: Gürses S, Leuven KU (eds.) Technical report WP2 – D2.1, SBO Security and Privacy for Online Social Networks, Sept 2011

Acquisti A, Gross R (2006) Imagined communities: awareness, information sharing, and privacy on the facebook. Volume 4258 of Lecture notes in computer science. Springer, Berlin/Heidelberg

Baum F, Ziersch AM (2003) Social capital. J Epidemial Commun Health 57(5):320–323

Bauman Z (2001) Identity in the globalizing world. Soc Anthropol 9(2):121–129

Berlingerio M, Coscia M, Giannotti F, Monreale A, Pedreschi D (2011) Foundations of Multidimensional Network Analysis, Advances in Social Networks Analysis and Mining (ASONAM), International Conference, pp. 485–489. doi:10.1109/ASONAM.2011.103

Bhagat S, Cormode G, Krishnamurthy B, Srivastava D (2010) Prediction promotes privacy in dynamic social networks. In: Proceedings of the 3rd conference on online social networks, WOSN’10, Berkeley. USENIX Association

boyd d (2008) Taken out of context: American teen sociality in networked publics. Master’s thesis, University of California boyd d, Ellison NB (2007) Social network sites: definition, history, and scholarship. Journal of Computer-Mediated Communication 13(1):210–230

boyd d, Hargittai E (2010) Facebook privacy settings: who cares? First Monday 15(8)

Boyd D, Heer J (2006) Profiles as Conversation: Networked Identity Performance on Friendster, System Sciences, 2006. HICSS '06. Proceedings of the 39th Annual Hawaii International Conference, 3:59c. doi:10.1109/HICSS.2006.394

boyd D, Marwick AE (2011) Social privacy in networked publics: teens attitudes, practices, and strategies. In: Proceedings of the A DECADE IN INTERNET TIME: OII SYMPOSIUM ON THE DYNAMICS OF THE INTERNET AND SOCIETY, 21–24 September 2011, University of Oxford.

Bumgarner B (2007) You have been poked: exploring the uses and gratifications of facebook among emerging adults. First Monday 12(11)

Calhoun G (1998) Community without propinquity revisited: communications technology and the transformation of the urban public sphere. Sociol Inq 68(3):373-397

Carminati B, Ferrari E, Viviani M (2012) A multi-dimensional and event-based model for trust computation in the social web. In: Proceedings of the 4th international conference on social informatics – socInfo 2012, Lausanne, 5–7 Dec 2012. Volume 7710 of Lecture notes in computer science. Springer, pp 323–336

Carroll JM (2011) The neighborhood and the internet: design research projects in community informaics. Routledge, London

Castells M (1996) The rise of the network society. Oxford, Blackwell

Castells M (2009) Communication power. Oxford University Press, New York

Cha M, Mislove A, Gummadi KP (2009) A measurement-driven analysis of information propagation in the flickr social network. In: Proceedings of the 18th international conference on world wide web, WWW'09, Madrid. ACM, New York, pp 721–730

Choi JP, Fershtman C, Gandal N (2010) Network security: vulnerabilities and disclosure policy. J Ind Econ 58(4):868–894

Delanty G (2003) Community. Routledge, New York

Díaz ID, Ralescu A (2012) Privacy issues in social networks: a brief survey. In: Greco S, Bouchon-Meunier B, Coletti G, Fedrizzi M, Matarazzo B, Yager R (eds) Advances in computational intelligence, Brisbane. Volume 300 of Communications in computer and information science. Springer, Berlin/Heidelberg, pp 509–518

Donath J, Boyd D (2004) Public displays of connection. BT Technol J 22(4):71–82

Dwyer CA, Hiltz SR (2008) Designing privacy into online communities. Social science research network working paper series, Nov

Dwyer C, Hiltz S, Passerini K (2007) Trust and privacy concern within social networking sites: a comparison of facebook and mySpace. In: Proceedings of the 13th Americas conference on information systems (AMCIS), Keystone, Aug 9–12, 2007

Ellison NB, Steinfield C, Lampe C (2011) Connection strategies: social capital implications of facebook-enabled communication practices. New Med Soc 13(6):873–892

Fang L, Kim H, LeFevre K, Tami A (2010) A privacy recommendation wizard for users of social networking sites. In: Proceedings of the 17th ACM conference on computer and communications security, CCS'10, Chicago. ACM, New York, pp 630–632

Fogel J, Nehmad E (2009) Internet social network communities: risk taking, trust, and privacy concerns. Comput Hum Behav 25(1):153–160

Fong PW (2011) Relationship-based access control: protection model and policy language. In: Proceedings of the 1st ACM conference on data and application security and privacy, CODASPY'11, San Antonio. ACM, New York, pp 191–202

Frye NE, Dornisch MM (2010) When is trust not enough? The role of perceived privacy of communication tools in comfort with self-disclosure. Comput Hum Behav 26(5):1120–1127

Giddens A (1990) The consequences of modernity. Polity, Cambridge

Goffman E (1959) The presentation of self in everyday life. Anchor Books, New York

Granovetter M (1983) The strength of weak ties: a network theory revisited. Sociol Theory 1:201–233

Gross R, Acquisti A (2005) Information revelation and privacy in online social networks. In: Proceedings of the 2005 ACM workshop on privacy in the electronic society, WPES'05, Alexandria. ACM, New York, pp 71–80
Hampton K, Goulet LS, Rainie L, Purcell K Social networking sites and our lives. Technical report, June 2011. Available from http://www.pewinternet.org/Reports/2011/Technology-and-social-networks.aspx
Hauben M, Hauben R (1997) Netizens: on the history and impact of Usenet and the Internet. IEEE Computer Society Press, Los Alamitos
Hewitt JP, Shulman D (2010) Self and society: a symbolic interactionist social psychology. Pearson
Hoadley CM, Xu H, Lee JJ, Rosson MB (2010) Privacy as information access and illusory control: the case of the Facebook News Feed privacy outcry. Electron Commer Res Appl 9(1):50–60. (Special issue: Social Networks and Web 2.0)
Hu H, Ahn GJ, Jorgensen J (2011) Detecting and resolving privacy conflicts for collaborative data sharing in online social networks. In: Proceedings of the 27th annual computer security applications conference, ACSAC'11, Orlando. ACM, New York, pp 103–112
James, Webb S (2008) A large-scale study of mySpace: observations and implications for online social networks. In: ICWSM'08, Seattle, vol 8
Joinson AN (2008) Looking at, looking up or keeping up with people? Motives and use of Facebook. In: Proceedings of the SIGCHI conference on human factors in computing systems, CHI'08, Florence. ACM, New York, pp 1027–1036
Joinson AN, Houghton DJ, Vasalou A, Marder BL (2011) Digital crowding: privacy, self-disclosure, and technology. In: Trepte S, Reinecke L, (eds) Privacy online. Springer, Berlin/Heidelberg, pp 33–45
Kairam S, Brzozowski M, Huffaker D, Chi E (2012) Talking in circles: selective sharing in Google+. In: Proceedings of the 2012 ACM annual conference on human factors in computing systems, CHI'12, Austin. ACM, New York, pp 1065–1074
Kivisto P (2003) Key ideas in sociocology. Pine Forest Press, New Delhi
Krasnova H, Kolesnikova E, Günther OG (2009) "It won't happen to me!": self-disclosure in online social networks. In: Proceedings of the 15th Americas conference on information systems, AMCIS 2009, San Francisco, Aug 6–9, 2009. Association for Information Systems, pp 343
Lampe C, Ellison N, Steinfield C (2006) A face(book) in the crowd: social searching vs. social browsing. In: Proceedings of the 2006 20th anniversary conference on computer supported cooperative work, CSCW'06, Banff(Alberta). ACM, New York, pp 167–170
Lampe C, Ellison NB, Steinfield C (2008) Changes in use and perception of Facebook. In: Proceedings of the 2008 ACM conference on computer supported cooperative work, CSCW'08, San Diego. ACM, New York, pp 721–730
Lampinen A, Tamminen S, Oulasvirta A (2009) All my people right here, right now: management of group co-presence on a social networking site. In: Proceedings of the ACM 2009 international conference on supporting group work, GROUP'09, Sanibel Island. ACM, New York, pp 281–290
Lampinen A, Lehtinen V, Lehmuskallio A, Tamminen S (2011) We're in it together: interpersonal management of disclosure in social network services. In: Proceedings of the SIGCHI conference on human factors in computing systems, CHI'11, Vancouver. ACM, New York, pp 3217–3226
Lenhart A, Madden M (2007) Teens, privacy & online social networks. How teens manage their online identities and personal information in the age of MySpace. Available from http://www.pewinternet.org/pdfs/PIP_Teens_Privacy_SNS_Report_Final.pdf
Li N, Li T, Venkatasubramanian S (2007) t-closeness: privacy beyond k-anonymity and l-diversity. In: Proceedings of the ICDE'07: IEEE 23rd international conference on data engineering, Istanbul, pp 106–115
Liben-Nowell D, Kleinberg J (2003) The link prediction problem for social networks. In: Proceedings of the 12th international conference on information and knowledge management, CIKM'03, New Orleans. ACM, New York, pp 556–559

Liu Y, Gummadi KP, Krishnamurthy B, Mislove A (2011) Analyzing Facebook privacy settings: user expectations vs. reality. In: Proceedings of the 2011 ACM SIGCOMM conference on internet measurement conference, IMC'11, Berlin. ACM, New York, pp 61–70
Livermore CR, Setzekorn K (2009) Social networking communities and E-Dating services: concepts and implications. IGI Global, Hershey
Lü L, Zhou L (2011) Link prediction in complex networks: a survey. Physica A Stat Mech Appl 390(6):1150–1170
Lyon L (1986) The community in urban society. Waveland Press, Dawley
Machanavajjhala A, Kifer D, Gehrke J, Venkitasubramaniam M (2007) L-diversity: privacy beyond k-anonymity. ACM Trans Knowl Discov Data 1(1):24
Madden M (2012) Privacy management on social media sites. Technical report, Pew Research Center,Feb2012.Availablefrom:http://www.pewinternet.org/Reports/2012/Privacy-management-on-social-media.aspx.
Marwick AE, Boyd D (2011) I tweet honestly, I tweet passionately: twitter users, context collapse, and the imagined audience. New Med Soc 13(1):114–133
Miconi A The big G and the never-ending story of knowledge monopolies. Mediascapes J (1):158–160
Nguyen M, Bin YS, Campbell A (2012) Comparing online and offline self-disclosure: a systematic review. Cyberpsychol Behav Soc Networking 15(2):103–111
Nisbet R (1976) The quest for community. Free Press, New York
Pariser E (2011) The filter bubble: what the internet is hiding from you. The Penguin Press, New York
Parks MR (2011) Social network sites as virtual communities. In: A networked self: identity, community, and culture on social network sites, Routledge, New York/London
Pierson J, Heyman R (2011) Social media and cookies: challenges for online privacy. Emerald J 13(6):30–42
Preibusch S, Hoser B, ürses SG, Berendt B (2007) Ubiquitous social networks – opportunities and challenges for privacy-aware user modelling. In: Proceedings of the workshop on data mining for user modelling at UM 2007, 2007. Available from: http://vasarely.wiwi.hu-berlin.de/DM.UM07/Proceedings/DM.UM07-proceedings.pdf
Raynes-Goldie K (2010) Aliases, creeping, and wall cleaning: Understanding privacy in the age of Facebook. First Monday, 15(1)
Rheingold H (2000) The virtual community: homesteading on the electric frontier. MIT, Cambridge
Rosen J (2001) Out of context: the purposes of privacy. Soc Res 68(1):209222
Samarati P (2001) Protecting respondents' identities in microdata release. IEEE Trans Knowl Data Eng 13(6):1010–1027
Scott JP (2012) Social network analysis. Sage, Borgatti
Solove DJ (2001) Privacy and power: computer databases and metaphors for information privacy. Stanf Law Rev 53(6):1393–1462
Special WP, Li-Barber KT (2012) Self-disclosure and student satisfaction with Facebook. Comput Hum Behav 28(2):624–630
Sprecher S, Treger S, Wondra JD, Hilaire N, & Wallpe K (2013) Taking turns: Reciprocal self-disclosure promotes liking in initial interactions. Journal of Experimental Social Psychology, 49, 860–866
Stutzman F (2006) An evaluation of identity-sharing behavior in social network communities. iDMAa J 3(1)
Stutzman F, Kramer-Duffield J (2010) Friends only: examining a privacy-enhancing behavior in facebook. In: Proceedings of the SIGCHI conference on human factors in computing systems, CHI'10, Atlanta. ACM, New York, pp 1553–1562
Thelwall M (2008) Social networks, gender, and friending: an analysis of myspace member profiles. J Am Soc Inf Sci Technol 59(8):1321–1330
Tönnies F, Loomis CP (1957) Community & society (gemeinschaft und gesellschaft). Michigan State University Press, East Lansing

Vitak J (2013) Measuring privacy in online spaces: considerations for users attitudes, disclosures, and network composition. In: Measuring networked social privacy workshop, CSCW13, Feb 23–27, 2013, San Antonio (to appear)

Wellman B, Haase AQ, Witte J, Hampton K (2001) Does the internet increase, decrease, or supplement social capital? Social networks, participation, and community commitment. Am Behav Sci 45(3):436–455

Xue M, Karras P, Chedy R, Kalnis P, Pung HK (2012) Delineating social network data anonymization via random edge perturbation. In: Proceedings of the 21st ACM international conference on information and knowledge management, CIKM'12, Maui. ACM, New York, pp 475–484

Zhang G, Jacob EK (2012) Community: issues, definitions, and operationalization on the web. In: Proceedings of the 21st international conference companion on world wide web, WWW'12 companion, Lyon. ACM, New York, pp 1121–1130

Zhang S, Jiang H, Carroll J (2011) Integrating online and offline community through Facebook. In: International conference on collaboration technologies and systems (CTS), 2011, Philadelphia, pp 569–578

Zheleva E, Getoor L (2008) Preserving the privacy of sensitive relationships in graph data. In: Proceedings of the 1st ACM SIGKDD international conference on privacy, security, and trust in KDD, PinKDD'07, San Jose. Springer, Berlin/Heidelberg, pp 153–171

Applying Security Lessons Learned to Human Computation Solving Systems

Dan Thomsen

This chapter looks at the security issues that arise when using human computation systems to solve problems that no one has solved before. Researchers have spent decades on computer security research and yet surprisingly the biggest factor impacting security issues remains economics. Researchers know how to build secure systems, but cannot develop high assurance software fast enough to keep up with the feature race that shapes modern IT products. Techniques that attempt to crowdsource formal verification may reduce the time it takes for formal assurance, but formal assurance of any kind adds an extra step that slows time to market. The first product to market often has tremendous payoffs in terms of capturing market share. Today the rich feature environment and integration of millions of lines of code into even a simple application have made "security" mean getting hacked less than your competitors. True security means good architecture to control, but more importantly, good architecture to understand the flow of data in a system. You cannot secure what you do not understand. When looking at using a crowd of humans to solve problems, unique security issues arise because the developers must understand how humans impact security.

Money in both the terms of development costs, and reducing the time to market to capture early market share, shaped today's security mechanisms. Human computation systems must build on those flawed mechanisms. However human computation faces unique security challenges, and there is a chance at this early stage to think deeply about these issues and get them right. However, the pessimist in me says follow the money. Money and economics will shape the security philosophies that emerge for human computation unless the groundwork for good security and good architecture gets created early.

D. Thomsen (✉)
SIFT LLC, USA
e-mail: dthomsen@sift.net

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_71,

Driving security choices for human computation systems with economics may seem odd for an emerging science where all the successful examples run on volunteer contributions, but money defines the critical pieces of data targeted for malicious activity. Security analysis identifies these critical pieces of data as assets to be protected. The economics of computer security ensure that you cannot protect all the assets, so when you have limited funds the wisest approach suggests investing security to cover the most valuable assets. This chapter discusses the computer security basics, and then how that reflects on the assets of a human computation system and what protection they need.

Computer security consists of three aspects; confidentiality, integrity, and availability. Confidentiality means ensuring only the right people see they data. Integrity means ensuring only the right people or processes can modify the data. Availability means ensuring that people can access they data they need when they need it. These same aspects apply to human computation, but there are some unique problems arising for human computation. Researchers have postulated many different models of human computation, but by definition all these model involve humans aiding the computation. Humans have very different security properties and behaviors that will influence necessary changes.

For example, consider confidentiality. Often, proprietary information from several different stakeholders may occupy the same computer. A straight-foreword case involves business competitors sharing multi-million dollar high performance computers optimized for running complex models. Competitors can share this expensive resource securely because the computing service can wipe the computer completely blank before receiving the next competitor's model. Unfortunately, if we contemplate an analogous system that incorporates humans as computational elements, each of whom operates on proprietary data, we have no control on what aspects of the data the human participants may remember. Thus to use human computation safely for problem solving, each business would require its own population of human solvers to guarantee their secrets remain secret. This could result in stiff competition to recruit the best human solvers. Of course humans can keep secrets if motivated to by agreements, or laws. Even then, people often disclose little pieces of information they personally deem public or unimportant. If a competitor can aggregate all these little disclosures they might learn something about their competitor. A business does not have to worry about its computers making such self-directed decisions. Human computation systems confidentiality mechanism must address the fact that humans do not always act in predictable ways.

Humans also have different models of data integrity. Integrity ensures that people and processes only modify data in well-defined and understood ways. Computers are very good at following precise rules for modifying data. On the other hand humans almost always put in their own cognitive biases. For example, human solvers tasked with culling data with a specific rubric may systematically develop their own rubric culling more data than desired. Wikipedia serves as an example, where contributors change text descriptions to fit their own biases. Audit logs and history mechanisms can always track these changes, and technology can rollback the changes, but someone must know there was an erroneous change to start with. One can easily envision human computation systems that evolve to contain the cognitive biases of the dominant sub-population of human solvers. When the system hopes to incorporate multiple viewpoints to cover different possible solutions, bias creep could shrink the number of viewpoints. Standard computer systems only contain the biases of their developers. Human computation systems will contain a myriad of biases. Attackers may change the computers behavior by modifying software, but the computer can be purged and start again with a clean state. Malicious users could attack the integrity of the human solvers directly, consciously and maliciously embedding biases in the solver population to prevent finding a solution.

Availability ensures that computer resources and data are available when needed. Providing availability for different parties using a shared resource remains a challenging problem. Denial of service attacks require little sophistication, but have required service providers to greatly expand their processing power to deal with them. In a human computation system with a large enough crowd spanning the globe there could always be a population of humans available. The question remains are they the correct humans for the problem being solved? Do they have the skills and knowledge to contribute quality effort? Also just because a human solver has been assigned a task does not mean they are actually performing it. Incentives may motivate humans to contribute in a timely fashion, but there are no guarantees. If incentives work to motivate contributors, what if a competing system has better incentives? How does one maintain a crowd of solvers if competitors are willing to pay more for solver services?

Human computation may require an additional security consideration beyond confidentiality, integrity and availability. Human computation systems consist of large distributed systems with crowds of human solvers each of whom may have a different world view and agenda. A security rule or mechanism that the system designers have carefully thought about and implemented to ensure proper behavior may seem arbitrary and whimsical to a given human solver that does not understand why the rule exists. Thus, human computation systems may benefit greatly by adding a forth security aspect of "why" to the system, that provides a rationale for system rules.

In high assurance systems, the rationale for every critical security decision exists in formal arguments to prove the software functions correctly. But this "why" never gets passed along to the human users, and in fact may be too complex for the users to understand. Without having a "why" component, security mechanisms could demotivate human solvers by making the system seem burdensome for no purpose.

So far people who understand the system create the security policies. They know why each security decision was made. Human computation distributes the system to many different humans, most of whom know nothing about the goals and reasons for the security policy. Dictating rules to humans volunteering their time to contribute will not be as successful as explaining to humans that the imposed measures protect the critical assets of the system from compromise for stated reasons. If people think a security rule or mechanism is arbitrary they will bypass it when they personally think the benefits outweigh the risks. However, since they do not understand the whole system and all its goals, even well intentioned solvers will make choices that compromise the developer's vision, putting the assets at risk. Human computation requires a security environment that motivates compliance, not defiance.

Human Computation Assets

What are the assets in a human computation environment? For example, solvers might add a file to a specific repository, or they may simply post on an Internet forum. Each of these interactions leaves a trace on the final solution, and becomes part of the critical assets the human computation system must protect.

The organization sponsoring the human computation environment shapes the security solution. Deciding on, implementing, and enforcing security falls squarely on the shoulders of the sponsors. The sponsors benefit from finding a solution, so security compromises that hinder finding a solution directly impacts the sponsors.

A strawman list of human computation assets includes:

- Solutions—solutions emerge from human solver interactions. They cannot be destroyed because the human solvers can recreate them, but they can be stolen.
- Problem specifications—A concise problem specification that motivates human contributors to donate their time is critical to find a solution
- Contributions to a solution—any human solver interaction that moves closer to a solution
- Contributions that do not lead to a solution—these interactions have value because they document parts of the solution space that have been explored and eliminated. Losing them might mean others would invest time exploring the same space again
- The human solvers themselves—they represent the most valuable asset, as no solution will be found without them
- The human computation solution environment—an environment that supports massive collaboration and that can produce solutions to unsolved problems has tremendous value
- Rewards for human solvers—any reward that motivates human solvers will be desired and face security threats by people that want reward, but do not want to do the work.

Intellectual Property

As the asset list for a HC-based problem solving system shows, the intellectual property includes the problem specification, the solution and all the contributions by human solvers. Not all of this IP has the same intrinsic worth. The likelihood of finding a solution increases the more the solvers share information. The more information you share the less control you have over it. Human solvers will need a bare minimum of information simply to get started. Protected IP that no one sees does not help produce a solution, since information has no value unless it allows a human to make a better decision.

Consider, for example, the problem statement for a project that was successful because it found a solution. In this case the problem statement served as an effective marketing campaign that got the right human solvers interested and engaged in solving a problem. If the sponsors did not share the problem statement, chances are the right solvers never heard of the problem. Sharing information provides critical momentum for the project's success.

From a computer security point of view, sharing with another human represents letting the cat out of the bag. Technology cannot put the cat back in the bag, or even allow a peek at the cat in the bag without letting the cat out. Once information transfers to a human mind, that human can duplicate the information and bypass any technology. In the world of government security they have created a procedure to address this fact based on clearances. A security clearance represents a contract between two parties to share the information in controlled and predefined ways. A security clearance represents trust between two parties. For the most part trust in clearances works, but it can also fail spectacularly in cases such as wikileaks (Keller 2011).

Human computation environments may need to create an agreement that parallels government clearances. It might be as simple as a non-disclosure agreement, or it could be a complex set of clearances that allows different solvers to see different pieces of IP. Whatever the agreement winds up being, it must have some teeth, some penalty for the human solvers that break the trust. Breaking a non-disclosure agreement could be resolved in a court of law, but proving the amount and value of information disclosed may make such a court case hard to win, or simply too cost prohibitive to ever enforce. Other penalties might include ostracizing violators from the site or other reputation degradation penalties for violators. Reputation penalties require associating the solver's real world identity to the human computation environment to ensure the penalty actually penalizes the violator.

In the case of altruistic goals, sharing the IP maybe considered a good thing, so no agreement is needed. But IP that solves a technical problem may be repurposed to solve other related problems. For example, a solver could learn something from the computation system and use that to start a competing product. The sponsors would harvest no benefit from that product, but maybe they would accept that risk to allow them to make progress on their core problem. The sponsors must decide if they would rather reap rewards from partial or tangential solutions, or if they want a free exchange to increase the chance of finding a solution.

The Crowd

The human solver crowd represents the most valuable commodity for the sponsors. Without human solvers the solution stagnates. Jane McGonigal has postulated an engagement economy that competes for the eyes and brains of humans to join into human computation environments (McGonigal 2011). Many collaboration sites exist that never got the minimum number of people involved to make progress toward the goal. Solvers typically can change their minds and switch alliances to other sites, or simply decide to no longer invest their time in a specific site.

For example, suppose a person did not like the altruistic goal of a human computation environment. How can they prevent the sponsors from reaching that goal? They can push all the human solvers away from the site. They could try doing it with a negative advertisement campaign, but a more subtle and effective approach would be to pose as a legitimate solver and cause the system to crash or become unstable. If people think a web site is unstable, or the site appears to drop their work they will choose to stop investing their time in that site, even if they still believe in the altruistic goal.

Another attack on the crowd of human solvers would be a well-placed trolling attack. Comments like, "You call that a logical argument?" or, "That will never work!" can derail a collaborative effort. If human solvers feel no one appreciates their contributions, or if associating with the site makes them feel bad they will stop interacting with the site. If the site allowed anonymous posting, you could envision a malicious bot that randomly posts negative comments on forum threads, which we will call a robo-troll attack. With no human interaction by the attacker, the overall sentiment of the site becomes negative and the attrition rate of legitimate solvers will climb.

A more subtle robo-troll attack involves posting legitimate, but dumb answers. The software could use even poor natural language processing software to create posts that sound like they are related, but that make no sense. These poor posts will waste legitimate human solvers time reading and responding to them. Eventually they may feel they contributions are falling on deaf ears and drop out of the project.

Tying human solvers to their real world identity will curb robo-troll attacks because at the very least a single human must register the account. Unfortunately, once the account is created the malicious attacker could install a robo-troll that continues to post around the clock greatly magnifying the amount of damage a single attacker can do. This implies human computation environments will need a reputation system that eventually silences people that continuously make unconstructive contributions. Such censorship must be clearly explained to the other solvers so they do not feel the site has become draconian.

Reward

Many human computation environments build motivation by rewarding the human solvers for their participation. The reward may be as simple as an in game reputation score, or it may be a large monetary rewards for winning a contest. Any reward that motivates human solvers will be coveted by people who want the reward without doing the work.

What types of rewards are there?

- Money
- Reputation
- Altruistic
- In environment rewards
- Education/knowledge

Money

A human computation environment may use money in a variety of ways. The amount can range from micro-payments to large cash payouts for winning a contest. The sponsors determine the rewards based on the behavior they need to motivate, and the available budget.

Sites like Amazon Mechanical Turk provide micro-payments for doing tasks requiring human insights. Micro-payments can motivate people with low earning potential, or provide additional incentive for doing a worthwhile task, such as participating in a scientific experiment. Usually the sponsors have large numbers of simple tasks like categorizing images. At any one time the sponsor can check the work of the person and decide if they are doing the job correctly. Often these solvers must first perform a qualifying task that establishes the trust relationship between the worker and the sponsor. During the qualifying task the sponsor can check the work against expected responses to ensure the workers simply isn't picking random answers. Often sponsors then have multiple people perform the task and compare the tasks to check for people who pass qualification and then revert to random guessing.

Some micropayments are used to elicit opinions from humans. Sponsors cannot check a person's opinion for correctness. People trying to earn more micro-payments faster could create programs to answer opinion questions randomly. This attack, which I will call robo-pundits, generalizes to all opinions systems. In the case of Amazon mechanical Turk the on-line pseudonym is tied to the persons real world identity to allow the robo-pundit to collect the micro-payments. However, robo-pundits could make many shell web accounts that funnel to the same bank account. In this case you would have to follow the money to ensure you have only one human associated with each web account.

We haven't seen a large impact of robo-pundits, because to earn a lot of money still requires a lot of human intervention and opinion systems pay only small rewards. It is clear that people will exploit this avenue of attack when suitable rewards exist. Already many product review sites have been tainted by people willing to use their real world reputation to extol the virtues of a product for a micro-payment. By applying robo-pundit technology these people could greatly increase their reward.

In the case where sponsors offer real money for completing some task, such as a contest to find a specific solution, the security posture changes. For large prizes the sponsors must scrutinize submitted solutions to ensure they satisfy the win criteria. Stealing solutions in contests has already been seen in the U.S. State Department "Tag Challenge" by attacking the reputation of other teams to steal their crowd of followers (Rahwan et al. 2013). Interestingly, combining ideas to solve a problem provides a valid way to solve problems, but the originating human solvers will find other people benefiting from their efforts detrimental to motivation. In these situations people or teams will keep their research and work secret to prevent theft. Hoarding insightful information will hinder finding solutions. So the sponsor must carefully set the reward criteria to reward the behavior they want, and provide the necessary security to protect solvers efforts.

Depending on the problem domain several solutions might provide viable protection to the sponsors. First the sponsors could strip off the domain-specific jargon and try to have the solvers address the generic problem. The problem remains if the solver can solve the general problem, they probably possess enough intelligence to apply the solution to different domains. The second approach protects the solution by breaking the problem into many smaller problems. When the problems assigned have little context, the chances of a solver seeing how the pieces fit into the solution shrink. Sometimes, this may remove some cognitive biases that will allow the solver to see a solution, but often it will handicap the solver because they won't have enough context to find a solution. Fold-it provides an example where the shrinking context could protect the final solution without handicapping the solvers. Fold-it provides a game where the crowd manipulates protein structures in three dimensions to determine how they will fold (Cooper et al. 2010). Here the crowd uses its understanding of manipulating objects in three dimensions, but that does not translate to how the shape of the resulting protein interacts with other proteins. Solvers may recognize the protein being folded which may reveal some information the sponsors would like to keep secret. Striking a balance between disclose and protecting assets will be a constant balancing act for problem solving environments.

Human solvers will attempt to maximize their monetary reward, and unwary sponsors may be surprised at how clever the solvers are at circumventing the intent of the rules simply to win the reward. The sponsors should adopt proven reward functions or even run small contests with smaller rewards to see if anyone can find loopholes in the reward criteria. One safety net clause to put in the contest rules may simply state that a valid solution must meet the intent of the contest as defined by the sponsors. Legitimate solvers will probably not be concerned about such a statement.

Reputation

If a human builds up a reputation for a pseudonym on a human computation site, an attacker can potentially steal the identity and reap the benefits. However, when tied to the real world identity the victim can prove they are who they say they are through conventional means, like passports and fingerprints.

People will attempt to steal reputation as well. In the 1990 one of the early firewalls, Sidewinder, hosted a contest to break the security of its firewall (Thomsen 1995). The prize was a custom leather jacket with the Sidewinder logo on back. As the contest moderator I was surprised how many people claimed to have hacked the firewall and received the jacket even though no one ever met the victory criteria. Claiming victory without producing the jacket costs the reputation stealer nothing, but the cost of a custom leather jacket would be a small cost for someone hoping to establish his reputation in the hacker community. This reinforces the idea that contest sponsors must protect solver reputations to avoid de-motivating them.

Altruistic

Contests that motivate solvers to achieve altruistic goals like curing cancer provide the solver an intrinsic reward that cannot be stolen. Or can it? What if a malicious person published a false account of the same cancer curing technology being used to create a bio-weapon? The solver may loose the feeling of accomplishment for aiding the effort to find a cure.

Humans can delude themselves into thinking they have contributed. Maybe the person simply created an account on the collaboration site, or simply talked to someone else who had, and because of that they felt good when a solution was found. In the end, the sponsors will still have gotten a solution and some person will have felt better about himself for no reason. While this may seem benign in the case when the system found a solution, consider the case when the site never motivates enough solvers to actually perform work and thus no solution is found. Sponsors must clearly define what constitutes a solid contribution and advertise it to potential solvers. Altruistic rewards require proactive protection by the sponsors just as much as monetary rewards.

In Environment Rewards

"In environment" rewards represent unique, often digital, goods used in human computation environments as rewards. Suppose for example that the Farmville game (a game about planting crops) was actually a serious human computation system with a purpose of finding optimal crop rotations. The special edition seed planter given out to those that participated in March represents a badge of accomplishment intended to reward participation over a specified period of time, but also it represents something that people value because it helps them play the game. Possessing such an item reinforces putting in the work to get the reward.

In game digital rewards like this cost the sponsors very little and can provide significant motivation to solvers to engage in behaviors the sponsors believe will result in a solution. When the rewards only live in the collaboration environment the sponsors can create sufficient security to ensure that the no one can counterfeit the digital goods. In the very least proper auditing of solver behavior could reveal whether the solver earned the reward or not.

Some organizations allow the digital goods to go beyond the collaboration environment. Mozilla has a project at openbadges.org that allow organizations to create badges people can display, and which interested parties can authenticate to ensure the badges' legitimacy. Such portable digital goods provide more motivation for solvers because they can use their digital goods in more places. Such portability requires a long-lived infrastructure to provide authentication. Future rewards might consider code fragments that provide some utility. Consider for example instead of an animated gif, an interactive gif as a reward that a solver can put on her own site.

For either reward, the sponsor must protect the solver's value by preventing others from copying the image or code fragment for display on their own web site when they did not earn the reward.

Education

One reward comes as a natural by-product of solving hard problems: education. If the problem requires human insight to solve, there is a good chance the human will learn something in the process. This includes insights into the target problem and unique problem solving skills. Educational rewards cannot be stolen, and they cannot be earned without the solver doing the work.

Summary

The goal of freely sharing information to solve problems directly conflicts with the system's ability to protect intellectual property without creating some sort of user agreement to mitigate the risk. Robo-troll attacks present a new kind of attack specific to human computation environments designed to erode the number of human solvers on a project; the project's most valuable asset.

Many of these attacks point to solutions that do not allow anonymous human solvers, but tie the solvers to their real world identities. This allows for real world punishment for rule breakers, but also reduces the number of robo-troll attacks. When tied to a real world identity, reputation remains a reward that costs the sponsors little and cannot be stolen or sold.

Overall the magic of the human computation comes from bringing many people together, at their convenience to solve a problem. While it would be nice to create a cyber-utopia where people cooperate freely to solve the problem, unfortunate aspects of the real world will assert themselves to interfere with this goal. Fortunately, many of the solutions to these real world problems can also be applied in cyber space. For example, if the sponsors know who is contributing and in what ways, they will be able to execute both punishments and rewards that have an impact. Follow the money may seem cynical, but it provides the best insight on where the problems will emerge in the new area of human computation solution environments.

References

Cooper S et al (2010) Predicting protein structures with a multiplayer online game. Nature 466(7307):756–760

Keller B (2011) Dealing with Assange and the WikiLeaks secrets. New York Times

McGonigal J (2011) Reality is broken: why games make us better and how they can change the world. Penguin Pr, New York

Rahwan I et al (2013) Global manhunt pushes the limits of social mobilization. Computer 64:68–75

Thomsen DJ (1995) The sidewinder challenge—results so far. Electron Cipher IEEE Secur Priv Newslett

Part IX
Impact

The Impact of Human Computation

Pietro Michelucci

Surely you have heard of Pandora, who according to Greek mythology was the first woman on Earth. Perhaps even more famous is the container she was gifted by the gods and instructed to never open. Of course, ultimately, Pandora succumbed to her curiosity and let escape all of the evils of the universe before she managed to replace the lid. But maybe you were not aware of this important detail concerning the fate of her container: *hope remained sealed within*.

Technology can be used for good or evil. It is hard to imagine that this will ever change. And as long as creativity exists, new technologies will emerge. So when we find ourselves at the brink of transformational capabilities, should we quickly close the lid? Should we turn away from opportunity only to discover that, like Pandora who released the evils from her locus of control, we have unleashed inevitable technologies, allowing them to be self-driven? Or should we instead blaze forward conscientiously, embracing new hopes while doing our best to anticipate and mitigate the challenges and abuses that may arise?

This section seeks to initiate the latter. That is, herein, we consider carefully and speculate wildly about the implications and possible futures of human computation, an evolving paradigm in which humans and machines combine in new ways to exhibit unprecedented capabilities. We do this because we recognize the potentially transformative effect of human computation on individuals, society, culture, humanity, and our biosphere.

We further envision that the world is shifting toward an idea economy, in which human innovation is a natural resource that is amplified by human computation. At the same time, we see human computation as a vehicle for stability—a unifying mechanism that builds relationships and bridges cultures. But we cannot help also imagining it as a mechanism for competition and adversarial advantage. And we can't neglect, of course, the growing pains—the known vulnerabilities and unanticipated risks associated with an emerging and disruptive technology.

P. Michelucci (✉)
ThinkSplash LLC, Fairfax, Virginia, USA
e-mail: pem@thinksplash.com

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_72,

The chapters in this section exemplify the broad observed and expected impact of human computation examining it through various filters, including distributed intelligence, evolution, pervasiveness, culture, conflict, and social progress. These chapters have been written by scientific visionaries, each applying a distinct context and analysis to the assessment of human computation, bravely speculating so that we may have a deeper understanding of human computation, a greater awareness of its risks, and a glimpse into the future.

The first chapter, scribed lucidly by Francis Heylighen, plausibly introduces the notion of distributed intelligence in terms of collective problem solving efficacy that is technologically augmented. In considering the expanding global network of connectivity, he conjures the metaphor of a global brain, which serves the executive functions of a planetary superorganism. The key functions are to detect problems in the world and then generate and coordinate relevant solutions. To make this more concrete, Dr. Heylighen alludes to methods for measuring such intelligence.

Rife with creativity, the following chapter, by Theodore Pavlic and Stephen Pratt, lends credibility to this notion of emergent behavior by exposing parallels between eusocial insect behavior and extant human behavior, taking open source software development as a case study. These parallels are then extrapolated to project a thought-provoking evolution of humanity as a superorganism, and consider its possible futures.

Jonathan Lawhead and Daniel Estrada make the notion of emergent behavior even more concrete by proposing a human computation experiment designed to elicit self-organization behavior. They leverage augmented reality to endow participants with virtual scent trails that manifest in the collective experience. With the addition of a simple incentive, this leads to a game that tests the plausibility of an economic model that has attention as its currency.

In a fascinating treatise on cumulative culture, Paul Smaldino and Peter Richerson reveal the cultural conditions that give rise to complex technology and those that threaten it. In particular, they describe the benefit of social learning on innovation and the effects of social connectivity within and among populations on the dissemination and persistence of complex technology.

Next, Juan Pablo Hourcade and Lisa Nathan discuss the current and future impact of human computation on armed conflict from a socio-technical perspective. In addition to identifying potential risks, they focus on novel forms of interaction among humans toward peaceful objectives and social good.

In her chapter on human computation and sustainability, Bonnie Nardi describes the instrumental role that information technology has played in advancing social progress. She reminds us that our technical infrastructure critically depends upon waning natural resources, the absence of which could lead to societal collapse. This leads to speculation about the role of human computation in addressing sustainability issues and, consequently, preserving social progress.

In the final chapter of this section and hence the book, the editor introduces a manifesto to bring together the interests of a loosely bound community of like-minded *human computation scientists*—investigators and practitioners that seek to understand the behavior of distributed networks of human and machine agents and to engineer them in novel ways toward useful capabilities. Toward that end, the editor sets an agenda that responds to the goals of accelerating innovation and advancing human computation conscientiously as a formal discipline.

Prepare for a journey into our brave new future.

From Human Computation to the Global Brain: The Self-Organization of Distributed Intelligence

Francis Heylighen

Introduction

The present chapter wishes to investigate the wider context of human computation, viewing it as merely one approach within the broad domain of distributed human-computer symbiosis. The multifarious developments in the "social" Internet have shown the great potential of large-scale collaborative systems that involve both people and the various information and communication technologies (ICT) that process, store and distribute data. Here, I wish to explore this development in the broadest sense, as the self-organization of a distributed intelligence system at the planetary level—a phenomenon that has been called the "global brain".

To get there, I will first define and illustrate the fundamental concept of distributed intelligence. Then I will review how such an intelligent network emerges and grows through a combination of self-organization and design. Finally, I will sketch some potential applications of the anticipated global brain.

Human-Computer Complementarity

The rationale for human computation is that people have certain intrinsic skills that are difficult to reproduce in computer programs. A computation system that requires those skills must therefore include people as information-processing agents. Thus, in human computation, people and computers are supposed to work together synergetically, the one complementing the other.

F. Heylighen (✉)
Global Brain Institute, Vrije Universiteit Brussel, Brussel, Italy
e-mail: fheylighen@gmail.com

P. Michelucci (ed.), *Handbook of Human Computation*, 897
DOI 10.1007/978-1-4614-8806-4_73, © Springer Science+Business Media New York 2013

The reason for this complementarity lies in the fact that humans and computers process information in very different ways. Computers excel at accurately storing and retrieving discrete items, such as numbers or strings of characters. Human long-term memory, on the other hand is a network of associations that is continuously being modified by selective strengthening, weakening, adding or combining of memory traces. As a result, people not only forget much of what they observed, but they are strongly biased in what they recall, and sometimes even "remember" things that never happened (Loftus and Pickrell 1995). Thus, human memory is very unreliable compared to computer memory.

The problem gets even worse when people need to manipulate data, which happens in their *working memory*. This is the human equivalent of computer RAM. Working memory, however, cannot sustain more than some four items simultaneously (Cowan 2001). Therefore, most people are unable to make any but the most trivial calculations without the help of pen and paper. Computers, on the other hand, are virtually unlimited in the amount of items they can manipulate, and do not make mistakes when retrieving stored information.

This unreliability of human memory is compensated by the fact that the neural networks that make up our brain are very effective at learning associations between different experiences, and thus uncovering subtle patterns in information (McLeod et al. 1998). Moreover, the brain is remarkably reliable in the *recognition* of patterns similar to patterns experienced before, even while being poor at *recall*, i.e. retrieving exact data. Recognition is so robust because the newly activated pattern can be very different from the one stored in memory, but still the activation spreads through a myriad of learned associations until it activates memories that are in some way related to the new perception. This prepares the mind to anticipate features of that pattern analogous to features experienced before, a form of "intuition" that computers generally lack.

Moreover, human cognition is *situated* and *embodied* (Anderson 2003; Clancey 1997; Clark 1998): we continuously interact with our environment via exquisitely sensitive and sophisticated sensory organs and muscle systems, which have evolved over billions of years. This provides our brain with a very high-bandwidth channel for input, output and feedback, allowing it to learn the high-dimensional, fine-grained patterns and correlations that characterize the real world with all its complexities and dynamics. Thanks to this immediate, real-time coupling between brain and outside world we learn not only to recognize subtle patterns, but to perform precisely coordinated actions. Indeed, the fine-grained sensory feedback we constantly get allows us to automatically perform the kind of complex manipulations that are so difficult for robotic devices.

This on-going interaction has provided people with a lifetime of real-world experience, getting them to know subtle relations between millions of phenomena, variables and stimuli. The resulting knowledge is nearly impossible to implement in a computer program, as most of it is too fuzzy, holistic and context-dependent to be exteriorized in the form of symbols and rules. The difficulty of formalizing such knowledge is known in AI as the "knowledge acquisition bottleneck" (Wagner 2006). It is one of the reasons that information technologists have turned to systems that include *human computation*: letting people perform those tasks that are too difficult for a computer program, while using computers to do the tasks that are difficult or tedious for people.

Distributed Intelligence

Human computation is one among a variety of paradigms that study how people supported by information and communication technologies are able to solve more problems together than when working in isolation. These approaches have been variously called man–machine interaction, human-computer symbiosis, computer-supported cooperative work, social computing, crowdsourcing, collective intelligence, and wisdom of crowds. Different labels or metaphors tend to emphasize different aspects of this synergetic interaction, while ignoring other aspects. For example, the original human computation metaphor sees individuals as computational components performing specific subroutines within a clearly defined overall program, thus viewing them as subordinate to a technological system (Nagar 2011). Computer-supported cooperative work, on the other hand, takes the opposite stance, seeing the technology as subordinate to the human interaction, while collective intelligence a priori ignores technology, even though its practical implementations almost always rely on some kind of information technology.

My aim here is to look for the most general collaborative framework, implying a minimal bias about what kind of activity is performed by whom or by what. A good starting point can be found in the concepts of *information processing* and *distributed cognition* (Heylighen 2012a; Nagar 2011). While most commonly associated with computer technology, information processing has been extensively used to analyze how organizations solve their problems (Galbraith 1974; Tushman and Nadler 1978)—with or without computers. Neural network models illustrate how information processing in the brain is *distributed* (Rumelhart and McClelland 1986): different neurons deal simultaneously with different aspects of the information, while aggregating their results into a comprehensive interpretation. Such collaboration between different components of the brain inspired Minsky (1988) to conceive of the mind as a "society" of interacting agents. The analogy works both ways, though: human society itself is in a number of respects similar to a brain, since it consists of agents that together solve problems that are too difficult for the agents individually. Thus, the distributed information processing perspective is applicable at all levels, from neural circuits, via the brain and organizations, to society as a whole. The principle is simply that a collective of collaborating agents can process more complex information more extensively than any individual member of that collective.

The last step we need to reach the notion of distributed cognition is to observe that physical objects or tools too can function as information processing agents. The simplest tools, such as books, merely store information for later use, thus compensating for the unreliability of human memory. Other tools, such as telephones, can transfer information from one agent to another across distances. Yet other tools,

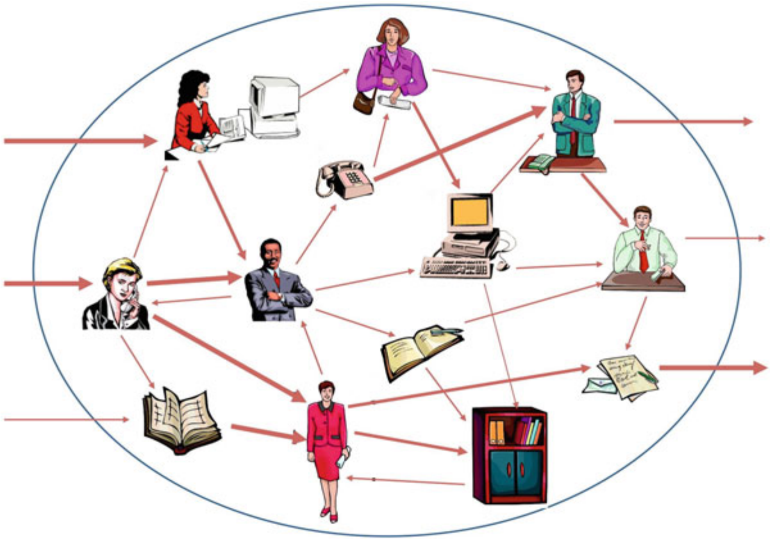


Fig. 1 A depiction of an organization as a distributed cognitive system, i.e. a network of humans and artifacts that store, process, and propagate information along the links in the network. The thickness of an *arrow* represents the intensity of the communication across the corresponding link. Incoming *arrows* represent input to the system (its perception of the environment), outgoing *arrows* its output (its action on the environment)

such as sensors, cameras or recorders, can capture external information. The most sophisticated tools—as exemplified by modern ICT—register, store, transfer and process information. Integrate such tools with human agents into a coordinated system or organization and the result is *distributed cognition* (Fig. 1): acquisition, propagation and processing of information across a heterogeneous network of people and artifacts (Dror and Harnad 2008; Hollan et al. 2000; Hutchins 2000).

Distributed cognition as originally conceived by Hutchins (1995, 2000) is basically a description of an existing situation: social systems have always used external aids for propagating and processing information. What the newer approaches, such as human computation, aim at is to use information technologies to make such distributed processing much more powerful, focused and efficient, i.e. more intelligent. Let us then call this new endeavor *distributed intelligence* (Fischer 2006). Intelligence can be defined as the ability to solve problems, or more generally tackle challenges. The more problems a system can solve, or the more quickly it can solve them, the more intelligent it is. Distributed intelligence, then, means the ability to solve problems collaboratively, by integrating the contributions from a broad assembly of human and technological agents (Heylighen 2012a). The wider the variety of skills that the different agents contribute, and the better the coordination between their contributions, the higher the distributed intelligence of the system they collectively form. The present paper wishes to investigate the future of distributed intelligence: how are distributed intelligence technologies likely to develop and to affect society at large? To answer that question, we must first understand how distributed intelligence emerges from its components.

Self-Organization

Distributed intelligence can be understood as the coordinated activity of a collective of agents (human or technological) that process and propagate information between them. In formal organizations, such as firms, computer systems, or administrations, such coordination is normally the result of *design* (Galbraith 1974; Tushman and Nadler 1978). This means that some person or group of people has developed a scheme that specifies which information is to be processed by which agent, and how the output of that process is then sent for further processing to one or more other agents. Such schemes take the form of computer programs, organizational charts or workflow diagrams.

However, as everybody who has worked in an organization knows, such a scheme only captures a small part of the actual information flow. Most communication follows informal channels, which together form a social network. A social network is formed by links of acquaintanceship, friendship or trust, which are built up through the personal encounters and experiences of the people in the group. In other words, a social network is not imposed by central design, but emerges through decentralized *self-organization*. If we zoom out and consider increasingly large distributed cognitive systems, we will notice that imposed organization plays an increasingly small role, while spontaneous networks become increasingly more important. The reason is simply that the more complex the system, the more difficult it becomes to completely specify the rules about which component is to work with which other component in which way. If we compare the poor results of central planning in communist societies with the effectiveness of the "invisible hand" of the market, then we can only conclude that self-organization must be the major driver of coordination in a system as complex as society.

Self-organization is not just the foundation on which social systems are built. Its power is increasingly being harnessed for building technological systems. Here too, designers are confronted with a complexity bottleneck: as soon as the number of components and their interactions become too large and/or too variable, explicit design or "programming" of the system becomes infeasible. That is why computer scientists and engineers are now exploring self-organizing solutions to the problem of how to coordinate a variety of interacting software and/or hardware components (Bartholdi et al. 2010; Dressler 2008; Elmenreich et al. 2009).

Self-organization is perhaps most critical in the Internet, which is the most complex socio-technological system that presently exists. It is simply impossible to make a rational design for how the different websites and services on the Internet should be connected, because no one knows exactly which services exist and what they can do. Moreover, thousands of new pages, forums and applications appear every day, seeking their place within an anarchic and highly competitive network of linked information sources. Thus, the topology of cyberspace is changing so rapidly that no central authority can ever hope to control it.

How does self-organization work? At the most basic level, every evolutionary process uses *trial-and-error*, a mechanism that can be described more accurately as *blind-variation-and-selective-retention*. If you do not know how to fit things together, then you try a variety of combinations. You then eliminate the ones that do not work (errors), and select the others for retention. This process is iterated: the retained solutions are again modified, producing some variants that work better and are therefore retained, some that work worse and are therefore rejected. If you continue this iteration long enough, you are bound to end up with something much better than what you started out with. This process can be speeded up with the help of positive feedback: amplifying or multiplying the "good" solutions in proportion to their fitness, so as to increase the average quality of your starting material for the next iteration, but without losing the necessary variety. This is the mechanism underlying both biological evolution and its application to computation as implemented e.g. by genetic algorithms (Booker et al. 1989).

The same kind of positive-feedback enhanced iteration occurs in self-organizing systems, with the difference that there is no external fitness criterion that distinguishes what to keep from what to reject. It is rather the system as a whole that determines what survives and what is eliminated. The selected variations are the ones that are adapted to their environment. But in the system as a whole, the environment of a component is constituted by the other components (or agents) it interacts with. Fitness is thus intrinsic to the system: it emerges through the mutual adaptation or co-evolution of the system's components. An interaction between two agents is fit when it is beneficial, so that the agents are inclined to continue it. If the interaction is not beneficial, then there is no reason to maintain it, and the link between the agents will be eliminated. Thus, natural selection here is in the first place a selection of links between components or nodes in the network. The same component may fit in well with certain agents, but not with others. To find out where it fits best, it needs to try out various links, keeping (or strengthening) the good ones and eliminating (or weakening) the less good ones. This is the same mechanism that underlies learning in the brain: useful links (as embodied by synapses connecting neurons) are reinforced; less useful ones are weakened, and eventually cut.

The Self-Organization of Distributed Intelligence

Let us now apply this self-organizing dynamics to heterogeneous networks of cognitive agents, i.e. people and ICT systems. Human computation systems are examples of such heterogeneous networks, albeit that their organization is largely designed or programmed. At the level of the Internet as a whole, however, size and heterogeneity increase to such a degree that design must make place for self-organization via *selective linking*. A simple illustration of how this happens is bookmarking: when a person surfing the web encounters a particularly interesting or useful page, such as a weather forecasting service, a search engine, or an overview of the domain in which the person is interested, then that person will store a link to that page in the browser, as a "favorite" or "bookmark". This makes it easy for the person to come back frequently to that page. Here, a stable link is created between a human and an ICT agent.

A link between two human agents is created when one person meets another one—face-to-face or on the web—and finds that person interesting enough to add him to her list of "contacts" in some social network application, such as Facebook or LinkedIn. This link now makes it easy for the first person to directly pass on information to the second one. A connection between two ICT agents is established when a hyperlink is made from one webpage or website to another one, or when one ICT system (say, the Facebook platform) starts to exchange data with another one (say, the Skype calling service).

In all these cases, links that are successful, in the sense that the agents benefit from them, will survive and be reinforced, while links that are useless or counter-productive will be forgotten and eventually erased. For example, your link to site A may turn out to be particularly useful, and therefore you give it a more prominent place, making the one to the less user-friendly site B redundant, so that you eventually remove it. Similarly, you may from time to time remove "contacts" that turn out to be tiresome, while upgrading others to the status of "friends" or "collaborators". This on-going variation and selection of links makes the network as a whole evolve towards an increasingly efficient or "intelligent" organization. This is analogous to the way the neural networks in our brain learn how to respond more intelligently to the problems they encounter.

The intelligence of this distributed system can be understood through the paradigm of *challenge propagation* (Heylighen 2012a). A problem, question, message or opportunity constitutes a *challenge* for one or more agents: it incites the agent to act, i.e. to respond in a way that may solve the problem, answer the query, reply to the message, or seize the opportunity. A challenge in this sense is a generic term for a piece of information that carries value for an agent, and that therefore can motivate the agent to process the information in order to extract that value. Challenges can be positive (acting on them gives you benefit: *opportunities*) and/or negative (not acting on them makes you lose benefit: *problems*). Dealing with challenges is therefore a straightforward generalization of solving problems.

To measure the intelligence of a distributed network, we can then try to establish its capacity to effectively process challenges. Normally, different agents have different skills in dealing with challenges. For example, computers excel in making complex calculations, while people excel in understanding spoken language. Different people and different computer agents have further their own special abilities, so that our network as a whole will present a wide range of finely grained skills and expertise. A complex challenge (say, global warming) has a large number of aspects that each require different skills. The problem now is to *distribute* the different challenge components across the different agents so as to make sure the challenge as a whole is dealt with in an efficient way. This is the basic problem of coordination, which includes *division of labor* (who deals with what challenge component?), *workflow* (where does a component go after it has been partially dealt with?), and *aggregation* (how are all the finished pieces of work assembled?) (Heylighen et al. 2013; Heylighen 2013) (see Fig. 2).

Perhaps surprisingly, such distributed coordination can self-organize relatively easily across the Internet, via the mechanisms of stigmergy (Heylighen 2007a;

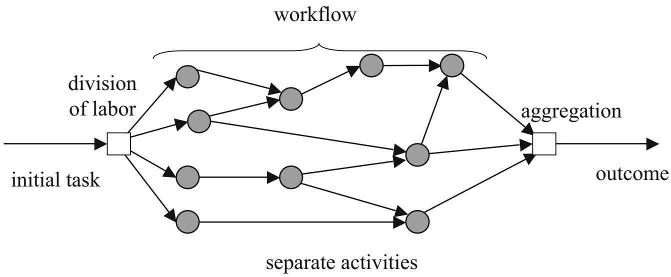


Fig. 2 An illustration of coordination, in which an initial task is split up in separate activities performed by different agents (division of labor), which are followed by other activities (workflow), and whose results are assembled into a final product (aggregation). *Grey circles* represent individual agents performing activities. *Arrows* represent the "flow" of challenges from one agent to the next

Heylighen et al. 2013) and challenge propagation (Heylighen 2012a). A good illustration can be found in the different open source communities developing complex software without central supervision, and in Wikipedia, the online encyclopedia created and maintained by millions of volunteer contributors. *Stigmergy* is an implicit coordination mechanism whereby a challenge left by an agent in a workspace that is shared with other agents stimulates those agents to continue dealing with that challenge (Parunak 2006). For example, a paragraph added to a Wikipedia page by one person may incite a second person to add some extra details, a third one to add a reference for the new material, and a fourth one to edit the text so as to make it more readable. The reference may then be checked and more accurately formatted by a software agent. In this case, challenges are spontaneously addressed by subsequent agents as mediated by the shared workspace (in this case the Wikipedia website). In the case of *challenge propagation*, the workflow is initiated by the agents themselves. An example is an email message sent and forwarded with comments by different people, a "post" in a social network or forum that is reposted to other forums, or a task that is proposed by a crowdsourcing system to people interested to work on it.

In both cases, challenges can travel more or less efficiently across the network of agents and workspaces until they find an agent able and willing to deal with them, and then continue their journey along other agents dealing with the remaining aspects. This allows complex challenges to be resolved in a distributed manner, by harnessing the collective intelligence of the different components (human and technological) of the network. Presently, my research group is developing a mathematical model of this process, in order to investigate precisely how the distributed intelligence of the network increases as it selectively strengthens or weakens its links (Heylighen et al. 2012). The distributed intelligence measure is simply the degree to which challenges are resolved by the networked agents as compared to the same group of agents without connections.

The Global Brain

What happens when such a self-organizing distributed intelligence network grows to encompass the planet (as the Internet already does)? The result can perhaps best be understood with the help of the metaphor of a *global brain* (Bernstein et al. 2012; Goertzel 2002; Heylighen 2008; Mayer-Kress and Barczys 1995). The global brain can be seen as the nervous system of the *planetary superorganism* (De Rosnay 2000; Heylighen 2007b; Stock 1993). This is the "living system" (Miller 1995) formed by all people on this planet together with their artifacts and technologies. The task of its brain is to gather and process information about the situation of the world and all its people, find solutions to any problems it detects, and incite and coordinate actions to deal with those challenges (cf. Helbing et al. 2012). This is similar to the task of the human brain, which gathers information through its sensory organs, processes that information in order to evaluate the situation, then reflects about strategies to deal with the challenges it finds, and finally implements those strategies by sending signals to the muscles so as to direct and coordinate their actions. A secondary task of both human and global brain is to learn from its experiences by reinforcing the successful links in its network (and weakening the others). This allows it to develop ever more detailed and accurate knowledge about itself and the environment in which it lives, and thus to become ever better at dealing with the challenges it encounters.

We should expect the problem-solving abilities of the global brain to be orders of magnitudes larger than that of any single individual, organization, or computer system. This is because all people and computers collectively have access to immensely more knowledge and processing capacities than any of them individually (Heylighen 2012b). The only requirement to efficiently harness this collective intelligence is coordination. This can be expected to self-organize, as illustrated by both empirical observations (Heylighen 2013; Woolley et al. 2010) and simulations (Elmenreich et al. 2009; Heylighen et al. 2012). However, self-organization at a scale as large as the world obviously needs time, as countless iterations of the variation and selective reinforcement process must take place, and as any provisionally "fit" result will need to be updated as soon as a new agent or technology appears on the scene. Thus, all components of the global network will continue to co-evolve at a rapid pace, increasing their degree of coordination, efficiency and intelligence in the process, but in a manner so complex that we cannot predict it in any detail.

It is impossible to say at what moment this process will have produced the equivalent of a global brain, since distributed intelligence is a continuously growing and evolving measure of coordination, not a phenomenon that either is or is not present. Thus, we cannot "detect" the presence or absence of a global brain, but we can conceivably measure the increase in distributed intelligence of the global network. In our mathematical model (Heylighen et al. 2012), we have developed one such quantitative measure, and suggested some methods to gather the necessary empirical data to test its evolution in the real world—but these are very preliminary results.

The self-organization of the global brain could in principle be accelerated by complementing it with thoughtful design. As we start to better understand processes such as self-organization, distributed cognition, collective intelligence and human-computer complementarity, we may be able to avoid some of the trial-and-error search, and develop systems that produce coordinated information processing more quickly and more reliably. For example, inspired by their insights into collective intelligence, Bernstein et al. (2012) have suggested methods for "programming" the global brain, that is, devising schemes that steer a heterogeneous collective of people and computers towards the solution of particular problems—but these too are very preliminary. Further methods are likely to be discovered through research in human computation, crowdsourcing, ontology development, and related fields. However, no single system, method or program will ever be able to capture the immense size and complexity of our planetary network. Therefore, we must resign ourselves to the fact that we will never be able to fully control the process. Perhaps the most promising overall strategy is what has been called "guided self-organization" (Helbing 2012; Prokopenko 2009): developing schemes, programs, institutions or environments that stimulate, facilitate and to some degree steer the self-organization of the global brain towards what appear to be the most fruitful directions, while leaving enough freedom for the system to explore a variety of unforeseen approaches. But to achieve that, we must first of all better understand what the global brain would be able to do, and especially what we want it to do.

Some Implications of the Global Brain

Now that we have a better grasp of how a global brain-like system would emerge, let us try to sketch some of its potential benefits for society. In principle, the Global Brain should help us to tackle any individual or collective challenge, by providing us with a vast reservoir of knowledge, sensory data, information processing capacity, and ability to incite coordinated action.

A first domain that would profit from these superhuman abilities is the economy. The market is the collective system of transactions that helps supply to match demand, and thus to satisfy the public's need for products and services. A traditional market is rather inefficient, requiring a huge infrastructure of middlemen, specialized organizations such as stock exchanges and auctions, and communication channels. The Internet already allows such transactions to take place much more quickly and transparently, with less cost and effort. This strongly reduces friction, making the economy more efficient so that demand can be satisfied more rapidly, more accurately, and at a lower cost. The global brain will not only facilitate communication between suppliers and clients, but help buyers to find the best value (e.g. through shopping agents to recommend and find items and compare prices), and help sellers to get the best price (e.g. through auctioning systems, targeted advertisements, and the ability to reach the "long tail" of customers with very specific requirements).

The net effect will be that growth and productivity increases, while inflation and economic instability decrease. Moreover, there will be less waste because of unsold items or goods shipped far away when there is demand around the corner. More generally, a distributed intelligence system allows us to take into account collective costs and benefits (what economists call "externalities"), such as pollution, noise or public health, which are borne by society as a whole rather than by the parties in the transaction. These costs and benefits can be transparently incorporated by a smart software system into the price of the transaction, in the form of an automatically deduced tax or added subsidy. In this way, interactions are directed towards those that are collectively most beneficial, while avoiding the complexity, bureaucracy and rigidity that tend to accompany such interventions in a centralized political system.

The global brain can moreover help eliminate conflicts. It in principle provides a universal channel through which people from all countries, languages and cultures of this world can converse, as already happens through a variety of forums and social media. This makes it easier to reduce mutual ignorance and misunderstandings. Distributed intelligence systems have already been designed that help large groups to discuss and resolve differences of opinion, while thus devising integrated strategies to solve complex problems such as global warming (Faieta et al. 2006; Iandoli et al. 2009). The greater ease with which good ideas can spread over the whole planet and the collective improvement on those ideas will make it easier to reach global consensus about issues that concern everybody. The free flow of information will make it more difficult for authoritarian regimes to plan suppression or war. The growing interdependence will stimulate collaboration, while making war more difficult. The more efficient economy will indirectly reduce the threat of conflict, since there will be less competition for scarce resources.

Of course, communication alone cannot solve all the problems that threaten our planet: in the end, people will have to agree on concrete policies to tackle e.g. global warming or poverty. Yet, the global brain can support not only the process of devising and reaching consensus on an effective plan of action, but also the practical implementation of that plan. For example, combating infectious diseases or pollution will require extensive monitoring of the number of infections or concentration of polluting gases in different regions. Information collected by local observers or by electronic sensors can directly enter the global brain, be processed to reveal underlying trends, and be forwarded to the people or institutions most ready for taking direct action.

Similarly positive effects can be conceived in domains as diverse as health, wellbeing, democratic participation, sustainable development, work productivity, disaster prevention and relief, education, research, innovation, industrial production, traffic, logistics, and ecosystem management (Heylighen 2002, 2007b; Heylighen et al. 2013). There seems to be no end to the potential applications of a distributed intelligence system at the world level. Many of these applications are already becoming apparent in the present Internet, but their beneficial effect is held back by the general confusion, information overload and uncertainty that accompanies the present explosion in new technologies and functions (Heylighen et al. 2013). It is to be expected that the overall benefits will multiply as the network becomes more streamlined and intelligent, and the agents using it more coordinated in their activities. Then, only the sky will be the limit to what a global brain can achieve…

References

Anderson ML (2003) Embodied cognition: a field guide. Artif Intell 149(1):91–130

Bartholdi JJ III, Eisenstein DD, Lim YF (2010) Self-organizing logistics systems. Annu Rev Control 34(1):111–117. doi:10.1016/j.arcontrol.2010.02.006

Bernstein A, Klein M, Malone TW (2012) Programming the global brain. Commun ACM 55(5):1. Retrieved from http://cci.mit.edu/publications/CCIwp2011-04.pdf

Booker LB, Goldberg DE, Holland JH (1989) Classifier systems and genetic algorithms. Artif Intell 40(1–3):235–282

Clancey WJ (1997) Situated cognition: on human knowledge and computer representations. Cambridge University Press, Cambridge

Clark A (1998) Embodied, situated, and distributed cognition. In: A companion to cognitive science. Blackwell Publishers, Oxford, pp 506–517

Cowan N (2001) The magical number 4 in short-term memory: a reconsideration of mental storage capacity. Behav Brain Sci 24(01):87–114

De Rosnay J (2000) The symbiotic man: a new understanding of the organization of life and a vision of the future. Mcgraw-Hill, New York. Retrieved from http://pespmc1.vub.ac.be/books/DeRosnay.TheSymbioticMan.pdf

Dressler F (2008) A study of self-organization mechanisms in ad hoc and sensor networks. Comput Commun 31(13):3018–3029

Dror IE, Harnad SR (2008) Cognition distributed: how cognitive technology extends our minds. John Benjamins, Amsterdam

Elmenreich W, D'Souza R, Bettstetter C, de Meer H (2009) A survey of models and design methods for self-organizing networked systems. Self-Organizing Syst, pp 37–49

Faieta B, Huberman B, Verhaeghe P (2006) Scalable online discussions as listening technology. In: Proceedings of the 39th annual Hawaii international conference on system sciences—volume 01, HICSS'06. IEEE Computer Society, Washington, DC. p 15.3, doi:10.1109/HICSS.2006.427

Fischer G (2006) Distributed intelligence: extending the power of the unaided, individual human mind. In: Proceedings of the working conference on advanced visual interfaces, AVI'06, ACM, New York, pp 7–14, doi:10.1145/1133265.1133268

Galbraith JR (1974) Organization design: an information processing view. Interfaces 4(3):28–36. Retrieved from http://interfaces.journal.informs.org/content/4/3/28.short

Goertzel B (2002) Creating internet intelligence: wild computing, distributed digital consciousness, and the emerging global brain. Kluwer Academic/Plenum Publishers, New York

Helbing D (2012) Managing complexity. In: Helbing D (ed) Social self-organization, understanding complex systems. Springer, Berlin/Heidelberg, pp 285–299. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-24004-1_15

Helbing D, Bishop S, Conte R, Lukowicz P, McCarthy JB (2012) FuturICT: participatory computing to understand and manage our complex world in a more sustainable and resilient way. Eur Phys J Spec Top 214:11–39. Retrieved from http://adsabs.harvard.edu/abs/2012EPJST.214…11H

Heylighen F (2002) The global brain as a new utopia. In: Zukunftsfiguren. Suhrkamp, Frankurt. Retrieved from http://pespmc1.vub.ac.be/papers/GB-Utopia.pdf

Heylighen F (2007a) Why is open access development so successful? Stigmergic organization and the economics of information. In: Lutterbeck B, Baerwolff M, Gehring RA (eds) Open source Jahrbuch 2007. Lehmanns Media, Berlin, pp 165–180. Retrieved from http://pespmc1.vub.ac.be/Papers/OpenSourceStigmergy.pdf

Heylighen F (2007b) The global superorganism: an evolutionary-cybernetic model of the emerging network society. Soc Evol Hist 6(1):58–119. Retrieved from http://pcp.vub.ac.be/papers/Superorganism.pdf

Heylighen F (2008) Accelerating socio-technological evolution: from ephemeralization and stigmergy to the global brain. In: Globalization as evolutionary process: modeling global change, Rethinking globalizations. Routledge, London, p 284

Heylighen F (2012a) Challenge propagation: a new paradigm for modeling distributed intelligence (no 2012–2001). GBI working papers. Brussels, Belgium. Retrieved from http://pcp.vub.ac.be/papers/ChallengePropagation.pdf

Heylighen F (2012b) A brain in a vat cannot break out: why the singularity must be extended, embedded and embodied. J Conscious Stud 19(1–2):126–142. Retrieved from http://pcp.vub.ac.be/Papers/Singularity-Reply2Chalmers.pdf

Heylighen F (2013) Self-organization in communicating groups: the emergence of coordination, shared references and collective intelligence. In: Massip-Bonet À, Bastardas-Boada A (eds) Complexity perspectives on language, communication and society, understanding complex systems. Springer, Berlin, pp 117–149. Retrieved from http://pcp.vub.ac.be/Papers/Barcelona-LanguageSO.pdf

Heylighen F, Busseniers E, Veitas V, Vidal C, Weinbaum DR (2012) Foundations for a mathematical model of the global brain: architecture, components, and specifications. Global Brain Institute working papers no 2012–05, Brussels. Retrieved from http://pespmc1.vub.ac.be/papers/TowardsGB-model.pdf

Heylighen F, Kostov I, Kiemen M (2013) Mobilization systems: technologies for motivating and coordinating human action. In: Peters MA, Besley T, Araya D (eds) The new development paradigm: education, knowledge economy and digital futures. Routledge, London. Retrieved from http://pcp.vub.ac.be/Papers/MobilizationSystems.pdf

Hollan J, Hutchins E, Kirsh D (2000) Distributed cognition: toward a new foundation for human-computer interaction research. ACM Trans Comput Hum Interact (TOCHI) 7(2):174–196

Hutchins E (1995) Cognition in the wild, vol 262082314. MIT press, Cambridge. Retrieved from http://books.google.be/books?id=AfupQgAACAAJ&dq=Cognition+in+the+Wild&hl=en&sa=X&ei=YjLWT6PGLaOn0QWwyqiBBA&redir_esc=y

Hutchins E (2000) Distributed cognition. In: Smelser NJ, Baltes PB (eds) International encyclopedia of the social and behavioral sciences. Elsevier Science, Amsterdam

Iandoli L, Klein M, Zollo G (2009) Enabling on-line deliberation and collective decision-making through large-scale argumentation: a new approach to the design of an internet-based mass collaboration platform. Int J Decis Support Syst Technol 1(1):69–92

Loftus EF, Pickrell JE (1995) The formation of false memories. Psychiatr Ann 25(12):720–725

Mayer-Kress G, Barczys C (1995) The global brain as an emergent structure from the worldwide computing network, and its implications for modeling. Inf Soc 11(1):1–27

McLeod P, Plunkett K, Rolls ET (1998) Introduction to connectionist modelling of cognitive processes. Oxford University Press, Oxford

Miller JG (1995) Living systems. University Press of Colorado, Niwot

Minsky M (1988) The society of mind. Simon & Schuster, New York

Nagar Y (2011) Beyond the human-computation metaphor. In: Privacy, security, risk and trust (PASSAT), 2011 IEEE third international conference on and 2011 IEEE third international conference on social computing (SocialCom), pp 800–805. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6113219

Parunak HVD (2006) A survey of environments and mechanisms for human-human stigmergy. In: Weyns D, Parunak HVD, Michel F (eds) Environments for multi-agent systems II. Springer, Berlin, pp 163–186

Prokopenko M (2009) Guided self-organization. HFSP J 3(5):287–289. doi:10.1080/19552068.2009.9635816

Rumelhart DE, McClelland JL (1986) Parallel distributed processing. University of California Press, San Diego

Stock G (1993) Metaman: the merging of humans and machines into a global superorganism. Simon & Schuster, New York

Tushman ML, Nadler DA (1978) Information processing as an integrating concept in organizational design. Acad Manage Rev 3:613–624. Retrieved from http://www.jstor.org/stable/10.2307/257550

Wagner C (2006) Breaking the knowledge acquisition bottleneck through conversational knowledge management. Inf Resour Manage J (IRMJ) 19(1):70–83. Retrieved from http://www.igi-global.com/article/information-resources-management-journal-irmj/1286

Woolley AW, Chabris CF, Pentland A, Hashmi N, Malone TW (2010) Evidence for a collective intelligence factor in the performance of human groups. Science 330(6004):686–688. Retrieved from http://www.sciencemag.org/content/330/6004/686.short

Superorganismic Behavior via Human Computation

Theodore P. Pavlic and Stephen C. Pratt

In a future world with pervasive Human Computation (HC), there may be profound effects on how humanity functions at multiple levels from individual behaviors to species-wide changes in evolutionary development. What would such an HC-shaped human society look like? This hypothetical society would be the result of successful adaptations that provide both increased benefit to the high-level facilitators of large-scale computations as well as sufficient incentives to individuals to participate in those computations. In nature, the eusocial insects (Wilson 1971) are a living outcome of similar multi-level selective pressures. Modern-day colony-living honeybees, wasps, and ants descended from a solitary ancestor in which daughters sacrificed their own chance at reproduction to help their mother have more offspring. Despite the apparent reproductive costs, sociality succeeded due to the benefit of indirect reproduction through helping relatives, as well as the competitive advantage enjoyed by cooperative groups. Colony size and complexity expanded over evolutionary time, eventually producing elaborate societies in which reproduction is centralized in a single mother queen, and all other tasks (e.g., brood care, waste management, foraging) are distributed among specialized groups of effectively sterile workers. In these modern colonies, each task group functions like a specialized colony-level organ—the queen acts as the colony's gonads, the nurse workers act as its womb, a waste-management team provides excretory function, foragers seek and find food, and a food-processing team acts as a gut that receives, stores, and distributes food to the rest of the colony. Consequently, the eusocial insect colony is often called a *superorganism* (Hölldobler and Wilson 2009) composed of individual organisms functioning together to support the activities of the

T.P. Pavlic (✉) • S.C. Pratt
School of Life Sciences, Arizona State University, Tempe, AZ, USA
e-mail: tpavlic@asu.edu; Stephen.Pratt@asu.edu

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_74, © Springer Science+Business Media New York 2013

colony as a whole. Even if HC does not result in physiological specializations in humans, it is possible that humanity shaped by HC will evolve analogous specialized organizational structures or even worker castes. Thus, the decentralized superorganismic behavior of eusocial insect colonies can be a window into the future of *Homo sapiens*. It can both provide design support to technology-mediated Human Computation and highlight the risks that emerge in the formation of such collaborative groups.

Existing Human Imitations of Eusocial Insect Society

There are already signs that the trajectory of Human Computation is following that of eusocial-insect evolution. To make this comparison, we use parlance from evolutionary biology to characterize different forms of HC. In particular, when we say that one form is more "primitive" (or less "derived") than another, we mean that it better resembles ancestral versions. This relationship is not necessarily temporal; a more derived species can exist at the same time as a more primitive species. Likewise, as different forms of HC evolve in parallel, some will show more signs of innovation than others.

Perhaps the most primitive form of distributed HC is open-source software (OSS). We say that OSS is primitive because, although it differs notably from software development by a single individual or proprietary software team, the code contributed by each individual of an OSS team is not significantly different in form to the code developed in more traditional settings. In OSS development, individuals share their source code with the Internet at large, and other skilled developers join the effort to maintain and extend the codebase. In many cases, although work is distributed across a team, the key priorities remain consistent with the goals of the founding developer who remains in contact with the team on Internet forums or mailing lists. If that founder leaves the team (either explicitly by announcement or implicitly by prolonged absence), a successor may be appointed. Alternatively, elite members of the remaining team may assert themselves as new creative directors of the project. This process can involve conflict between these elites until agreement on a future direction is established. Alternatively, even when the founding developer is still present, some individuals may leave the group and create a new branch of the software that eventually becomes independent. Moreover, developers of any rank may choose to switch efforts to other unrelated projects at any time.

This process of leadership evolution, conflict, reproduction, and group change is not unlike the development of certain more evolutionarily primitive social insect species. Like developers in an OSS team, members of these primitive societies retain many morphological and behavioral features of their solitary cousins; like OSS teams, however, these societies have also evolved special structures that facilitate superorganismic specialization. Here, we develop this comparison by focusing on the polistine wasps and ponerine ants.

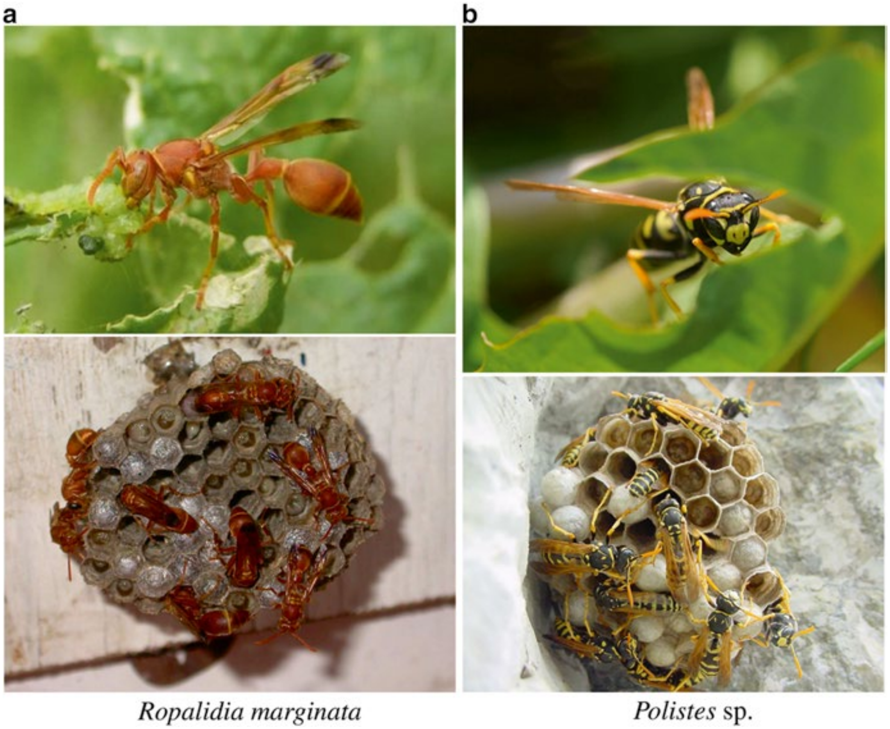


Fig. 1 Primitively eusocial wasps of the polistine (paper wasp) subfamily (Photo credits to: K. Chandrasekhara (*top* photo in (**a**)); Anindita Bhadra (*bottom* photo in (**a**)); Thomas Bresson (*top* photo in (**b**)); Fabio Brambilla (*bottom* photo in (**b**))

Paper Wasps and the Evolution of Open-Source Software

First, we consider projects where each developer retains a freedom of action that resembles that of a paper wasp (Fig. 1). Unlike more derived social insects, where the evolution of specialized morphological castes prevents workers from founding new colonies, a paper wasp can leave her nest at any time to start or join a new one (Bhadra and Gadagkar 2008; Nonacs and Reeve 1993, 1995; Reeve et al. 2000; Shakarad and Gadagkar 1995). In the same way, the developers we focus on here can leave an OSS project at any time to join existing projects or start their own new projects.

Life Histories of Nests and Software Within Open-Air Copyleft Ecosystems

In an experiment with the paper wasp *Ropalidia marginata* (Fig. 1a), Shakarad and Gadagkar (1995) observed a wide variety of nest histories, summarized in Fig. 2.

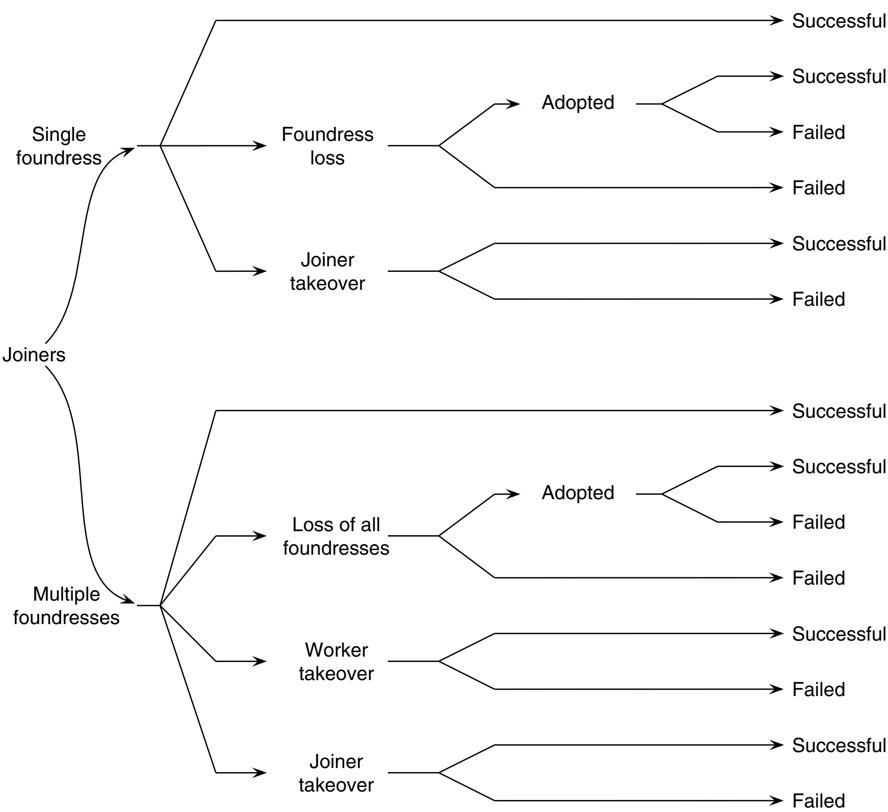


Fig. 2 Life history trajectories of paper wasp nests. The paths depicted are a simplified reproduction of results from an experiment by Shakarad and Gadagkar (1995) with *Ropalidia marginata*. Every path describes the history of at least one nest observed in the experiment. A nest is considered to be successful if it produces adult offspring. A wasp that assists in the construction of a nest is called a "foundress" of the nest. Wasps that assist in the rearing of brood but do not lay any eggs are called "workers." Wasps that join a nest after its construction are called "joiners." A "takeover" event is when the single egg layer of a nest (i.e., the "queen") is usurped by another who then becomes the nest's new egg layer

In about one fifth of cases, a single wasp builds the nest, lays eggs, and feeds and protects her brood as they mature to adulthood. In the remaining cases, a team of wasps founds a nest together and shares these tasks. In either case, there is only one active egg layer at a time even though all foundresses have the ability to lay eggs; this egg layer is the so-called "queen" and parallels the role of an OSS team leader who provides high-level architectural direction for the software. After a nest is founded, the queen can be usurped by existing nest workers or by new individuals who join the nest and seize control over egg laying. Whether laying eggs or not, foundresses can leave a nest to join a new nest, and any queen that has been usurped always leaves her original nest. Due to these losses, nests can also become orphaned, with no remaining workers to care for any surviving brood. Similarly, OSS software projects with publicly available codebases hosted on third-party web sites (e.g., Google Code, GitHub, BitBucket, SourceForge) may become orphaned by their founding developers. Just as orphaned nests are sometimes adopted by newly arriving wasps, orphaned software projects can be adopted by new developers who never have any contact with the original developers. The nests of *R. marginata* are perennial and may survive long after the original foundresses have left the nest. Likewise, OSS software projects hosted on third-party services can have lifetimes far longer than the average time each individual developer commits to the project.

Keeping this nest–project parallelism in mind, the evolution of OSS work-sharing structures is likely similar to that of primitive sociality in paper wasps. In the early days of OSS, the source code for small, mostly non-commercial software was made available for public distribution by individual developers. Like wasps joining an existing nest, other developers could make incremental improvements without the large time investment needed to build the entire project from scratch. Like wasps, these developers could join and leave projects at will. In nature, such wandering wasps are very likely to find new nests to join because nests are physically accessible to the open-air environment. In the OSS ecosystem, so-called "copyleft" licensing schemes create a similar open atmosphere. The salient features of copyleft licensing (St. Laurent 2004) are that source code must be distributed with projects and that derived work must inherit the license. Consequently, copyleft OSS projects beget more copyleft OSS projects, and each copyleft project provides interested wandering developers an opportunity to see, interact with, and even re-distribute modified forms of the project's software code. Increasingly powerful collaborative software version control systems, like Git (Loeliger and McCullough 2012), and large source-code hosting providers, like GitHub, act like new man-made structures on which wasp-like developers can build nest-like projects, and developers can easily move from project to project.

Moreover, just like the turnover of egg-laying individuals in paper-wasp nests, the focal individual associated with a project can change over time. Some projects will fail due to abandonment, but some abandoned projects will later be resurrected by new developers. Still, even when a project has an active developer base, it may fail to attract widespread attention and can be superseded by other functionally similar but unrelated projects. Moreover, a long-lived successful project must attract sufficient interest from other strong developers to withstand the loss of its original founders. In both paper-wasp nests and OSS projects, an open environment for mobile individuals that have the ability to work alone or in teams generates dynamical trajectories similar to those depicted in Fig. 2.

Leadership Maintenance: Queens, Nests, and Internet Forums

Despite their name, the queens of highly derived social-insect species have little-to-no role in managing the activity of workers. In these species, once a colony is established, its queen is only responsible for laying eggs. This level of decentralization is extreme even for present-day state-of-the-art examples of Human Computation. In HC-primitive OSS teams, elite leaders still naturally emerge and help to facilitate the

Fig. 3 *Polistes fuscatus* (Photo credits to: Ettore Balocchi (left photo); Ken Thomas (middle and right photos))

synchronization discussed earlier in the Analysis portion of this book. Likewise, the queens of primitively eusocial paper wasps not only lay eggs but play an active role in coordinating colony activities. Furthermore, the mechanisms these leader queens use to regulate activity are remarkably similar to the strategies available to OSS team leaders via Internet-enabled communication. In studies with the paper wasp *Polistes fuscatus* (Fig. 3), removal of a queen from her nest led to colony activity that was strongly depressed, and workers became far less synchronized (Reeve and Gamboa 1983, 1987). Moreover, when the queen was chilled to make her totally inactive and yet still observable by her workers, colony activity was even further depressed.

In the case of *P. fuscatus* and many other paper wasps, the queen's coordinating role depends on her use of aggression to stimulate activity, and a successor queen can be predicted from a dominance hierarchy (Deshpande et al. 2006; Pardi 1948; Reeve and Gamboa 1987; West-Eberhard 1969). This top-down leadership structure seems more characteristic of large proprietary business software projects that are driven by company profit, developer salary, and managerial rank. However, some paper wasp species manage nest coordination in a distributed way that seems more similar to OSS teams. In *Ropalidia marginata* (Fig. 1a), there is very little observed aggression, no dominance hierarchy, and no known way to predict the line of queen succession (Bhadraa et al. 2007; Bhadra and Gadagkar 2008). Computer simulation further shows that observed levels of coordination cannot be maintained via direct wasp-to-wasp interactions, and there is evidence that the queen instead makes her presence known by continuously depositing a non-volatile pheromone, or chemical signal, directly onto the nest (Bhadraa et al. 2007). Each deposit of this pheromone would be perceivable only by nearby wasps (i.e., it would not spread throughout the nest), and its effect would fade over time as the pheromone signal decayed. Thus, the queen and her pheromone are like an OSS developer moving from one public Internet forum to another posting messages and code patches that are observable to many other team members even after the developer leaves the forum. For large projects distributed over a wide geographic area, such indirect coordination is the rule. Project leaders can confront individuals directly and privately, but one-on-one communication with the project lead is not feasible even for small teams.

In both wasps and OSS teams, reliance on indelible and informative observable signals facilitates changes of leadership. When the original queen is lost from a wasp colony, a new queen emerges swiftly (i.e., within minutes) without contention (Bhadra and Gadagkar 2008). This lack of conflict is evidence that the putative nest pheromone is an honest signal of fertility that sufficiently suppresses egg laying in other workers; otherwise, candidate queens would initially compete to demonstrate reproductive dominance. Similarly, when leaders of OSS teams are active, their presence is observable and their competence can be measured by reviewing the comments they make and the software patches they commit to public repositories or submit to public mailing lists for group review. The absence of an OSS leader is palpable, and a competent replacement emerges quickly without contention because of the transparency of the entire group's participation in the project. Just as potential new queens can demonstrate fertility honestly through production of pheromone, new OSS leaders can demonstrate competence honestly through the team's awareness of their recent contributions to the project. In the case of *R. marginata*, fast succession has been adapted to tropical, aseasonal climates where queen replacement is frequent (Deshpande et al. 2006); in the case of open-source software, swift succession is necessary to maintain the energy and momentum of the project.

Primitively Eusocial Ponerine Ants, OSS Teams, and Technology-Mediated Leadership

To continue the analogy with HC-primitive open-source software, we focus on superorganismic characteristics that stem from evolving caste systems in primitive eusociality. Whereas the paper wasps represented early OSS projects staffed by developers with similar capabilities, these more evolved eusocial societies will represent larger, more-modern OSS teams with a subset of individuals whose small or very specialized skillset puts leadership out of reach. Members of this class of developers must necessarily associate themselves with a leader. Consequently, when a new leader emerges and initiates a new project derived from the original project, she may bring with her a team of developers whose interests are more aligned with her vision than the original leadership. If the daughter project is sufficiently different from the parent (e.g., the Songbird media player and Thunderbird mail client were each derived from a codebase originally intended for the Firefox web browser), the two projects will not compete with each other. However, some competition is unavoidable (e.g., the Pentadactyl daughter and Vimperator parent extensions for Firefox which now are in direct competition for developers and audience).

Although paper wasps form social colonies with reproductive division of labor (i.e., a single egg-laying queen and a worker caste), they are referred to as being *primitively* eusocial because workers and queens are essentially indistinguishable, and workers retain reproductive capabilities (Wheeler 1986; Wilson 1971). That is, primitively eusocial workers are not apparently very different from their solitary ancestors. Higher (i.e., more derived) levels of eusociality are characterized by the addition of specialized worker castes that assist the reproductive caste but cannot themselves reproduce (Hölldobler and Wilson 1977, 1990, 2009; Wheeler 1986).

Similarly, we refer to OSS as a primitive example of Human Computation because each member of the team is almost indistinguishable from a classical software developer. More derived versions of HC are marked by individuals that lose the ability to do similar work in isolation. To understand the evolution of these more derived cases, we now focus on social-insect species that show more specialization than the paper wasps.

Specialized castes are found in all the major eusocial insect taxa, but the most striking differences between workers and reproductives are seen in the ants and termites (Peeters and Ito 2001). This extreme differentiation in these groups is because they, unlike the wasps and bees, combine a flightless worker caste with a winged reproductive caste. The flightless workers are well adapted to their terrestrial ecological niche, while the flying abilities of reproductives allow them to disperse far from their natal nest in search of a diverse gene pool for mating. In many ant and termite species, queens are further specialized to take on the unique tasks of independent colony foundation, when they must build a nest and rear the first generation of workers without any parent-colony support.

Thus, we now shift our focus to primitively eusocial ants from the subfamily Ponerinae. If the paper wasps are like small teams of software developers who could each start a new project entirely on their own, ponerine ants are like larger projects that include some developers with only the skills or interests to work on specialized sections of a project initiated by someone else. Like the paper wasps, these ant colonies contain individuals who could potentially be queen; however, they also have many ants that can only function as workers (Hölldobler and Wilson 1990, 2009; Peeters and Ito 2001; Wheeler 1986). We discuss here how ants in the reproductive class maintain their elite status and control over colony direction. In particular, we focus on two control mechanisms in ant colonies that are similar to technology-mediated solutions seen in OSS teams.

Source Control and the Mutilation of Reproductive Organs

In the course of a large open-source-software project, new project leaders may emerge from within the team's elite members. Due to the openness of the codebase, one industrious individual may start to usurp ownership by rapidly reshaping large portions of the source code. As other workers on the team modify the resulting code, they become committed to it, and a reversal to the basal code becomes less likely. To prevent this, many OSS projects restrict direct access to the codebase to very few individuals. New code is instead posted in so-called "patches" to mailing lists where it can be reviewed and responded to by other developers; the keepers in charge of the codebase can then accept or reject each patch. If these keepers become unresponsive for a long time, others who have been maintaining clones of the codebase can assert themselves as the new masters. Thus, the access limitations on these source-code repositories provide technological mediation of source control, and the ability to create clonal repositories with different access limitations lets source control evolve.

This technologically mediated process is not unlike reproductive succession in the ponerine ant genus *Diacamma* (Fig. 4a) (Baratte et al. 2006; Cuvillier-Hot et al.

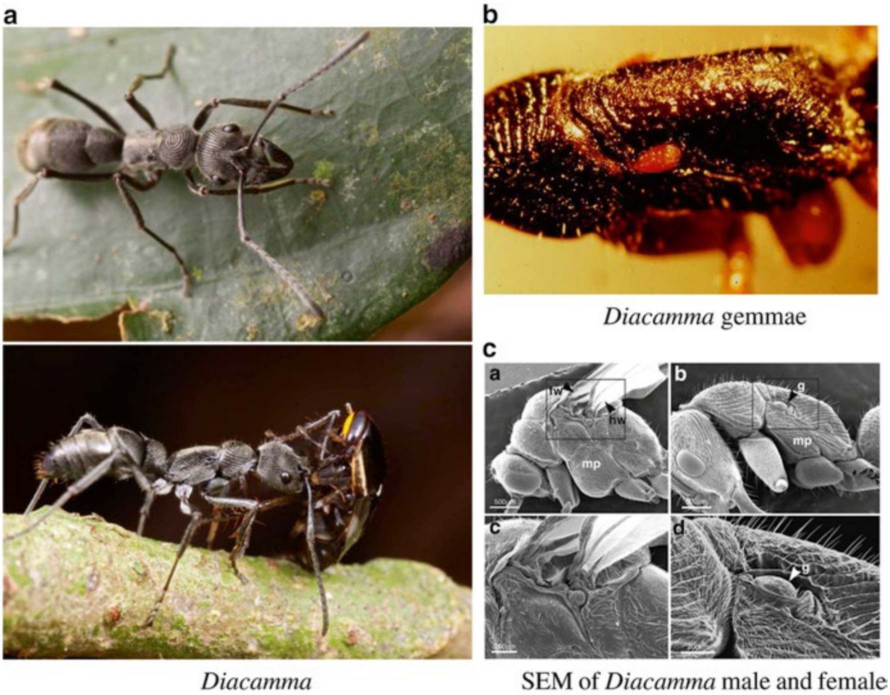


Fig. 4 Ants of the ponerine genus *Diacamma*. Shown in (**b**) is a *top* view of the thorax of an unmutilated female; the gemmae are *orange* appendages homologous to the forewings of a male (Gotoh et al. 2005). Scanning electron micrographs of a male (left column) and a female (right column) are shown in (**c**) where the *bottom* row is a magnified version of the boxes shown in the *top* row. Instead of gemmae, males have wings and large flight muscles that facilitate dispersal from the nest to mate with sexually viable female workers of other colonies. As in most ants, relative to the ant's body, the male head is significantly smaller than the female head (Photo credits to: Steve Shattuck (photos in (**a**)); Alfred Buschinger (photo (**b**)); Gotoh et al. (2005) (photo (**c**)))

2002; Gotoh et al. 2005). In these queenless ants, all individuals are physiologically capable of mating and producing offspring; however, a single mated worker, called the "gamergate" (Peeters 1991b), assumes the role of an egg-laying queen, suppressing reproduction in all other workers in a peculiar way. When each young worker first emerges from her cocoon, she bears a pair of small thoracic appendages called gemmae. The gamergate immediately mutilates the gemmae (Figs. 4b, c), irreversibly preventing the new worker from becoming reproductively viable. When the gamergate eventually dies, the first young worker to emerge without being mutilated will immediately take on the role of gamergate, mutilating all other young workers around her. After the new gamergate mates and begins to lay eggs, she will be accepted by the existing workers as the sole reproductive (Cuvillier-Hot et al. 2002), and the colony's workforce will become increasingly composed of her daughters (André et al. 2001). This process parallels the project-control scheme described above, with gemmae playing the role of access-restriction technologies, and reproduction playing the role of codebase development.

Fig. 5 A worker of the ponerine ant genus *Dinoponera*. These ants can be over 1 inch in length (Photo credit to Alex Wild)

Reputation and the Maintenance of Hierarchies

Especially when public Internet forums are involved, regulation of control of an OSS project can also involve reputation. A new developer proposing a major change can simply be shamed by a respected developer once, and other workers on the team will cease to consider any major new directions from that individual. Even if the shamed developer manages to insert new code, low-ranking individuals may revert those changes with extreme prejudice. Reputation staining is catalyzed by communications technology, such as mailing lists or Internet forums. A similar kind of communication-mediated control occurs in reproductive policing by some ponerine ants of the genus *Dinoponera* (Fig. 5) (Monnin et al. 2002). As in *Diacamma*, workers of these ants can mate and become gamergates. However, workers are not mutilated upon emergence from their cocoons and thus retain reproductive potential throughout their lives. Colonies nonetheless form a dominance hierarchy topped by a single alpha gamergate that monopolizes egg laying and does no other work in the colony. Just beneath her in the hierarchy is a caste of beta workers who do not lay eggs but also do very little work. In an OSS team, if a lead developer leaves the project, she will be succeeded by another who takes over architectural and leadership tasks. Similarly, if the *Dinoponera* alpha gamergate dies, a beta worker will take over and become the new sole egg layer of the colony.

Both OSS teams and *Dinoponera* colonies experience leadership challenges. A beta worker may engage the alpha in sequences of fighting, chasing, and trampling brood. During relatively calm periods within these sequences, the alpha will smear a chemical onto the beta, who then becomes the target of other low-ranking workers who seize and physically immobilize the challenger for several days or weeks. When finally released, she loses her rank in the hierarchy and continues her life as a worker (Monnin et al. 2002). This chemical smearing process is similar to the public shaming an upstart developer might receive from a well-respected lead developer in a public forum. After such exchanges, other developers may cease to entertain new feature suggestions by the shamed developer, who will be reduced to contributing only through the day-to-day maintenance of the established codebase. Message threads on active Internet forums become diluted into obscurity just as chemical signals gradually disperse and become imperceptible; however, the damaging exchange between developers leaves an indelible mark on the rest of the team. Thus, ants and humans have both evolved analogous mechanisms to demote middle managers who seek ascension out of turn.

Resource Limitations in Colonies and Software Teams: Alternative Reproduction Strategies

In section "Paper Wasps and the Evolution of Open-Source Software", we explored how individual paper wasps leaving one nest could start a new nest from scratch. Now that we have shifted our focus to include colonies with large worker castes who may follow an emigrating reproductive, we can also consider the phenomenon of colony fission. Open-source-software teams can bifurcate as well; an individual developer can start a new project and take with her a sizable proportion of resources from the old project, including both team members and cloned code. Fission can be deleterious if the resulting smaller projects compete for developer resources as well as users. Even when competition is not a concern, inheriting the codebase of the initial project also means inheriting bugs, vulnerabilities, or outdated legacy structures that hinder future growth. In other cases, fission provides new per-developer opportunities by reducing team size. It also allows for software frameworks to move into new application spaces (e.g., a useful framework for a popular web browser is quickly adapted into an electronic-mail client). Fission faces similar costs and benefits in eusocial insect colonies, and consequently it is favored under only certain ecological constraints. Hence, we now consider how similar environmental conditions lead to similar foundation patterns in OSS projects and social-insect colonies.

Background: The Multiple Ways to Found a Project

Ant colonies typically reproduce by sending out specialized winged individuals called *alates* (Hölldobler and Wilson 1990), analogous to software developers with the ability to start and nurture a new project in isolation. Unlike workers (all of which are female), alates come in male and female forms that mate in flight after dispersing from their natal nest. The males die shortly thereafter, but the females go on to found new colonies of which they become the queen. They use sperm stored from their matings to fertilize eggs, most of which develop into sterile workers that build and defend the nest, collect food, and nurture further generations of workers. In this way, the colony grows until it is large enough to produce its own reproductive offspring. Thus, alates may be viewed like software architects that are prolific sources of ideas for software projects but must build a team of other developers (i.e., the workers) to actually implement those ideas.

Ant colony formation is typically done in isolation—a newly mated queen excavates a small nest and cloisters herself within it, rearing her first brood of workers by metabolizing stored fat and muscle. This process is analogous to a developer who leaves an old project and uses her personal time and resources to start a new OSS project that will hopefully grow and attract additional help. Sometimes, however, queens without sufficient energy stores must leave the newly excavated nest to take on the dangerous task of foraging. Likewise, the monetary income of an OSS developer likely comes from an outside occupation that prevents full-time commitment to the nascent OSS project. In the ants, alternative colony formation strategies have evolved that mitigate the high cost of independent formation (Hölldobler and Wilson 1977; Krebs and Rissing 1991; Molet et al. 2008; Rissing et al. 2000, 1989). For example, unrelated queens sometimes join forces to start a new nest together so that the burden can be shared (Hölldobler and Wilson 1977; Pollock and Rissing 1985). In software development, a small team of capable developers can similarly join forces to reduce individual workload. In a more extreme solution, seen in many ponerine ants, colonies simply abandon reproduction by female alates (André et al. 2001; Molet et al. 2008; Peeters and Ito 2001). Instead, the colony splits in two, with each segment including a flightless queen accompanied by a large retinue of workers that help her to found a new colony (Peeters 1991a,b; Heinze 1998). In the same way, a developer who chooses to start a new project can attract members of her prior projects. By bringing with her a ready-made team, she may surrender full control over the direction of the new project and will have to spend more time managing these human resources. In both the ant and OSS cases, foundation by large teams of workers inherited from a parent project reduces how often new projects are formed. In general, ant colonies and OSS projects face very similar costs and benefits to different forms of foundation, and they have evolved similar reproductive strategies to mitigate the costs and capitalize on the benefits.

Intellectual Property and Inbreeding in Lieu of New Project Foundation

In some ants, something analogous to intellectual property (IP) has led to a reduction in the occurrence of fission and an increase in inbreeding. These ants are much less like a typical OSS team aiming for new marketable features, and more like a commercial software team with the sole purpose of maintaining an existing one-of-a-kind proprietary project. That is, a nest of these ants is like a software package made for a very specialized purpose targeting a small set of high-value clientele. For example, the software that manages inventory and interfaces with cashiers at a large retail chain may be highly customized for that particular chain. It evolves over time with the customer's needs, but it retains the aesthetic characteristics of much earlier versions of the software (e.g., keyboard-only terminal-mode applications that look relatively unchanged over several decades although the operating systems they run within become increasingly stylized). These products persist because the market is very small and controlled by few developers. Such software can be maintained by a small team of developers who become dependent upon the longevity of the product as their individual talents stagnate. If any developer leaves to join another project, she brings little with her, either because of legal restrictions or because the resources

Fig. 6 Ants of the ponerine genus *Harpegnathos*. These evolutionarily primitive ants hunt for live prey that they can spear with their pointed mandibles and then paralyze with their sting. They are also known for their ritualistic aggression displays between mated workers who compete to become an egg layer and fill the vacuum left by an expired queen (Photo credits to: Kalyan Varma (left); Steve Shattuck (right))

from the old project are too specialized to be of use anywhere else. If a key leader leaves, she will likely be replaced from within because no outsider would be familiar with the extremely specialized codebase.

The ant genus *Harpegnathos* (Fig. 6) has an unusual life history that shares many features of these proprietary software teams (Peeters et al. 2000; Hölldobler and Wilson 2009). These ants might be expected to reproduce by colony fission, like many of their fellow ponerines (Baratte et al. 2006; Cuvillier-Hot et al. 2002; Peeters and Ito 2001). In those other species, workers (or worker-like queens) mate in the nest with alate males that fly in from other colonies (e.g., André et al. 2001). They then leave to found new colonies, accompanied by a retinue of fellow workers. In this way they combine the benefits of outbreeding with the assistance of their parent colony to quickly achieve large group sizes. *Harpegnathos saltator* has the physiological capability to pursue this strategy (Liebig et al. 1998), but colony fission has never been observed in nature (Hölldobler and Wilson 2009). Colonies instead produce many alate queens that disperse to form new colonies in isolation. Sexually capable workers that remain in the nest may also mate, but they tend to do so with their brothers rather than with alate males from other colonies (Peeters and Hölldobler 1995). Furthermore, they do not leave the nest to found new colonies. What results are persistent colonies that remain small and experience reduced genetic diversity due to inbreeding. These behaviors appear to be driven by the highly elaborate nest structures these ants build to resist frequent flooding in their native Indian habitat (Peeters and Hölldobler 1995; Peeters et al. 1994). When the founding queen dies, a daughter gamergate inherits the valuable nest and continues to maintain and improve it. Just as a palace is passed down to noble mated cousins in a royal dynasty, this process can continue forever in principle. Consequently, nests observed in nature accumulate extremely elaborate constructions despite only containing a small number of workers at any one time (Peeters et al. 1994; Hölldobler

Fig. 7 Stingless bees of the genus *Trigona* (Photo credits to: José Reynaldo da Fonseca (*left*); James Niland (*middle* and *right*))

and Wilson 2009). In short, these colonies produce gamergates not for colony fission, like other ponerines, but instead, to retain family resources despite frequent queen turnover.

This process is not unlike software projects with proprietary IP components that may prevent project replication. Alate-like developers that leave the project to work elsewhere cannot bring technology with them. Moreover, the longevity of the project is benefited by maintaining a stock of skilled workers that have experience with the proprietary IP. Like gamergates, new leaders are promoted from workers already within the project. After the death of a *Harpegnathos* queen or gamergate, the upward mobility of mated workers to replace her is usually accompanied by ritualized aggression (Hölldobler and Wilson 2009), which can also be seen during similar transitions in human organizations. Thus, in both software teams and *Harpegnathos* colonies, the existence of assets that cannot be duplicated leads to small groups with much internal turnover, even when such policies reduce diversity within the group. These assets must be extremely valuable in order to sustain these patterns despite the costs.

Using Shareable Resources to Accelerate New Colony Formation

Software projects that make use of intellectual property still contain open components that can be re-purposed in other projects. For example, a technology company may produce new hardware products that make internal use of a popular open-source-software operating-system platform. The OSS platform may be augmented by proprietary hardware drivers as well as improvements to open-source modules that ensure a certain marketable specification for the product. Those open-source components can be used on other projects even by competitors.

In a similar way, a colony of stingless bees (Fig. 7) can provide both personnel as well as physical materials in support of a new daughter colony that is the product of fission (Inoue et al. 1984; Peeters and Ito 2001). In *Trigona laeviceps*, for example, a colony sends a worker team to scout nearby for an empty cavity to house a daughter colony. Once found, workers carry building material from the old nest to the new site. As the amount of materials transported to the new nest is a negligible fraction of the stock at the old nest, this process is much like producing a forked repository from the open portions of an existing OSS project. After the new nest is

Fig. 8 Honeybees from the genus *Apis*. In photo (**c**), an *Apis mellifera* colony makes a tree branch its temporary home; meanwhile, a decentralized selection process goes on among scouts that search for a nest cavity that will eventually become the colony's new home (Photo credits to: Jon Sullivan (left photo (**a**)); Louise Locker (right photo (**a**)); Gideon Pisanty (photo (**b**)); Nancy McClure (photo (**c**)))

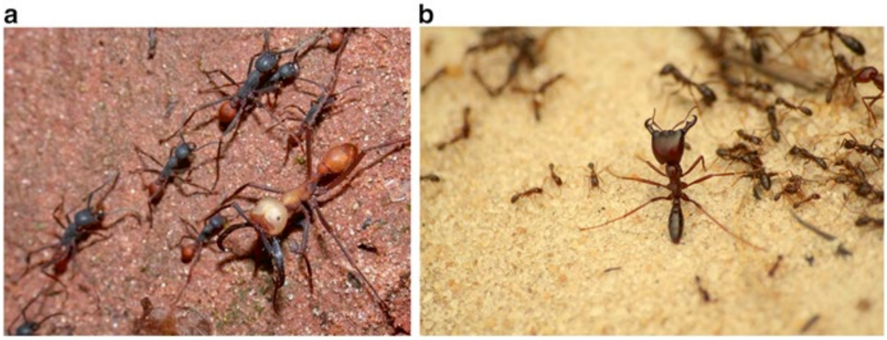
prepared, a swarm of workers and a virgin queen fly there from the old nest. Some of these will return to the old nest, much like software developers choosing to rejoin their original project after not finding interesting opportunities in the new project. The new queen mates and starts populating the colony with her daughters. Material transport from the parent nest continues for a short period, but the daughter colony eventually achieves full independence. This nest-foundation process is not unlike the genesis of an open-source project sanctioned by and based on a project with proprietary roots (e.g., the way Mozilla emerged from Netscape). The daughter project may accept contributions from the original project, but it continues with a new developer community and has a different direction than its more commercial relative. Nevertheless, because of prevailing similarities between the two projects, they may compete for user attention. Likewise, because daughter colonies of stingless bees are so close to their mother nest, there is a chance that they will compete for the same resources. Close distance helps to facilitate quick construction of new projects, but it also presents sustainability issues for their co-existence.

Colony reproduction in honeybees (Fig. 8) shows an alternate path that leverages the aid of the mother colony while preventing future competition (Peeters and Ito 2001; Seeley 1995). Honeybees also reproduce by fission, but unlike stingless bees, they bequeath the old nest and roughly half the workers to a virgin queen. The old queen and the remaining workers leave the nest and settle at a temporary location (Fig. 8c) from which they carry out the decentralized process of finding a new home beyond the competitive reach of the old nest. Once the swarm's scouts reach a decision quorum at a candidate site (Seeley 2010), the bees fly there and build a new nest. This process is much like a team of developers who leave a project they founded after it matures, entrusting it to a set of younger developers so that the founders can start a new project. To reduce competition, the emigrating developers choose a new application area and may sign non-disclosure or non-competition agreements. This process ensures survival of the old project and uses its stability to mitigate the difficulties that young developers may have taking over as leaders. This process also naturally allocates those developers with proven success to nascent projects that will benefit from that experience.

Leveraging Diversity in Large, Long-Lasting Projects

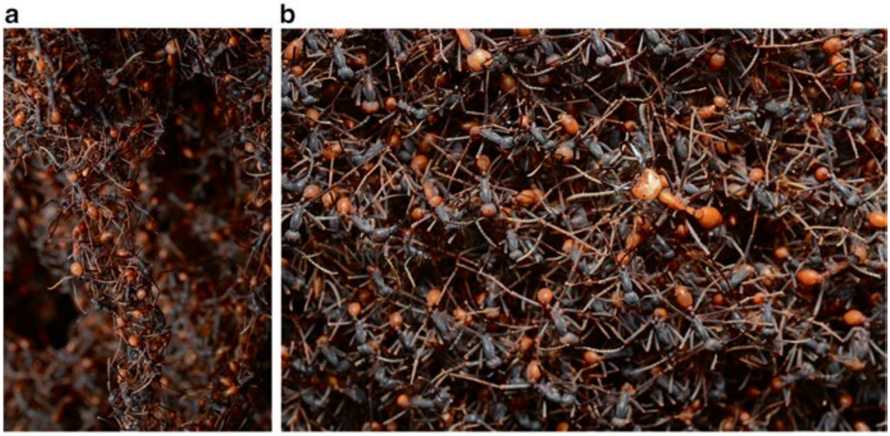
So far, we have discussed how fission alleviates some of the challenges of starting new projects. By inheriting workers and other resources, new projects immediately inherit momentum and a workforce to maintain that momentum. However, fission is not without costs. As mentioned above, daughter and parent projects must disperse far enough away from each other to prevent significant future competition. Still, even when dispersal is guaranteed, fission can proliferate deleterious parasites. In software teams, these parasites may take the form of vulnerabilities, viruses, inefficient code, deprecated protocols, or ineffective team members that slow group productivity. Similar risks exist in ants, and those species that use colony fission have also evolved mechanisms to reduce those risks.

To illustrate how ants manage the risks of fission, we contrast two groups of army ants-one that reproduces exclusively by fission and one that sometimes uses fission and other times relies on independent colony foundation. This examination is partly meant to show parallels between ant-colony and software foundation and partly meant to illustrate the dangers of building software derived from the codebase of another project. Two well-known army-ant genera are the legionary ants (Fig. 9a, *Eciton*) and the driver ants (Fig. 9b, *Dorylus*) (Gotwald 1995). Studying the similarities and differences between these two genera gives insights into the costs, benefits, and maintenance of fission. In contrast to the evolutionarily primitive ponerine ants discussed in section "Primitively Eusocial Ponerine Ants, OSS Teams, and Technology-Mediated Leadership", the army ants are a derived species with a variety of worker castes (Hölldobler and Wilson 1990; Gotwald 1995). However, they have evolved a unique nest structure that allows for added mobility that has led to the re-emergence of reproduction by fission. Rather than excavating nests or living in pre-formed cavities, army ants link their bodies together to form living nest structures consisting of hundreds of thousands of workers (Fig. 10) (Anderson et al.



Legionary ants (*Eciton*) Driver ants (*Dorylus*)

Fig. 9 Army ants. Multiple army ant worker castes are shown in each photo (Photo credits to: Alex Wild (photo (**a**)); James Niland (photo (**b**)))



Eciton bridge *Eciton* bivouac surface

Fig. 10 Army ant (*Eciton*) structures. Colonies of *Eciton* do not nest in cavities or excavate nests in the ground. Instead, they link their bodies into large bivouacs that act as mobile nests (Anderson et al. 2002). Despite being constructed entirely of living colony members, the inside of the bivouac contains sufficient structure for chambers and division of labor based on position within the nest. When *Eciton* colonies reproduce by fission, bivouacs split, each one taking a queen to continue reproduction of workers after dispersal (Photo credits to Geoff Gallice)

2002). These bivouacs are well suited to the ants' nomadic lifestyle, facilitating their frequent and rapid emigrations.

In theory, if the competitive costs of fission can be reduced, then it should be observed more often in nature. In software terms, daughter projects will be more successful if they do not compete for users with the original parent project. For example, the Mozilla Application Suite produced daughter projects Firefox and Thunderbird which then competed with their parent for users. However, because the web browser, Firefox, and the e-mail client, Thunderbird, were different application types, they did not compete with each other. When application or geographic boundaries can prevent competition, more fission might be expected. Likewise, fission has evolved in some army-ant species because these highly mobile colonies can reduce its competitive costs. However, not all army ants reproduce this way, and those that do have had to evolve additional mechanisms to reduce other costs of fission not related to competition.

In the New World army ant genus *Eciton*, fission is the only method available for colonies to reproduce (Hölldobler and Wilson 1990; Peeters and Ito 2001). In contrast, many army-ant species of the genus *Dorylus* do not use fission even though they nest in mobile bivouacs very similar to those of *Eciton*. Instead, they produce alate queens that disperse from the nest and found new colonies on their own (Hölldobler and Wilson 1990). Thus, while fission reduces the burden on a daughter colony and nest mobility reduces the competitive costs of fission, *Dorylus* has not evolved this form of reproduction to the degree that *Eciton* has. Likewise, despite the wide availability of code and developers in the OSS ecosystem, new projects are periodically started from scratch, and OSS libraries are often re-factored or totally re-invented. Thus, in both ants and software development, there must be other fission-related costs to overcome.

One such cost is enhanced parasite transmission. In ants like *Dorylus*, there is low risk of transmission from a mother to a daughter colony because the parasites must infect the single alate founding the new colony. In *Eciton*, there is no such bottleneck; any of the tens of thousands of workers that join the new colony may harbor parasites. Other social features can ameliorate these risks. In *Eciton*, for example, high levels of task specialization may isolate worker groups from one another and their parasites. Multiple mating by the colony's queen occurs with high frequency (Denny et al. 2004; Kronauer et al. 2007; Palmer and Oldroyd 2000; Tarpy et al. 2004); this can increase genetic diversity and thus the likelihood of the colony containing individuals with heritable resistance to any given parasite. For *Eciton*, an additional source of genetic diversity is colony fusion (Kronauer et al. 2010; Schneirla 1971). When a colony's queen dies, the orphaned workers follow and eventually join other colonies that still have a queen (Schneirla 1949; Schneirla and Brown 1950).

Parallels of these fission-related costs and prophylactics can be found in OSS projects. Long-lived software projects can also accumulate deleterious "bugs" and vulnerabilities due to code stagnation, and legacy components that either depend on deprecated protocols or have prohibitive operational constraints. These problems multiply when such projects are cloned to generate the seed of a new project. Even without inherited code, developer teams can accumulate deleterious or deprecated practices. Just as for ants, foundation of software projects from scratch by single developers prevents these problems. However, independent foundation is not practical for very large projects; inevitably, teams of developers build off of existing code repositories and make use of well-known libraries. To resist infection, large OSS projects must be generated by a diverse developer community that, in the aggregate, is immune to systematic deficiencies. In particular:

- A single developer who contributes to a wide range of the codebase may introduce the same vulnerable code (e.g., buffer overflows or dangling pointers) to multiple unrelated parts of the project. The spread of this code can eventually be limited and the vulnerable code repaired, but the full extent of how far the deleterious code has spread may not be known. So, just as a high level of task specialization in an army-ant colony reduces susceptibility, there is value for developers to limit the scope of their contributions and specialize on small components of larger projects.
- For large OSS projects to have high longevity, organizers of such teams must promote the regular incorporation of new ideas and new developers. For ants, novelty comes from multiple mating and adoption of orphaned workers. For developers, novelty comes from continued training and incorporating new workers from outside projects. As new developers gain access to old code, additional dangerous yet subtle vulnerabilities can be found and fixed. This increased developer diversity is similar to new genetic variation that prevents the spread of an extant infection.
- Commonly used OSS libraries and utilities are the result of combining the prior two points. That is, as a section of developers becomes compartmentalized in order to prevent the spread of infections, the subcomponent they write can become its own open-source project in order to gain the attention and additional diversity of more contributors.

These OSS practices are in stark contrast to the *Harpegnathos*-like projects described in section “Resource Limitations in Colonies and Software Teams: Alternative Reproduction Strategies”. For those cases, to ensure longevity of a project depending on valuable proprietary intellectual property, teams have to be kept small, stay isolated from outside influence, and generate new leaders from within the team. These practices are both impractical and ill-advised in an open-source-software project.

Future Human Imitations of Eusocial Insect Society

We have demonstrated parallels between evolutionarily primitive Human Computation, like open-source software, and primitively eusocial insects, like polistine wasps and ponerine ants. Assuming that similar pressures will continue to guide the evolution of HC, then we speculate that its future forms will share similar characteristics with more derived eusociality. At the point at which Human Computation takes on superorgansmic qualities, computations will be decentralized to the point of being leaderless. Some human participants of these computations may specialize at particular tasks, but their participation in different projects will be self guided. It will be more common that participants have some ability to do a variety of tasks, and a participant may choose at any time to switch from one task to another even before completing the prior task. Often, two ongoing tasks will be in opposition to each other, and one participant will actively and unknowingly undo the work of another. Additionally, at any given time, a large proportion of potentially active participants may be idle. In general, all participants will be entirely ignorant of the collective progress toward any particular goal. Moreover, the computational strengths of these collaborations will come from the network of interactions between individuals as opposed to the individuals themselves; each individual participant will only have marginal importance. Counter intuitively, efficient and robust computation will emerge because of, not in spite of, these properties.

Just as any one species of social insect has been specially adapted for its natural environment, different tasks and interaction mechanisms will be matched to different kinds of problems. Rather than being explicitly designed, this mechanism-to-problem matching will evolve naturally from existing technologically mediated interaction networks. That is, with increasing digital connections between electronically augmented participants, there will be increasing potential for networks to do work. Just as increasing temperature can lead to phase changes in matter, increasing network potential can lead to a sudden and emergent computational ability in a group of interacting individuals. The most familiar phase changes in matter are so-called "first-order" changes that are marked by abrupt shifts in observable physical properties, like volume or density. For example, as a fluid moves through a first-order phase transition from liquid to gas, it will become a mixture of some parts that are liquid and some parts that are gas; consequently, the phase transition will be accompanied by violent boiling. However, higher-order phase transitions also exist, and these are continuous in observable properties. Under special conditions, there can be a continuous higher-order phase transition from liquid to gas which does not involve a violent mixture of the two phases; instead, the whole fluid simultaneously shares properties of both phases. In the case of HC networks, it is likely that the transition to superorganismic computation will be of this latter kind. Moreover, as we will show, there are signs that some networks are already near the continuous transition region—exhibiting early transitory signs of superorganismic computation.

In the remaining section of this chapter, we give examples of superorganismic computation in highly derived eusocial insects and speculate about parallels with future Human Computation. When possible, we highlight existing technologically mediated human organizational structures that share properties with these natural insect systems.

Oligogyny and Leaderlessness: Competitors that Share the Same Workers

After colony foundation, a queen plays little role in coordinating the activities of her colony. Her main responsibility is to produce new workers and reproductives. Whereas the natural lifespan of a worker may be on the order of months, a queen can live for years or even decades. Despite this relatively long life, she carries no seniority; she is largely at the mercy of her workers. Thus, a queen is less like a leader than a captive wealthy donor who has no choice but to continue funding her captors.

Fig. 11 Meat ants (*Iridomyrmex*). These omnivorous ants are found in Australia, where they form large colonies and scavenge for a wide variety of foods including large animal carcasses (Photo credits to Steve Shattuck)

As described in section "Background: The Multiple Ways to Found a Project", a colony can be founded by multiple unrelated queens that may then continue to co-exist after colony foundation (Hölldobler and Wilson 1977, 1990; Pollock and Rissing 1985). Continuing the analogy with donors funding a large-scale HC project, this so-called "polygyny" might be thought of as multiple donors pooling their resources to better support a common goal. However, in functioning eusocial insect colonies, standing queens in the same colony can be antagonistic rivals. This special form of polygyny, known as "oligogyny," is seen in the meat ant *Iridomyrmex purpureus* (Hölldobler and Carlin 1985) (Fig. 11), and in *Camponotus ligniperdus* (Gadau et al. 1998), a species of carpenter ant (Fig. 12). While queens in oligogynous colonies are hostile to each other, their workers tolerate all of the queens and each other. Moreover, they form a barrier between the queens, eliminating dominance behavior and allowing all queens to produce brood. Consequently, workers in oligogynous colonies show relatively low levels of relatedness. Although workers from different colonies may be hostile to each other (Gadau et al. 1998), workers from an extant colony will adopt a newly inseminated queen (Hölldobler and Carlin 1985; Hölldobler and Wilson 1990). Thus, the genetic variation among workers does not come from fusion with other colonies or initial foundation by multiple queens but instead from continual adoption of newly inseminated queens.

Internet marketplaces, like Amazon Mechanical Turk (MTurk) (Amazon.com, Inc. 2005), are presently some of the most advanced examples of crowd-sourced Human Computation, and they are much like oligogynous ant colonies. On MTurk, a class of human requesters makes monetary payment available to a class of human workers who can choose to complete tasks designed by the requesters. In principle, the requester class may contain multiple business competitors that each use MTurk as a source of shared computational power—it is as if competing car makers produced vehicles using the same manufacturing line. The MTurk interface prevents any requester from directly impeding the progress of another requester while allowing all workers the opportunity to complete tasks of any and all requesters. Thus, just as queens in an oligogynous ant colony perform the important task of replenishing the work force, requesters replenish the payments that are necessary for human workers to do work. Moreover, just as the crowded colony buffers the queens from ever discovering each other, the MTurk interface prevents interactions between

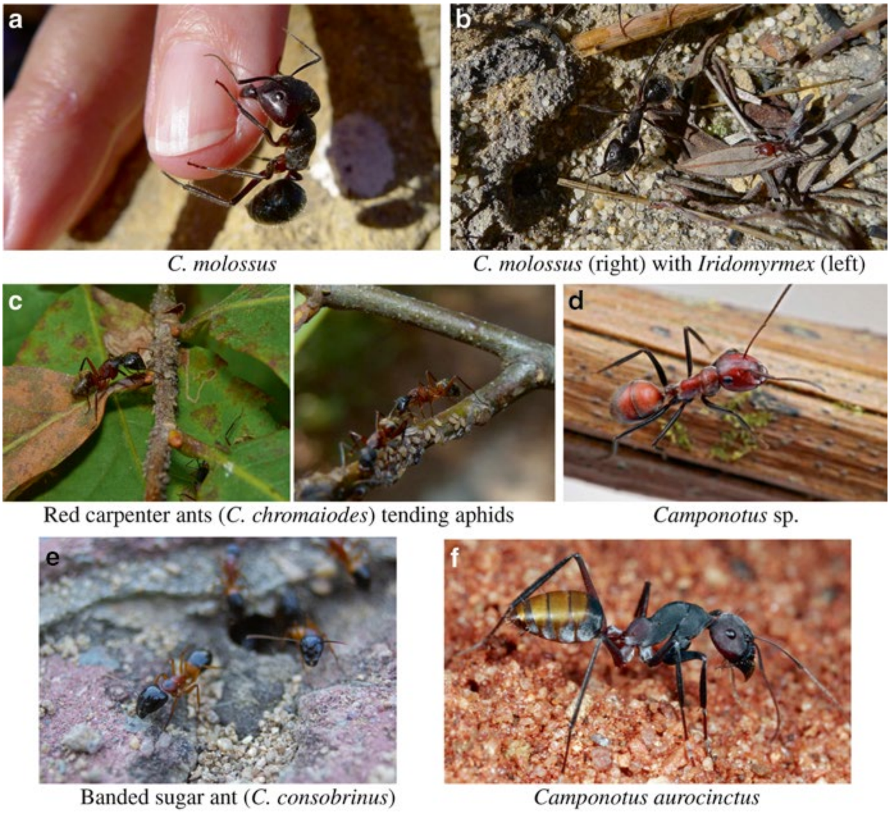


Fig. 12 Carpenter ants (*Camponotus*). This diverse genus of ants nests in hollowed-out cavities in wood, explaining their common name. Some species have distinct morphological worker castes determined by their environment during development; these castes differ in both morphology (e.g., size) as well as behavior. Like other ants of the formicine subfamily, *Camponotus* primarily defend themselves by biting and spraying acid as opposed to using a sting. Consequently, researchers who collect *Camponotus* ants in the field cannot use mouth aspirators because it could lead to inhaling large quantities of the irritant. At least one species, *C. saundersi*, possesses large mandibular glands filled with a sticky corrosive secretion; the ant can then contract abdominal muscles in a suicidal act that ruptures these glands and sprays this immobilizing secretion onto its attacker. Some *Camponotus* ants, like the ones shown in (**c**) and (**e**), have been likened to aphid ranchers; not only do they forage on secretions from the aphids, but they protect the aphids from predators and periodically relocate them much like a human rancher protects and herds cattle (Photo credits to: John Tann (photos (**a**), (**b**)); John Beetham (photos in (**c**)); Steve Shattuck (photos (**d**) and (**f**)); Ryan Wick (photo (**e**)))

requesters. In nature, not every inseminated queen will be lucky enough to be adopted by an existing colony. Likewise, on MTurk, not every requester will be fortunate enough to benefit from the collective action of the workforce. Thus, as described at the start of this section, it is the decentralized network of workers that both provides computational power and selects the problems for which that computational power will be used.

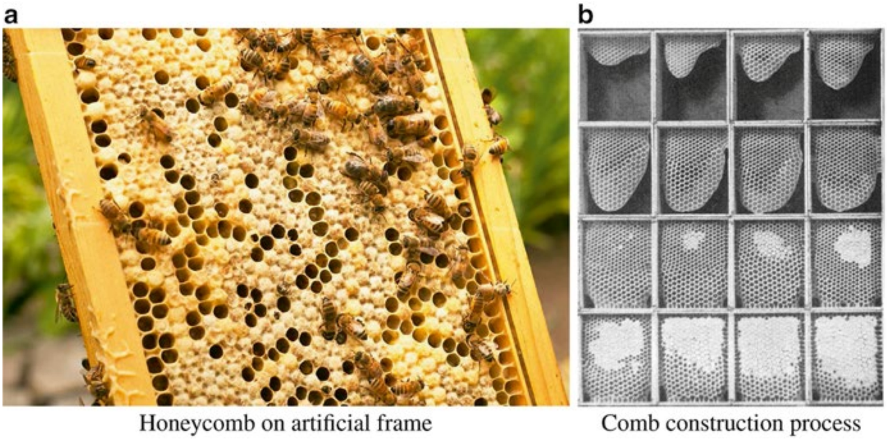


Fig. 13 Honeybee (*Apis mellifera*) comb. Shown in (**b**) are several stages of comb construction (Photo credits to: David Goehring (photo (**a**)); Beach and McMurry (1914) (photo (**b**)))

Still, despite its workforce being decentralized, the MTurk mechanism itself is a centralized bottleneck that is notably distinct from an ant colony. In the future, it can be assumed that MTurk will be replaced by a truly decentralized network of peer-to-peer software that both buffers requesters from interfering with each other and allows workers to self allocate to different tasks entering the network. The behaviors that facilitated this level of decentralization in ants arose randomly and were favored by natural selection because they led to emergent and efficient task allocation. Likewise, the peer-to-peer software that will facilitate similar structures for Human Computation will likely emerge randomly due to the efforts of a few empowered developers, see widespread adoption by a decentralized population of requesters and workers, and then be self sustained by massive activity levels.

Decentralized Harmony Through Individual Contention

Now that we have discussed how networks can self allocate tasks to connected workers, we shift to considering how tasks might interact or even interfere with one another. Honeybees, *Apis mellifera*, construct nests out of wax secreted from glands in their abdomens that they mold into large combs. Each comb consists of a regular array of hexagonal cells (Fig. 13) that are used to store honey and pollen, and also serve as cradles for rearing new female workers, male drones, and virgin queens. Characteristics of each cell, particularly its size, are specialized for its target contents. Thus, the comb must be constructed so that the relative proportions of each type of cell match the particular foraging environment and sex-allocation strategy of the colony. The construction of this properly proportioned comb is a highly decentralized process in which hundreds of bees contribute to the construction of each cell (Pratt 2004). Individual bees often appear to work at cross purposes, with one bee applying wax that is removed seconds later by another bee. Indeed, the construction of a cell can be followed by its complete destruction before it even is used (Cargel and Rinderer 2004). From this description alone, this process seems arbitrary and capricious and possibly inferior to the blueprinted construction of human buildings; however, it somehow consistently leads to recognizable, elegant, and functional structures in nature.

Recently, roboticists have taken an interest in synthesizing large groups of robots that function like honeybees to assemble collections of heterogeneous parts into desired configurations without central control. These robots might be found scurrying around a factory floor, tirelessly converting raw materials at one end to products at the other. Alternatively, microscopic versions might be injected into a human patient to actively regulate the proper proportions of cholesterols in the blood. The need for decentralized control is especially apparent in the microscopic case, where the robots will operate without external control and without sophisticated communication abilities. Matthey et al. (2009) used simple chemical reaction networks (CRNs) to generate local interaction rules for robot teams that guarantee construction of desired quantities of different products. We summarize some key results of that model to show how apparent contention, like that observed in honeycomb construction, may be necessary to ensure proper function at the level of the collective.

In the target application of Matthey et al., robots move randomly around a two-dimensional arena that is cluttered with parts of different types. For simplicity, we assume there are three different part types, A, B, and C, that can be combined to make two different conglomerate products, $A\,B$ and $B\,C$. These two different part assembly plans can be written

$$A + B \rightarrow AB \qquad \text{and} \qquad B + C \rightarrow BC. \tag{1}$$

However, parts cannot assemble themselves. It is the role of the robots to pick up the parts, find other robots carrying other parts, and then assemble the conglomerates. Thus, if we let A, B, and C represent types of parts that are currently in motion on a robot, we can introduce corresponding types, a, b, and c, to represent stationary parts waiting to be found and loaded onto an unburdened robot. If unburdened robots are themselves considered to be a fictitious part type R, then we can augment the assembly plan in Eq. (1) with

$$R + a \rightarrow A \quad \text{and} \quad R + b \rightarrow B \quad \text{and} \quad R + c \rightarrow C.$$

So a stationary part of type a encounters a robot of type R, and the two combine to become mobile part A. That mobile part A eventually combines with another mobile part B to become a mobile conglomerate $A\,B$ and a liberated robot that is free to find other stationary parts to pick up. Thus, the complete assembly plan is

$$\begin{aligned} &R + a \rightarrow A \quad \text{and} \quad R + b \rightarrow B \quad \text{and} \quad R + c \rightarrow C \\ &A + B \rightarrow AB + R \qquad \text{and} \qquad B + C \rightarrow BC + R. \end{aligned} \tag{2}$$

In this scenario, robots encounter parts and other robots at an average rate that is a function of the robot speed and the relative geometries of the robots, parts, and the arena. Thus, the random process of robots picking up and assembling parts is not unlike the random process of gas molecules finding and reacting with each other. Based on this reasoning, Matthey et al. convert the assembly plan in Eq. (2) to the chemical reaction network

$$
\begin{aligned}
&R+a \xrightarrow{k_A} A \quad \text{and} \quad R+b \xrightarrow{k_B} B \quad \text{and} \quad R+c \xrightarrow{k_C} C\\
&A+B \xrightarrow{k_{AB}} AB+R \qquad \text{and} \qquad B+C \xrightarrow{k_{BC}} BC+R
\end{aligned}
\tag{3}
$$

where each reaction rate $k_i \triangleq e_i p_i$ is the product of e_i, which is the mean encounter rate between any pair of the corresponding reactants, and p_i, which is the probability that the corresponding reactants will carry out the reaction after the encounter. Although each encounter rate is a function of the environment, the reaction probability can be picked a priori before dispatching the robot swarm. Thus, the programmer has the freedom to choose these reaction probabilities in order to control the reaction rates. In principle, the resulting system can have an equilibrium distribution of entities that is predictable from the theory of continuous-time Markov processes. This distribution will be parameterized by the reaction probabilities, and so the swarm can be "programmed" to reach a target distribution by choosing the corresponding set of probabilities. However, if only forward (i.e., constructive) reactions are possible, then this stable equilibrium distribution will not exist. In order to generate a stable equilibrium of conglomerates, reactions must be reversible, as in the final assembly reaction network

$$
\begin{aligned}
&R+a \xrightarrow{k_A} A \quad \text{and} \quad R+b \xrightarrow{k_B} B \quad \text{and} \quad R+c \xrightarrow{k_C} C\\
&A+B \xrightarrow{k_{AB}} AB+R \qquad \text{and} \qquad B+C \xrightarrow{k_{BC}} BC+R\\
&AB \xrightarrow{k_{Ab}} A+b \qquad \text{and} \qquad BC \xrightarrow{k_{Bc}} B+c\\
&AB \xrightarrow{k_{aB}} a+B \qquad \text{and} \qquad BC \xrightarrow{k_{bC}} b+C
\end{aligned}
$$

which is identical to Eq. (3) augmented with several spontaneous reverse (destructive) reactions that reduce assembled conglomerates (e.g., $A\ B$) back into mobile and unmoving parts (e.g., A and b) that will then be free for future forward (constructive) reactions. Whereas the forward reaction rates are manipulated through setting reaction probabilities, the reverse reaction rates reflect a programmed timeout on each robot; after carrying a conglomerate object for sufficiently long, the robot breaks the object into its constituents. By tuning the tension between the forward and reverse reaction rates, the decentralized random collective process will maintain a precise balance between the average numbers of $A\ B$ and $B\ C$ conglomerates. In other words, because the decentralized process provides no feedback to individuals about the global number of conglomerates, there is no way to inhibit the construction of a particular conglomerate when a surplus develops. However, because of the reverse reactions, a surplus of that conglomerate will catalyze its own reduction; the greater the surplus, the greater the propensity of reverse reactions to reduce the surplus. Thus, even without individual-level feedback, the collective is able to regulate properties of the ensemble.

Although simultaneous construction and destruction seems counterproductive, it is possible that it reduces the amount of centralized coordination necessary for a given distributed task. In fact, Livnat and Pippenger (2006) make the argument that, due to physiological limitations, internal conflict is actually the *optimal* strategy even within a single human brain. Thus, in highly decentralized instances of HC in the future, individual-level tasks may necessarily undo the apparent progress of other individual-level tasks. Furthermore, some individuals within a collective computation may come into direct one-on-one conflict with other individuals, as occurs between ant workers in some species (Hart and Ratnieks 2001). This apparent local conflict, however, will ensure progress toward the collective goal.

Individual Ignorance Reduces Collective Cognitive Overload

Not only is an ant colony highly decentralized, but its work is completed by ants that are ignorant of colony-level objectives as well as their role in achieving those goals. Army ants are a clear example of global effectiveness emerging from individual ignorance. These ants are named for their group raiding behavior, in which large swarms of foragers flush out and capture insect prey (Fig. 14). The raiding groups form long branching columns guided by chemical trails, along which they return prey to the nest. These species-typical branching patterns of raiding columns emerge without any individual ant possessing any information whatsoever about their existence. In fact, the ants are virtually blind and navigate entirely by following the chemical pheromones left by their nestmates (Gotwald 1995). Distinct branching patterns emerge from interactions between the ants' simple rules for responding to pheromones and the distinctive spatial distributions of the different prey types used by each species (Franks et al. 1991).

So, despite the similarity between these raiding groups and a human military column, none of the individual ants is an "army of one." Each raider is entirely dependent upon being a part of the raiding team. Consequently, several species of ants that are the victims of army ant raids have evolved a simple but effective defense—evacuate, disperse, wait for the invasion to end, and then move back into the original nest (Lamon and Topoff 1981; Smith and Haight 2008). For army ants that specialize on other ant colonies, a successful raid depends on prey being densely concentrated and is largely ineffective when a target colony disperses. A particularly dramatic (although somewhat artificial) illustration of individual army ant ignorance is the formation of so-called "circular mills." This can occur when the head of a foraging column is induced to double back and encounter its tail, leading the ants to rotate continuously in a circle until they either die of exhaustion or escape the mill (Schneirla 1944; Brady 2003; Delsuc 2003). These mills reflect the ants' total dependence on following the chemical trails laid by preceding ants. Their lack of any other navigational mode prevents them from realizing that they are moving in circles or that the chemical signal they are following was actually deposited by the ants that are following them.

Fig. 14 Army ant (*Eciton burchellii*) raiders carrying captured brood from a wasp nest back to their home bivouac. Army ants are named for this characteristic group foraging behavior. A large team of foragers marches away from their home bivouac in a column formation that, in some species, can bifurcate multiple times to form large branching structures. Foragers from the column flush out insect prey or invade the nests of other social insect to take their brood. They retrieve their prey to the bivouac along the same foraging column (Photo credit to Geoff Gallice)

Ignorance as Enforced Independence

It may be tempting to suggest that less-ignorant army ants would make for more successful colonies that are immune to prey evacuations and deathly ant milling. However, in other ant species where it is easier to test the connection between the individual and the colony, ignorance has been shown to be adaptive. The underlying reasons are related to the requirement of independence among group members for the "Wisdom of Crowds" (Surowieckie 2004), as discussed in the Algorithms portion of this book. Ants of the genus *Temnothorax* (Fig. 15) are very small crevice dwellers that can be induced to migrate into a credit-card-sized artificial nest consisting of a cavity in a balsa wood slat sandwiched between two microscope slides. When a homeless *Temnothorax* colony is given the choice of several artificial nests, it will reliably choose one based on a variety of criteria (Visscher 2007), including entrance size (smaller is better) and cavity illumination (darker is better). This colony-level choice does not depend on individual ants visiting all options and comparing them. It emerges instead from a decentralized process that aggregates the assessments of many scouts, few of which visit multiple sites (Pratt 2005a). However, when a single ant is isolated and required to make this choice on her own, she is capable of doing so (Sasaki and Pratt 2011, 2012). This makes it possible to compare the decision-making performance of individual ants and whole colonies.

Fig. 15 Painted *Temnothorax rugatulus* ants next to an artificial nest consisting of balsa wood sandwiched between two microscope slides; the second ant from the left is holding a brood item in her mandibles. A colony of several hundred *Temnothorax* ants may reside in a crevice formed from a hollow acorn or a small crack in a rock. Under a microscope, the ants can be immobilized and painted with four color marks so that individuals can later be uniquely identified during behavioral experiments. Consequently, their small size allows for detailed observations of how individuals contribute to colony-level decisions. In the past, *Temnothorax* ants were classified in the genus *Leptothorax*, which is the name used to refer to them in the Foundations section of this book (Photo credit to Takao Sasaki and James S. Waters)

For example, Sasaki and Pratt (2011) showed that individuals are vulnerable to the "decoy effect", a form of irrational decision making. This effect is well known in humans, where it is evoked in the presence of two target options that pose a trade-off between important attributes. If a third "decoy" option is added that is clearly inferior to only one of the two targets, it can greatly increase the preference for that target, even though the decoy itself is never chosen. Sasaki and Pratt found that individual ants were strongly influenced by the decoy, but colonies were immune to its effect. This immunity is potentially important to colony fitness, as sensitivity to irrelevant decoys is not consistent with a decision maker maximizing fitness.

The key advantage of colonies over isolated ants appears to be the relative ignorance of individuals in the colony setting. Because each worker visits only one site, this ensures that option assessment is truly independent, a basic requirement for the Wisdom of Crowds. A lone ant, in contrast, must do all of the cognitive work of comparing multiple options that vary discordantly in several attributes. To do so, she likely relies on simplifying decision heuristics that work most of the time but leave her vulnerable to systematic errors like the decoy effect. In the colony setting, comparison is distributed over all of the colony's scouts, thus relieving any single ant of the burden of processing all available information.

This burden sharing also allows colonies to handle more data than a single ant can. When presented with a simple choice between one good and one poor nest, colonies and individuals are similarly effective at choosing the better option Sasaki and Pratt (2012). When the challenge is increased by presenting eight candidate nests—half good and half poor—colonies continue to do well, but individual performance plummets to no better than random. In humans, this effect is known as “cognitive overload”—the ability to make a good choice is impaired by the number of choices. For individual ants, the problem appears to be that they attempt to process more information than they have the cognitive capacity to handle. For whole colonies, the distributed process of nest-site choice reaches a conclusion before many individual ants have had time to visit more than one or two sites. Thus, although lone individuals have the ability to directly compare multiple options and choose between them, that ability is significantly less effective than the decentralized process that aggregates assessments of individuals that have only experienced one option. The colony’s collective wisdom emerges from individual ignorance.

A similar advantage of individual-level ignorance is seen in nest-site selection by honeybees (Visscher 2007; Seeley 2010) and may be a general feature of collective decision making by insect societies. Thus, it appears that the evolution of eusociality has led to a decrease, not an increase, in individual awareness. Likewise, advances in Human Computation may ironically correspond to a reduction in the role or awareness of each individual involved in the computation. For example, the reCAPTCHA system acquired by Google in 2009 (Google 2009) coerces large teams of humans to unknowingly digitize books, street numbers, and other images of text while simultaneously verifying to a third party that they are human. The system works by presenting two images of text, one of which is a known word that has been obscured and another that is unknown text taken from some source of interest to Google. In order to gain access to the third party, the human has to properly input the known text; however, because she does not know which field is her entry key, she is forced to also lend her computational skills temporarily to Google. The system capitalizes on the ignorance that comes about through her lack of awareness. Like the decentralized *Temnothorax* colony that makes a decision too quickly for any scout to visit multiple candidate nest sites, the system is designed to prevent her awareness from impeding the progress of the distributed computation.

Automatic and Ubiquitous Collective Computation: Global Brains

This notion of distributed ignorance is also consistent with the emergence of self-selecting computations discussed in section “Oligogyny and Leaderlessness: Competitors that Share the Same Workers” and earlier in this chapter. For example, as the level of automatic electronic personal instrumentation increases via smartphones or Internet-enabled automobiles, unprecedented amounts of data about the current state of the world will be immediately available to very wide audiences. Software applications are already being developed for augmented-reality devices (introduced in the Techniques and Modalities section of this book). These devices effectively implement artificial sensory modalities that allow real-time perception of aggregated data (Jenkins 2013), like seeing a virtual "chemical trail" recording the history of pedestrian traffic on real pavement.

How to induce humans to use these technologies is discussed in detail in the Participation section of this book. It would not be unprecedented for games developed today to lead to more practical applications afterward. It would also not be surprising if HC applications could be disguised as games, especially if those applications are motivated by noble causes, like scientific exploration. Moreover, either due to mechanism design (Mas-Colell et al. 1995; Osborne and Rubinstein 1994; Feigenbaum and Shenker 2002) that rewards participation or just because of convenience, there may be an emergence of always-on software that continuously samples aspects of the environment and relays anonymous data to a network of others using that software. In fact, something similar already occurs as smart phones gather and aggregate traffic data from their mobile hosts; this data collection certainly goes on while navigation applications are running, but it may also occur at other times by always-on social-networking applications that automatically "check in" periodically (e.g., Google Latitude, Foursquare). In such systems, data sources are ignorant of how their data are used by various consumers. Moreover, normally accepted principles of locality are violated as individuals make decisions primarily based on stimuli from far-flung sources. As these individuals make decisions in parallel based on related data, the group as a whole appears to make colony-like aggregate decisions that may share properties with how *Temnothorax* colonies choose a new nest.

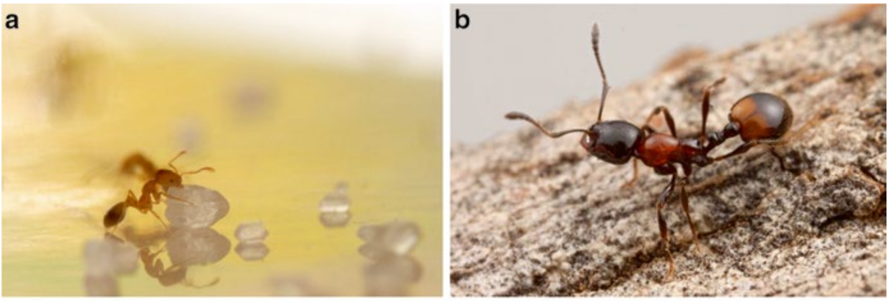
The result is not unlike the HC-induced "global brain" that is discussed in the opening chapter of this section of the book. However, there is an important difference between a global brain and a real brain, in terms of the independence of their constituent parts. Although the brain appears to be a decentralized collective of neurons, its parts are physically co-located. So there are added difficulties in ensuring that the real brain aggregates truly *independent* assessments. A global brain, on the other hand, is like a *Temnothorax* colony whose scouts each see only one of many candidate nests. Like the ants, the decentralized agents within these global brains are forced to make independent assessments. Consequently, they may be *qualitatively* superior to real brains, because they can aggregate independent assessments of parallel aspects of a challenging problem and thus avoid the cognitive traps associated with non-independence.

There are negative as well as positive consequences to automatic collective computation. For example, augmented-reality devices may help people with similar interests synchronize in time and space so that it is easier for them to meet. However, these devices may also allow for unrelated criminals interested in robbing the same bank to find each other and pool resources. Even if such a team does not formally meet, the augmented-reality traces that accumulate in the shared paths that they travel may help any one of them to find vulnerabilities more easily. Moreover, because distributed information persists over time and coordination is implicit, there will be little ability to detect any deleterious shared computation until after the bank is robbed. In fact, researchers funded by law enforcement agencies are already using fictitious games in MTurk (Amazon.com, Inc. 2005) to accurately characterize human deviations from rationality (Yang et al. 2012). These data are then used to build random patrol schedules that minimize the probability that a watchful adversary will be able to game the schedule and smuggle contraband into sensitive areas. It seems inevitable that criminal organizations will someday use the same methods to design optimal adversarial schedules to maximize patrol vulnerabilities. At the moment, marketplaces like MTurk are bottlenecks for gathering the requisite data for such research. However, as Human Computation becomes decentralized, it is not clear how to control access to its potential. Superorganisms are marvels of nature, but they can also be invasive pests.

When Ants Fail

One of the best-known instances of collective decision making in ants is pheromone-trail following, which is also discussed briefly in the Foundations section of this book. This behavior has inspired a trail-laying-inspired metaheuristic optimization algorithm known as Ant Colony Optimization (ACO) (Bonabeau et al. 1999; Birattari et al. 2002; Dorigo et al. 2006). While ACO mimics ant chemical signaling within simulated parameter spaces, roboticists have gone further and implemented true chemical-trail following on mobile robots (e.g., Sharpe and Webb 1996; Svennebring and Koenig 2003; Fujisawa et al. 2008). Not surprisingly, software applications currently in development for augmented-reality systems achieve collective network computation by some form of simulated trail laying meant to induce human agents to behave like the virtual ants in an ACO algorithm (e.g., Jenkins 2013). Despite its inclusive name, ACO caricatures only a subset of ants found in nature. Moreover, natural trail-laying has been tuned by natural selection for specific environments. When trail-laying ants are induced to complete tasks under laboratory conditions that differ from their natural environment, colonies can fail to make good decisions. These failures not only highlight weaknesses of distributed decision making via trail laying, but they show that complex systems in general can be maladaptive and need to be specially tuned for particular contexts. In this section, we also describe how other ants have evolved decentralized behaviors to solve similar problems without the use of chemical trails. These alternatives have their own strengths and weaknesses. Thus, there is much to be learned from mixing different decentralized strategies when appropriate. In general, the future success of widespread Human Computation will likely come from architectural diversity and not hegemony.

Collective decision making based on pheromone trails is well illustrated by the foraging behavior of the Pharaoh ant (*Monomorium pharaonis*; Fig. 16a). When a scout finds food, she recruits other ants to it by laying a chemical trail back to the nest (Hölldobler and Wilson 1990; Sumpter and Beekman 2003). Recruits follow this trail to the food and may reinforce it by adding more pheromone, with a strength that depends on the quality of the food source. Reinforcement makes the trail still more attractive to further recruits, generating a positive feedback loop. If trails are



Pharaoh's ant (*M. pharaonis*) with sugar   *M. rubriceps*

Fig. 16 *Monomorium* ants. The genus *Monomorium* is diverse and widespread. The small (2 mm) ant in (**a**) is a typical Pharaoh ant (*M. pharaonis*), a worldwide indoor pest species that has spread from tropical to temperate zones by human commerce. The slightly larger and considerably more colorful Australian *M. rubriceps* is shown in (**b**). Although it is generally uncommon in ants (Heinze and Keller 2000), *M. rubriceps* and some other Australian *Monomorium* species (but not *M. pharaonis*) can produce both winged and wingless "intermorphic" queens from the same colony (Fersch et al. 2000; Buschinger 2011) (Photo credits to: Julian Szulc (Szulc 2011) (photo (**a**)); Steve Shattuck (photo (**b**)))

laid to multiple food sources simultaneously, the colony's foragers will eventually concentrate on a single trail to the best source (or a randomly chosen one if sources do not vary in quality). This happens because ants have a threshold-like response to pheromone concentration that amplifies even small differences in attractiveness between trails. The pheromone also decays over time so that trails to less competitive sites eventually fade away. Moreover, the same process leads the ants to settle on the shortest path between the nest and the best food source (Beckers et al. 1990; de Biseau et al. 1991; Beckers et al. 1992a,b; Camazine et al. 2001). This process has inspired methods for solving optimization problems that must pick the best of a wide variety of parameter combinations. Simulated ants move around the parameter space and leave virtual trails according to the subjective value of the parameters discovered. The simulated recruitment process prioritizes search effort so as to maximize the chance of finding the optimal parameter combination without having to test all possible combinations (Dorigo et al. 1996, 1999; Bonabeau et al. 1999; Birattari et al. 2002; Dorigo et al. 2006).

A distinguishing characteristic of trail following is that it is *decisive*; it is pathologically rare for trail-following process to come to a split decision, and this property holds for both the differential-equation models of trail following as well as real ants foraging in controlled experiments (Sumpter and Beekman 2003). Even when there is only a small difference in quality between options, trail following coalesces on one option relatively quickly. However, as the difference in quality between options becomes small, the outcome of the decision becomes more reliant on the initial bias in the scouting team than on the actual quality difference between options. That is, the decision becomes a social cascade driven by popularity rather than the efficient independent assessment discussed in the Algorithms section of this book.

Consequently, trail-following ants are poor at adapting to changing environments (Beckers et al. 1990; Nicolis and Deneubourg 1999; Camazine et al. 2001). If given two feeders of equal concentration, they will randomly commit to one of them and will be locked into that choice until the colony satiates and foraging stops. Not only will the colony be ignorant of any augmentation of other feeders, but it will be unable to quickly adapt to reductions in quality at its chosen feeder due to inevitable depletion effects. Consequently, classical trail following is not suitable for all environments. The success of simulated trail following in optimization problems is in great part because the value landscape over the parameter space is fixed over time. If trail following was used via Human Computation to, for example, find the least crowded restaurant in a city, the collective choice could quickly become the most crowded restaurant before negative feedback could re-allocate incoming diners to another option.

Of course, trail-laying algorithms can be altered to reduce the chances of such deleterious positive-feedback popularity cascades. In fact, it has recently been discovered that the trail-laying big-headed ants (*Pheidole megacephala*) can adaptively track changes in feeder quality during experiments (Dussutour et al. 2009). Moreover, the temporal characteristics of the shift in foraging allocation after a change in feeders are captured by a model that adds a certain amount of noise to each scout's choice of foraging route. The added noise ensures that a significant fraction of scouts continues to visit apparently suboptimal sites. If this pool of uncommitted scouts is sufficiently large, it can dislodge a highly reinforced trail so that the colony can switch to a site that gains comparative advantage over time. However, the optimal level of noise varies with how frequently the environment changes. So even this improved trail following must be specially tuned for each environment.

Ants are a diverse group, and many species rely on recruitment methods very different from pheromone trails. These other methods can also support collective decision making but can lead to different decision dynamics and outcomes. For example, the *Temnothorax* ants described in section "Individual Ignorance Reduces Collective Cognitive Overload" use "tandem running" to recruit to rich food sources or potential new homes (Franks and Richardson 2006; Hölldobler et al. 1974; Pratt 2005b). In a tandem run, a successful scout individually leads a single follower from the nest to the target location (Fig. 17). In particular, after finding food, a forager returns to the nest and releases a "calling" pheromone (Möglich et al. 1974) that usually attracts a single follower. The leader–follower pair then leave the nest together. The leader moves toward the discovered food by roughly a body length and then stops and waits for the follower to make physical contact with her rear end. Meanwhile, the follower usually sweeps her head from side to side as she closes the distance between her and the leader. Once she touches the leader, the process repeats until both reach the food item. At that point, one or both of them can return to the nest and start a new tandem run. However, after the tandem run, the two ants may take different paths back to the home nest, and future visits to the food item by either ant may be along different paths.

The resulting colony-level behavior is qualitatively different from trail following in a number of ways. Whereas unanimous agreement is expected in trail laying, tandem running can support persistent non-trivial allocations of foragers across

Fig. 17 *Temnothorax rugatulus* during a tandem run. Here, the leader (*right*) waits for her follower (*left*) to make physical contact. At that point, the leader moves forward a small distance and repeats the process until both ants reach the destination of interest (i.e., a candidate nest site or an item of food). The leader can make her presence known to the follower through chemical communication, but chemical trails are not used for navigation. Moreover, both the leader and the follower may take different paths on subsequent visits. Thus, the destination is encoded within the "memory" of each ant (Photo credit to Takao Sasaki and James S. Waters)

multiple food items. This property is because tandem runs do not have the same step-like increase in effectiveness with recruitment effort that is seen in pheromone trails. The strongly non-linear relationship in pheromone-trail recruitment magnifies small chance differences in exploitation, driving the ants toward exclusive use of only one option—the option with the strongest trail. The effectiveness of tandem runs, on the other hand, is linear in recruitment effort; as long as there is a pool of potential recruits at the old nest, each additional tandem run is expected to increase the arrival rate of new ants by the same amount. So when multiple food sources are discovered, there is exploitation of all of them. Furthermore, if the probability of initiating tandem runs depends on food quality, the colony will distribute its foraging force across the food sources according to their quality (Shaffer et al. 2013). For similar reasons, tandem running can adapt more quickly to changing environments, such as the discovery of a good food source after the colony has already begun exploiting a mediocre one. In this situation, trail-laying ants may be trapped by their already established trail, which will outcompete any nascent trail at the new source (Beckers et al. 1990; Detrain and Deneubourg 2008). Tandem runs, on the other hand, can always divert some foraging effort to the new source, initiating a process of positive feedback that will eventually overtake the original source (Shaffer et al. 2013).

Thus, tandem running is a dynamic resource allocation strategy adapted for simultaneous exploitation of multiple foraging sites. In optimization heuristics inspired by trail laying, regions of the parameter space are virtually stained in a way that is globally visible, yet decaying. That globally visible staining is able to re-prioritize the search for the best set of parameters. With Human Computation in mind, tandem running is analogous to re-distributing a pool of human computers among a set of problems based on need. As was discussed in the Infrastructure and Architecture section of this book, humans may be viewed as computational resources that need to be allocated efficiently to different problems. Problems that have high computational need should lead to more recruitment of additional help. However,

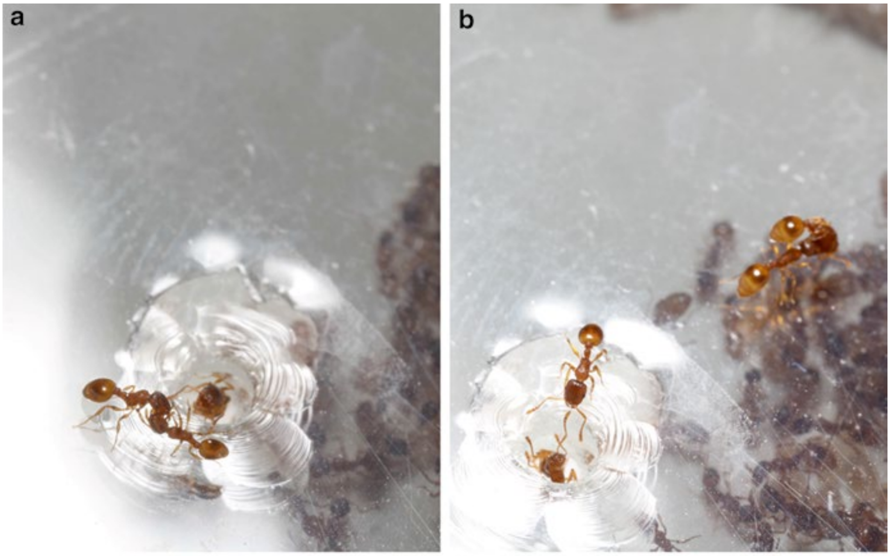


Fig. 18 Social transport by *Temnothorax rugatulus*. Transports bring the bulk of the colony to its new home during colony emigration. Once an ant begins transport to a site, she ceases to lead tandem runs there. Carried ants are generally the in-nest workers and brood that do not travel on their own outside the nest and will not be able to return to the old site. Thus, the switch to transport marks the "commitment" of a scout to the candidate site. In (**a**), a committed scout near the entrance of its old nest initiates transport with another ant who reciprocates by adopting a position suitable for being carried. In (**b**), the pair are shown moving away from the old nest toward the new nest (Photo credit to Takao Sasaki and James S. Waters)

rather than sending a global advertisement to attract large increases in work force, a random individual could be selected for a single advertisement, with the probability of delivering the advertisement increasing with some local measure of the need for additional help. So rather than the problem with the greatest need taking all of the computational resources, as in the trail-laying process, computational resources are allocated to all problems simultaneously and are proportioned according to need.

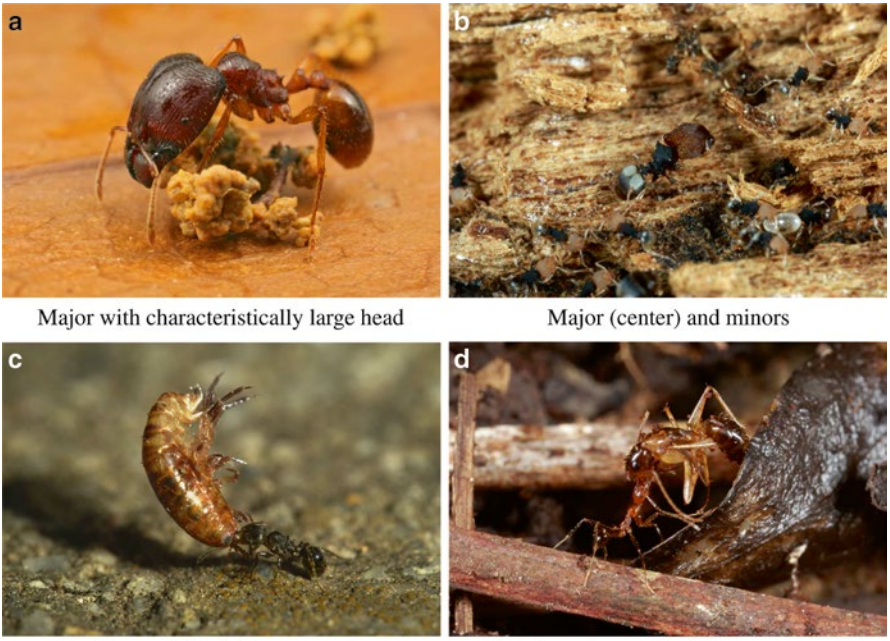
As discussed in section "Individual Ignorance Reduces Collective Cognitive Overload", *Temnothorax* colonies frequently have to choose the best of a set of candidate nests. This task is not well suited to a resource allocator like tandem running by itself. When its nest is destroyed, the colony has to assess the relative quality of new candidate homes and then move the colony into the single best one. This assessment process is similar to a foraging task as it requires scouts to search the environment for different opportunities. During the initial assessment process, scouts make use of tandem running to gradually allocate the scouting team to different nests in proportion to nest quality. In order to convert this distribution into consensus on a single site, the ants add a non-linear component in the form of a quorum rule. As soon as one site achieves a minimum number of adherents, its scouts switch from slow tandem runs to faster direct transports (Fig. 18) (Pratt et al. 2002; Pratt 2005b).

In this recruitment method, the scouts repeatedly travel to the old site and use their mandibles to lift up nestmates (including brood items and the queen) and rapidly carry them to the new nest. This switch accelerates migration, allowing the colony to move into the first site to reach a quorum before any other site has done so. The chosen site is likely to be the best one, because the tandem-run phase apportions scouts according to site quality. Thus, reliance on a quorum rule increases the likelihood of consensus on the best site. By using this rule contingently, colonies can match decision outcomes to context, achieving either consensus or allocation as appropriate to each setting (i.e., nest-site selection and foraging, respectively). The decentralized decision-making processes in future Human Computation systems may similarly need to optimally mix different kinds of linear and non-linear recruitment for different contexts so that computations are sufficiently fast, able to respond to environmental changes, and robust to individual errors.

A Diversity of Unforeseen Futures

In this chapter, we have attempted to establish parallels between Human Computation and eusociality so as to speculate about a future human superorganism that emerges via HC. Given the tremendous diversity in the social insects, it is clear that we have left a great deal out. A few notable omissions include:

Division of labor: We have not discussed how the division of labor within worker castes is established and maintained (Beshers and Fewell 2001; Gordon 1996; Hölldobler and Wilson 2009; Richardson et al. 2011; Dornhaus et al. 2008; Robson and Beshers 1997; Calderone and Page 1996; Tofts and Franks 1992). A variety of ants have distinct morphological castes specialized for different tasks (Hölldobler and Wilson 1990, 2009). These polymorphic ants include the army ants described in sections “Leveraging Diversity in Large, Long-Lasting Projects” and “Individual Ignorance Reduces Collective Cognitive Overload”, the carpenter ants described in section “Oligogyny and Leaderlessness: Competitors that Share the Same Workers”, the widespread *Pheidole* genus (Fig. 19), the leafcutter ants (*Atta*; Fig. 20) described below, and the well-known fire ants (*Solenopsis*; Fig. 21) (Hölldobler and Wilson 1990, 2009; Tschinkel 2006).As already mentioned in the Foundations section of this book, even in species without physical polymorphism, individual workers show strong tendencies to specialize on particular tasks (Beshers and Fewell 2001; Hölldobler and Wilson 2009). Moreover, a worker may change her specialization as she ages (Franks and Tofts 1994; Hölldobler and Wilson 2009; Tofts and Franks 1992; Calderone and Page 1996; Schofield et al. 2011; Tripet and Nonacs 2004). Polymorphism, workforce symmetry breaking, and age-induced changes in specialization are all issues that could be relevant to a future with widespread Human Computation, and there is much research into the mechanisms that drive this so-called “polyethism” in social insects. For brevity, we focus here on “age polyethism,” the age-related



a Major with characteristically large head
b Major (center) and minors
c Carrying a relatively large crustacean
d Carrying a larger dead *Camponotus* ant

Fig. 19 *Pheidole* ants. The genus *Pheidole* is widespread and diverse. Most species have two distinct worker classes-"minor" and "major" workers. The major workers have distinctively large heads; their large mandibles are used in colony defense or to break up large pieces of food. (Photo credits to Steve Shattuck)

division of labor. There are several possible mechanisms that explain how these temporal changes may occur in social insects, and similar mechanisms might be found in HC systems. For example:

- Developmental programs may trigger age-related changes in worker behavior (Calderone and Page 1996). Similarly, a human teenager emerging from adolescence will experience neural or hormonal changes that may lead to a shift in digital behavior. Thus, some age-related division of labor in HC systems may be a shadow of the human developmental program. This program has been shaped by natural selection, and so it is tempting to consider whether a future with ubiquitous HC could act as an additional selective pressure on human development.
- Age-related changes in specialization may also be driven by fatigue. For example, as leaf-cutter ant mandibles wear, the workers switch to carrying the leaf fragments cut by workers with sharper mandibles (Schofield et al. 2011). Likewise, if a complicated visual classification task is distributed across a bank of human classifiers, individuals that specialize on small features may have to shift to different tasks after years of eye strain.

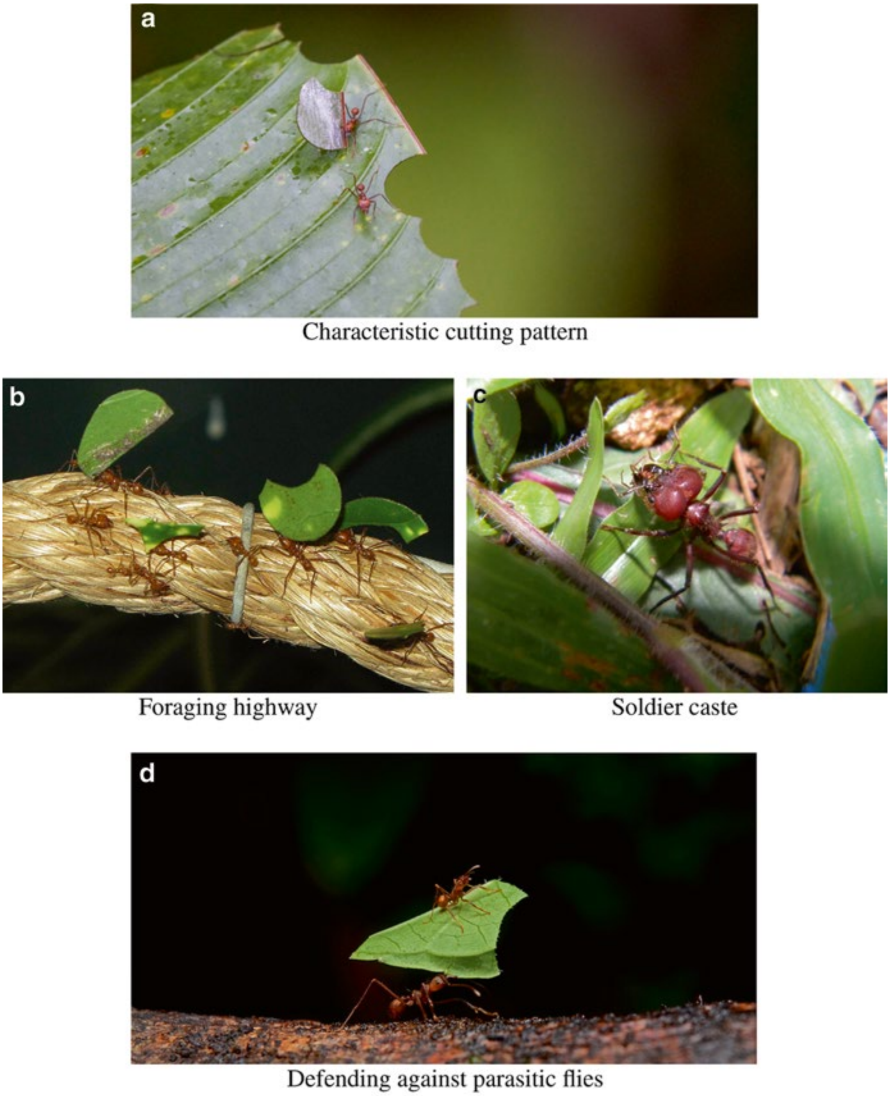


Fig. 20 Leaf-cutter ants (*Atta cephalotes*). One of the several castes of leaf-cutter ants uses its sharp mandibles in a scissor-like motion to cut leaves in the pattern shown in (**a**). These leaves are then carried back along foraging highways, like the one in (**b**). When the leaves reach the colony, they are used to nourish a fungus garden that grows inside the ants' nest. The ants then feed on the fruiting bodies produced by the fungus. The fungus grown by leaf-cutter colonies is not found elsewhere in nature; it is a monoculture passed down from mother colony to daughter alate, and the ants maintain its health with by applying chemicals similar to pesticides in human agriculture. Unlike the big-headed *Pheidole* solders, leaf-cutter solders like the one in (**c**) aggressively defend the colony. Additionally, as shown in (**d**), workers of the smallest caste ride on top of leaves and defend against parasitic flies that can lay eggs within the body of the otherwise vulnerable ant carrying the leaf. Leaf-cutter ants also aggressively fight their own trash-handling workers to prevent them from re-entering the nest and contaminating the fungus (Hart and Ratnieks 2001). While some of the *Camponotus* ants discussed earlier are called "ranchers" due to their management of aphid herds, fungus-growing ants like these are sometimes called "farmers" (Photo credits to: Matt MacGillivray (photo (**a**)); Adrian Pingstone (photo (**b**)); Maximilian Paradiz (photo (**c**)); Geoff Gallice (photo (**d**)))

Fig. 21 The red imported fire ant (*Solenopsis invicta*). Originally from South America, this ant has become a worldwide pest. They are highly invasive, predatory, and can damage agricultural crops either by injuring plants or by killing natural pollinators. A resilient species, they can form large floating colonial rafts to withstand floods (Anderson et al. 2002; Haight 2006; Mlot et al. 2011) (Photo credit to Scott Bauer)

- Alternatively, shifts in task preference may be an emergent property (Franks and Tofts 1994; Tofts and Franks 1992). Tasks in a typical ant colony have an orderly spatial distribution, with nursing taking place at the brood pile near the nest center, food processing just outside the center, nest maintenance and defense at the periphery, and foraging outside. Ants always start their adult life on the brood pile. If they follow a simple rule of always moving away from the center when they perceive a lack of available work, then they will tend to follow a task sequence that mirrors the spatial layout of tasks in the nest, with brood care at the start and foraging at the end. Similarly, if workers can select different HC applications from a relatively static list on an application marketplace, the newest workers will likely choose the most popular applications near the top of the list. With their entry, the workers already engaged in those applications will detect less work availability. Some of those experienced workers will then switch to less well-known applications from farther down on the popularity list. This process will yield a division of labor based on the amount of experience with the system. If these HC systems become widespread and adopted for life starting at an early age, the most experienced human computers will also be the oldest. Consequently, there will be an age-related division of labor driven not by developmental program but by the dynamics of work availability.

Age polyethism is complex and likely results from a mixture of causes (Tripet and Nonacs 2004). In general, HC system designers should be cognizant of the expected demographics of their workforce. Moreover, to account for developmental changes, task-allocation strategies should be adaptable based on the performance of each individual.

Trophallaxis: Although we discussed peer-to-peer software briefly, we did not highlight its possible relationship with trophallaxis in social insects. Trophallaxis is the direct transfer of food among colony members (Wheeler 1918), and it may serve a variety of different functions. For example, many colonies have a high number of apparently inactive workers who are sustained with food shared by their nestmates (Dornhaus et al. 2008; Gordon 2010). Their function, if any, remains unclear, but they may serve as a labor reserve. Likewise, a given HC workforce might retain more workers than are generally necessary, to deal with occasional bursts of high demand. If workers are rewarded only when immediately productive, then a buffer of idle workers will not be sustainable, as the unrewarded workers will leave the pool. The longevity of the project might be reduced if large bursts of work cannot be effectively dispatched, and so it is in the interest of the workers that are consistently receiving rewards to share some of those rewards with idle workers. These peer-to-peer incentive transfers are a kind of HC trophallaxis, and they help to artificially inflate the standing workforce so it is better equipped to handle occasional bursts of work. Alternatively, the work itself can be the substance moving via HC trophallaxis between workers. If the task received by one worker can be partitioned and re-distributed, then many workers are able to stay active at one time while keeping the system well under its total capacity.

Interaction networks: Additionally, we have not given adequate attention to how networks and interaction rates regulate behavior in a social-insect colony (Bonabeau et al. 1998; Fewell 2003; Gordon et al. 1993, 2008; Pinter-Wollman et al. 2011; Pratt 2005b; Gordon 2010; Waters and Fewell 2012). Much of the decentralized ability of colonies to complete tasks is regulated by topological and temporal properties of networks of interacting workers. Simple behavioral rules based on interaction rates can explain much of the self organization observed in social-insect colonies. These rules and structures can serve as inspiration for building HC networks that have sufficient potential for a phase transition into superorganismic computation. For example, notable similarities exist between the social graphs of ant colonies and natural regulatory networks (Waters and Fewell 2012). In fact, ant network topology shares more in common with biological regulatory networks than with social networks. Consequently, when designing networks to facilitate HC, it may be a mistake to catalyze connections along social directions; efficient computation might be better assisted by enforcing regulatory network motifs.

Traffic patterns and flow control: In sections "Individual Ignorance Reduces Collective Cognitive Overload" and "When Ants Fail", we discussed how ants

Fig. 22 Characteristic fork in a foraging trail of *Leptogenys* (Photo credit to Steve Shattuck)

make use of trails for navigation and recruitment. As hinted in the Foundations section of this book, ant trails and traffic management on them is a much richer topic than we have presented here, and aspects we have not discussed could potentially provide useful inspiration for protocols that facilitate future Human Computation. For example:

- Trail-laying ants like *Leptogenys processionalis* and the marauder ant *Pheidologeton diversus* have characteristic branching patterns in their foraging trails (Fig. 22) (Ganeshaiah and Veena 1991; Moffett 1988). These bifurcation patterns are non-random and may result from the finite range of chemical communication between foragers (Ganeshaiah and Veena 1991). The resulting topological pattern seems to be an efficient structure for exploring a large area with relatively short total trail length. Trail-inspired search heuristics might be informed by these branching patterns. Moreover, the putative mechanism that forms these trails shows again how an apparent limitation (i.e., finite communication range) is adaptive when tuned to generate useful patterns.
- Leaf-cutter ants (*Atta*; Fig. 20) build elaborate and well-maintained highway systems on which they transport leaves to feed their underground fungus gardens. As the highways connect the central nest directly to cutting sites on trees, they have a natural branching pattern. Consequently, these ants have developed leaf flow-control mechanisms that depend on first saturating the highway with unloaded workers. Outgoing ants choose whether to carry leaves back to the nest based on their interaction rate with loaded incoming

Fig. 23 *Aphaenogaster cockerelli*. This ant is feeding on fig paste that has been presented during an experiment in the field. The *orange* coloring on her head, body, and legs is paint that has been applied to track her nest origin during the experiment. Historically, *Aphaenogaster cockerelli* was called *Novomessor cockerelli* (Photo credit to Jessica D. Ebie)

ants; these rules help to regulate leaf flow despite variation in highway branch width (Burd 2000; Dussutour et al. 2004; Farji-Brener et al. 2010; Fourcassié et al. 2010). These mechanisms demonstrate how even the idle workers in a decentralized system may actually serve an information-related purpose.

Polydomy: Given that Human Computation will likely be distributed over a large geographic area, investigations of "polydomy" in ants may be relevant to understanding future HC systems. Polydomy, or occupation of multiple nest sites by a single colony (Partridge et al. 1997; Schmolke 2009; Hölldobler and Carlin 1989; Smith et al. 2011), is also discussed in the Foundations section of this book. Polydomy presents a number of interesting problems in decentralized control. For example, colonies of *Aphaenogaster cockerelli* (Fig. 23) have only one queen but typically occupy multiple nests. If the queen dies, workers respond to her absence by developing their ovaries and laying eggs that produce alate males (Hölldobler and Carlin 1989; Smith et al. 2011). To suppress worker reproduction while she is alive, the queen somehow signals her presence to workers in all of the colony's nests, even though she can only reside in one of them. How she does so remains unknown.

The benefits of polydomy itself are not well understood, but there is some evidence that it increases colony foraging success (Schmolke 2009). This explanation appears to parallel how an Internet Content Distribution Network (CDN) improves the quality of service delivered to an audience dispersed around the globe. In particular, if food is randomly scattered throughout an environment, foragers in a polydomous colony face a lower transport burden than a monodomous colony. Just as a CDN distributes content so it can be close to consumers, a polydomous colony distributes its nests so they can be close to their food sources. Likewise, HC systems may need a similar dispersion to facilitate parallel clusters of co-located human computers. Within each subnetwork, human participants will have fast access to the data being processed as well as to their peers in the network. In fact, Internet gaming communities and high-speed stock-market flash traders already show some signs of polydomy-like optimization to maximize efficiency. Moreover, teams of co-located humans that have direct physical access to data may be the natural HC extension of a co-located server farm sharing direct access to an important resource.

As more HC projects compete for human computational talents, there may be additional lessons to learn from polydomy. The polydomous ant *Aphaenogaster cockerelli* is also known to form teams that collectively retrieve large food items (Hölldobler et al. 1978; Berman et al. 2010, 2011; Kumar et al. 2013). They can form these teams by recruiting local assistance (i.e., they need not return to the nest to form a carrying team) (Hölldobler et al. 1978). Doing so allows more rapid retrieval of large prey than the alternative method of cutting it up in place and retrieving the pieces. Combined with polydomy, which ensures shorter transport distances, this strategy may reduce the time prey spends above ground where it can be captured by competitors. Rather than comparing humans in an HC-system to individual ants, it may be useful to think of the humans as the prey items being collected by the HC colonies. If humans self subscribe to different HC projects that have tasks that are disguised as on-line games with rewards that improve with decreasing delay, then it will be in the best interest of the HC-organizer to distribute multiple "nest entrances" nearest to potential sources of HC talent. Otherwise, competing projects will better attract the attention of the self-assorting human computers.

It is not possible for one chapter to completely capture the rich set of social insect model systems for a variety of distributed phenomena. We have leveraged this great diversity as a bank of examples that each might represent one particular future of Human Computation. However, like modern social-insect fauna, it is more likely that a wide variety of different kinds of HC will co-exist simultaneously. The resulting computational ecosystem is difficult to picture.

Such a future might be beyond the "technological singularity" predicted by the futurist Kurzweil (1999). He suggests that after some point in time, humans will "transcend biology" (Kurzweil 2005) and create computers that "exceed human intelligence" (Kurzweil 1999). In some ways, this vision is consistent with the superorganismic phase transition we described at the start of section "Future Human Imitations of Eusocial Insect Society", albeit the imagery seems to be more abrupt than the continuous higher-order phase transition that we picture. The fuzzier transition that we described is more in line with the "mitochondrial singularity" recently suggested by microbiologist Slonczewski in order to predict the role of humans in a post-singularity world (Ghose 2013; Slonczewski 2013a,b). As we look to the evolution of eusociality for lessons, she looks to the evolution of the mitochondria within our cells. These organelles have the highly specialized task of providing power to each cell, but their ancestors were once free-living bacteria that performed all of the general functions of a living cell. Eventually, some of those ancestral bacteria managed to embed themselves within another free-living cell, and the pair became symbiotic partners. Gradually, the mitochondrial ancestors gave up their other roles and became a specialized organelle. Nevertheless, mitochondria today retain some of their past identity – each one contains its own separate DNA and is passed directly from mother to offspring without any modifications outside of the occasional random mutation. Slonczewski pictures a similar fate for humans – as humans augment their abilities with computers, the result is a symbiotic relationship. However, if computer intelligence eclipses that of humans, the human side of the symbiosis will gradually lose its intellect in favor of specializing on other support functions.

When focusing on a future driven by *Human* Computation, the Kurzweil–Slonczewski picture seems lacking because it neglects the fact that aggregate digital intelligence may largely depend on synergistic connections between ignorant but still cognitive individuals. Thus, we think it is informative to look to recent work of astrobiologists like Walker et al. to re-define life in terms of its information processing ability (Walker et al. 2012; Walker and Davies 2013). In their view, the transition from non-living collections of particles to a living aggregate must go through a corresponding transition from bottom-up causality to top-down causality. That is, before the transition, the aggregate behavior is a simple combination of independent actions by the constituents; causality points "up" from local to global. After the transition, the behaviors of the constituents lose much of their independence and instead are clearly responsive to signals found in the aggregate; causality points "down" from global to local. When we consider the co-existence of multiple forms of HC that each compete for humans to participate in the computation, the humans seem less like workers in an ant colony and more like morsels of food that are the prizes in competitions between multiple co-existing colonies. As humans transition from independent engines of computation to digital nutrients for computational networks, it seems as if HC goes through a corresponding transition from bottom-up to top-down causality. Humans will not be the mitochondria of this post-singularity world. Instead, they are digital food that sustains emergent decentralized artificial life.

Acknowledgements Thanks to Bert Hölldobler for connecting us to this interesting speculative project. The writing of this chapter was supported by the National Science Foundation (award 1012029). Images in this chapter that were not already in the public domain were used either under the explicit permission of the image owner or according to a CC BY 2.0 (Creative Commons 2013a), CC BY 2.5 (Creative Commons 2013b), or CC BY 3.0 (Creative Commons 2013c) license.

References

Amazon.com, Inc. (2005) Amazon mechanial turk. http://www.mturk.com/

Anderson C, Theraulaz G, Deneubourg JL (2002) Self-assemblages in insect societies. Insectes Sociaux 49(2):99–110. doi:10.1007/s00040-002-8286-y

André JB, Peeters C, Doums C (2001) Serial polygyny and colony genetic structure in the monogynous queenless ant *Diacamma* cyaneiventre. Behav Ecol Sociobiol 50(1):72–80. doi:10.1007/s002650100330

Baratte S, Cobb M, Peeters C (2006) Reproductive conflicts and mutilation in queenless *Diacamma* ants. Anim Behav 72(2):305–311. doi:10.1016/j.anbehav.2005.10.025

Beach CB, McMurry FM (eds) (1914) The new student's reference work for teachers, students and families. FE Compton, Chicago

Beckers R, Deneubourg JL, Goss S, Pasteels JM (1990) Collective decision making through food recruitment. Insectes Sociaux 37(3):258–267. doi:10.1007/BF02224053

Beckers R, Deneubourg JL, Goss S (1992a) Trail laying behaviour during food recruitment in the ant *Lasius niger* (L.). Insectes Sociaux 39(1):59–72. doi:10.1007/BF01240531

Beckers R, Deneubourg JL, Goss S (1992b) Trails and U-turns in the selection of a path by the ant *Lasius niger*. J Theor Biol 159(4):397–415. doi:10.1016/S0022-5193(05)80686-1

Berman S, Lindsey Q, Sakar MS, Kumar V, Pratt SC (2010) Study of group food retrieval by ants as a model for multi-robot collective transport strategies. In: Proceedings of robotics: science and systems, Zaragoza

Berman S, Lindsey Q, Sakar MS, Kumar V, Pratt SC (2011) Experimental study and modeling of group retrieval in ants as an approach to collective transport in swarm robotic systems. Proc IEEE 99(9):1470–1481. doi:10.1109/JPROC.2011.2111450

Beshers SN, Fewell JH (2001) Models of division of labor in social insects. Annu Rev Entomol 46:413–440. doi:10.1146/annurev.ento.46.1.413

Bhadra A, Gadagkar R (2008) We know that the wasps 'know': cryptic successors to the queen in *Ropalidia marginata*. Biol Lett 4(6):634–637. doi:10.1098/rsbl.2008.0455

Bhadraa A, Iyera PL, Sumanaa A, Deshpandea SA, Ghosha S, Gadagkar R (2007) How do workers of the primitively eusocial wasp *Ropalidia marginata* detect the presence of their queens? J Theor Biol 246(3):574–582. doi:10.1016/j.jtbi.2007.01.007

Birattari M, Di Caro G, Dorigo M (2002) Toward the formal foundation of ant programming. In: Dorigo M, Di Caro G, Sampels M (eds) Proceedings of the third international workshop on ant algorithms (ANTS 2002), Brussels, pp 39–72. doi:10.1007/3-540-45724-0_16

Bonabeau E, Dorigo M, Theraulaz G (1999) Swarm intelligence: from natural to artificial systems. Oxford University Press, Oxford

Bonabeau E, Theraulaz G, Deneubourg JL (1998) Group and mass recruitment in ant colonies: the influence of contact rates. J Theor Biol 195(2):157–166. doi:10.1006/jtbi.1998.0789

Brady SG (2003) Evolution of the army ant syndrome: the origin and long-term evolutionary stasis of a complex of behavioral and reproductive adaptations. Proc Natl Acad Sci USA 100(11):6575–6579. doi:10.1073/pnas.1137809100

Burd M (2000) Foraging behaviour of *Atta cephalotes* (leaf-cutting ants): an examination of two predictions for load selection. Anim Behav 60(6):781–788. doi:10.1006/anbe.2000.1537

Buschinger A (2011) Queen polymorphism in an Australian ant, *Monomorium* cf. *rubriceps*Mayr, 1876 (Hymenoptera: Formicidae). Myrmecol News 15:63–66

Calderone NW, Page RE (1996) Temporal polyethism and behavioural canalization in the honey bee, *Apis mellifera*. Anim Behav 51(3):631–643. doi:10.1006/anbe.1996.0068

Camazine S, Deneubourg JL, Franks NR, Sneyd J, Theraulaz G, Bonabeau E (2001) Self-organization in biological systems. Princeton University Press, Princeton

Cargel RA, Rinderer TE (2004) Unusual queen cell construction and destruction in *Apis mellifera* from far-eastern Russia. J Apic Res 43(4):188–190

Creative Commons (2013a) Attribution 2.0 Generic (CC BY 2.0). http://creativecommons.org/licenses/by/2.0/

Creative Commons (2013b) Attribution 2.5 Generic (CC BY 2.5). http://creativecommons.org/licenses/by/2.5/

Creative Commons (2013c) Attribution 3.0 Unported (CC BY 3.0). http://creativecommons.org/licenses/by/3.0/

Cuvillier-Hot V, Gadagkar R, Peeters C, Cobb M (2002) Regulation of reproduction in a queenless ant: aggression, pheromones and reduction in conflict. Proc R Soc B 269(1497):1471–2954. doi:10.1098/rspb.2002.1991

Delsuc F (2003) Army ants trapped by their evolutionary history. PLoS Biol 1(2):e37. doi:10.1371/journal.pbio.0000037

de Biseau JC, Deneubourg JL, Pasteels JM (1991) Collective flexibility during mass recruitment in the ant *Myrmica sabuleti* (Hymenoptera: Formicidae). Psyche 98(4):323–336. doi:10.1155/1991/38402

Denny AJ, Franks NR, Powell S, Edwards KJ (2004) Exceptionally high levels of multiple mating in an army ant. Naturwissenschaften 91(8):396–399. doi:10.1007/s00114-004-0546-4

Deshpande SA, Sumana A, Surbeck M, Gadagkar R (2006) Wasp who would be queen: a comparative study of two primitively eusocial species. Curr Sci 91(3):332–336

Detrain C, Deneubourg JL (2008) Collective decision-making and foraging patterns in ants and honeybees, vol 35. Academic, Amsterdam/Boston, pp 123–173. doi:10.1016/S0065-2806(08)00002-7

Dornhaus A, Holley JA, Pook VG, Worswick G, Franks NR (2008) Why do not all workers work? colony size and workload during emigrations in the ant *Temnothorax albipennis*. Behav Ecol Sociobiol 63(1):43–51. doi:10.1007/s00265-008-0634-0

Dorigo M, Birattari M, Stützle T (2006) Ant colony optimization: artificial ants as a computational intelligence technique. IEEE Comput Intell Mag 1(4):28–39. doi:10.1109/MCI.2006.329691

Dorigo M, Di Caro G, Gambardella LM (1999) Ant algorithms for discrete optimization. Artif Life 5(2):137–172. doi:10.1162/106454699568728

Dorigo M, Maniezzo V, Colorni A (1996) Ant System: optimization by a colony of cooperating agents. IEEE Trans Syst Man Cybern Part B: Cybern 26(1):29–41. doi:10.1109/3477.484436

Dussutour A, Beekman M, Nicolis SC, Meyer B (2009) Noise improves collective decision-making by ants in dynamic environments. Proc R Soc B 276(1677):4353–4361. doi:10.1098/rspb.2009.1235

Dussutour A, Fourcassié V, Helbing D, Deneubourg JL (2004) Optimal traffic organization in ants under crowded conditions. Nature 428:70–73. doi:10.1038/nature02345

Farji-Brener AG, Amador-Vargas S, Chinchilla F, Escobar S, Cabrera S, Herrera MI, Sandoval C (2010) Information transfer in head-on encounters between leaf-cutting ant workers: food, trail condition or orientation cues? Anim Behav 79(2):343–349. doi:10.1016/j.anbehav.2009.11.009

Feigenbaum J, Shenker S (2002) Distributed algorithmic mechanism design: recent results and future directions. In: Proceedings of the 6th international workshop on discrete algorithms and methods for mobile computing and communication, Atlanta, pp 1–13. doi:10.1145/570810.570812

Fersch R, Buschinger A, Heinze J (2000) Queen polymorphism in the australian ant *Monomorium* sp.10. Insectes Sociaux 47(3):280–284. doi:10.1007/PL00001715

Fewell JH (2003) Social insect networks. Science 301(5641):1867–1870. doi:10.1126/science.1088945

Fourcassié V, Dussutour A, Deneubourg JL (2010) Ant traffic rules. J Exp Biol 213(14):2357–2363. doi:10.1242/jeb.031237

Franks NR, Gomez N, Goss S, Deneubourg JL (1991) The blind leading the blind in army ant raid patterns: testing a model of self-organization (Hymenoptera: Formicidae). J Insect Behav 4(4):583–607. doi:10.1007/BF01048072

Franks NR, Richardson T (2006) Teaching in tandem-running ants. Nature 439:153. doi:10.1038/439153a

Franks NR, Tofts C (1994) Foraging for work: how tasks allocate workers. Anim Behav 48(2):470–472. doi:10.1006/anbe.1994.1261

Fujisawa R, Dobata S, Kubota D, Imamura H, Matsuno F (2008) Dependency by concentration of pheromone trail for multiple robots. In: Dorigo M, Birattari M, Blum C, Clerc M, Stüzle T, Winfield AFT (eds) Proceedings of the 6th international conference on ant colony optimization and swarm intelligence (ANTS 2008), Brussels

Gadau J, Gertsch PJ, Heinze J, Pamilo P, Hölldobler B (1998) Oligogyny by unrelated queens in the carpenter ant, *Camponotus ligniperdus*. Behav Ecol Sociobiol 44(1):23–33. doi:10.1007/s002650050511

Ganeshaiah KN, Veena T (1991) Topology of the foraging trails of *Leptogenys processionalis*: why are they branched? Behav Ecol Sociobiol 29(4):263–270. doi:10.1007/BF00163983

Ghose T (2013) Human takeover by machines may be closer than we think. http://science.nbcnews.com/_news/2013/05/07/18109236-human-takeover-by-machines-may-be-closer-than-we-think

Google (2009) reCAPTCHA. http://www.google.com/recaptcha
Gordon DM (1996) The organization of work in social insect colonies. Nature 380:121–124. doi:10.1038/380121a0
Gordon DM (2010) Ant encounters: interaction networks and colony behavior. Princeton University Press, Princeton
Gordon DM, Holmes S, Nacu S (2008) The short-term regulation of foraging in harvester ants. Behav Ecol 19(1):217–222. doi:10.1093/beheco/arm125
Gordon DM, Paul RE, Thorpe K (1993) What is the function of encounter patterns in ant colonies? Anim Behav 45(6):1083–1100. doi:10.1006/anbe.1993.1134
Gotoh A, Sameshima S, Tsuji K, Matsumoto T, Miura T (2005) Apoptotic wing degeneration and formation of an altruism-regulating glandular appendage (gemma) in the ponerine ant *Diacamma* sp. from japan (Hymenoptera, Formicidae, Ponerinae). Dev Genes Evol 215(2):69–77. doi:10.1007/s00427-004n-0456-7n
Gotwald WH Jr (1995) Army ants: the biology of social predation. Cornell University Press, Ithaca
Haight KL (2006) Defensiveness of the fire ant, *Solenopsis invicta*, is increased during colony rafting. Insectes Sociaux 53(1):32–36. doi:10.1007/s00040-005n-0832-y
Hart AG, Ratnieks FLW (2001) Task partitioning, division of labour and nest compartmentalisation collectively isolate hazardous waste in the leafcutting ant *atta cephalotes*. Behav Ecol Sociobiol 49(5):387–392. doi:10.1007/s002650000312
Heinze J (1998) Intercastes, intermorphs, and ergatoid queens: who is who in ant reproduction? Insectes Sociaux 45(2):113–124. doi:10.1007/s000400050073
Heinze J, Keller L (2000) Alternative reproductive strategies: a queen perspective in ants. Trends Ecol Evol 15(12):508–512. doi:10.1016/S0169-5347(00)01995-9
Hölldobler B, Carlin NF (1985) Colony founding, queen dominance and oligogyny in the australian meat ant *Iridomyrmex purpureus*. Behav Ecol Sociobiol 18(1):45–58. doi:10.1007/BF00299237
Hölldobler B, Carlin NF (1989) Colony founding, queen control, and worker reproduction in the ant *Aphaenogaster (=Novomessor) cockerelli*. Psyche 96(3–4):131–151. doi:10.1155/1989/74135
Hölldobler B, Möglich M, Maschwitz U (1974) Communication by tandem running in the ant *Camponotus sericeus*. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 90(2):105–127. doi:10.1007/BF00694481
Hölldobler B, Stanton RC, Markl H (1978) Recruitment and food-retrieving behavior in *Novomessor* (Formicidae: Hymenoptera): I. chemical signals. Behav Ecol Sociobiol 4(2):163–181. doi:10.1007/BF00354978
Hölldobler B, Wilson EO (1977) The number of queens: an important trait in ant evolution. Naturwissenschaften 64(1):8–15. doi:10.1007/BF00439886
Hölldobler B, Wilson EO (1990) The Ants. Harvard University Press, Cambridge
Hölldobler B, Wilson EO (2009) The Superorganism: the beauty, elegance, and strangeness of insect societies. WW Norton, New York
Inoue T, Sakagami SF, Salmah S, Yamane S (1984) The process of colony multiplication in the Sumatran stingless bee *Trigona laeviceps*. Biotropica 16(2):100–111
Jenkins C (2013) [APP] Swarm! on Glass: The attention economy, glamified. http://livingthrugłass.com/app-swarm-on-glass-the-attention-economy-gamified/
Krebs RA, Rissing SW (1991) Preference for large foundress associations in the desert ant *Messor pergandei*. Anim Behav 41(2):361–363. doi:10.1016/S0003-3472(05)80487-7n
Kronauer DJC, Johnson RA, Boomsma JJ (2007) The evolution of multiple mating in army ants. Evolution 61(2):413–422. doi:10.1111/j.1558-5646n.2007.00040.x
Kronauer DJC, Schöning C, d'Ettorre P, Boomsma JJ (2010) Colony fusion and worker reproduction after queen loss in army ants. Proc R Soc B 277(1682):755–763. doi:10.1098/rspb.2009.1591
Kumar GP, Buffin A, Pavlic TP, Pratt SC, Berman SM (2013) A stochastic hybrid system model of collective transport in the desert ant *Aphaenogaster cockerelli*. In: Proceedings of the 16th ACM international conference on hybrid systems: computation and control, Philadelphia
Kurzweil R (1999) The age of spiritial machines: when computers exceed human intelligence. Viking, New York

Kurzweil R (2005) The singularity is near: when humans transcend biology. Viking, New York

Lamon B, Topoff H (1981) Avoiding predation by army ants: defensive behaviours of three ant species of the genus *Camponotus*. Anim Behav 29(4):1070–1081. doi:10.1016/S0003-3472(81)80060-7

Liebig J, Hölldobler B, Peeters C (1998) Are ant workers capable of colony foundation? Naturwissenschaften 85(3):133–135. doi:10.1007/s001140050470

Livnat A, Pippenger N (2006) An optimal brain can be composed of conflicting agents. Proc Natl Acad Sci U S A 103(9):3198–3202. doi:10.1073/pnas.0510932103

Loeliger J, McCullough M (2012) Version control with git: powerful tools and techniques for collaborative software development, 2nd edn. O'Reilly, Cambridge

Mas-Colell A, Whinston MD, Green JR (1995) Microeconomic theory. Oxford University Press, New York

Matthey L, Berman S, Kumar V (2009) Stochastic strategies for a swarm robotic assembly system. In: Proceedings of the 2009 IEEE international conference on robotics and automation, Kobe, pp 1953–1958

Mlot NJ, Tovey CA, Hu DL (2011) Fire ants self-assemble into waterproof rafts to survive floods. Proc Natl Acad Sci U S A 108(19):7669–7673. doi:10.1073/pnas.1016658108

Moffett MW (1988) Foraging dynamics in the group-hunting myrmicine ant, *Pheidologeton diversus*. J Insect Behav 1(3):309–331. doi:10.1007/BF01054528

Möglich M, Maschwitz U, Hölldobler B (1974) Tandem calling: a new kind of signal in ant communication. Science 186(4168):1046–1047. doi:10.1126/science.186.4168.1046

Molet M, Baalen MV, Peeters C (2008) Shift in colonial reproductive strategy associated with a tropical-temperate gradient in *Rhytidoponera* ants. Am Nat 172(1):75–87. doi:10.1086/588079

Monnin T, Ratnieks FLW, Jones GR, Beard R (2002) Pretender punishment induced by chemical signalling in a queenless ant. Nature 419(6902):61–65, doi:10.1038/nature00932

Nicolis SC, Deneubourg JL (1999) Emerging patterns and food recruitment in ants: an analytical study. J Theor Biol 198:575–592. doi:10.1006/jtbi.1999.0934

Nonacs P, Reeve HK (1993) Opportunistic adoption of orphaned nests in paper wasps as an alternative reproductive strategy. Behav Process 30(1):47–59. doi:10.1016/0376-6357(93)90011-F

Nonacs P, Reeve HK (1995) The ecology of cooperation in wasps: causes and consequences of alternative reproductive decisions. Ecology 76(3):953–967. doi:dx.doi.org/10.2307/1939359

Osborne MJ, Rubinstein A (1994) A course in game theory. MIT, Cambridge

Palmer KA, Oldroyd BP (2000) Evolution of multiple mating in the genus *Apis*. Apidologie 31(2):235–248. doi:10.1051/apido:2000119

Pardi L (1948) Dominance order in *Polistes* wasps. Physiol Zool 21(1):1–13

Partridge LW, Partridge KA, Franks NR (1997) Field survey of a monogynous leptothoracine ant (Hymenoptera, Formicidae): evidence of seasonal polydomy? Insectes Sociaux 44(2):75–83. doi:10.1007/s000400050031

Peeters C (1991a) Ergatoid queens and intercastes in ants: two distinct adult forms which look morphologically intermediate between workers and winged queens. Insectes Sociaux 38(1): 1–15. doi:10.1007/BF01242708

Peeters C (1991b) The occurence of sexual reproduction among ant workers. Biol J Linn Soc 44(2):141–152. doi:10.1111/j.1095-8312.1991.tb00612.x

Peeters C, Hölldobler B, Moffett M, Musthak Ali TM (1994) "wall-papering" and elaborate nest architecture in the ponerine ant *Harpegnathos saltator*. Insectes Sociaux 41(2):211–218. doi:10.1007/BF01240479

Peeters C, Hölldobler B (1995) Reproductive cooperation between queens and their mated workers: the complex life history of an ant with a valuable nest. Proc Natl Acad Sci U S A 92(24): 10,977–10,979

Peeters C, Ito F (2001) Colony dispersal and the evolution of queen morphology in social Hymenoptera. Annu Rev Entomol 46:601–630. doi:10.1146/annurev.ento.46.1.601

Peeters C, Liebig J, Hölldobler B (2000) Sexual reproduction by both queens and workers in the ponerine ant *Harpegnathos saltator*. Insectes Sociaux 47(4):325–332. doi:10.1007/PL00001724

Pinter-Wollman N, Wollman R, Guetz A, Holmes S, Gordon DM (2011) The effect of individual variation on the structure and function of interaction networks in harvester ants. J R Soc Interface 8(64):1562–1573. doi:10.1098/rsif.2011.0059

Pollock GB, Rissing SW (1985) Mating season and colony foundation of the seed-harvester ant, *Veromessor pergandei*. Psyche 92:125–134. doi:10.1155/1985/87410

Pratt SC (2004) Collective control of the timing and type of comb construction by honey bees (*Apis mellifera*). Apidologie 35:193–205. doi:10.1051/apido:2004005

Pratt SC (2005a) Behavioral mechanisms of collective nest-site choice by the ant *Temnothorax curvispinosus*. Insectes Sociaux 52(4):383–392. doi:10.1007/s00040-005-0823-z

Pratt SC (2005b) Quorum sensing by encounter rates in the ant *Temnothorax albipennis*. Behav Ecol 16(2):488–496. doi:10.1093/beheco/ari020

Pratt SC, Mallon EB, Sumpter DJ, Franks NR (2002) Quorum sensing, recruitment, and collective decision-making during colony emigration by the ant *Leptothorax albipennis*. Behav Ecol Sociobiol 52(2):117–127. doi:10.1007/s00265-002-0487-x

Reeve HK, Gamboa GJ (1983) Colony activity integration in primitively eusocial wasps: the role of the queen (*Polistes fuscatus*, Hymenoptera: Vespidae). Behav Ecol Sociobiol 13(1):63–74. doi:10.1007/BF00295077

Reeve HK, Gamboa GJ (1987) Queen regulation of worker foraging in paper wasps: a social feedback control system (*Polistes fuscatus*, Hymenoptera: Vespidae). Behaviour 102(3/4): 147–167

Reeve HK, Starks PT, Peters JM, Nonacs P (2000) Genetic support for the evolutionary theory of reproductive transactions in social wasps. Proc R Soc B 267(1438):75–79. doi:10.1098/rspb.2000.0969

Richardson TO, Christensen K, Franks NR, Jensen HJ, Sendova-Franks AB (2011) Ants in a labyrinth: a statistical mechanics approach to the division of labour. PLoS ONE 6(4):e18416. doi:10.1371/journal.pone.0018416

Rissing SW, Pollock GB, Higgins MR, Hagen RH, Smith DR (1989) Foraging specialization without relatedness or dominance among co-founding ant queens. Nature 338(6214):420–422. doi:10.1038/338420a0

Rissing SW, Johnson RA, Martin JW (2000) Colony founding behavior of some desert ants: geographic variation in metrosis. Psyche 103(1–2):95–101. doi:10.1155/2000/20135

Robson SK, Beshers SN (1997) Division of labour and 'foraging for work': simulating reality versus the reality of simulations. Anim Behav 53(1):214–218. doi:10.1006/anbe.1996.0290

Sasaki T, Pratt SC (2011) Emergence of group rationality from irrational individuals. Behav Ecol 22(2):276–281. doi:10.1093/beheco/arq198

Sasaki T, Pratt SC (2012) Groups have a larger cognitive capacity than individuals. Curr Biol 22(19):R827–R829. doi:10.1016/j.cub.2012.07.058

Schmolke A (2009) Benefits of dispersed central-place foraging: an individual-based model of a polydomous ant colony. Am Nat 173(6):772–778. doi:10.1086/598493

Schneirla TC (1944) A unique case of circular milling in ants, considered in relation to trail following and the general problem of orientation. Am Mus Novit 1253:1–26

Schneirla TC (1949) Army-ant life and behavior under dry-season conditions. 3. the course of reproduction and colony behavior. Bull Am Mus Nat Hist 94(1):1–82

Schneirla TC (1971) Army ants: a study in social organization. WH Freeman, San Francisco

Schneirla TC, Brown RZ (1950) Army-ant life and behavior under dry-season conditions. 4. further investigation of cyclic processes in behavioral and reproductive functions. Bull Am Mus Nat Hist 50(5):263–354

Schofield RMS, Emmett KD, Niedbala JC, Nesson MH (2011) Leaf-cutter ants with worn mandibles cut half as fast, spend twice the energy, and tend to carry instead of cut. Behav Ecol Sociobiol 65(5):969–982. doi:10.1007/s00265-010-1098-6

Seeley TD (1995) The wisdom of the hive: the social physiology of honey bee colonies. Harvard University, Cambridge

Seeley TD (2010) Honeybee democracy. Princeton University Press, Princeton

Shaffer Z, Sasaki T, Pratt SC (2013) Linear recruitment leads to allocation and flexibility in collective foraging by ants. Anim Behav (in press)

Shakarad M, Gadagkar R (1995) Colony founding in the primitively eusocial wasp, *Ropalidia marginata*(Hymenoptera:Vespidae).EcolEntomol20(3):273–282.doi:10.1111/j.1365-2311.1995.tb00457.x

Sharpe T, Webb B (1996) Simulated and situated models of chemical trail following in ants. In: Pfeifer R, Blumberg B, Meyer JA, Wilson SW (eds) Proceedings of the fifth international conference on simulation of adaptive behavior (SAB96), North Falmouth

Slonczewski J (2013a) Mitochondrial singularity. http://www.antipope.org/charlie/blog-static/2013/03/mitochondrial-singularity.html

Slonczewski J (2013b) Mitochondrial singularity. http://ultraphyte.com/2013/03/25/mitochondrial-singularity/

Smith AA, Haight KL (2008) Army ants as research and collection tools. J Insect Sci 8:71. doi:10.1673/031.008.7101

Smith AA, Hölldobler B, Liebig J (2011) Reclaiming the crown: queen to worker conflict over reproduction in *Aphaenogaster cockerelli*. Naturwissenschaften 98(3):237–240. doi:10.1007/s00114-011-0761-8

St Laurent AM (2004) Understanding open source and free software licensing. O'Reilly, Beijing/Sebastopol

Sumpter DJT, Beekman M (2003) From nonlinearity to optimality: pheromone trail foraging by ants. Anim Behav 66(2):273–280. doi:10.1006/anbe.2003.2224

Surowieckie J (2004) The wisdom of crowds: why the many are smarter than the few and how collective wisdom shapes business, economies, societies and nations. Doubleday, New York

Svennebring J, Koenig S (2003) Trail-laying robots for robust terrain coverage. In: Proceedings of the 2003 IEEE International conference on robotics and automation (ICRA '03), Taipei, vol 1, pp 75–82. doi:10.1109/ROBOT.2003.1241576

Szulc J (2011) Galeria :: Cyfrowo :: Macro :: Mrówki. http://foto.julian.net.pl/gallery3/digital/Mocro-fotografia/HomeAnts

Tarpy DR, Nielsen R, Nielsen DI (2004) A scientific note on the revised estimates of effective paternity frequency in *Apis*. Insectes Sociaux 51(2):203–204. doi:10.1007/s00040-004-0734-4

Tofts C, Franks NR (1992) Doing the right thing: ants, honeybees and naked mole-rats. Trends Ecol Evol 7(10):346–349. doi:10.1016/0169-5347(92)90128-X

Tripet F, Nonacs P (2004) Foraging for work and age-based polyethism: the roles of age and previous experience on task choice in ants. Ethology 110(11):863–877. doi:10.1111/j.1439-0310.2004.01023.x

Tschinkel WR (2006) The fire ants. Belknap, Cambridge

Visscher PK (2007) Group decision making in nest-site selection among social insects. Annu Rev Entomol 52:255–275. doi:10.1146/annurev.ento.51.110104.151025

Walker SI, Cisneros L, Davies PCW (2012) Evolutionary transitions and top-down causation. In: Proceedings of the thirteenth international conference on the simulation and synthesis of living systems (Artificial Life 13), vol 13, East Lansing, pp 283–290. doi:10.7551/978-0-262-31050-5-ch038

Walker SI, Davies PCW (2013) The algorithmic origins of life. J R Soc Interface 10(79):20120869. doi:10.1098/rsif.2012.0869

Waters JS, Fewell JH (2012) Information processing in social insect networks. PLoS ONE 7(7):e40337. doi:10.1371/journal.pone.0040337

West-Eberhard MJ (1969) The social biology of polistine wasps. Miscellaneous publications, vol 140. Museum of Zoology, University of Michigan, Ann Arbor

Wheeler WM (1918) A study of some ant larvæ, with a consideration of the origin and meaning of the social habit among insects. Proc Am Philos Soc 57(4):293–343

Wheeler DE (1986) Developmental and physiological determinants of caste in social Hymenoptera: evolutionary implications. Am Nat 128(1):13–34

Wilson EO (1971) The insect societies. Belknap Press, Cambridge

Yang R, Fang F, Jiang AX, Rajagopal K, Tambe M, Maheswaran RT (2012) Designing better strategies against human adversaries in network security games. In: Proceedings of the 11th international conference on autonomous agents and multiagent systems (AAMAS '12), Valencia, vol 3, pp 1299–1300

Gaming the Attention Economy

Daniel Estrada and Jonathan Lawhead

Natural Human Computation

Human computation (HC) involves the creation of mixed organic-digital systems to solve difficult problems by outsourcing certain computational tasks to the human brain. However, we can distinguish between HC approaches that require a user to engage with a specific (and arbitrary) program or system, and HC approaches that simply leverage a user's normal activity to compute the solutions to complex problems. We call this latter approach *natural human computation* (NHC). An instance of HC is *natural* when the behavior necessary for carrying out the proposed computation is already manifest in the system.

Eusocial insect colonies are models of natural computation (Dorigo et al. 2000; Gordon 2010). The information processing potential of ant colonies emerges from the small-scale, everyday interactions of individual ants: everything individual ants do is computationally significant, both for the management of their own lives and for the colony's success. This alignment between individual and colony-level goals means that ant colonies need not direct the behavior of individual ants through any sort of top-down social engineering. The queen issues no royal decrees; insofar as she has any special control over the success of the colony, that control is a product of her influence on individual colony members with whom she comes into contact. The sophisticated information processing capabilities of the colony as a whole are a product of each ant obeying relatively simple local interaction rules—those local

Our sincere thanks to Pietro Michelucci for his prompt, helpful, and encouraging comments on drafts of this paper. His patience and assistance in the production of this paper has been invaluable.

D. Estrada
University of Illinois, Urbana-Champaign, Champaign, USA
e-mail: djestrada@gmail.com

J. Lawhead (✉)
Columbia University, New York, USA
e-mail: reality.apologist@gmail.com

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_75,

interaction rules, however, allow an aggregate of ants to influence each others' behavior in such a way that together, they are capable of far more complicated computing tasks than individual colony members would be on their own. Crucially, the computational power of the colony *just is* the concerted action of individual ants responding to the behavior of other ants: any change in the colony's behavior will both be a result of and have an impact on the behavior of colony members. In this sense, natural ant behavior is both *stable* and *natural:* the computing activity of the colony can't disrupt the behavior of colony members out of their standard behavior routines, since those standard behavior routines *just are* the computing activity of the colony. The stability of this behavior can in turn support a number of additional ecological functions. The regular harvesting of individual bees not only supports the activity of the hive, but also solves the pollination problem for flowers in what we might call "natural bee computing"[1] which piggybacks on the behavior. NHC approaches take these natural models of computation as the paradigm case, and seek to implement similar patterns in human communities.

We have sketched a definition for NHC in terms of *stable* and *disruptive* computation, and turn now to discuss these concepts directly. Disruptive computation requires a *change* in an agent's behavior in order to make their performance computationally significant. Human computation is increasingly *stable* as its impact on agent behavior is reduced. Describing an instance of human computation as "natural" is not itself a claim that the *human activity* is stable or disruptive, since NHC techniques can be used to extract computationally significant data in both stable and disruptive contexts. Rather, describing an instance of HC as natural makes the more limited claim that the computation in question was not *itself* a source of disruption. We introduce the vocabulary of stability and disruption to clearly articulate this aspect of NHCs.

It may be instructive to compare NHC and gamification (Deterding et al. 2011; McGonigal 2011) as strategies for human computing. Gamification makes an HC task more palatable to users, but often alters user behavior in order to engage with the computational system. In contrast, NHC systems transparently leverage existing behaviors for computation. For instance, reCAPTCHA (von Ahn et al. 2008; von Ahn and Dabbish 2008) repurposes an existing task (solving text-recognition puzzles to gain access to a website) to solve a new problem (digitizing books for online use). This pushes HC to the background; rather than explicitly asking users to participate in the solution of word recognition problems, it piggybacks on existing behavior. Gamification is not always disruptive in the sense used here; in some cases described below gamification techniques can serve to *stabilize* (rather than *disrupt*) the dynamics of systems to which they are applied. This suggests that we need a more robust vocabulary to map the conceptual territory.

[1] Of course, bees and flowers achieved this stable dynamic through millions of years of mutualistic interaction; as we discuss in "Developing the Attention Economy", we expect any HC technique to require some period of adaptation and development.

Table 1 A two-dimensional model of human computation

	Stable	Disruptive
Emergent	American Idol predictions	Yelp
Engineered	Zombies Run	FoldIT

Michelucci (this volume) distinguishes between "emergent human computation" and "engineered human computation." Emergent HC systems analyze uncoordinated behavior from populations to do interesting computational work, while engineered HC systems might be highly designed and coordinated for specific computing needs. We see natural human computation as a concept that is complementary to but distinct from Michelucci's distinction. The defining characteristic of NHC is the potential for extracting additional computational work from human activity without creating additional disturbances in that behavior. This definition makes no assumptions about the degree to which these behaviors have been designed or coordinated for particular computing functions. In fact, we assume that natural human behavior involves organizational dynamics that cut across Michelucci's distinction. NHC systems like Swarm!, described in "Introducing Swarm!" below, can be understood as a method for discerning natural organizational patterns as a potentially fruitful source of human computation.

We're thinking about NHC in terms of the impact a computing task has on the behavior of its computers; NHC tasks introduce minimal disruptions to existing behavior. In contrast, Michelucci's distinction isn't concerned with the impact HC has on its agents. Rather, it is concerned with the performance of the computing task in question. Emergent cases of computing are where the goal is best accomplished by passively analyzing agents for specific computational results, more or less independent of other aspects of their behavior. Engineered systems require increasingly coordinated activity to achieve computational results. For these reasons, we consider Michelucci's distinction to be a system-level or "top-down" perspective on computing tasks, while the stable/disruptive distinction is an agent-level or "bottom-up" perspective on the same tasks. Or to cast the issue in techno-phenomenological terms: Michelucci is taking a designer's perspective on human computing, where purposes (functions, tasks, goals, ends) are *imposed* on a computing population; on the other hand, we're interested in the user's perspective, where the generation and pursuit of purposes is a constitutive aspect of one's ongoing committed engagement with the world.

It is worth reiterating that the sense of "natural" being articulated cuts across the categories represented in Table 1 below. We can think of these categories as defining the axes of a continuous space of possible computing systems. Claiming that a given system is emergent and disruptive (for instance) is to locate within this space. However, claiming that a given instance of human computation is *natural* is to point out a very different sort of fact about the system. In the context of human computation, *naturalness* is something like an indexical, describing words with use-relative content like "here" or "now." Rather than giving an absolute location in the space defined by the distinctions discussed above, calling an instance of HC "natural" is to assert a fact about the HC system *relative* to the current state of the computational substrate. A NHC might be engineered, emergent, disruptive, or stable to some greater or lesser degree; the ascription of naturalness depends only on a comparison between the system's state *now* and the state that would be necessary for performing the desired computation. The distinctions between emergent, engineered, stable, and disruptive HC systems can be more clearly illustrated if we consider a few representative examples. An absolute attribution of naturalness in any of these cases is not possible, as "naturalness" is an index to a user-relative state. As such, the following examples contain no direct appeal to "naturalness", since the degree of naturalness for some HC process may vary between individual users with distinct behavioral profiles. Using Yelp in deciding on some service, or using ZR to motivate your run, will integrate naturally into the usage patterns of some users and may be more disruptive in the lives of others.

Consider the following cases:

Emergent/Stable: HC systems are emergent when they exploit uncoordinated behavior in a population, and they are stable when that computing goal is met without further disruption. reCaptcha has already been mentioned as an example of HC that falls in this quadrant. A more illustrative example can be found in Ciulla et al. (2012), which describes modeling approaches to the Twitter datastream that successfully anticipate the results of a recent American Idol voting contest. In this study, users Tweeted their thoughts on the contest of their own accord,[2] without coordination and independently of their potential use in predictive speculation, and so meets the definition of emergent. Solving the prediction task required no additional input from the users beyond this existing social behavior, and so also meets the definition of stable.

Engineered/Stable: Engineered computing tasks are highly coordinated and designed for specific computing purposes. These designs can be stable in our sense when the computation fits existing patterns of behavior rather than creating new ones. BOINC's successful @HOME distributed computing projects (Anderson 2004) are familiar examples of stable computing strategies, using spare processor cycles for useful computational work without occupying an additional computational footprint. For a more explicitly gamified example, consider the 2012 exercise motivation app called "Zombies Run".[3] Zombies Run (ZR) is designed to work in tandem with a player's existing exercise routine, casting her as a "runner" employed by a post-apocalyptic settlement surrounded by the undead. The game's story is revealed through audio tracks rewarding player for gathering supplies, distracting zombies, and maneuvering through the dangerous post-apocalyptic wasteland, all accomplished by monitoring a few simple features of the user's run.

[2] We ignore for the sake of the example any potential feedback from advertising or other systems that reinforce tweeting behavior surrounding the American Idol event.

[3] From the UK-based Six to Start. https://www.zombiesrungame.com/

The app motivates runners to continue a routine they've already developed, using tools already appropriated in that behavior; the app isn't designed to help users to start running, it is designed to help them *keep* running. This is a defining feature of engineered/stable systems: while they are the product of deliberate design, the design's primary effect is to reinforce (rather than alter) existing patterns of behavior. While ZR players aren't (necessarily) performing any particularly interesting computing function, the app provides a clear example of how a highly designed, immersive app can nevertheless be stably introduced into a user's activity.

Emergent/Disruptive: A computational state is *disruptive* when implementation would involve a significant reorientation of the behavior and/or goals of the agents under consideration. This can occur in emergent computing contexts where individuals are acting independently and arbitrarily. Yelp.com is a popular web-based service that compiles crowd-sourced reviews of local businesses and services. These reviews are used to compute a rating of a given service based on search criteria. And indeed, solving this computing problem itself changes the activity of the population: Luca (2011) finds that the a one-star rating increase amounts to a 5–9 % increase in revenue. In other words, the self-directed, emergent activity of Yelp reviewers is disruptive to the behavior of the dining community, effectively redirecting a portion of them to services with higher ratings. It may be supposed that Yelp's disruptive status is a consequence of feedback from the HC system being used to guide the decisions of future diners. However, Zombies Run provides an example where feedback on HC behaviors can reinforce those behaviors with little disruption. This suggests that Yelp's economic impact involves more than providing feedback on the HC task; it reflects something about the specific computations performed by the system. We will return to this point in "Naturally Optimizing the Economy".

Engineered/Disruptive: FoldIT is a puzzle-solving game in which the puzzles solved by players are isomorphic to protein folding problems (Khatib et al. 2011). FoldIT is a paradigm case of gamification: it makes a HC task more palatable to the users, but significantly disrupts their behavior in the process by demanding their focus on the game. FoldIT is engineered in the sense that the task has been deliberately designed to provide computationally significant results, and disruptive in the sense that the task is a departure from the behavior in which players otherwise engage.

The above examples are offered in the hopes of making clear a complex conceptual landscape that serves as the backdrop for the discussion of natural human computing. A full discussion of the dynamics of purposive human behavior is beyond the scope of this paper, but we understand our contributions here as a step in that direction. Despite the perspectival dimensions of "naturalness," we can talk sensibly about designing natural human computing systems that leverage existing HC work in minimally disruptive ways. We turn now to describe a NHC system that demonstrates these features.

Introducing Swarm!

Swarm!, an application under development for Google Glass,[4] is an implementation of NHC methods for solving a class of economic optimization problems. Swarm! uses the GPS coordinates of players to construct a location-based real time strategy game that users can "play" simply by going about their everyday routines. Individual cognitive systems have limited resources for processing data and must allocate their attention (including their movement through space and time) judiciously under these constraints. Therefore, we can interpret the data gathered by Swarm! as a NHC solution to the task of attention management: Swarm! generates a visualization of aggregate human activity as players negotiate their environments and engage objects in their world (Fig. 1).

The Swarm! engine is designed as a basic NHC application: it's a game that's played just by going about your normal routine, frictionlessly integrating game mechanics into a player's everyday life. Swarm![5] simulates membership in a functioning ant colony, with players assuming the role of distinct castes within one colony or another. Players are responsible for managing their own resources and

Fig. 1 A person wearing Google Glass

[4] Glass is a wearable computer designed and manufactured by Google. The Glass headset features a camera, microphone with voice commands, optical display, and a touch-sensitive interface. It duplicates some limited functions of a modern smartphone, but with a hands-free design. Figure 1 depicts a user wearing a Google Glass unit.

[5] Complete game bible can be found at http://www.CorporationChaos.com.

contributing to the resource management of the colony. Swarm! data is visualized as colorful trails on a map card presented on request to the Glass display in order to engage the resulting behavior. These trails are designed so they cannot be used to locate or track any individual uniquely. Instead, we're interested in the broader patterns of behavior: where do players spend their time? When is a certain park most likely to be visited? When and where do players from two different neighborhoods cross paths most frequently?

Swarm! Mechanics

Ant behavior is coordinated through purely local interactions between individuals and a shared environment without any central direction (Dorigo et al. 2000). Individual ants exchange information primarily through direct physical contact and the creation of pheromone trails. Pheromone trails, which can be used to indicate the location of resources, warn of danger, or request help with a tricky job, are temporary (but persistent) environmental modifications laid down by individual that help ants coordinate with each other and organize over time to manage colony needs.

Swarm! adopts the pheromone trail as its central mechanic. By moving around in physical space, players lay down "trails" that are visible through the in-game interface as colorful lines on a map. These trails encode context-specific information about the history and status of user interactions around a location. Just like real-world ants, Swarm! trails are reinforced by repeated interaction with a region of space, so the saturation of trails in a particular location represents the degree of activity in that location. Trails also encode some information about in-game identity, but the focus of Swarm! is on impersonal aggregate data and not unique player identification. Since trails are semi-persistent and fade slowly with time, the specific time that a player passed a location cannot be deduced by looking at the map. Players also have the option to define "privacy zones" around their homes and other sensitive areas where Swarm! data collection is prohibited.

Swarm! gameplay is styled after many popular resource collection games, with central goals revolving around finding enough food to stay alive, disposing of trash ("midden"), and defending the colony from incursions by rivals. However, Swarm!'s central innovation is its emphasis on self-organized dynamic game maps and frictionless player interaction. Player interactions result primarily from trail crossings: when one player crosses the trail laid down by another player, an appropriate context-dependent event is triggered. Note that this activity does not require players to be present simultaneously at one location. Trails laid down by users decay gradually over time, and require reinforcement to sustain. Thus, crossing the trail of a rival ant means that ant (or possibly several ants from the same colony) have reinforced this trail within the decay period. In other words, all player activity is rendered on the map as "active" and will trigger engagements and events specific to those interactions.

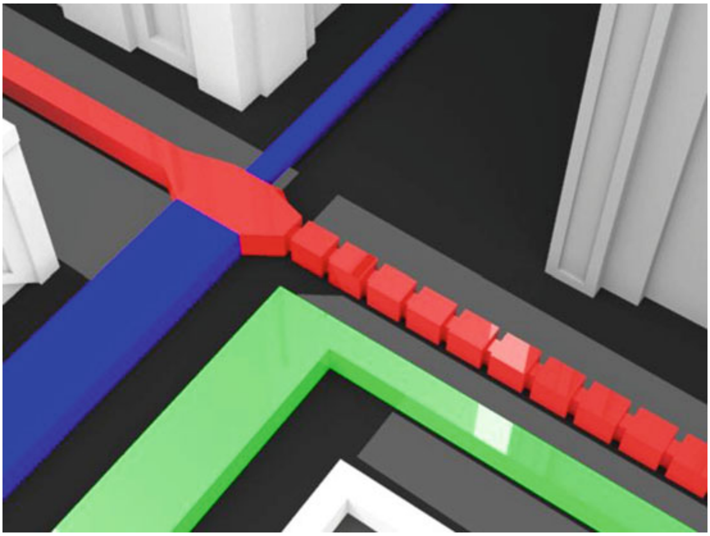


Fig. 2 Our player Eve (indicated by the lower *green trail* that makes a right angle) considers a regular interaction at a busy intersection with a hostile colony (indicated by the bumpy *red trail*), which imposes caste-specific effects on a region (Image credit: Kyle Broom)

Players also have caste-specific abilities to augment the structure of the game map. These abilities are triggered by more in-depth interaction with a location—for instance, spending an extended amount of time in the same public place, or taking some number of pictures of an important game location. Each caste has a unique set of strengths, weaknesses, and abilities that affect the range of in-game options available to the player. These augmentations can provide powerful bonuses to members of a player's colony, hinder the activities of rivals, or alter the availability of resources in the area. Strategic deployment of these abilities is one of the most tactically deep and immersive aspects of Swarm! gameplay.

For illustration, consider the following in-game scenario (Fig. 2). Suppose a player (call her Eve) consistently moves through territory that is controlled by an enemy colony—that is, she crosses a region that is densely saturated with the trails of hostile players. Moving through this region has a significant negative impact on Eve's resource collection rate, and unbeknownst to Eve (who doesn't like to be bothered by game updates) this penalty has been adversely affecting her contributions to her colony for weeks, keeping her at a relatively low level than where she might be otherwise. However, suppose that 1 day Eve decides to actively play Swarm!. Upon downloading the latest game map she observes the impact this region has had on her collection rate. Swarm!'s game mechanics reward this attention to detail, and allow Eve to do something about it. When Eve photographs the locations that are controlled by a rival colony, she creates an in-game tag that calls attention to her predicament and provides caste-specific in-game effects that potentially offset the impact of the rival colony's trail. In other words, her action (taking a picture) has produced an in-game structure that warps the map and partially ameliorates the penalty that she would otherwise suffer. This in-game structure might attract other active players to the territory to build more structures that further magnify these changes. In this way, close attention to (and interaction with) the game map is rewarded, while casual players are still able to contribute meaningfully to the overall game dynamic.

This reveals an important aspect of Swarm! related to the distinctions drawn in "Natural Human Computation". Although the game is designed to passively harvest aggregate user behavior, it also incentivizes the curation of that data allowing for active user engagement. Thus, some users may experience Swarm! as unobtrusive and stable, with computation occurring largely in the background, while others may enjoy significant disruptions as they actively play the game. Moreover, the two might interact with each other through in-game mechanics around shared spaces without either player being aware of the other's presence. When Eve tags a highly trafficked area of the map with her picture, she is highlighting an attractor[6] in *both* the physical space and the game space. Those attractors emerge naturally in the behavior of some Swarm! players, and Eve's active engagement with the trails further augments the map to highlight the relevance of those attractors. These attractors can in turn coordinate others to further document and engage an area, filling out the digital profile of regions that are of central use in human social behaviors, and effectively turning Swarm! players into an engineered team of self-directed, self-organized content curators. Every Swarm! player's behavior is thus influenced both by the structure of the game map, and the structure of the game map is influenced by the behavior of Swarm! players. However, since the initial structure of the Swarm! game map is dictated by the antecedent behavior of Swarm! players, this mechanic only serves to reinforce extant patterns of behavior.

The resulting model highlights patterns of natural human behavior that can be directly harvested for computational work. For instance, consider the problem of locating a working electrical outlet at the airport.[7] Traditional resource distribution structures (like the financial markets or public regulatory structures) have until now failed to provide enough incentive to curate a digital outlet location map for wide public use, despite its potential value to both customers (who may wish to charge their electronics while they wait for a connecting flight), and the airport businesses (who might be able to draw customers and control the flow of airport patrons by advertising their location). Online databases like Yelp work well for services that have existing advocates, like restaurant owners, who can represent those interests by responding and reacting to Yelp reviews, but little incentive exists for a curation task

[6] An *attractor* is just a location or state in a system toward which nearby states or locations tend to be "sucked." Minimum-energy states in mechanical system are commonly attractors. For instance, in a system consisting of a marble confined to the inside of a mixing bowl, the state in which the marble is at rest at the bottom of the bowl is an attractor: no matter where you start the marble, it will eventually end up at rest at the bottom of the bowl. For an accessible introduction to the language of attractors and dynamical systems theory, see Strogatz (2001) and Morrison (2008).

[7] Credit goes to Robert Scoble for raising the example during a recent conversation about Swarm!.

like this. On the other hand, with suitable resolution Swarm! provides an immediate visual representation of the activity of airport patrons that allows for intuitive predictions about where the outlets might be: look for clustering behavior near walls. Moreover, Swarm! rewards active players for tagging public spaces with pictures and notes that fill in details of the interaction history at that location. The result is an NHC method for computing a solution to the problem of finding electrical outlets without the need for natural advocates or market representation to explicitly engineer this behavior.

This example has Swarm! players uncovering the use-value of objects which have been under-represented by other records of social value, and it has accomplished this without creating any additional demand on social behaviors. Perhaps a close analog is the use of GPS requests for identifying traffic congestion (Taylor et al. 2000), but the game mechanics of Swarm! generalizes the approach for a broad range of human activities. We turn now to a general discussion of the strategies described above.

NHC Applications of Swarm!?

Consider the mechanic described in "Swarm! Mechanics" for modifying the game map by taking and tagging pictures. A strategically-minded Swarm! player will not use this ability at just any location (Rashid 2006; Ames and Naaman 2007); rather, she will study the structure of local trails over the course of a few days, and engage with the map in a tactically-optimal location—that is, a location that already experiences heavy traffic of the right sort. In this way, the Swarm! map will become a fairly detailed representation of patterns of player engagement with the real world; locations that are naturally highly trafficked will become increasingly important, and thus increasingly saturated with trails and in-game structures.

The fact that interesting locations in the game tends to mirror the interesting locations in the real world is central to Swarm!'s design. While Swarm!'s mechanics might well have some influence on the behavior of more strategically-minded players, that influence will *remain* a mirror of the aggregate pre-game behavior of the community, and thus a useful starting point for NHC data collection about use behavior. Ingress, a somewhat similar augmented reality game developed by Niantic Labs for Android mobile devices (Hodson 2012), makes for an instructive contrast case. Ingress features two in-game "teams" (Enlightened and Resistance) involved in attempts to capture and maintain control of "portals," which have been seeded by Google at various real-world locations. Players take control of a portal by visiting the location (sometimes in cooperation with other players), and remaining there for a set amount of time. Players may also "attack" portals controlled by the opposing team through a similar location-based mechanic.

Notice the difference between tracking the behavior of Ingress players and tracking the behavior of Swarm! players. Despite both games featuring similar location-based mechanics, the fact that Ingress' portals—the significant in-game attention attractors—have been seeded by the game's designers renders the activity of Ingress players a poor proxy for their natural, out of game behavior, and thus a poor proxy for NHC data collection. In contrast, Swarm! players create the structure of the map themselves, and the strategically optimal approach to modifying it involves reinforcing existing patterns of behavior. The structure of the Swarm! map reveals at a glance sophisticated facts about the natural attention patterns of Swarm! players. It is this fact that makes Swarm! an important first step toward a mature NHC application.

Transitioning Swarm! from a NHC-oriented game to a real NHC application will involve tightly integrating Swarm!'s mechanics with real-world tasks. We suggest that Swarm!'s existing mechanics might be easily tied in to a service like Craigslist.org. Craigslist is a popular and free web-based service facilitating the exchange of good and services that run the gamut from used cars and furniture to prospective romantic encounters—all of which are organized geographically and easily searchable. The Swarm! platform, with its built-in mechanics for tracking location, activity, and experience could serve as a platform for visualizing Craigslist service requests and evaluating the results of the transaction. If successful, such a system would allow for a self-organized, entirely horizontal resource and labor management system for its users. Such integration would be a large step toward turning Swarm! into the sort of robust economic HC application that we discuss in "Developing the Attention Economy".

Consider the following hypothetical future in-game scenario: Eve, our intrepid player from "Swarm! Mechanics", has access to a Craigslist-like service integrated with an advanced version of Swarm!, and this service informs her (on request) about posts made by other players in her immediate geographical region. With access to this information, Eve can decide whether or not to accommodate the requests of other players in her vicinity. Suppose, for instance, that Eve notices a posting near her home base requesting a 40 W CFL light-bulb to replace a bulb that just burned out. Eve was targeted with the request because her patterns of behavior repeatedly cross paths with the requesting user; depending on how sophisticated the service has become, it might even recognize her surplus of light bulbs. In any case, Eve knows that she has several matching bulbs under her kitchen sink, and considers using the bulb to gain experience and influence within Swarm!. Eve notices that the specified drop point is on her way to work, and agrees to drop the bulb by as she walks to the subway. Perhaps the dropoff is coordinated by each party taking a picture of the object using QR codes that signal drop off and receipt of the object. Upon completion, this transaction augments player statistics within Swarm! to reflect the success of the transaction. As a result, Eve's public standing within the player community increases, just as it would have if Eve had participated in a coordinated attempt to seize a food source for her colony. Her increased influence within game environment might increase the chances that her next request for a set of AA batteries is also filled.

This mechanic creates an environment in which contributing to the welfare of other Swarm! players through the redistribution of goods and services is rewarded not monetarily, but through the attraction of attention and the generation of influence and repute. The attention attracted by the request is converted into user experience upon completion of the task, allowing the user's behavior to have a more significant impact on the dynamics of the game. Again, this mechanic helps to blur the line between in-game and out-of-game interactions: the in-game world of Swarm! is a distillation and reflection of the everyday out-of-game world of Swarm!'s players. Eve's history as a Swarm! player disposed to help other players in need might be intuitively presented to other members of her colony through features of her trail. When Eve makes a request for aid other players will be more disposed to respond in kind.[8]

Although our examples have focused on minor transactions of relatively little significance, the game mechanics described here suggest a number of important principles for designing HC systems that harvest the computational dynamics of natural human activity, and the profound impacts these applications might have on a number of vitally important human activities, including education, politics, and urban development. We focus the remaining discussion on economic considerations.

Naturally Optimizing the Economy

We can think of the global economy as being a certain kind of HC system in which the problem being computed involves the search for optimal (or near-optimal)[9] allocations of raw materials, labor, and other finite resources ("the economic optimization problem"). This approach to economic theory is broadly called "computational economics" (see e.g. Velupillai 2000; Norman 1996), and it takes economic theory to be an application of computability theory and game theory. Historically, some economists have argued that a free capitalist market composed of minimally constrained individual agents (and suitable technological conditions supporting their behavior) provides the most efficient possible economic system (Hayek 1948). We shall conclude our paper with a discussion of NHC applications as an alternative approach for tackling the economic optimization problem.

Kocherlakota (1998) argues that money is best thought of as a "primitive form of memory" (*ibid.* p. 2). That is, money is a technological innovation that provides a medium for a limited recording of an agent's history of interactions with other agents. On this view, rather than being an intrinsic store of value or an independent medium of exchange, money is merely a way to record a set of facts about the past.

---

[8] The influence of perceptions of fairness on economic interactions is an increasingly well-studied phenomenon among economists and psychologists. For a comprehensive overview, see Kolm and Ythier (2006), especially Chap. 8 (Fehr and Schmidt).

[9] The definition of "optimal" is disputed, but the discussion here does not turn on the adoption of a particular interpretation. In general, recall that solving the economic optimization problem involves deciding on a distribution of finite resources (labor, natural resources, &c.). Precisely which distribution counts as "optimal" will depend on the prioritization of values. A robust literature on dealing with conflicting (or even incommensurable) values exists. See, for example, Anderson (1995), Chap. 13 of Raz (1988), and Sen (1997).

Kocherlakota argues that this technological role can be subsumed under "memory" in a more general sense, and that while access to money provides opportunities for system behavior that wouldn't exist otherwise, other (more comprehensive) kinds of memory might do all that money does, and more: "…in at least some environments, memory [in the form of high quality information storage and access] may technologically dominate money" (*ibid.* p. 27).

If this characterization is correct, then solving the economic optimization problem involves accomplishing two distinct tasks: identifying precisely *what* information should be recorded in economic memory, and we must devising ways to store and manipulate that information. We might understand Yelp as recording user accounts of a service that attempts to meet these memory challenges. Yelp users leave comments, reviews, and ratings that provide a far more detailed and relevant transaction history with customers than is represented by the relative wealth of the business as a market agent. Luca (2011) finds not only that these reviews have an impact on revenue, but that impact is strengthened with the information content of the reviews, suggesting one place where money may be showing evidence of domination by rich sources of memory.

Swarm! offers a natural approach for meeting the same challenges, in which NHC is leveraged to help solve the economic optimization problem without introducing new economic frictions. This computational work is accomplished through the recording of trails that represents incremental changes in the use history of that location. As Swarm! maps become increasingly detailed and populated they likewise come to function as an effective representation of the attention economy (Simon 1971; Weng et al. 2012) in which the saturation of trails around an object approximates a quantitative measure of the value of objects relative to their use.[10] We treat this measure as the aggregate "use-value" of the object (Vargo et al. 2008), and argue that a model of the use-value of objects allows for novel NHC-based solutions to a variety of standard problems in the optimization of economic systems. A full articulation of the attention economy is not possible here, but we will provide a sketch of one possible implementation using the Swarm! framework.

Developing the Attention Economy

Recall the central mechanic of Swarm!. GPS data about players' movement patterns are aggregated, whether or not a player is actively engaged with the game. Strategically-minded players are rewarded for tagging and modifying the map in a way that gives rise to a detailed reflection of how all Swarm! players use the space covered by the map. The data collected by a Swarm!-like application has the potential to encode many of the facts that might otherwise be encoded less explicitly. Monetary transaction records act as proxy recordings for what we have called

[10] As opposed to value relative to *exchange*. See Marx (1859).

use-value. The mechanics of Swarm! suggest a way to measure use-value directly by recording how economic agents move through space, how their movement is related to the movement of others, what objects they interact with, the length and circumstances of those interactions, and so on. By tracking this data, we can transform the everyday activities of agents into records of what those agents value and to what degree. This is the "high quality information storage and access" that Kocherlakota suggests may come to technologically dominate currency as economic memory. Still, a number of practical challenges must be surmounted before a NHC/AE based approach to solving the economic optimization problem is realistically viable.

Any implementation of an attention economy in which the economic optimization problem is solved with NHC will clearly involve data collection on a scale that goes far beyond what's possible in Swarm! or with Google Glass, as the mere tracking of gross geospatial position will not record enough information to (for instance) assay the value of individual material objects like pens and lightbulbs. Swarm! is an incremental step in that direction, with the more modest and technologically feasible goals of acclimating people to regular engagement with AE platforms, and with developing the social norms appropriate to the demands of an AE. The structure of human communities is strongly coupled to the technology available during their development. Absent major catastrophes, the sort of ubiquitous computing and social norms necessary for the implementation of an AE will continue to develop in tandem.

Indeed, the success of AE in some sense depends on the development of social customs and attitudes to compensate for the more invasive social coordination technologies that dominated the Industrial Age, which are almost universally characterized by the establishment of hierarchical institutions of control. In such a system, power is concentrated in the hands of the very few, to be executed within very narrow channels of operation. For the disenfranchised, finding ways to circumvent or usurp this power is often a more attractive than accumulating power through so-called "legitimate" means—especially as the powerful increasingly protect their positions through deliberate corruption and abuse, thereby weighting the system heavily against "fair play". In other words, enterprising opportunists looking for success in systems with limited hierarchical control have a disproportionate incentive to "game the system", or exploit loopholes in the rules in ways that give them a disproportionate advantage. Preventing the exploitation of such loopholes requires an ever increasing concentration of power, creating greater incentives to break the system, and greater costs for failing in those attempts. Social customs discouraging such behavior must be imposed from the top, often with violence, as a means of retaining control, since these customs are not reinforced from below.

In contrast, the AE describes a self-organizing system without hierarchical control or concentrations of power, because the rules for operating within the system also support the success of the system as a whole, and so are supported from the bottom without need for top-down enforcement. In other words, the impulse to game an attention economy can be actively encouraged by all parties, since individual attempts to gain a disproportionate advantage within the system simultaneously reinforce the success of the system overall. Recall from "Swarm! Mechanics", when Eve snaps a picture of a highly trafficked block. This apparently self-interested act to improve her own in-game resource collection rate is simultaneously a contribution to the economic optimization problem, and is therefore reinforced by her colony's goals. Of course, Eve is not only rewarded by pursuing self-interested goals; potentially everything Eve does in an attention economy is computationally significant for her community, and therefore her community can support Eve in the pursuit of any goals she wishes without worrying about how her actions might upset the delicate balance of power that supports institutional control. In an attention economy, Eve is not rewarded to the extent that she appeals to existing centers of power; instead, she is rewarded to the extent that her participation has an impact on the development of her community (see also, Rashid et al. 2006).

We conclude by mentioning some design considerations inspired by Swarm! for building an "Internet of Things" that facilitates the use of NHCs for managing the attention economy. Most obviously, Swarm! is a step toward the creation of pervasive, universally accessible, comprehensive record of the relationship between agents, locations, and objects. As we have said, widespread location and identity tracking of at least *some* sort is essential for the implementation of a true AE. This is a major design challenge in at least two senses: it is a technical engineering challenge, and a social engineering challenge.

The solution to the first challenge will still require technological progress; we do not yet have ubiquitous distribution of the sort of computing devices that would be necessary to implement the fine-grained level of data collection that a real AE would require. In addition to aggregate movement patterns, an AE platform will need to track patterns in the relationships between agents and physical objects. Sterling (2005) introduces the term "spime" to refer to inanimate objects that are trackable in space and time, and broadcast this data throughout their lifetimes. Objects that are designed to live in an attention economy must track more than just their own location and history: they must be able to track their own use conditions, and change state when those use conditions have been met. This will require objects to be sensitive not just to their own internal states, but also to the states of the objects (and agents) around them: this is the so-called "Internet of Things" (Atzori et al. 2010). There is already some precedent for very primitive functions of this sort. Consider, for instance, the fact that modern high-end televisions often feature embedded optical sensors to detect ambient light levels, and adjust backlighting accordingly for optimal picture quality. We can imagine expanding and improving on that kind of functionality to develop (say) a television that mutes itself when the telephone rings, pauses when you leave the room, or turns itself off when a user engages deeply with another object (for instance a laptop computer) that's also in the room. These examples are relatively mundane, but they are suggestive of the sort of industrial design creativity and integration needed to design AE-optimized artifacts.

Swarm! approaches this design challenge by imposing some novel clustering methods represented by the caste and colony system. The colony system is a geographical constraint designed to cluster colony members to ensure that they aren't spread so thin as to undermine the game dynamics. The caste system is a design constraint on the patterns of user activity, and allows users to tell at a glance the functional results of some possible sequence of engagements without knowing too many details about other players. This latter feature is inspired directly by ant colonies, and is important to the organizational dynamics of an AE. In particular, it gives contexts in which it is appropriate for certain agents to have disproportionate influence on some computing task, thereby carving out emergent hierarchies and cliques. The AE/NHC platform is thus applicable to the solution of non-economic social problems, and can be leveraged to help compute solutions to other legal, political, and social puzzles.

As an illustration of how NHCs might be applied to the distribution and management of resources and labor, consider the transaction history for some arbitrary object X. If this record has been reliably maintained on a user-per user basis, it might serve as the basis for resolving disputes about ownership, rights of use, and other coordination problems traditionally settled by legal and political frameworks. If I have years of history driving a specific car on Wednesday mornings, and the use record shows you driving this car some particular Wednesday morning, then absent some explanation this appears to be a disturbance in use patterns. This information might itself be enough to warrant a complaint through official channels and initiate the machinery of the public justice system to account for this disturbance. In other words, a well-maintained record of the use history of an object might serve as a foundation for NHC solutions to political and legal disputes, and provides a framework for dealing naturally with apparent cases of "stealing" without requiring anything like the disruptive technologies of property, contracts, and other legal frictions.

This is the real heart of the AE/NHC approach to economic optimization: the NHC acts entirely upon data about local patterns of attention, use, and interaction without significantly disturbing the behavioral patterns that generate the data. Rather than indirectly recording facts about my contribution to (or value of) some object or process in monetary memory, which requires its own set of social conventions and techniques to maintain, those facts are recorded *directly* in the history of my relationship to the object or process. We suggest that careful management of those facts, combined with a distributed NHC framework, might allow for a far more efficient economic system than any money-based system.

We've given a characterization of the shape and character of the first of the two design challenges we mentioned above: the technical engineering challenge. While solving this challenge is central to the implementation of the AE, we should not overlook the importance of solving the second challenge either. While technological advances are important, so are advances in the relationship between humans, technology, and society at large. Just as the dissemination of other major, epoch-defining technologies (like the automobile or the telephone) were accompanied by a certain degree of widespread anxiety and social disruption, we expect that the adoption of the ubiquitous computing platforms required for AE implementation (and their concomitant changes in social practice) will be associated with some unrest as society acclimates to some of the underlying changes. In this respect, Swarm! is more than just an experiment in designing a NHC application—it is an attempt to give society at large a chance to experience the artifacts and socio-cultural practices required for a well-managed AE. The more time we have to grapple with those issues as a community, the smoother the transition to the future will be.

References

Ames M, Naaman M (2007) Why we tag: motivations for annotation in mobile and online media. In: Proceedings of the SIGCHI conference on human factors in computing systems, ACM, pp 971–980

Anderson E (1995) Value in ethics and economics. Harvard University Press, Cambridge

Anderson DP (2004) Boinc: a system for public-resource computing and storage. In: Grid computing, 2004. Proceedings of Fifth IEEE/ACM international workshop. IEEE, pp 4–10

Atzori L, Iera A, Morabito G (2010) The internet of things: a survey. Comput Netw 54(15): 2787–2805

Ciulla F, Mocanu D, Baronchelli A, Gonçalves B, Perra N, Vespignani A (2012) Beating the news using social media: the case study of American Idol. EPJ Data Sci 1(8):1–11

Deterding S, Sicart M, Nacke L, O'Hara K, Dixon D (2011) Gamification. Using game-design elements in non-gaming contexts. In: PART 2—Proceedings of the 2011 annual conference extended abstracts on human factors in computing systems. ACM, pp 2425–2428

Dorigo M, Bonabeau E, Theraulaz G (2000) Ant algorithms and stigmergy. Future Gen Comput Syst 16(8):851–871

Fehr E, Schmidt K (2006) The economics of fairness, reciprocity and altruism—experimental evidence and new theories. In: Kolm S, Ythier J (eds) The handbook of the economics of giving, altruism and reciprocity, vol 1. Elsevier, London, pp 616–691

Gordon DM (2010) Ant encounters: interaction networks and colony behavior. Princeton University Press, Princeton

Greene K, Thomsen D, Michelucci P (2012) Massively collaborative problem solving: new security solutions and new security risks. Secur Inform 1(1):1–17

Hayek F (1948) Individualism and economic order. The University of Chicago Press, Chicago

Hodson H (2012) Google's Ingress game is a gold mine for augmented reality. New Sci 216(2893):19

Khatib F, Cooper S, Tyka MD, Xu K, Makedon I, Popović Z, Baker D, Players F (2011) Algorithm discovery by protein folding game players. Proc Natl Acad Sci 108(47):18949–18953

Kocherlakota N (1998) Money is memory. J Econ Theory 81(2):232–251

Luca M (2011) Reviews, reputation, and revenue: the case of Yelp.com (no 12–016). Harvard Business School, Cambridge

Marx K (1859) A contribution to the critique of political economy, International Publishers, New York, 1979

McGonigal J (2011) Reality is broken: why games make us better and how they can change the world. Penguin Books, New York

Morrison F (2008) The art of modeling dynamic systems: forecasting for chaos, randomness and determinism. Dover Publications, New York

Norman A (1996) Computability, complexity, and economics. Comput Econ 7(1):1–21

Rashid AM, Ling K, Tassone RD, Resnick P, Kraut R, Riedl J (2006) Motivating participation by displaying the value of contribution. In: Proceedings of the SIGCHI conference on human factors in computing systems, ACM, pp 955–958

Raz J (1988) The morality of freedom. Oxford University Press

Sen A (1997) Maximization and the act of choice. Econometrica 65(4):745–779

Simon H (1971) Designing organizations for an information-rich world. In: Greenberger M (ed) Computers, communication, and the public interest. The Johns Hopkins Press, Baltimore, pp 37–52

Sterling B (2005) Shaping things. MIT Press, Cambridge

Strogatz S (2001) Nonlinear dynamics and Chaos: with applications to physics, biology, chemistry, and engineering. Westview Press

Taylor MA, Woolley JE, Zito R (2000) Integration of the global positioning system and geographical information systems for traffic congestion studies. Transport Res Part C Emerg Technol 8(1):257–285

Vargo SL, Maglio PP, Akaka MA (2008) On value and value co-creation: a service systems and service logic perspective. Eur Manage J 26(3):145–152
Velupillai K (2000) Computable economics. Oxford University Press, New York
Von Ahn L, Dabbish L (2008) Designing games with a purpose. Commun ACM 51(8):58–67
Von Ahn L, Maurer B, McMillen C, Abraham D, Blum M (2008) reCAPTCHA: human-based character recognition via web security measures. Science 321(5895):1465–1468
Weng L, Flammini A, Vespignani A, Menczer F (2012) Competition among memes in a world with limited attention. Sci Rep 2:335

Human Cumulative Cultural Evolution as a Form of Distributed Computation

Paul E. Smaldino and Peter J. Richerson

Introduction

This chapter, like most if not all the other chapters in this book, was written and edited on a digital computer. That computer can perform incredible feats of numerical computation at blindingly fast speeds, store massive amounts of data, and be used as a tool for everything from writing to music production to scientific analysis to communication. The abilities of a digital computer, however, are insignificant next to the computational power of the network of human beings, their communication infrastructure, and the accumulated knowledge tapped into by those individuals responsible for building it. No single human being knows how to build a modern computer from scratch. Indeed, no one knows how to build a computer mouse, or a lead pencil, or many of the complex tools we rely on for modern living from scratch (Read 1958; Ridley 2010). For that matter, hardly any of us know how to make simple two-strand twisted string from local raw materials, something that practically every adult once knew how to do. A key factor that enables us as human beings to solve complex problems and achieve a level of dominance over a wide variety of environments from the desert to the arctic to the deep ocean is not simply our individual big brains, but our capacities for extreme sociality, to cooperate and learn from one another, and our ability to build on previous knowledge and to accumulate culture.

It is often said that the human brain is like a computer. It processes information, takes in input, produces output, stores and retrieves memory. Groups of people,

P.E. Smaldino (✉)
Center for Advanced Modeling in the Social, Behavioral, and Health Sciences, Johns Hopkins University, Baltimore, MD, USA
e-mail: paul.smaldino@gmail.com

P.J. Richerson
Department of Environmental Science and Policy, University of California, Davis, Davis, CA, USA
e-mail: pjricherson@ucdavis.edu

P. Michelucci (ed.), *Handbook of Human Computation*, 979
DOI 10.1007/978-1-4614-8806-4_76,

then, are like supercomputers. They can process in parallel, allocate resources, and divide tasks to produce faster and better solutions to problems than lone individuals can manage, and that are more than the sum of the individuals' abilities were they working separately (Smaldino in press). For example, Woolley and colleagues presented small groups with a number of tasks requiring different types of collaborations to solve them (Woolley et al. 2010; see also Woolley and Hashmi this volume). Their results showed, firstly, that groups' performance between tasks were correlated, pointing to an emergent "collective intelligence" for each group configuration, and secondly, that a group's performance was uncorrelated with the intelligence of its individual members, but rather stemmed from their ability to communicate in a understanding and democratic fashion.

If groups of humans are like a supercomputer, then what of human *cultures*, which store and process the cumulative innovations and collaborations of generations of individuals? Cumulative culture allows human societies to act as *super-duper computers*. Humans are unique in the animal kingdom for our tremendous capacity to learn from one another. Our relatively fast and accurate imitation and willingness to teach others allows us to acquire complex skills without having to reinvent them for ourselves. Individuals sometimes improve upon the skills they have acquired and these improvements can be passed on to those who learn from the inventor. Furthermore, we are smart shoppers in the marketplace of ideas. We selectively adopt innovations from others that work better or whose use is correlated with success. Human social learning cumulatively ratchets up technology and innovation, providing groups with progressively better solutions to the problems they encounter (Tennie et al. 2009; Boyd et al. 2011). By such means Stone Age bowyers produced bows that modern engineers find to be approximately optimal designs (Allely et al. 1992). From this perspective, the human-based genetic algorithm (Kosorukoff 2001; Grier, this volume), a computational technique in which human users are involved in both judging the fitness of problem solutions as well as suggesting novel solutions, is simply an application of evolutionary processes that have been driving human innovation since the ancient hominins began to make multi-part tools.

In this chapter, we will first discuss how social learning can increase the fitness of a population by allowing cultural innovations to accumulate. We will then discuss the importance of population size and social connectivity on maintaining those innovations, with a focus on the fragility of human computational systems to sudden isolation or population loss. We end with a consideration of the implications of our discussion on the design of human computation systems in the future.

Roger's Paradox: Why Social Learning Is Not Enough

In a complex world in which decisions must often be made quickly and in which skills may be difficult to acquire, individual trial-and-error learning can be overly costly in terms of time, cognitive capacity, and the potential consequences of a poor decision. A hunter-gatherer learning on his own may spend months trying to construct a hunting apparatus or, much worse, misread animal tracks and be eaten by a

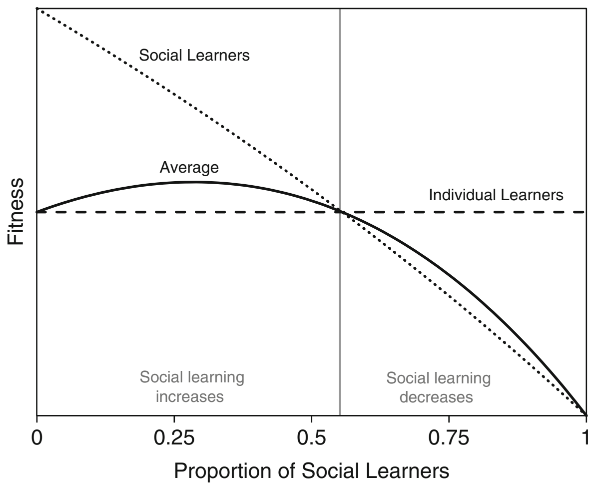


Fig. 1 Rogers' model. Social learners have higher fitness than individual learners when rare, and their invasion briefly increases the average fitness of the population. However, as the proportion of social learners in the population increases, the population stabilizes at a mixed equilibrium (*vertical line*) in which the average fitness is identical to that of a population of individual learners (Adapted from Rogers (1988))

predator. Social learning—which includes knowledge or behaviors acquired via teaching, imitation, or social influence (e.g., you learn to play video games because your friends all hang out and play Xbox)—helps individuals to gain important skills without the time costs and risks associated with individual, trial-and-error learning. It may therefore seem obvious that the adoption of a strategy of social learning should benefit a population—increasing its fitness, in the language of evolutionary biology. Yet, it turns out that if the only benefit of social learning is to avoid the costs of trial-and-error, then social learning strategies will be indeed adopted but will *not* increase the fitness of the population.

This apparent paradox was demonstrated by Rogers (1988) via a simple mathematical model. Suppose a hypothetical world in which there are two possible behaviors, either of which individuals can adopt to help them survive and reproduce. Suppose also that the environment changes periodically between two states, and that in each of these states a different behavior yields a fitness advantage over the other behavior. Finally, suppose that individuals in this world fall into two categories of learner, and that learning strategies are passed on from parents to their children. *Individual learners* always learn the optimal behavior for the environment, but at a heavy cost. *Social learners* choose an individual from the previous generation at random and adopt that individual's behavior. Because they simply copy the behavior of another individual, social learners avoid the cost of individual learning.

When social learners are rare, they will have higher fitness than individual learners, and their prevalence in the population will therefore increase. However, as social learners become more common, a social learner becomes increasingly likely to learn her behavior from a fellow social learner. The state of the environment is not stable, so social learners risk learning an out-of-date behavior, and transmitting that incorrect behavior to social learners in subsequent generations. Thus, as the prevalence of social learners increases, their fitness begins to fall, until it again reaches the fitness of individual learners (Fig. 1). In other words, we should expect social learning strategies to evolve in a social species, but the introduction of social learning behavior does not, in and of itself, increase the average fitness of populations of that species. Boyd and Richerson (1995) studied variations on Rogers' model, including ones in which social learners could identify and preferentially learn from individual learners and ones in which there were more than two behaviors. They showed that Rogers' results were robust: social learning on its own does not increase the mean fitness of a population.

Cumulative Culture

One way of getting around Roger's paradox is if social learning can improve the efficiency of individual learning. For example, suppose individuals sample multiple behaviors, and choose one if it is obviously better than the others. If there is an insufficiently clear signal, and one behavior is not obviously superior, then the individual chooses a random behavior through imitation. This strategy of "conditional social learning" has been shown to increase the mean fitness of the population relative to a population of individual learners under a variety of conditions (Boyd and Richerson 1995; Enquist et al. 2007; Ehn and Laland 2012). In this and related scenarios, behaviors can always be hypothetically learned by individuals through trial and error. Social learning can be adaptive if it hastens the spread of beneficial behaviors. Indeed, this seems to be the primary benefit of social learning in non-human apes (Tennie et al. 2009). Nonetheless, human groups have utilized social learning to make extremely large adaptive gains, much larger than would be facilitated solely through the increased spread of beneficial behaviors. *The key to human success is the spread of incremental innovations based on existing behaviors—often quite complex behaviors—which naïve individuals would be unable to learn on their own*. Human social learning based on teaching and imitation is so efficient that a Stone Age bowyer could acquire an advanced technology like a bow that already incorporated the hard-won innovations of dozens if not hundreds of his ancestors before he contemplated innovations that might improve it still further.

Humans are the most successful vertebrate species on the planet. We have managed to conquer a vast range of environments from the desert to the tropics to the Arctic. We have modified our environments to facilitate our survival and our expansion. At the dawn of the agricultural revolution 10,000 years ago, the entire human population across the globe was approximately five million (Keinan and Clark 2012). Today, over 37 million people live in the Tokyo metropolitan area alone. We have achieved this success through the process of *cumulative culture*. By this, we mean that learned information and behaviors are reliably transmitted and improved upon, such that those improvements can in turn be transmitted through learning.

Cumulative culture allows individuals to build on previous adaptations. To exploit this, humans have evolved psychological mechanisms, heuristics, and biases that facilitate the acquisition of useful knowledge and behaviors (Henrich and McElreath 2003; Tomasello et al. 2005; Herrmann et al. 2007; Hill et al. 2009). This includes learning from individuals who have demonstrated success either directly or by proxy, indicated by traits like status or prestige (Henrich and Gil-White 2001). Symbolic communication additionally allows for transmission of behaviors and practices without direct observation. Stories, myths, and moral doctrines play an important role in the cultural transmission of norms of social behavior as well as useful information about survival (Chudek and Henrich 2011). Moreover, the spread of organizational norms through narratives, religions, and other social institutions have enabled the creation and transmission of emergent group-level traits that rely on social organization and division of labor (Henrich and Boyd 2008; Smaldino in press). The capacities of human populations for the storage and transmission of collective information constitute a feat of evolutionary computing unmatched in the natural world.

Why Culture Is Common but Cultural Evolution Is Rare

Since cumulative culture has made our species so outstandingly successful, why didn't this capacity evolve long ago and lead to many highly cultural species? After all, most "killer" adaptations like internal skeletons, camera-type eyes, and efficient flight have been around for hundreds of millions of years and characterize many lineages. At least two answers are possible. First, cumulative culture would be very difficult to get started if the capacity for it is costly, as our large brains suggest it is (Aiello and Wheeler 1995). The problem is that it takes many large brains operating over many generations to evolve complex cultural adaptations. The first individuals to pay the cost for a capacity for cumulative culture would find no useful complex traits to imitate and hence would get no fitness payoff to cover the overhead of the capacity (Boyd and Richerson 1996). Second, the kind of environment that makes complex cumulative culture useful may be of very recent vintage. Theoretical models suggest that cumulative culture is most useful in moderately variable environments, especially environments that vary on time scales too short for genes to track. Cultural evolution is rapid compared to genetic evolution and can thus generate adaptations to more ephemeral environmental changes than can genes (Perreault 2012). We have known for decades that the earth's climates became much more variable during the Plio-Pleistocene—from 5.3 million years ago to about 11.5 thousand years ago—than they were during the preceding 60 million years (Zachos et al. 2001). Until recently, however, the best data described climate variation on time scales too long to favor costly culture. As better paleoclimate data has come available, we have begun to resolve climate variation at the millennial and submillennial scale that in theory should favor a capacity for cumulative culture. This variation appears to have been increasing over the last few glacial cycles (Loulergue et al. 2008), in rough parallel to human brain size increase and the increasing sophistication of stone tools. In other words, paleoclimate data for the last several hundred thousand years shows a steady increase in environmental variability that appears to be tracked by the emergence and spread of cumulative culture.

Innovation, Environment, and Population Size

The ability to socially learn and accumulate culture provides opportunities for the expansion of humans' ability to problem-solve. However, the ability to learn and adapt does not necessarily lead to the runaway growth of cultural innovation. As a prime example, anatomically modern humans first appeared in Africa between 160 and 200 kya (kya = thousand years ago; McBrearty and Brooks 2000; White et al. 2003; McDougall et al. 2005), and had spread to most of the habitable parts of the globe between 90 and 40 kya (Ambrose 1998; Ray et al. 2005). Yet the appearance of agriculture and the comparatively rapid growth of culture, technology, and population size only occurred around 11 kya. Why is this the case? It seems unlikely that it took somewhere between 30 and 150,000 years for humans to happen upon the idea of domesticating crops and adopting stationary (as opposed to nomadic) lifestyles. Richerson et al. (2001) have argued that the drier, highly variable, low CO_2 world of the last glacial period would have been unfavorable for the evolution of agriculture until about 11,000 years ago, which is in fact when agriculture began to develop. It is true that anatomically modern humans were present in Africa during the last interglacial period without developing agriculture, and our understanding of events in Africa leading up to anatomical moderns spreading out of Africa around 50 kya is still rudimentary (Richerson et al. 2009). Nevertheless, the case is strong that it was the right mix of biological and cultural preparedness and worldwide climatic factors that triggered the widespread adoption of agriculture, sowing the seeds of modern civilization.

Even in the relatively stable global climate of the last 11,000 years, cultural innovations have required the right social and environmental circumstances to thrive and evolve. One of the best markers of cumulative cultural evolution is the presence of complex technology. Before most of the world was connected by webs of communication and commerce, tremendous variability could be found in the complexity of each culture's toolkits. What factors determine the limits of a cultural population's technological complexity?

Population Size and Connectedness Predict Technological Complexity

Cumulative culture allows for complex technologies to be maintained and transmitted across generations. As a result, humans have developed technologies that have allowed them to survive and flourish in wide ranges of environments. Spears and kayaks are useful for fishing in marine environments, but would be quite difficult for naïve individuals to make and exploit on their own. Inuit populations living in Arctic climates learned to make warm skin clothing, build sleds, and breed and train dogs for sled travel. All of these cultural adaptations were essential for their long-term survival in the icy climate of the frozen North. The creation of all of these technologies depends on the ability to transmit and maintain complex and cumulative information and behaviors. However, learning complex behaviors and technologies is not easy. For example, not everyone is a great model for learning. In a generation, there may be only one or two great kayak builders. Moreover, not everyone has the ability or the proclivity to learn complex skills. From among many students, a gifted teacher may have only a few gifted pupils who can go on to teach others to the same high standard. How might these factors influence the maintenance and transmission of complex cultural adaptations?

Using a mathematical model, Henrich (2004) showed that if (1) there is individual variation in learning ability, and (2) students are usually less skilled than their teachers, then more complex skills will require larger population sizes in order to be maintained. Powell et al. (2009) then extended Henrich's model and showed that it is not necessarily the absolute size of the population that matters, but rather the number of effective teachers available. Thus, contact with other groups can compensate for a given group's small population size. On the other hand, this theory implies that smaller and more isolated groups should have less complex technology than larger and more connected groups. Kline and Boyd (2010) analyzed fishing and marine foraging toolkits from ten small-scale societies in Oceania and found that, as the models predicted, population size was the best predictor of toolkit complexity, and also that higher rates of contact with other groups were associated with greater toolkit complexity, especially in relatively small groups.

When Disaster Strikes

If larger and more connected populations are associated with increased technological complexity, what happens if there is a catastrophic event that suddenly shrinks the population or isolates a group from outside contact? Several documented cases suggest that this can lead to a loss of previously held technologies (Boyd et al. 2011). A well-known example is the case of the Tasmanians (Diamond 1978; Henrich 2004; Davidson and Roberts 2009). Isolated from mainland Australia after the seas began to rise at the end of the last glacial period, humans on Tasmania were stranded for about 8,000 years on an island that could not sustain more than a few thousand people, and remained isolated from the rest of the world until their first contact with Europeans in the late eighteenth century (Pardoe et al. 1991). Over the next several thousands of years following their isolation, the Tasmanians lost a number of previously held technologies, such as bone tools and fishing hooks and the ability to make cold-weather clothing, which were nevertheless maintained in Aboriginal communities on the Australian mainland. Moreover, the archaeological record on Tasmania points to a gradual loss of technology following their isolation. For example, the record indicates that between 8,000 and 5,000 years ago the Tasmanian diet was heavily dependent on fish. However, the presence of fish in their diet (as seen in the archaeological record) was declining by 5,000 years ago and had completely disappeared by 3,800 years before present, even though the relative proportions of other elements in the Tasmanians' diet did not shift much. Other technologies that appeared on the mainland after the Tasmanian separation may have never arisen on the island at all. For example, boomerang use was widespread on the mainland by the time of European contact, and the remains of boomerangs dating back to over 11,000 years ago have been found in a peat bog in South Australia, but none have ever been found in Tasmania (Davidson and Roberts 2009).

Technological Complexity in the Modern World

The past two centuries have seen exponential growth in both the size of the human population on Earth and the complexity of our technology. Each new technological innovation has built upon previous or contemporary technologies, and those innovations have spread with startling speed. Consider that a person alive today born 100 years ago would have witnessed the inventions of—just to name a few—the television, the transistor radio, modern plastics, the jet airplane, the electric guitar, the microwave, the credit card, the remote control, the compact disc, personal computers, cell phones, the internet, GPS, DNA fingerprinting, Prozac, Viagra, gene sequencing, smart phones, and unmanned drone aircraft. Consider also that in that time, the world population has not only quadrupled, but has become vastly more connected. Three-fourths of the world's people now have access to cell phones (World Bank 2012). Social media such as Facebook, Twitter, and Google + allow information, ideas, and norms to spread faster and wider than ever before. The research university and other formal educational and R&D organizations spread technical information to ever larger numbers of students and support ever larger numbers of research scientists, design engineers, manufacturing specialists and maintenance technicians.

Access to advanced technologies is extremely widespread. However, the knowledge and ability to create innovative new technologies rests in the hands of rather few individuals in each separate field. Moreover, many technologies require complicated collaborations between individuals with different skill sets and access to resources—a prerequisite fulfilled by large numbers of people with access to communication networks and substantial wealth. As our population grew and became more interconnected, our capacities to invent and sustain complex technologies increased. However, the maintenance of those technologies is not guaranteed. For example, it may be that the number of highly trained engineers necessary to sustain (much less advance) modern highly complex technology is quite large. If the global human computer were to suffer some kind of setback (due to a political, economic, or environmental shock), the loss of individual experts or the reduced communication between them might further exacerbate the original setback, much as a shrinkage of population size and/or contact caused the loss of complex technology on Tasmania and on remote Pacific islands. As an example from a complex society, consider also the setbacks to technology and knowledge in the former Western Roman Empire after its political collapse. Depopulation and the fragmentation of the formerly unified polity and economy resulted in declines in literacy and the hollowing out of civil engineering skills so advanced by the Romans. Old buildings and roads fell into ruin and new construction did not recover to Roman standards until the Late Medieval period.

How Fragile Is the Global Human Computer?

Could a disaster characterized by a sudden loss of either population or connectedness really lead to a global loss of technology? Maybe not. The world is extremely well connected. This may create redundancies and plasticity that could prevent loss of technology, provided the damage was not too extensive. Consider an analogy to the human brain. Damage to the left hemisphere of the cerebral cortex, such as caused by a stroke or traumatic brain injury, can cause difficulties in the production or comprehension of language (aphasia). However, if a patient speaks more than one language, she may show differential patterns of damage and recovery between languages, making the patient more likely to regain or retain language use in at least one language (Goral et al. 2002). Increased social connections whereby similar technologies are produced via slightly different pathways may increase the robustness of complex technologies. We also have large information reserves in the forms of books and internet databases. Although the Romans had books, the printing press had not yet been invented, and as such dissemination of information was limited. Merlin Donald (1991) has argued that literacy and numeracy, leading to the external storage of information, is one of the great advances (of three) in the origin of the modern mind. The unaided human brain has a limited memory and a limited ability to handle quantitative calculations. Literacy and numeracy not only relieve these limitations but also increase connectedness via books and other forms of written communication. Mass literacy and inexpensive mass media following the invention of the printing press greatly multiplied the number of people who could participate in advancing and spreading technology and other innovative ideas.

Nevertheless, it is possible that technology may be lost should disaster strike. Much of our specialized knowledge is collected by institutions, and that knowledge could rapidly vanish. Skilled people can die, books can be burned, and computers can wear out. Cumulative culture creates infrastructures that facilitate the persistence and growth of technologies and innovations. The maintenance of modern medicine, for example, leans on the infrastructure of the medical school system as much as it does on the availability of information. If young would-be doctors didn't have anywhere to train, it would be difficult for them to become as skilled as today's highly trained doctors regardless of the persistence of medical textbooks. If the population necessary to maintain a particular technology were to suffer a loss, it is not clear that recovery would be swift.

Even if a technology *is* lost, knowledge of the technology's existence and its associated benefits may help to recover it. The archaeological evidence shows that for thousands of years following their separation from the mainland, the Tasmanians'

Table 1 Culturally evolved mechanisms for improving the efficacy of cumulative culture

Mechanism	Examples
Development of off-line storage of information	Literacy, numeracy, and cheaper media such as clay tablets, paper, and electronic storage
Improvements in the dissemination of information	Cheap printing, lending libraries, internet search engines
Improvements in the capacity of individuals to innovate and the number of individuals prepared to learn	Mass education, specialized scientific and technical occupations, specialist textbooks and journals
Development of institutions designed to favor innovation	Intellectual property rights, research universities, lavish government support for basic and applied research, "Silicon Valley culture"

diet involved large quantities of fish. When European's visited the island, however, the islanders were astonished at their success in pulling fish out of the ocean, but refused offers of fish to eat (Davidson and Roberts 2009). This taboo on eating fish almost certainly arose after the islanders stopped fishing, along with their loss of the technology to make bone fishing hooks. Importantly, the taboo would have also dissuaded potential innovators from re-inventing fishing equipment, or prevented such technology from spreading should it have arisen. While this taboo may have helped to concentrate efforts on the acquisition of land-based sources of food, it also potentially halted technological innovation. In contrast, some of the more recent technological innovations in the modern world were driven by science fiction writers' visions of what could be. It is possible that, in the aftermath of a technology loss, recovery could be aided by visions of what once was.

The Future of Innovation

Is it possible for human computation to improve the operation of cumulative culture, so that we may avoid catastrophic technology loss altogether? We can draw insight from how it has been improved in the past, focusing on the last ten millennia and particularly on the most recent centuries and even decades. Table 1 details some of these mechanisms. Contemporary web-based efforts to speed the rate of innovation (e.g. Spigit.com, Innocentive.com, Google Scholar) are based on further improvements along the lines outlined in Table 1. Electronic storage is so cheap that we can aspire to have the sum total of human knowledge stored in electronic media. Even today a large fraction of that total is available to those of us with access to the web and a research library with electronic journal subscriptions. Top universities are offering free online courses and Wikipedia has authoritative micro-courses on a host of topics. Given that much innovation involves novel combinations of ideas, the ability to rapidly access a wider variety of ideas and skills should increase the rate of innovation.

Two limitations to the power of the internet to help speed innovation remain. First, intellectual property issues must be solved. Book and journal publishers have so far prevented rapid, inexpensive access to all potentially useful information, even in the realm of academic publishing where authors expect to be paid mainly in the form of prestige rather than money. Companies that crowdsource innovation, like Spigit and Innocentive, restrict access to their services to a closed community to protect intellectual property rights. Complicating the issue, the protection of intellectual property rights likely provides a key incentive for innovators in the modern world of global interconnections. Users of information will often be strangers who will not be inclined to bestow prestige rewards on innovators, much less material rewards. Some small, simple, and automated per-view or per-download royalties would compensate creators of group-beneficial innovations, who would not otherwise benefit from their work.

Second, as Polanyi (1966) argued, much knowledge is "tacit"—the fingertip feel for things that is very difficult, if not impossible, to reduce to print or pictures. For example, to educate scientists we still rely on the highly personalized, labor intensive PhD system pioneered in Germany in the early nineteenth century precisely because it allows for the transfer of tacit knowledge. Similarly, firms expect to have to inculcate their organization's ethos into technical and management hires. And one might reasonably entertain the possibility that most undergraduate students prefer to attend residential universities, perhaps not entirely for the parties. Thus, the need to transmit tacit knowledge is enduring, and as such, seems to act as a fundamental limit on the rate of cumulative cultural evolution.

As we proceed into the future, what it means to improve our ability to generate and maintain cultural innovations may grow considerably less straightforward. The pace of evolution is set by the rates of dissemination and innovation. Human communities are now integrated with digital devices that are themselves excellent learners, ratcheting up our ability to innovate. With more people producing more technology with more variation than ever before, culture is evolving at increasing rates. A concern, however, is that modern cumulative culture operates much more rapidly in some areas than in others. Technology that can be adopted by an individual, for example, can often spread faster than institutions that require coordination among many individuals.

Which problems we solve with our technology will depend partly on what we perceive to be the salient problems. Culture influences how we perceive the world (Smaldino and Richerson 2012). If our problems continue to be framed in terms of the accumulation of wealth and power, then the human computational engine will surely continue to apply its problem-solving powers in that sphere. That strategy, however, will very likely lead to a catastrophe with the potential to disable the infrastructure of the global human computer that made it possible. As Nardi writes in her eloquent chapter (this volume), "There is no energy cornucopia waiting for us to tap into; we live on a specific planet, with specific resources. We are in the process of using up those resources." A more promising direction, in terms of prolonging the existence of our impressive culture achievements, is to attempt to guide the evolutionary forces that shape cultural change toward those individual mindsets and social and legal institutions that promote forethought and sustainability.

Summary

Cumulative culture is the engine that drives the remarkable power of the global human computer. It enables societies to act as super-duper computers by ratcheting up technological and cultural innovations. Once culture can accumulate, the ability of a society to maintain and spread complex technologies is directly related to the size of the population and its connectivity with other populations. Larger and more connected societies can maintain more complex technologies. This also means that sudden isolation or a drop in population size can lead to a loss of technology. The modern world maintains highly complex technology requiring the interactions of many varied, superbly trained individuals. A catastrophic loss in terms of either life or connectivity has the potential to trigger what some may consider an equally devastating loss of technology. While our interconnected society may have the resources to avert such a second-order crisis, our best bet in the face of a loss of technology is to retain knowledge of its existence in our collective memory and to continue developing the kinds of human computational tools that have served us in the past. As the accelerated pace of our cultural evolution comes at the cost of increased resource use, however, we may need to focus on shifting the evolutionary forces that guide the norms and institutions that define the psychological state space of problems and solutions. It would be a shame to damage the global human computer through a product of its own doing.

References

Aiello LC, Wheeler P (1995) The expensive-tissue hypothesis: the brain and the digestive system in human and primate evolution. Curr Anthropol 36:199–221

Allely S, Baker T, Comstock P, Hamm J, Hardcastle R, Massey J, Strunk J (1992) The traditional Bowyer's Bible. Lyons Press, Guilford

Ambrose SH (1998) Late Pleistocene human population bottlenecks, volcanic winter, and differentiation of modern humans. J Hum Evol 34:623–651

Boyd R, Richerson PJ (1995) Why does culture increase human adaptability? Ethol Sociobiol 16:125–143

Boyd R, Richerson PJ (1996) Why culture is common but cultural evolution is rare. Proc Br Acad 88:73–93

Boyd R, Richerson PJ, Henrich J (2011) The cultural niche: why social learning is essential for human adaptation. Proc Natl Acad Sci 108:10918–10925

Chudek M, Henrich J (2011) Culture–gene coevolution, norm-psychology and the emergence of human prosociality. Trends Cogn Sci 15:218–226

Davidson I, Roberts DA (2009) On being alone: the isolation of the Tasmanians. In: Crotty M, Andrews DA (eds) Turning points in Australian history. University of New South Wales Press, Sydney, pp 18–31

Diamond JM (1978) The Tasmanians: the longest isolation, the simplest technology. Nature 273:185–186

Donald M (1991) Origins of the modern mind: three stages in the evolution of culture and cognition. Harvard University Press, Cambridge

Ehn M, Laland K (2012) Adaptive strategies for cumulative cultural learning. J Theor Biol 301:103–111
Enquist M, Eriksson K, Ghirlanda S (2007) Critical social learning: a solution to Rogers' paradox of non-adaptive culture. Am Anthropol 109:727–734
Goral M, Levy ES, Obler LK (2002) Neurolinguistic aspects of bilingualism. Int J Biling 6:411–440
Henrich J (2004) Demography and cultural evolution: how adaptive cultural processes can produce maladaptive losses – the Tasmanian case. Am Antiq 69:197–214
Henrich J, Boyd R (2008) Division of labor, economic specialization and the evolution of social stratification. Curr Anthropol 49:715–724
Henrich J, Gil-White FJ (2001) The evolution of prestige: freely conferred deference as a mechanism for enhancing benefits of cultural transmission. Evol Hum Behav 22:165–196
Henrich J, McElreath R (2003) The evolution of cultural evolution. Evol Anthropol 12:123–135
Herrmann E, Call J, Hernández-Lloreda MV, Hare B, Tomasello M (2007) Humans have evolved specialized skills of social cognition: the cultural intelligence hypothesis. Science 317: 1360–1366
Hill K, Barton M, Hurtado AM (2009) The emergence of human uniqueness: characters underlying behavioral modernity. Evol Anthropol 18:187–200
Keinan A, Clark AG (2012) Recent explosive human population growth has resulted in an excess of rare genetic variants. Science 336:740–743
Kline MA, Boyd R (2010) Population size predicts technological complexity in Oceania. Proc R Soc Lond B 277:2559–2564
Kosorukoff A (2001) Human based genetic algorithm. IEEE Trans Syst Man Cybern 5:3464–3469
Loulergue L, Schilt A, Spahni R, Masson-Delmotte V, Blunier T, Lemieux B, Barnola J-M, Raynaud D, Stocker TF, Chappellaz J (2008) Orbital and millennial-scale features of atmospheric CH4 over the past 800,000 years. Nature 453:383–386
McBrearty S, Brooks AS (2000) The revolution that wasn't: a new interpretation of the origin of modern human behavior. J Hum Evol 39:453–563
McDougall I, Brown FH, Fleagle JG (2005) Stratigraphic placement and age of modern humans from Kibish, Ethiopia. Nature 433:733–736
Pardoe C, Bowdler S, Brace CL, Plomley NJB, Turner CG, Wolpoff MH (1991) Isolation and evolution in Tasmania [and comments and reply]. Curr Anthropol 32:1–21
Perreault C (2012) The pace of cultural evolution. PLoS One 7:e45150
Polanyi M (1966) The tacit dimension. Doubleday, Garden City
Powell A, Shennan S, Thomas MG (2009) Late Pleistocene demography and the appearance of modern human behavior. Science 324:1298–1301
Ray N, Currat M, Berthier P, Excoffier L (2005) Recovering the geographic origin of early modern humans by realistic and spatially explicit simulations. Genome Res 15:1161–1167
Read L (1958) I, pencil: my family tree as told to Leonard E. Read. In: The freeman; reprinted by The library of economics and liberty. Retrieved 27 Feb 2013 from http://www.econlib.org/library/Essays/rdPncl1.html
Richerson PJ, Boyd R, Bettinger RL (2001) Was agriculture impossible during the Pleistocene but mandatory during the Holocene? A climate change hypothesis. Am Antiq 66:387–411
Richerson PJ, Boyd R, Bettinger RL (2009) Cultural innovations and demographic change. Hum Biol 81:211–235
Ridley M (2010) The rational optimist: how prosperity evolves. HarperCollins, New York
Rogers AR (1988) Does biology constrain culture? Am Anthropol 90:819–831
Smaldino PE (in press) The cultural evolution of emergent group-level traits. Behav Brain Sci
Smaldino PE, Richerson PJ (2012) The origins of options. Front Neurosci 5:50
Tennie C, Call J, Tomasello M (2009) Ratcheting up the ratchet: on the evolution of cumulative culture. Philos Trans R Soc B 364:2405–2415
Tomasello M, Carpenter M, Call J, Behne T, Moll H (2005) Understanding and sharing intentions: the origins of cultural cognition. Behav Brain Sci 28:675–735

White TD, Asfaw B, DeGusta D, Gilbert H, Richards GD, Suwa G, Howell FC (2003) Pleistocene *Homo sapiens* from Middle Awash, Ethiopia. Nature 423:742–747

Woolley AW, Chabris CF, Pentland A, Hashmi N, Malone TW (2010) Evidence for a collective intelligence factor in the performance of human groups. Science 330:686–688

World Bank (2012) Mobile phone access reaches three quarters of planet's population. Retrieved 27 Feb 2013 from http://www.worldbank.org/en/news/press-release/2012/07/17/mobile-phone-access-reaches-three-quarters-planets-population

Zachos JC, Shackleton NJ, Revenaugh JS, Palike H, Flower BP (2001) Climate response to orbital forcing across the Oligocene-Miocene boundary. Science 292:274–278

Human Computation and Conflict

Juan Pablo Hourcade and Lisa P. Nathan

Introduction

Technological developments play a critical role in arenas of war, from increasing or decreasing their likelihood, to changing the nature of warfare, to affecting recovery. Writing, cartography, advances in mathematics, and the rise of robotics are examples of such developments. Human computation is a relatively new, but already significant contributor to this area. In this chapter, we explore the topic of human computation and conflict from the perspective of trying to diminish the damaging impacts of war. In particular, we focus on how human computation may influence, positively or negatively, efforts to prevent, de-escalate, and recover from armed conflict in the early twenty-first century.

Our approach reflects our interest in minimizing the human and financial cost of warfare. We draw from recent empirical research on the causes of war at a societal level, reconciliation efforts at the community level, and drivers of empathy and compassion at a personal level. We proceed by highlighting a range of work from the area of human-computer interaction (HCI) and the diverse disciplines and research streams that it brings together (e.g., persuasive computing, crisis informatics, collapse informatics, etc.). These exemplar projects weave together research and practice, exploring the diverse ways that human computation might help (or hinder) efforts to avoid, minimize and recover from armed conflict. We also suggest research questions that scholars embarking on these projects are well positioned to address. In these discussions we focus on human computation from an HCI

J.P. Hourcade (✉)
University of Iowa, Des Moines, USA
e-mail: juanpablo-hourcade@uiowa.edu

L.P. Nathan
University of British Columbia, Vancouver, Canada
e-mail: lisa.nathan@ubc.ca

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_77,

perspective. We highlight interactive technologies and techniques that facilitate novel forms of human interaction in support of explicit objectives.

We stress the need for work in this area to be explicitly socio-technical in approach. We strive to avoid technological deterministic positioning, recognizing the influence of policy and praxis, values and valuing in this complex human arena. The features of computational tools do not stand in isolation, they are enmeshed in complicated modern ecosystems. In other words, we believe that attention must be paid to the social and cultural context within which specific human computation technologies will be used. In this way, the impact of these tools can be better understood and objectives are more likely to be achieved in specific contexts.

Empirical Research and Conflict

During the past 15 years, solid empirical research has significantly enriched discussions around the causes of armed conflict. Prominent scholars, such as Frances Stewart and Paul Collier have used 50 years of human development data from the United Nations to understand the conditions that make it more likely that a country will be part of a civil or international war in the near future (Stewart 2002; Collier 2007). Collier, in particular, has done extensive statistical analyses of these data sets. They found private motivation (i.e. financial profit) to be a primary motivator of armed conflict, often linked to trade in primary commodities. Other causes include a failure of the social contract, rapid economic decline and social unrest, environmental stress, certain forms of ethnic and religious fractionalization, a high proportion of young men in the population, partially democratic governments, and geographic and historical factors (Stewart 2002; Collier 2007; DeRouen and Goldfinch 2005; Nardi, this volume). These researchers have also identified factors that reduce the likelihood of armed conflict. These factors include being a fully democratic country, and having a better educated population (Stewart 2002; Collier 2007; DeRouen and Goldfinch 2005). Collier, in particular, found that each additional year of education for the general population reduces the risk of civil war by about 20 % in a country (Collier 2007). A recent issue of Science magazine has explored additional factors with respect to their impact on armed conflict including the role of gender, climate change, empathy, racism, and modern forms of communication such as Twitter (Riddihough et al. 2012). In our discussion of human computation, we present how some of the key factors (e.g., political engagement, health, social stability, connectedness, education, empathy, access to modes of communication) are-and will likely continue to be-affected by human computation. A more detailed discussion of these factors can be found in a paper by Hourcade and Bullock-Rest (2011).

In addition to societal factors, psychologists and neurologists have provided novel insights into how empathy and compassion work in individual minds and across groups. In his recent book *The Science of Evil*, Simon Baron-Cohen discusses the brain's empathy circuit and how damage to these parts of the brain affect empathy. In doing so, he also discusses the elements that activate the circuit, including touch, gaze, seeing someone else do or experience something, and recognizing emotions (Baron-Cohen 2011). These are critical to our discussion of human computation as technologies that do not let us see other people or understand how they feel can prevent empathy and compassion from manifesting during periods of critical communications.

Addressing the Precursors to Conflict

In the following sections the reader will find many examples of measures that appear promising in term of addressing the causes of war, reflecting our stance that preventing conflict is far more cost effective than attempting to ameliorate its effects once begun.

Politics, Democracy and Political Engagement

Given that fully democratic nation states are less likely to participate in armed conflicts than less democratic states, studying how human computation technologies are leveraged in political campaigns is a pertinent topic. This is one area where the effects of tools using human computation techniques are already noticeable, with political campaigns attempting to leverage new mobilization strategies. For example, in the United States, the 2012 presidential campaign in support of Barack Obama used combinations of web and mobile apps along with a massive database that made it very easy to distribute work for the campaign and to aggregate individual decisions and actions into a coherent whole. These tools were used to organize volunteers, manage personal visits and phone calls in real time, even enabling volunteers to make quick, targeted phone calls when they only had a few minutes available. They even matched volunteers with the people they would contact based on common life experiences (Lohr 2012).

Social media and user-generated content are also playing a role. One example is the successful online campaign to defeat the SOPA/PIPA legislation in the United States, which initially had large bipartisan support in Congress. While not entirely a grassroots campaign since it had the support of the likes of Wikipedia and Google, it provides an example of how online organizing can influence decisions by elected officials (Kravets 2012). These tactics are increasingly used by advocacy groups who prepare messages and ask supporters to mail them to their elected representatives through automated systems. Messages that may have taken significant effort or cost in writing and mailing a letter, or making a long distance phone call 20 years ago, are now “delivered” after only a few clicks. Thus, the convincing of politicians is quickly distributed to interested citizens who can quickly comment and exercise influence.

These efforts are spreading. Just north of the United States, extreme budget cuts proposed in 2013 motivated the grassroots development of the Budget ReACTion Project (BRP). BRP is an initiative designed to scaffold the ability of communities across Canada to work collaboratively to analyze the 2013 budget and how it will affect local services (www.openthebudget.ca). BRP is using crowdsourcing, a catch-all term for distributed problem solving. In this case algorithms outsource specific steps to human experts, engaging local policy wonks in a breakdown of the hundreds of pages of the budget proposal. It is hoped that by dividing up the budget and localizing the level of analysis, BRP will help communities across the nation-state come to a richer understanding of the likely influences of the recent budget proposal, so they can mobilize action more effectively.

A positive from some of the examples above is that human computation may provide ways for citizens to increase their civic engagement. Consider the Budget ReACTion Project discussed above and how the use of algorithms for dividing the work makes it easier for policy experts to participate in parsing government budget proposals, in turn this parsing makes it possible for local citizens to take informed action in response to these proposals. This may involve discussing political issues with neighbors, potentially engaging the brain's empathy circuit. These types of engagement may also lead to a better informed citizenry, something that has long been highlighted as a necessary component of a well-functioning democracy. Human computation also provides more opportunities for people to organize campaigns and develop informed proposals that make their way from the bottom up, increasing participation in the democratic process.

Obviously activism and political campaigns can also be used to start wars or stigmatize groups; consider the use of online "guerrilla tactics" to promote or demote content and opinion. Activist groups are training followers on these tactics so that when news stories about a topic of interest are presented in an online forum, followers are alerted and proceed to flood discussion areas related to the story with comments supporting their viewpoint, adding "likes" to friendly comments and "dislikes" to opposing comments (Hiar 2010). Even in widely read national news media, 100 online "soldiers" are able to manipulate their views into the "most liked" category for a given story, providing an inaccurate sense of popular opinion on a matter. These tactics could easily extend to efforts of intimidation, suppressing the opinions of those who have moderate opinions, and we can see this happening in online forums when things get particularly heated. Eventually, extremists are often the only ones left. Guerrilla tactics may also be used to quickly spread misinformation. Additionally, there is the possibility that human-computation could lead to more top-down approaches where people get used to repeating slogans or "talking points" coming from organizations or political parties (e.g., through the use of the Obama campaign's extensive voters database), without researching the information and reaching their own conclusions. There is no question that human computation tools in the political arena have the potential to be used to increase the chances of armed conflict.

One of the most compelling areas of research on political engagement explores using human computation to entice citizens to interact with each other around divisive political topics in constructive ways. One HCI-focused research group investigating this area is the University of Washington's Engage Project. This team of researchers is exploring the design of interactive, digital systems that encourage reflective engagement on charged topics, with a recent focus on voting (e.g., Freelon et al. 2011). Their Living Voters Guide aims to bring "… voters together to discuss and integrate their perspectives, in contrast to our media environment of divisive soundbites. It is a voters' guide that is co-created by everyone who participates. It evolves as citizens consider the tradeoffs for each measure."

There are further obstacles to supporting civil, civic debate when considering those living under more restrictive political regimes. In many contemporary nation-states stating the wrong opinion online can land an individual, and those who manage the site, in prison. How can human computing endeavors and associated policies help protect open discourse in politically fraught environments? The Voices from the Rwanda Tribunal project is investigating this space, designing a system to foster discourse in Rwanda, engaging the tensions related to citizens' abilities to express opinions related to the United Nations Criminal Tribunal for Rwanda in a highly constrained political climate (e.g., Nathan et al. 2011). This project underscores the need to take political conditions seriously as technological features that suggest but are unable to guarantee anonymity can have dire consequences for individuals who express ideas deemed unfavorable by the current political regime.

It is also important to keep in mind that requests for engagement in political activities could quickly become overwhelming, with ever urgent requests from advocacy groups haragunging citizens for their participation. There is a clear possibility that this could overwhelm people and lead them to disengage from all such activities. Research is needed so opportunities to engage are presented in a manner that helps individuals assess the importance of participation and its repercussions in relation to their interests, values and the wider political context.

Consumer Awareness

Primary commodities are often used to fund wars, especially civil wars waged in low-income regions of the world. Ross (2006) identified a causal relationship between oil, gas, and diamond wealth and the onset of civil war, with a marked increase in the likelihood of this happening starting in the 1970s. In addition, he found that conflict duration was linked to the amount of "contraband" possible, including gemstones, timber, and narcotics.

Primary commodities have to be sold somewhere and often end up in the hands of consumers who are unaware that they are indirectly funding a violent civil war halfway across the world. The eastern side of the Democratic Republic of Congo is a region where the insatiable demand for rare earth minerals needed to build contemporary digital tools continues to add fuel to horrific conflict. Simply stated, tantalum (recovered from ore minerals such as columbite and tantalite) incorporated into a new cellphone or laptop can lead to more weapons being purchased by violent warlords (Smith and Mantz 2006). There are increasing efforts to provide consumers with the ability to trace what they purchase. Many grocery stores and supermarkets specializing in "natural" products increasingly display the origins of their merchandise. Clothing companies (e.g., All American Clothing) are also moving to provide tracing of their supply chain, especially when they wish to highlight a product's origin in a particular country.

Human computation could play a role in bringing further transparency to product tracing. One way would be to enable people who work in different parts of the supply chain to provide testimonies or even live feeds of their workdays. The technology exists but the practices and policies of such broadcasting are still developing. Similarly human computation could play a role in providing unofficial tracing of products, enabling people working in the supply chain to provide "on the ground" testimonies and information on the products consumers may purchase, even if the manufacturing companies are not interested in supporting this type of information sharing. Buycott is an example of such an app, enabling consumers to scan barcodes and instantly learn about the companies involved in manufacturing a product (Buycott 2013).

Human computation could be a game changer in product traceability, enabling better informed consumers who may come to expect ready access to that information and distrust products that do not provide it. In turn, this could potentially deal a serious blow to the financing of civil wars, and to companies and organizations that directly profit from the hidden nature of the funding sources for these conflicts. Black lists of products, for example, could be provided by organizations consumers trust.

An unintended negative effect could occur because people who do decide to participate in "on the ground" reporting could risk their jobs and livelihoods. If employers discover that individual employees are revealing less savory practices, questionable conditions and connections to conflict, they may fire these employees. There is also the likelihood of misrepresentation, with companies either obscuring the true origins of products or the broader socio-cultural conditions in which the products were manufactured (similar to "green washing"). Human computation traceability endeavors require a high level of trust and oversight processes to be developed.

The ability to provide product tracing is still in its infancy and there is the need for investigations into robust ways of providing this data that incorporate human computation techniques. This may require generating an automated narrative of how a product got to you based on individual contributions from people throughout the supply chain. It would be useful for tools to be able to provide the right amount of information tailored to particular consumer interests (e.g., conflict prevention or ecological footprint) while enabling a quick, intuitive way of visualizing consumer impact (e.g., visual analytics).

Humanizing Connections

Convincing people to support a war often involves the demonization and dehumanization of "the other". This happens through the spectrum of actors in armed conflicts, from democracies to terrorist groups (Ivie 1980; Weimann 2004).

Demonization efforts are much harder to accomplish if many people are familiar with members of the other groups and are aware of their common humanity. The increasing worldwide availability of the Internet can make it easier than in years past to forge connections between people from widely different groups. Social media provides interesting platforms for these connections to happen in a massively distributed way.

A particularly compelling example of these connections comes from the use of early versions of social networking software on the island of Cyprus during the 1990s. At the time, people from the Greek and Turkish regions of the island were forbidden to communicate with each other and could only meet by travelling to other countries. Digital connectivity in the early years of the Internet provided an opportunity for dialog that likely played a role in the thawing of tensions between the two sides (Hourcade et al. 2012).

There have also been unanticipated opportunities for massive international collaborations. For example, consider musical compilations or movies put together by a large number of online contributors from around the world. Oscar-winning director Kevin Macdonald put together the feature film *Life in a Day* based on thousands of contributions from around the world (Macdonald 2011).

Finally, there could be opportunities to better understand events from the other side's point of view by engaging in discussions following rules that lead toward agreements. There are examples of co-narrating a conflict (none other than the Israeli-Palestinian) with multitouch interactive tabletops that could be taken to the online world (Zancanaro et al. 2012). The opportunities for creative collaborations with people we may never meet face-to-face will continue to increase and provide unprecedented worldwide humanizing connections.

Massive online connections with people from other groups can help humanize them and reduce the social distance between people from different backgrounds, cultures and value systems. Members of opposing sides may find commonalities in activities or interests that may take them away from thinking only about their differences. It certainly seems harder to support waging war, and potentially killing, the same people who helped you create something of beauty.

Online connections can also be used to affirm extreme views and communicate only with people who share one's ideas. It is easy nowadays for a racist to find a large number of people who will validate those ideas online, while 15 years ago in many places it would have been far more difficult to identify and engage with a handful of like-minded people. There is a danger of creating echo chambers for political views that human computation efforts can further amplify, pulling people away from moderation because they remove dissenting views.

The main challenge for future research in humanizing connections is in establishing connections when people do not previously know each other. While this is already happening to some degree, especially for projects that receive a lot of publicity, such as *Life in a day*, it could potentially happen much more often for a variety of creative activities that may benefit from international teams coming together online. A greater challenge is likely in providing people with incentives to join communities where there are respectful yet dissenting opinions, instead of joining communities that solely provide affirmation for existing views and positions. There may also be opportunities for research in designing massive online games that have peace instead of world domination as a goal. Role- playing games such as *Peacemaker* exist to challenge one's personal skills at conflict resolution, but there are certainly opportunities for taking these concepts to a larger scale.

Education

While it is unclear exactly why higher educational levels lead to a reduced likelihood of conflict, it is likely due to a combination of more knowledge about people and the world combined with greater economic opportunities that make entering an armed conflict a less palatable endeavor. Human computation is increasingly playing an important role in education (Beat et al., this volume).

Massive Online Open Courses (MOOCs) are arguably the most hyped phenomena across academia in 2013. These are usually free (hence "open") courses available online and intended for a massive audience. While these increasingly available offerings are primarily designed and targeted to a college-level audience, the impact of the approach could potentially be realized at other levels of education. For example, something that easily separates groups of people is language. MOOCs could be used for foreign language instruction in primary and secondary schools, helping fill a gap in countries where there may be a shortage of foreign language instructors. Such offerings could help shorten the amount of time needed for countries to make qualitative leaps in their ability to teach specific subjects.

These efforts would have much to learn from other large scale efforts connecting classrooms across the world. For example, the collarative project iPoPP (created by GlobalSchoolNet.org and eLanguages.org) is a global e-learning platform for multilingual, project-driven collaboration (http://www.globalschoolnet.org/ipopp/about-ipopp.html). It draws upon constructivist learning methodology and collaborative learning strategies to support projects amongst students from different language backgrounds. There are also less formal examples of multi-language, online collaborations that scaffold children teaching each other about their lives and activities (e.g., collaboratory blogs and YouTube videos) that further enrich the area of human computation.

The main positive about these endeavors is the sheer numbers of diverse peoples that may be reached. If motivated they may acquire valuable skills, learn novel points of view, open their minds, and develop critical thinking skills. The positives are more likely to occur in regions of the world that are far from educational resources, where people with the necessary background to teach may simply not be available.

Negatives may also come from these new technological practices in the form of decreasing the number of initiatives to improve educational institutions based within local communities. This may be of particular concern to regions of the world where insufficient resources are dedicated to education, and children may engage with instructors or others who are unaware of local norms, culture and concerns, risking further marginalization of these young people and decreasing their motivation to actively participate in the learning experience. There is growing evidence that college students may be given credit for MOOCs that do not deliver the same quality of education as in-person courses at institutions without enough funding to provide an adequate number of courses given their enrollment numbers (Gardner and Young 2013). This problematic trend may eventually work its way to all levels of education.

The biggest impact in terms of peace will only be accomplished if human computation efforts can contribute to educational levels in regions of the world where people are mostly illiterate. This means the difficulties of getting any type of MOOC to these settings would be significant in terms of developing supportive socio-technical infrastructures for student access. A compromise would be to use this technology to help accelerate the training of teachers, and do this while providing local content in a culturally appropriate way.

Poverty Reduction

Widespread poverty and a failure of the social contract can make joining an armed conflict more attractive, since the alternative is not particularly agreeable. Education and good governance could certainly help reduce poverty, but other initiatives that make use of human computation could provide additional critical help. An example that makes use of human computation approaches is peer-to-peer micro-financing. One of the best known examples is Kiva (kiva.org), a website that matches people with small amounts of money to lend, who are most likely in high-income regions of the world, with people in need of loans for small businesses in low-income regions of the world. As of March, 2013, after about 8 years of activity, Kiva had almost one million lenders and had given out over 400 million US dollars in loans to people in 67 different countries. Arguably the most significant statistic is that Kiva has achieved a 99 % repayment rate.

Kiva shows a positive example of how peer-to-peer micro-financing can significantly address poverty. Thousands of people have accessed its loans, individuals who likely would not have had other ways to invest in their educational endeavors or small businesses. At the very least, initiatives like Kiva can provide an alternative to predatory loans whereas taking a loan becomes cost effective, feasible and convenient. These services can also cut down on the need for “middle-men”, reducing the cost of giving a loan and receiving repayment.

Micro-financing is not immune to fraud, and while this has not yet hurt peer-to-peer micro-financing significantly, as the numbers increase, so will the likelihood of attracting scammers and hackers. The ability to scale auditing for large peer-to-peer micro-financing efforts could be costly and difficult. The failure of a highly visible initiative could cause significant damage to this approach to poverty reduction.

Substantially scaling up these efforts will require research on how to design systems that on a large scale can be trusted, reliable, transparent, and understandable by their users (both lenders and borrowers).

Ameliorative Actions

Once underway how might human computation efforts address the horrific realities of conflict, potentially easing the suffering of those involved? In particular, how might these efforts empower those most affected to aid in de-escalating war?

Citizen Journalism

With an ever increasing number of people carrying smartphones that can record video or broadcast it live, the amount of citizen journalism entering mainstream media has grown exponentially. Anyone carrying such a device can record events and quickly share them. This can potentially reduce the incentives for parties in conflict to use excessive violence as it is likely to be reported and may be difficult to deny or justify. This can prevent political disagreements from scaling into armed confrontations. For example, videos of an incident at the University of California, Davis where a police officer used pepper spray on demonstrators resulted in widespread condemnation, with the university later paying demonstrators 1 million US dollars to settle the lawsuit (Favate 2012). If the incident had not been video recorded, police officers could have more easily argued that the use of pepper spray was justified.

The past few years have seen a flurry of development as various organizations (NGO, non-profit and for-profit) are developing platforms to support citizen journalism. An example is Ushahidi Crisis Mapping (http://ushahidi.com) (see also Meier, this volume). Ushahidi was originally created to track post-election violence in Kenya in 2008, providing a free open-source tool for mapping and tracking information via text, email, twitter, and other forms of communication. Today it is used by political activists, but also made a transition into use for disaster relief—mapping damage following 2010s "Snowmaggedon" on the east coast of the US and the massive earthquakes in Haiti (Ushahidi 2013).

Another project, more focused on specifically supporting citizen journalism is Global Voices Online. Global Voices is a massive blogging community dedicated to bringing citizen reporting from around the world into one centralized location. Project members seek to create a credible source for citizen journalism in order to promote free speech and bring a more equal level of media attention to happenings around the world. Significantly, posts are translated into over 30 languages. (http://globalvoices.org)

Massive citizen journalism initiatives have the potential of preventing violence because of the difficulty of perpetrators to deny their involvement in the activities.

But it also has the potential to de-escalate war if citizens see the gory consequences of their support. For example, in the most recent conflict between Israel and Palestinians in the Gaza strip, both sides made efforts to immediately show the impact of the other side's strikes on civilians, and in particular on children (Mackey 2012). This helped put international pressure on both sides to reach a ceasefire. The same cannot be said of conflicts such as the civil war in Syria where media often has to be smuggled out of the country and tends to show consequences of violence without context (i.e. who was directly involved) as opposed to live or recorded video footage of violent events.

There is also the potential for citizen journalism techniques to be used to monitor ceasefires. Thousands of live camera feeds from a sensitive area could be used to instill trust between warring sides, and make it less likely for one side to break the ceasefire without it been clear they were to blame.

There can also be negatives to these developments. For example, seeing violence may increase the support for violence as people may deeply desire acts of revenge. Likewise violence from a small number of people belonging to one group, may be used to justify violence against an entire group. Hence, people who seek to paint a group in a negative light will have a much larger amount of material to support their views. An additional issue with having large numbers of video cameras constantly running is the loss of privacy and the potential for massive spying, especially in countries where political dissent is simply not tolerated.

The blessing and curse of citizen journalism is the overwhelming quantity and varying quality of information that can be generated related to a single incident. Finding relevant information and verifying this information leads to a host of research questions, visualization challenges and aggregation opportunities. Scaling up these systems so they can be robust and are difficult to subvert will be critical to ensure their effectiveness.

Dealing with Emergencies: Crisis Informatics

The term Crisis Informatics was coined by Chris Hagar and grew out of her work with farming communities in the United Kingdom affected by extreme quarantine measures during the Hoof and Mouth disease outbreak in 2001 (Hagar 2006; Hagar and Haythornthwaite 2005). Others have contributed significant work to this field, investigating socio-technical interactions that occur during times of extreme crisis with an eye towards developing ways to support the mitigation of suffering. Although Crisis Informatics research is not limited to the time period when a crisis is underway, some of the most compelling scholarship has contributed powerful analysis of activities during this phase of a crisis (e.g., Starbird and Palen 2011; Palen et al. 2011).

Specific Crisis Informatics related projects include the Humanitarian OpenStreetMap Team (HOT). HOT draws upon the wiki project OpenStreetMap to collaboratively map crises and disasters. Volunteers working remotely gather data based on satellite imagery or other available data sets. This information is used to inform and direct humanitarian responders on-location as they attempt to coordinate efforts quickly and efficiently. Areas where HOT mapping projects have occurred include the famine in Somalia, the Presidential Election crisis in Ivory Coast, and earthquake damage in Haiti (http://hot.openstreetmap.org/).

Crisis Informatics offers significant potential in human computing terms by providing insights into methods for combining local, on the ground expert knowledge, with distributed volunteer expertise from around the world, enhanced with access to open data sources.

Potential drawbacks to this field are related to the fact that many different players, with different motivations have access to open systems. Reports of crisis can bring aid, but can also bring those who wish to benefit from the period of chaos and destruction. There are many open research questions related to both technical and sociological conundrums that face Crisis Informatics related efforts. These include issues of scale, addressing the challenge of “bad actors”, prioritizing needs, and connecting needs with expertise and resources.

Recovery

Armed conflicts leave indescribable suffering and heartache in their wake. Post conflict states often experience a time when individuals are interested in pursuing activities related to justice and reconciliation efforts. The efforts are aimed at identifying and addressing wartime atrocities and grievances and ways for neighbors (whether at the nation or street level) to trust each other again. There is a growing realization that the goals of justice and reconciliation are not necessarily linked and require different approaches in order to avoid exacerbating harms and contributing to the likelihood of future conflict (Fletcher and Weinstein 2002).

Contemporary conceptualizations of *justice* and modern court systems require documentation, recorded evidence of wrongdoing to support investigations, indictments, and hearings. In order for affected individuals to believe that justice has been delivered, they need have some evidence that the judicial process was followed. Yet, *reconciliation* efforts require the development of mutual understanding and a rebuilding of trust. Activities that focus on attribution and blame do not contribute to reconciliation. Instead, reconciliation efforts require the development of discourse between the affected parties, speaking and being heard. Human computation developments, many already introduced above, appear promising in terms of supporting both justice and recovery efforts.

Archives and Justice

For centuries archivists have worked to preserve records, developing theory and practice related to critical attributes of records including issues of trust, authenticity and reliability. Archivists recognize the power of well managed records in battling revisionist histories and shifting technologies that threaten our ability to understand and learn from our past. Governments and organizations with a deep appreciation for the importance of documenting their history for the longer-term (decades and centuries rather than over months and years) hire archivists to manage their records. Yet for most of us the term archive is used as a verb because we are unaware of the deep body of scholarship that archivists draw upon when they engage in archival practice and study. To archive something in the colloquial sense simply means to save it for the very short term (i.e., hit the save button and perhaps put it in a folder). When it comes to efforts to support justice after conflict, nation-states and organizations who have had the ability to maintain records that hold up in court are far better positioned to see justice upheld.

Yet individuals, grassroots organizations, non-profits with low operating budgets and "seat of the pants" daily operations are not well positioned to contribute records considered trustworthy, reliable, and authentic to the courts. This is an area where human computation is beginning to contribute. While associated with earlier projects under the citizen journalism section that use citizen submitted video to share atrocities from around the world, WITNESS, a non-profit that uses video to document human rights abuses, also supports a media archive. The goal of the archive is to "…support of advocacy, prosecution of justice, truth telling, and the historical record. We believe that archives serve a critical role in human rights advocacy, by protecting and preserving evidence, restoring memory, ensuring the endurance of under-represented voices, and as a bulwark against impunity and forgetting" (http://witness.org/media-archive/about-the-collection). In large part, the professional archivists and digital preservation professionals that work for and with WITNESS aim to provide support for longer-term accountability by helping future generations understand some of the missing portions of official nation-state histories. The organization holds footage of atrocities as they occurred, victim and witness testimony and evidentiary submissions. These materials come from the people most affected by armed conflicts rather than nation-state bureaucracies. To ensure the sustainability of this project and others in this vein, significant resources and socio-technical, human computation research will be necessary.

These types of activities require increasingly complicated modes of verification and long-term stewardship as digital tools and processes become more complex. Modern digital tools have many amazing attributes (e.g., small, lightweight, inexpensive) but this technology is still in its infancy and we know that there are significant challenges to preserving bits and bytes along with their provenance over the longer term. Efforts to create trusted, digital repositories are underway, but there are still many unknowns when dealing with the digital record lifecycle, from point of creation to long-term preservation.

Digital Media and Reconciliation

There is a long and well accepted practice of supporting 'truth telling' activities within and between communities recovering from periods of extreme conflict (e.g., armed conflicts between nation-states, civil war, genocide). Stakeholders, whether perceived as victim, witness, or aggressor are asked to come forward and share their version of particularly traumatic events associated with the conflict. The belief stems from the idea that without acknowledging deep harms and different perspectives on these harms, recovery is more difficult and a return to conflict is likely.

Truth and Reconciliation Committees (TRCs) such as Canada's ongoing court sanctioned TRC provide a well established method for structuring truth telling events (Zalaquet 1997) often a first step in a longer stabilization process (Long and Brecke 2003). Canada's TRC is holding events across the country to acknowledge the extensive harms done to thousands of Aboriginal children, their families and their communities during the 130 year period when the federal government removed Aboriginal children from their homes and shipped them off to live in the (typically) inhumane conditions of Indian Residential Schools. Truth telling activities that occur in TRC sponsored events provide an opportunity for members of the various affected parties (in Canada's case government officials, former students, victims' family members, non-Aboriginal Canadians and members of religious organizations who helped run the schools) to come together and share their understandings of deeply distressing events.

How might the distributed, grassroots nature of social media applications and other more technically advanced human computation developments contribute to truth telling based reconciliation efforts, particularly in areas recovering from armed conflict? Recent work by Michael Best and his colleagues working in collaboration with Liberia's TRC provides initial answers to this question by carefully measuring and documenting the influence of rich digital media engagements, in this case involving the recording and viewing of video through a kiosk that toured the country, in support of the precursors to truth telling activities (Best et al. 2011). Although examples of media rich reconciliation efforts abound, the work in Liberia is distinct because they have attempted to empirically measure the influence of their digital media intervention.

Although endeavours in this area may appear similar to our earlier discussion of promoting human connections and empathy, it is critical to keep in mind that here we are considering human computation engagements with individuals, communities and nations that have undergone extreme trauma and are working through terribly painful memories. Although early work in the area of rich digital media and reconciliation efforts appears promising the obstacles are formidable in terms of establishing the infrastructure, both social and technical, needed to sustain and learn from initial efforts to promote reconciliation through rich digital engagement.

Human Computation and the Future of Warfare

Human computation is already taking an important place in modern warfare. From breaking up intelligence streams into pieces analyzed by thousands of people, to making key decisions in the ever-automated use of drones. For example, a drone may fly itself to an area of interest where a human operator makes a decision on what to target. This latter development, drones, and its extension into semi-automated robots is likely to bring about a significant shift in how war is waged and how participants experience conflict.

Drones provide many advantages to those using them. They greatly reduce the likelihood of casualties while providing better precision than indiscriminate use of bombing or artillery. In addition, they reduce the likelihood for mental illness associated with warfare since those controlling drones have less, if any, direct exposure to a high amount of violence.

On the negative side, drones and their robotic cousins make it much easier and less painful to enter into an armed conflict. With little risk for casualties, the human, financial, and political cost of war goes down. This approach to warfare also makes it easier for individuals to join the war effort. Driving to an office building to pilot drones or control robots is much less dangerous and life-changing than having to be deployed to a war zone. It also increases the distance between the drone pilot and those getting killed. As Lt. Col. Dave Grossman has mentioned, increasing physical distance makes it easier to kill (Grossman 1996), as the stimuli that might trigger our empathy circuit is much less likely to be felt.

Future use of these weapons could further reduce the presence of empathy in these situations, in particular if drones and robots are further automated so they make their own decisions on who to kill. An alternative to drones and robots that make decisions based on sets of rules (Arkin 2009) is to have human computation systems that bring more people in the loop to ensure that decisions to kill are well justified, and that the consequences of those decisions are known to those making decisions, and to the ones to whom they are accountable. In a democracy this would imply the right for citizens to know the human consequences of their country's actions in a war zone.

Conclusion

In this chapter we have discussed diverse ways in which human computation is currently affecting the likelihood of armed conflicts, and how it can impact de-escalation and recovery efforts. We have done this from the perspective that preventing, de-escalating, and recovering from armed conflicts is something to strive for, and that the opposite is not. We have touched on the likely future role of human computation in these areas. We have not intended to be exhaustive, but instead provide illustrative examples to lead to provocative discussions and to evoke ideas for future work. We have also tried to make clear that what matters the most is how human computation is used, and not so much what specific technologies are built.

Given our perspective, we believe human computation will yield the most positive results with respect to armed conflict when it is informed by research that helps us identify and mitigate factors that that make it more likely that armed conflict will occur, and to identify and strengthen the factors that make war less likely. In addition, human computation used to connect people in ways that engage the brain's empathy circuit will lower the possibility of violence or its recurrence.

The choices are there for us and for you on how to design the next generation of human computation systems, and how to use them. We invite you to think about peace when you consider these choices. Our world can be no brighter than the worlds we dream of.

Acknowledgements We would like to thank Natasha Bullock-Rest for her help with putting together this chapter.

References

Arkin R (2009) Governing lethal behavior in autonomous robots. CRC Press, New York

Baron-Cohen S (2011) The science of evil: on empathy and the origins of cruelty. Basic Books, Philadelphia

Best M, Long MJ, Etherton J, Smyth T (2011) Rich digital media as a tool in post-conflict truth and reconciliation. Media War Confl 4(3):231–249

Buycott (2013) Available at http://www.buycott.com/. Accessed 28 May 2013

Collier P (2007) Economic causes of civil conflict and their implications for policy. In: Crocker CA, Hampson FO, Aall PR (eds) Leashing the dogs of war: conflict management in a divided world. United States Institute of Peace Press, Washington, DC

DeRouen KR Jr, Goldfinch S (2005) Putting the numbers to work: implications for violence prevention. J Peace Res 42(1):27–45

Favate S (2012) UC Davis reaches $1M settlement with protestors over pepper spray. http://blogs.wsj.com/law/2012/09/27/uc-davis-reaches-1m-settlement-with-protestors-over-pepper-spray-incident/. Accessed 22 Mar 2013

Fletcher LE, Weinstein HM (2002) Violence and social repair: rethinking the contribution of justice to reconciliation. Hum Rights Q 24(3):573–639

Freelon D, Kriplean T, Morgan J, Bennett WL, Borning A (2011) Facilitating diverse political engagement with the living voters guide. J Inf Technol Polit 9(3):279–297

Gardner L, Young JR (2013) California's move toward MOOCs sends shock waves, but key questions remain unanswered. http://chronicle.com/article/California-Considers-a-Bold/137903/. Accessed 22 Mar 2013

Grossman D (1996) On killing: the psychological cost of learning to kill in war and society. Back Bay Books, Boston

Hagar C (2006) Using research to aid the design of a crisis information management course. Presented at ALISE SIG "Multicultural, Ethnic & Humanistic Concerns (MEH)." Information seeking and service delivery for communities in disaster/crisis, San Antonio

Hagar C, Haythornthwaite C (2005) Crisis, farming and community. J Community Inform 3: 41–52. http://ci-journal.net/viewarticle.php?id89&layouthtml. Accessed 22 Mar 2013

Hiar C (2010) How the tea party utilized digital media to gain power. http://www.pbs.org/mediashift/2010/10/how-the-tea-party-utilized-digital-media-to-gain-power301.html. Accessed 21 Mar 2013

Hourcade JP, Bullock-Rest NE (2011) HCI for peace: a call for constructive action. In: Proceedings of CHI 2011, ACM Press, New York, NY, pp 443–452

Hourcade JP, Bullock-Rest NE, Jayatilaka L, Nathan L (2012) HCI for peace: beyond tie dye. Interactions 19(5):40–47

Ivie RL (1980) Images of savagery in American justifications of war. Commun Monogr 47:279–294

Kravets D (2012) MPAA chief says SOPA, PIPA 'Are Dead,' But ISP warning scheme lives on. http://www.wired.com/threatlevel/2012/10/dodd-says-sopa-dead/. Accessed 22 Mar 2013

Lohr S (2012) The Obama Campaign's technology is a force multiplier. http://bits.blogs.nytimes.com/2012/11/08/the-obama-campaigns-technology-the-force-multiplier/. Accessed 21 Mar 2013

Long W, Brecke P (2003) War and reconciliation: reason and emotion in conflict resolution. MIT Press, Cambridge

Macdonald K (2011) Life in a day. http://youtu.be/JaFVr_cJJIY. Accessed 22 Mar 2013

Mackey R (2012) Palestinians and Israelis share images of dead and wounded children. http://thelede.blogs.nytimes.com/2012/11/15/palestinians-and-israelis-share-images-of-dead-and-wounded-children/. Accessed 22 Mar 2013

Nathan LP, Lake M, Grey NC, Nisen T, Utter RF, Utter E, Ring M, Kahn Z, Friedman B (2011) Multi-lifespan information system design: investigating a new design approach in Rwanda. In: Proceedings of the iSchool conference (iConference '11), Seattle, WA, pp 591–597

Palen L, Vieweg S, Anderson KM (2011) Supporting "Everyday Analysts" in time- and safety-critical situations. Info Soc J 27(1):52–62

Riddihough G, Chin G, Culotta E, Jasny B, Roberts L, Vignieri S (eds) (2012) Human conflict: winning the peace. Science 336(6083):818–819

Ross M (2006) A closer look at oil, diamonds, and civil war. Annu Rev Polit Sci 9:265–300

Smith JH, Mantz JW (2006) Do cellular phones dream of civil war? The mystification of production and the consequences of technology fetishism in the Eastern Congo. In: Kirsch M (ed) Inclusion and exclusion in the global arena, Routledge: New York, NY, pp 71–93

Starbird K, Palen L (2011) Voluntweeters: self-organizing by digital volunteers in times of crisis. In: ACM 2011 conference on computer human interaction (CHI 2011), Vancouver

Stewart F (2002) Root causes of violent conflict in developing countries. Br Med J 324:342–345

Ushahidi (2013) Ushahidi newsroom. http://ushahidi.com/about-us/newsroom/in-the-news. Accessed 22 Mar 2013

Weimann G (2004) www.terror.net How modern terrorism uses the internet. United States Institute of Peace, Washington, DC

Zalaquet J (1997) Truth commissions: a comparative assessment. Harvard Law School, Cambridge

Zancanaro M, Stock O, Eisikovits Z, Koren C, Weiss PL (2012) Co-narrating a conflict: an interactive tabletop to facilitate attitudinal shifts. ACM Trans Comput Hum Interact 19(3):Article 24

The Role of Human Computation in Sustainability, or, Social Progress Is Made of Fossil Fuels

Bonnie Nardi

The Probable Future

We in the Global North enjoy historically high levels of wealth and economic security. Our present abundance seems inevitable, deserved, stable. We do not believe our lives will ever be like those of people who lived during the Great Depression, or of struggling middle and lower classes in chronically economically depressed areas of the world. Yet a sober look at economic and environmental indicators strongly suggests that we are headed for a future of decreasing abundance. The goal of this chapter is to sketch a future of economic decline and discuss how we should prioritize computational resources to prevent the erosion of social gains achieved during the twentieth and twenty-first centuries. The argument is not about "saving the environment" or sustaining current lifestyles (which is impossible), but about sustaining and extending progressive social changes accrued during the period of industrial expansion. Human computation emerges as a positive force when collective human intelligence and technology are used together to solve problems and promote progressive changes (see Hourcade and Nathan, this volume). In this chapter, I make an argument for the likelihood of economic decline, and contend that information technology will serve an indispensable role in maintaining social progress. Technology has the capacity to help us defy historical patterns in which decline leads to regressive social trends in human relations.

Progressive change is built on what Clay Shirky calls a "cognitive surplus" (Shirky 2010). Shirky describes the cognitive surplus as abundant wealth that allows time for online participation such as crowdsourcing, writing fan fiction, game modding, and so on. But the notion of cognitive surplus is more general: wealth affords

With apologies to Tomlinson and Silberman (2012), of which more in a moment.

B. Nardi (✉)
University of California, Irvine, Irvine, USA
e-mail: nardi@ics.uci.edu

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4_78,

people the time and energy to do things other than meet basic needs. We have a lot of free time because our economic system is so productive. In this chapter I draw attention to one of the things some people have done with the cognitive surplus: develop and promote progressive social agendas. Some people spend their surplus watching television (up to several hours a day), but luckily for most of us, a persistent, energetic collection of various kinds of activists has been spending theirs looking out for our rights.

For most of human history, rights for workers, women, children, LGBTQ[1] persons, the disabled, the aged, the ill, and minority populations were unheard of. The dominant group (usually able-bodied men of the primary race/ethnicity) simply ran things. As an anthropologist I had the opportunity during the early days of my career to live in two such societies, one in Western Samoa and the other in Papua New Guinea. These were village-based societies with low levels of literacy, practicing agriculture with hand tools. Although communities in these cultures provided close social bonds of the sort that have eroded to some degree in industrial society, and the communities produced beautiful art, it was also true that women had no voice in governance, the disabled were ignored or ridiculed, and people with alternate sexual orientations were devalued. "Domestic violence" was not even a linguistic category of action because hitting women and children was seen as a natural mode of discipline. Old people, unproductive in a horticultural setting, were often isolated and untended as they grew feeble and sick.

Largely during the twentieth and twenty-first centuries, conditions changed as social activists addressed themselves to an Enlightenment agenda of progress, defined as equality for less equal groups. In industrial societies, workers and minorities were important groups for whom it was necessary to extend rights, in addition to women, the disabled, and so on.

We have not by any means solved the problems of inequality. Groups such as the mentally ill, homeless, and those addicted to drugs, are still often completely outside societal protections. We are a ways from true equality for all groups. Nonetheless, it is important that we recognize the immense progress that has been achieved. This progress is recent, tenuous, expensive to sustain, and far from stable. Looking to the future, equality is threatened in a scenario of economic decline because the cognitive surplus will be reduced as wealth is reduced. If we are economically stressed we will address ourselves to what will reasonably seem like more pressing problems such as food security, maintaining social order, providing shelter.

Is there a role for information technology in sustaining hard won gains in social equality? I believe there is. This chapter sketches probable causes for economic decline, followed by a discussion of what we know of "collapsed" societies historically, and how information technology might enable us to defy historical patterns. Both activist and technical activity will be necessary. Human computation should include using human cognitive capacity to understand how to deploy technical resources wisely, with compassion and social foresight—not only for instrumental

[1] Lesbian, gay, bisexual, transgender, queer.

purposes of efficiency and corporate profit. I argue that notions of human computation must, recursively, develop a clear sense of why we are using computation in the first place, understanding how it enhances human life. Vint Cerf recently called upon the ACM membership to "develop better tools and [a] much deeper understanding of the systems we invent" (2012). Cerf acknowledges that in its short history, computer science has transformed human experience, but he also notes that it has offered much less in terms of tools and practices for comprehending what it has unleashed. The call for the *Handbook of Human Computation* identifies "creativity, intuition, symbolic and logical reasoning" as central to human computation. These capacities derive from our lengthy sociobiological evolution from primitive humans to *homo sapiens sapiens*, but the speed with which we have only recently developed sophisticated information technologies along with a progressive social agenda, derive directly from the cognitive surplus.

The Wealth of Our Nations

Tomlinson and Silberman (2012) argue that "the cognitive surplus is made of fossil fuels." They remark that while Shirky takes the cognitive surplus as a given and seeks only to describe it, we must also understand how the cognitive surplus is possible, and why it occurred during the current historical era. Tomlinson and Silberman observe that our free time is not really quite so free:

> Both the free time that forms the "raw material" of the cognitive surplus and the technologies and practices of coordination that enable it to be treated as a single resource rely on huge technological infrastructures. These infrastructures are largely powered by fossil fuels.

So what will happen when we run out of fossil fuels? These fuels, in particular oil, are the most energy dense substances humanity has ever had at its disposal. One barrel of oil is the equivalent of about 25,000 hours of human manual labor (McKibben 2010). Hawken et al. (1999) observe that:

> Machines powered by water, wood, charcoal, coal, oil, and eventually electricity accelerated or accomplished some or all of the work formerly performed by laborers. Human productive capabilities began to grow exponentially. What took two hundred workers in 1770 could be done by a single spinner in the British textile industry by 1812.

And of course we have come a long way in efficiency since 1812.

But it is imperative to remember that fossil fuels are finite resources. Even disregarding the costs of environmental cleanup and health impacts the extraction and use of fossil fuels entail (see e.g., O'Rourke and Connolly 2003; U.S. National Research Council 2010; Epstein et al. 2011; IPCC 2012), the fact is that these resources are not forever. They will first become expensive, then prohibitively expensive, and then they will run out (see Hirsch et al.'s report for the US Department of Energy (2005)). Energy conservation, something we do not like to think much about, will be necessary.

Alternative sources of energy such as solar will be more fully utilized in the future. But alternative energies are no match for fossil fuels in terms of energy produced. Solar, for example, does not work well when the sun is not shining. In China, where solar energy is used much more widely than in the US, residents take short showers in the winter and put up with more discomfort than Americans and Europeans are used to (Gui, 2013, personal communication). All alternative energy sources rely on at least some fossil fuels for production and distribution (Zehner 2012). There is no energy cornucopia waiting for us to tap into; we live on a specific planet, with specific resources. We are in the process of using up those resources. O'Rourke and Connolly (2003) observe that going forward it will cost more to extract remaining fossil fuels, including escalating environmental and health costs:

> On– and off–shore exploration, drilling, and extraction activities are inherently invasive and affect ecosystems, human health, and local cultures. [Impacts] include deforestation, ecosystem destruction, chemical contamination of air and water, long–term harm to animal populations (particularly migratory birds and marine mammals), human health and safety risks for neighboring communities and…workers.

It seems likely that our reliance on fossil fuels will end in an economic decline to which we will have to adapt. This reality appears all but inevitable given several factors in addition to the finiteness of fossil fuels. First, we are doing little to alter current patterns of consumption; there is no real effort to conserve remaining resources. On the contrary, we are engaging in destructive, costly practices such as fracking to extract difficult-to-access oil and natural gas. Second, it is not feasible to expect that biofuels and other sources of alternative energy will be direct replacements for fossil fuels because their equivalencies to human labor are far below that of oil (Zehner 2012). Third, there are huge social costs to alternative energies; e.g., biofuels take land out of food production (Zehner 2012).

While it might seem that humans will once again pull the rabbit out of the hat in maintaining current levels of energy consumption through advances in technology, there are two things to consider. First, the price of the current prosperity of the Global North comes at the expense of the Global South. Our global society is one of massive inequality. Considerable global collapse already exists, once we look beyond the privileged countries of the West. Meadows et al. (1982) commented, "The view that global crises will occur in the future reflects a parochial, developed-world perspective. For two-thirds of the world's population, crises of scarce resources, inadequate housing, deplorable conditions of health, and starvation are already at hand." Our "success" as a populous species is deeply inequitable, and we can therefore expect increasing civil unrest with fewer resources with which to address it because armies, drones, and so forth, rely on fossil fuels. We can expect citizens in rapidly developing countries such as China to ramp up toward Western levels of consumption which will hasten the depletion of fossil fuels. Second, technological proposals like space-based solar farms are far in the future, if they are feasible at all. They would require great quantities of fossil fuels and would cost vast sums. Given that only 12 people have ever set foot on our nearest neighbor the moon (a long time ago), and that NASA's Mars Mission's most ambitious proposal for the near future is "the return of Martian soil and rock samples for studies in laboratories here on Earth" (NASA), it is an act of denial to suggest that we sit back and wait for technological fixes. It thus seems prudent to use some of our current cognitive surplus to ask how we can begin to design information technologies for a future of scarcity, and to engage in an exercise of prioritizing which computational resources we should guarantee in a situation of scarcity.

In this chapter I am particularly concerned with protecting social gains as the environmental dangers are well rehearsed. What could it mean to design for social sustainability? The most important point is that we must absolutely protect the global communication channels the internet has created. *Social gains in the twentieth and twenty-first centuries were made not at local or regional levels, but at national and international levels.* Historian Christine Stansell describes the global feminist movement and how it not only mobilized women but coalitions of diverse constituencies in various locales. For example, the abortion reform movement represents the efforts of "physicians, psychiatrists, and family planning professionals along with activists" (Stansell 2011). Although abortion reform predates personal digital technology, these gains were made with modern communication technologies, and the continuing battle to protect rights, which in the United States are always under siege, is waged in part with digital tools. Rapid progress on issues such as marriage equality and other LGBTQ concerns owes much to digital technology, as do other critical social struggles (Driver 2007; Gray 2009).

It might seem a no-brainer to advocate for a free internet. But how many of us really consider that even now the internet is vulnerable to bids for repressive government control in countries like China, and corporate control in countries like the US where issues such as net neutrality are far from settled? If, for example, corporations who own the infrastructure were to discount costs of connectivity to selected rich corporations that can afford to pay in volume, while charging the rest of us a premium, activists and ordinary citizens would suffer. As technologists we may feel that these decisions are outside our purview, *but they are in fact decisions made by technologists in corporations.* In this era of deregulation, government oversight is attenuated. The checks and balances of governance designed into the American Constitution (and similar documents in other countries) cannot operate if corporations assume governance. Lessig argues that "code is law" (2006), i.e., that the ubiquitous software systems underpinning commerce and communication dictate what we can and cannot do. Facebook can preserve everything it knows about you and use the information in ways it finds profitable. Amazon can offer cloud computing for vital services at low cost today but who knows what the pricing will be tomorrow? By contrast, telephony pricing was once strictly controlled by the government in order to offer universal service, and privacy protections for certain kinds of information such as health-related data were put in place before the era of deregulation. We must thus acknowledge that we are moving toward law outside democratic process. Corporations are tasked with ensuring profits, not promoting progressive social agendas (see Suarez-Villa 2012). It seems likely that a future of scarcity will make it even more tempting to increase profits by, for example, moving away from net neutrality. Thus human computation must consider how to protect and sustain a free internet. Proposals such as wireless texting and data transfer undergirded by locally controlled infrastructure should be explored and promoted (Michelucci, personal communication).

Learning from, Not Repeating, the Past

The urgency of sustaining free global communication in a future of scarcity is evident in the history contained in the archaeological record. Archaeological theories of collapse demonstrate that collapsed societies (such as the Maya, the Romans, and so on) lose complexity, *devolving to smaller scale units in smaller geographies* (Tainter 1990). When collapse occurs, the costs of governing wider areas become untenable, and social units shrink to smaller forms. It is precisely such smaller scale units (like the Samoans and Papua New Guineans I lived amongst) that assert rule by elites.

Smaldino and Richerson (this volume) note, "Larger and more connected societies can maintain more complex technologies." They comment on the fragility of connected societies: "Much of our specialized knowledge is collected by institutions, and that knowledge could rapidly vanish. Skilled people can die, books can be burned, and computers can wear out." It is only through protecting the strengths of modern information and communication technologies that connectedness, including broad coalitions of activists and citizens, can persist, uniting people to effect change and distribute control beyond small elites. In large-scale regimes of repression (such as the Soviets or the Nazis), elites maintained control by suppressing the free exchange of information and exerting stringent control over communication. Commentators such as Morozov (2013) observe that large corporations, which in the contemporary context have as much or more power as governments, are not subject to anything like the Freedom of Information Act. Are we moving toward systems in which we cannot question those who set policy? (There is some irony in the fact that Facebook, Google, etc. which traffic in information are themselves behind information firewalls.) In this historical moment of deregulation, as we cede control to corporations that furnish indispensable infrastructures without which the economy—indeed society itself—cannot operate, we must ask to what extent corporate policies protect social gains and promote continuing activism. And we must ask how we as citizens will influence those policies which operate in a universe largely outside democracy. Stansell (2011) says that feminism is "democracy's younger sister—an invocation of the linkages between progressive social forms and their necessary mutual reinforcement—as well as a reminder that protecting one involves protecting the other.

The history of social reform tells us that we do not want to return to the past, that nostalgia for simpler times is patently misplaced. It is in the current era of national and international communication and collaboration that we have rapidly won rights for the groups I discussed. Going forward, we need to use resources of human computation to prioritize sociotechnical projects to protect these rights. As Cerf said, it is important to develop a better understanding of the systems we invent, including their impacts on society. This prioritization is necessary as we envision a future of scarcity because the cognitive surplus will decline as we run out of fossil fuel. Time will be more precious. Levy (2007) invokes Thomas Aquinas to argue that time for reflection is a moral imperative, and that "self-destructive work-fanaticism" defeats efforts to live better. Without deliberately setting aside time for the most important social projects, it will be easy to fall into "work-fanaticism" that erodes the gains we have accumulated in the era of cognitive surplus.

Since our problems—including ongoing and predicted environmental damage—are global, it is essential that we sustain and promote empowered citizens of all kinds to work together to confront what will be very severe changes. Not only are rights for women, the disabled, and so on, critical for human dignity, they are crucial for empowering all people to address the massive, pervasive changes science tells us are imminent (see Greene; Hourcade and Nathan; Meier, this volume). Information technology has the capacity to empower formerly relatively powerless groups. For example, Wicks and Little (this volume) discuss ways in which people with serious illnesses make unique contributions to healthcare through participation in online forums. The authors note that people suffering stigmatizing diseases such as AIDS deployed communication technologies to organize and change the course of AIDS research. Information technologies have had a profound impact on society in extending new kinds of participation to formerly disempowered groups. A goal going forward is to recognize the fragility of sociotechnical systems that Smaldino and Richerson (this volume) point to, and the enormous potential of the collective intelligence embodied in human computation.

We will move to a new future that does not look like the past but also is quite different from the present. We will not have the economic abundance to sustain the way we live now. What will we give up? The amount of cheap consumerist junk that overflows our landfills will decrease. It is likely that we will travel less, eat more local foods, live closer to workplaces, perhaps even grow some of our own food. Proposals for edible offices (EO 2013), revivals of the ancient art of aquaponics (Rakocy et al. 2006), and urban chicken ranching may seem a little wild-eyed, but they are on the horizon (and present interesting computational problems). These changes constitute probable improvements to current ways of life. But I hope we do not give up our global network of information and communication technologies. Research areas such as crisis informatics (Starbird and Palen 2011; Al-Ani et al. 2012), collapse informatics (Tomlinson et al. 2012, 2013), and ICTD (information and communication technologies for development) (Sambasivan et al. 2010; Toyama 2010; Woelfer et al. 2011) are beginning to address how we will sustain connectivity in less than perfect conditions by studying and designing for current situations in which resources are stressed. We have much to learn from these efforts including designing digital technologies for unstable electrical grids, ensuring communication during emergencies, and orienting ourselves to plan ahead to mitigate and even forestall problems.

Coda

As I was working on this chapter, the power on most of my campus was knocked out for several hours (something that had not happened in the 10 years I have been at the University of California, Irvine). I wrote in the glow of my battery-powered laptop, mindful of the limited resource I on which I was now relying. As it happened, during the outage, Terry Winograd, an eminent scholar of human-computer interaction, was scheduled to give a talk to my department. We sat in a dim meeting room listening to Professor Winograd discuss his amazing life's work in which human computation has figured prominently. Professor Winograd had no slides because of the failed power, but his talk was an inspirational historical accounting of progress in human-computer interaction. Perhaps prophetically, *the internet was still working*—the university had decided that backup power would be allocated to connectivity during outages. As Professor Winograd spoke, we could tweet the event and some of the audience looked up things Professor Winograd was discussing, such as the old Eliza program with which some younger students were unfamiliar. It was a little warm and dark during the lecture, but we were enlightened! This occurrence was like a tiny visit to the future in which we will be making decisions such as: will it be slides and air conditioning or connectivity? The university had decided in advance on connectivity—surely the right choice given that had the emergency been more dire, communicating with the world and finding information would be the priorities. If we are to defy historical patterns of collapse in which social units devolve to more local forms affording less protection of progressive social agendas, we will be using the powers of digital technologies of information and communication to do so. Unlike the Maya and the Romans who did not have foresight attained through research in archaeology and history to guide them, we can assess likely future problems now, and plan for them. We understand that sustaining social gains rests on information and communication transmission at national and international scale, and we can prioritize resources in a future of scarcity just as my university prioritized internet connectivity.

The objective of this chapter has been to argue that social progress is made of fossil fuels. Once we realize the basis upon which this progress rests—and that it is not a given and it is not forever—we can plan to self-consciously expend resources to extend and maintain progressive social agendas. Net neutrality is one pertinent technological issue but there are many others including promoting broad-based computer science education to ensure that control of digital technology is not confined to technical elites, deciding who gets access to rare earth metals, encouraging citizen participation in the control of computing infrastructure, and continuing to develop innovative means of crowdsourcing to leverage whatever cognitive surplus we will have in the future. In short, at least some cycles of human computation should be used to plan for a future of scarcity in which economic decline will force us to use a smaller cognitive surplus wisely. This is just the sort of wicked problem that stands as a challenge to human computation which we can take on now, in an abundant present.

Acknowledgments I would like to thank the *Handbook* editor, Pietro Michelucci, for his thoughtful suggestions on improving this chapter, and his dedication to producing a *Handbook* with many voices. Sunny Gui provided useful background on life in China. Barton Friedland, Caitie Lustig, and Six Silberman gave careful readings of an earlier draft which produced helpful suggestions.

References

Al-Ani B, Mark G, Chung J, Jones J (2012) The Egyptian blogosphere during the revolution: a narrative of counter-power. In: Proceedings conference on computer-supported cooperative work, ACM Press, New York, pp 17–26

Cerf V (2012) Where is the science in computer science? Commun ACM 55(10):5

Driver S (2007) Queer girls and popular culture: reading, resisting, and creating media. Peter Lang, New York

EO (2013) http://www.biotope-city.net/gallery/edible-office-concept

Epstein P, Buonocore J, Eckerle K, Hendryx M, Stout B III et al (2011) Full cost accounting for the life cycle of coal. Ann N Y Acad Sci 1219:73–98

Gray M (2009) Negotiating identities/queering desires: coming out online and the remediation of the coming out story. J Comput Mediat Commun 14(4):1162–1189

Hawken P, Lovins A, Lovins LH (1999) Natural capitalism: creating the next industrial revolution. Little, Brown, Boston

Hirsch R, Bezdek R, Wendling R (2005) Peaking of world oil production: impacts, mitigation and risk management. U.S. Department of Energy, National Energy Technology Laboratory, Morgantown

Intergovernmental Panel on Climate Change (2012) Summary for policymakers. Managing the risks of extreme events and disasters to advance climate change adaptation. Cambridge University Press, Cambridge

Lessig L (2006) Code: and other laws of cyberspace, version 2.0. Basic Books, New York

Levy D (2007) No time to think: reflections on information technology and contemplative scholarship. Ethics Info Technol 9:237–249

McKibben B (2010) Eaarth: making a life on a tough new planet. Times Books, New York

Meadows D, Richardson J, Bruckmann G (1982) Groping in the dark: the first decade of global modelling. Wiley, New York

Morozov E (2013) To save everything, click here: the folly of technological solutionism. PublicAffairs, New York

NASA (2013) http://mars.jpl.nasa.gov/programmissions/overview/

O'Rourke D, Connolly S (2003) Just oil? The distribution of environmental and social impacts of oil production and consumption. Ann Rev Environ Res 28:587–617

Rakocy J, Masser M, Losordo T (2006) Recirculating aquaculture tank production systems: aquaponics: integrating fish and plant culture. Southern Regional Aquaculture Center Publication no 454

Sambasivan N, Cutrell E, Toyama K, Nardi B (2010) Intermediated technology use in developing communities. In: Proceedings of the SIGCHI conference on human factors in computing systems, ACM Press, New York, pp 2583–2592

Shirky C (2010) Cognitive surplus: creativity and generosity in a connected age. Penguin Press, New York

Stansell C (2011) The feminist promise: 1792 to the present. Modern Library, New York

Starbird K, Palen L (2011) Voluntweeters: self-organizing by digital volunteers in times of crisis. In: Proceedings of the SIGCHI conference on human factors in computing systems, ACM Press, New York, pp 1071–1080

Suarez-Villa L (2012) Globalization and technocapitalism. Ashgate, Surrey

Tainter J (1990) The collapse of complex societies. Cambridge University Press, Cambridge

Tomlinson B, Silberman MS (2012) The cognitive surplus is made of fossil fuels. First Monday. November

Tomlinson B, Silberman MS, Patterson D, Pan Y, Blevis E (2012) Collapse informatics: augmenting the sustainability & ICT4D discourse in HCI. In: Proceedings of the SIGCHI conference on human factors in computing systems, ACM Press, New York, pp 655–664

Tomlinson W, Blevis E, Nardi B, Patterson D, Silberman MS, Pan Y (2013) Collapse informatics and practice: theory, method, and design. Forthcoming in ACM Transactions on Computer-Human Interaction

Toyama K (2010) Human-computer interaction and global development. Now Publishers, New York

Woelfer J, Iverson A, Hendry D, Friedman B, Gill B (2011) Improving the safety of homeless young people with mobile phones: values, form and function. In: Proceedings of the SIGCHI conference on human factors in computing systems, ACM Press, New York, pp 1707–1716

Zehner O (2012) Green illusions: the dirty secrets of clean energy and the future of environmentalism. University of Nebraska Press, Lincoln

Human Computation: A Manifesto

Pietro Michelucci

Introduction

To live in the space of hope is to exist in an uncertain future.

We have become complacent in our circumstantial despair and now avert our eyes from the mounting challenges posed by the explosion of innovation in this digital age. Indeed, it is easier to believe that "powers that be" or even technology itself will deliver us beneficently from extinction. But should we accept on faith that all sovereign nations and rogue states have employed provably foolproof safeguards against unintended nuclear missile launches because if even the slightest chance of failure existed, the consequence would be so grave as to compel such safeguards? And what of Thomas Friedman's (1999) democratization of technology? The widespread availability of increasingly potent capabilities has empowered individuals and small groups with state-level capabilities. How does a people safeguard against ubiquitous omnipotence?

In the remainder of this introduction, we consider the growing potential for threats due to existing and emergent technologies, examine proposed strategies for managing them, and consider how Human Computation (HC), the study of humans as computational elements in a purposeful system, may be instrumental for mitigating future such risks. Following the introduction, we examine the maturity of human computation as both a practice and a discipline. This analysis informs a proposal for technical maturation as well as a formal definition of the field and its distinguishing qualities, all in service of accelerating research and ensuring responsible use of resultant capabilities. Though the ideas in this chapter may be informed by engagement with the HC community, this manifesto represents a personal perspective.

P. Michelucci (✉)
ThinkSplash LLC, Fairfax, USA
e-mail: hchandbook@gmail.com

P. Michelucci (ed.), *Handbook of Human Computation*, 1021
DOI 10.1007/978-1-4614-8806-4_79,

Fig. 1 A backpack for the US-manufactured Mk-54, a man-portable tactical nuclear weapon (Photo Source: http://en.wikipedia.org/wiki/Suitcase_nuke; licensed under Creative Commons attribution CC BY-SA 3.0)

The Democratization of Power

Consider that a sophisticated terrorist group could employ a single person with a "suitcase nuke" (See Fig. 1; Woolf 2010; Horrock 2001) to devastate the center of a metropolitan area (Bunn and Maslin 2011). The single greatest barrier to constructing such a weapon of terror is the acquisition of weapons-grade fissile material (Horrock 2001), such as highly enriched uranium (HEU). As it turns out, during the period from 1993 to 2007 the International Atomic Energy Agency reported 18 incidents of HEU trafficking (see Sanfilippo et al., this volume), some of which involved seizures of kilogram scale quantities (IAEA 2007).

Consider the polymerase chain reaction (PCR). This is a technique for high volume replication of DNA, the molecule that encodes the genetic programming for all known life, including most viruses. PCR requires a series of carefully calibrated temperature changes over a period of time. Such a process is enabled by a device called a "thermocycler", which is basically a high-precision, programmable oven. If you would like to try it in the comfort of your home, you can purchase a kit (Fig. 2) for $599.00 at Amazon.com. For safety, please replicate only harmless DNA.

Tiger by the Tail

Nuclear weapons management and genetically engineered pandemic viruses are among a growing list of *known* risks. What of the *unknown* risks? In their discussion

Fig. 2 A low-cost thermocycler

of cumulative culture as a collective memory for preserving and advancing technology, Paul Smaldino and Peter Richerson (this volume) aptly observe that no single human being today knows how to build a modern computer from scratch. This calls attention to our reliance on both communities of distributed knowledge and the infrastructure that supports the propagation of such knowledge, and hence our vulnerability to a breakdown of either. The situation is far worse. The insularity of expert knowledge has become such that even within a narrow field of study, the rate of advancement is so great that it is nigh impossible for researchers to maintain broad awareness of the intradisciplinary consequences of their work, not to mention the combinatorial explosion of disruptive possibilities that arises when new technologies are combined across fields. Thus, even with the most conservative policies in place, we could not presently appreciate the deep and thorough implications of our technological pursuits. Technology today is a tiger held by the tail.

An Aristotelian Oath

At the turn of the millennium, Bill Joy, the co-founder of Sun Microsystems, wrote a sobering Wired article (Joy 2000) in which he sought to rouse the rest of the world to the looming dangers of unchecked technological advancement, particularly in the areas of genetics, nanotechnology, and robotics (abbreviated "GNR"), with a focused concern about self-replication in all three domains. His answer to the existential threat has been one of relinquishment—that is, advocating that we simply give up certain perilous technological pursuits, and verify compliance by embracing a "strong code of ethical conduct".

Measures and Countermeasures

Joy's efforts to begin this conversation in earnest led to a panel discussion event at the Washington National Cathedral called "Are We Becoming an Endangered Species? Technology and Ethics in the Twenty First Century". In this discussion, Raymond Kurzweil (2001) countered Joy's noble, though perhaps unrealistic vision of consensual relinquishment by suggesting that we proceed gingerly: "…the only viable and responsible path is to set a careful course that can realize the benefits while managing the risks." In supporting this risk-management view, Kurzweil appealed to the observation that new technological threats do not arise in a vacuum, and that there is a commensurate coevolution of technological means to control them. He took computer viruses as a case study, observing that digital disease has remained in check due to the ebb and flow of measures and countermeasures (e.g., anti-virus software). From this, Kurzweil surmises that giving 15 billion dollars to NIH and NSF to spend on countermeasures to address new technology threats would go a long way toward keeping Joy's GNR risks in balance. And perhaps it would, though it has not been tried 12 years later.

Irreversible Disruption

Kurzweil's view on countermeasures, however, underplays the role of ecology, because it ignores that both technological risks and controls exist within a context that critically influences outcomes. Technology, humanity, and the planetary environment in which they coexist form a closed dynamical system. Such complex systems exist in an equilibrium state. As such, they exhibit sensitivity to bifurcation (e.g., Silvert 2002). That is, when a perturbation occurs that causes tolerances to be exceeded, there is a destabilizing and potentially irreversible effect. Though such a system will likely settle into a new equilibrium state, it may be qualitatively different than the former one. For example, there is a level of ionizing radiation above which most organisms cannot survive. Consider a new, runaway technology that irradiates the biosphere. If fatal levels of radiation are absorbed before a protective technology can be developed, then humanity faces extinction. The key point here is that as technologies extend their impact to a global scale, they are more likely to disrupt homeostatic factors in irreversible ways. In the microcosm of Kurzweil's computer viruses, there has always been the option to "cheat", that is to transcend the virtual domain within which these viruses are transmitted by physically disconnecting computers from each other. Such a cheat does not exist in the physical world.[1]

[1] In personal correspondence, Michael Witbrock has aptly observed that this is not strictly true; that if we were willing to transcend our planetary context, space colonization might afford a similarly cheat.

Can Machines Save Us?

How then do we mitigate the existential risk of irreversible disruption? Perhaps we leverage the inimitable power of the very emergent technologies we fear. For example, could we possibly build a machine that is smart enough to save us from our own undoing? And to what consequence? Raymond Kurzweil is well known for popularizing John von Neumann's notion of a technological singularity (Kurzweil 2006), the point in time at which machine-based intelligence will exceed human intelligence. This is often misconstrued to represent the demise of the humanity. In fact, "singularity" is a term borrowed from cosmology to refer metaphorically to a black hole's event horizon, beyond which nothing is knowable. The implication is that we cannot predict what life would be like after such an event. Most theories anticipating the near-term occurrence of a singularity are predicated on the belief that computer processing speed, in terms of calculations per second, is somehow tantamount to intelligence. In this view, a simple extrapolation of Moore's Law, which predicts a doubling of computational speed every 1.5 years, suggests that home computers will exceed the intelligence of humans by the year 2020. Kurzweil, however, acknowledges that processing speed is not enough—that to manifest increases in processing speed as superior intelligence, it will be necessary to build machine-based systems that emulate a precise physical model of the human brain. But is such a model truly within reach?

The Elusive Singularity

Our understanding of the human brain has increased dramatically over the past decade. We are developing a more detailed understanding of the role of glial cells as an adjunctive communication network to neurons and the existence of stigmergic hormonal processes used to communicate locally in the brain (see Larson-Prior, this volume). We, thus, increasingly view the brain as a complex system of intertwined systems. We are also now, for better or worse, able to use brain scans (fMRI) to measure consumer preference, detect lies, and recognize increasingly complex thought patterns. But recognizing patterns in the brain tells us no more about how those patterns formed than recognizing an animal species tells us about its complex ontogeny. Indeed, these advancements suggest, perhaps more than anything, that there is more to learn about the brain than we previously realized before it could be replicated in-silico (or the synthetic substrate du jour).

But even if we could embed a functional model of the human brain in a computer and run it a thousand or even a million times faster than biological brain, wouldn't it still think only as well as a human? In other words, wouldn't the complexity of its thought processes be the same and wouldn't its capacity for knowledge remain unchanged? Furthermore, there is no evidence to suggest that merely adding artificial neurons to such an artificial brain would make it smarter or that we would have any idea how to usefully connect it to other artificial brains to produce superhuman intelligence. This is not to say that machine-based intelligence will never exceed, in some manner, human intelligence, but rather that the key enabler of such advancement will likely not be processing speed or even brain replication, but rather a deep and sophisticated understanding of how intelligence manifests within a complex network such as the brain. Only then should we start to worry about machines saving us.

A Singularity with Humans-in-the-Loop

But perhaps now is the beginning of "then". Human computation represents the prospect of a different kind of technological singularity, and perhaps one that is more imminently attainable. Indeed, the opportunity exists today to sidestep the issue of replicating human intelligence in machines and turn our attention more directly to the study of methods by which unprecedented cognitive capabilities could be achieved through a carefully conceived combination of biologic human intelligence. In other words, we already have computational agents that are as smart as humans—they're called "humans". Let us then investigate in earnest how we might support the interaction of these agents within a technology-mediated infrastructure toward a degree of cognitive sophistication heretofore unseen. Indeed, there is preliminary evidence (see chapter "Organismic Computing", this volume) to suggest that, under the right circumstances, large groups of people can exhibit greater synergy than smaller groups. If we can identify and implement such circumstances in sustainable and purposeful ways, then perhaps we can induce a phase transition in humanity—a fundamental change in its collective capability without loss of individuality.[2]

When Technology Is a Solution

Bill Joy (2006) once said in a Ted Talk, "You can't solve a problem with the management of technology with more technology." To this we might add: "…unless it is problem-solving technology." Even if one does not fully embrace a speculative future with collective superhuman intelligence, there are many practical examples today of human computation technology being used to solve problems, some of which were potentially caused by technology in the first place. For example, Patrick Meier (this volume) reports on the use of Ushahidi, a crowd-powered crisis management system, to mitigate the damage caused by Hurricane Sandy and Typhoon Pablo. The frequency of intense storm systems such as these is believed to be increasing as a result of climate change (Knutson et al. 2010), which itself has been linked to democratization of external combustion engine technology and the

[2] This notion of a phase transition in humanity derives from the canonical notion of a physical phase transition, in which there is a change from one state to another without a change in composition.

resultant carbon dioxide emissions (Solomon et al. 2009). Human computation is also being used to improve outcomes in infectious disease (see Wicks and Little, this volume), for which technology has also been implicated as a cause (Breiman 1996). Indeed, the Haym Hirsh's essay "Human Computation in the Wild" (this volume) is rife with examples of crowd-enabled systems solving problems, even in the pre-digital age. As our understanding of human computation becomes more sophisticated and we gain experience in its application, it is not unreasonable to expect that it will become more prevalent in our arsenal of coping strategies.

What Worked for Linux

It is noteworthy that the "top-down" solutions to looming risks proposed by Bill Joy and Ray Kurzweil[3] originated from the Washington National Cathedral, though perhaps more figuratively than literally. Eric S. Raymond (1997) penned a catalytic essay called "The Cathedral and the Bazaar", about open source software (OSS) development. In this essay, Mr. Raymond extolled the virtues of bottom-up software development in which hundreds of disorganized software developers around the world ("the bazaar") volunteered small bits of time in piecemeal fashion, in what resulted ultimately in Linux,[4] arguably the most popular operating system in the world. What was most notable to Mr. Raymond was that such a distributed effort with so much left to chance could succeed so brilliantly where the traditional, top-down ("cathedral") model of software development had failed. Ironically, Pavlic and Pratt (this volume) have identified many parallels between human behavior in OSS and adaptive eusocial behavior in ants that endows them with emergent collective capabilities. Thus, it is a thesis of this manifesto that human computation (as a general class of organized distributed behavior) is the metaphorical bazaar to Joy's and Kurzweil's cathedral, and as such, may more robustly and adaptively address the existential risks of tomorrow and the practical issues of today.

A Plan for Conscientious Progress

It is one thing to speak evangelistically of progress and quite another to realize it. The remainder of this chapter serves as a proposal for the conscientious advancement of human computation as both a practice and a science. The next section

[3] Bill Joy and Ray Kurzweil are widely respected as technical luminaries of our times. It is only on the shoulders of these prescient giants that a context for advancing human computation is formulated herein.

[4] Linux underlies the Android operating system.

briefly outlines practical considerations for advancing the state of the art. This is followed by an analysis of human computation as a formal discipline, which forces some stakes into the ground. Finally, recognizing the inevitability of growth in this new field, we consider ways to improve the likelihood that the technology is developed and used responsibly.

Technical Maturity for Progress

We need repeatable methods. Due to the logistical complexity of human participation in human computation systems, we cannot simply employ extant software engineering methods to accomplish anything more than simple crowdsourcing. This is beginning to change (see Morishima, this volume), but in order to progress at a reasonable pace, putting more effort into HC research and less into HC engineering, we need a basic technical maturity. As it is, each novel manifestation of human computation requires a ground-up development effort. Thus, we will need the HC equivalent of a printing press in order for research to move beyond a geologic rate. The following is a representative list of technical desiderata that would be expected to enable a more mature HC practice.

Infrastructure

HC needs a technological state space, a persistent memory for HC systems that does not rely upon the fallible memory of humans. It would further benefit from an "always-on", generalized load-balancing architecture that is robust to the asynchronous and unpredictable availability of human agents. HC also needs service-oriented protocols that permit function calls to these asynchronous humans. Crowd Agents and related methods (see Lasecki and Bigham, this volume) constitute a significant advancement in this direction.

Programming Language

HC needs a development platform that includes an HC programming language, or at least new HC extensions to existing languages. It needs middleware with common classes of crowdsourcing algorithms, implementations of design patterns (see Greene's introduction to the Techniques and Modalities, this volume), and an associated API; and each platform should have associated open source software development projects to create and curate interface elements suited to human participation in HC systems. Ultimately, function calls should require only a specification of the information processing requirements of the human task (e.g., the input, expected output, processing time requirements, etc.); execution should be handled by platform-specific runtime modules that self-adapt to the interface affordances of the execution platform.

Integrated Development Environment (IDE)

HC needs a single integrated environment for development, debugging, performance testing, and execution. Generic, adaptable, and extensible IDEs exist today. Any one of them could be modified to serve as an interface for HC software development.

A three-phase debugging paradigm would help minimize the expense of utilizing actual human computational resources. In this paradigm, phase one debugging would involve farming out tasks to simulated human agents. This would enable a low-cost evaluation of basic system performance by simulating different degrees of variability in human response time and availability. In the second phase, a combination of machine and human agents could be employed in which the HC behaviors of a small proportion of real human agents would dynamically induce more human-like behavior in the machine agents. This phase would be suitable for testing the ability of the system to properly handle the expected information content returned by humans. The final phase of debugging would employ only human agents to ensure that the system would behave predictably in the context of both system performance and information processing. In this three-phase model, minimal use of human resources during testing would be assured by only advancing to subsequent phases of debugging when previous phases, which involve less human involvement, have passed without errors.

Toward a Common Framework

This brief exposition is not intended as a formal and precise specification for technical maturity, but rather to be suggestive of the kind of technologies and approaches that would lead to repeatability, modularity, code reuse, and cross-platform execution that is now commonplace in software engineering. A perusal of the Infrastructure and Architecture section (this volume) reveals that some of these pieces are already coming to fruition. Ultimately, it may be the binding of these pieces within a common development framework that gives rise to rapid HC development. However, the degree of community collaboration necessary for such technical coalescence may first require greater conceptual coalescence and maturity as a discipline. Indeed, that is the topic of the next section.

Toward a Discipline

In 1995, Donald Liles and his colleagues at the University of Texas in Arlington realized that Enterprise Engineering was beginning to distinguish itself from related fields, and took that as an opportunity to reflect on the significance and characterization of such an occurrence. This exercise in community self-reflection coalesced the views of such paradigmatic thinkers as Thomas Kuhn, Peter Keen, and Gavriel Salvendy into an elegant treatise of disciplinary emergence. According to Liles et al. (1996), a discipline represents a worldview, a community, and a set of practices that generate knowledge, which in turn further informs those practices. As with Enterprise Engineering, the emergence of Human Computation (HC) as a discipline is not an end; it is a process of reorganization to accommodate a distinct and increasingly prevalent new approach.

Liles et al. (1996) proposes a list of six defining characteristics for a discipline, which includes a focus of study, a world view or paradigm that binds the community, a set of reference disciplines from which the new field originated but now distinguishes itself, unique principles or practices, an active research agenda, and societal constructs, such as the deployment of education and promotion of professionalism. Herein we seek to describe the state of Human Computation according to those characteristics in order to better understand where HC stands today as an emerging discipline and to help inform its future course.

Focus of Study

According to Liles, disciplines emerge to solve new problems not addressed by existing disciplines. Thus, the focus of study stems from the fundamental question being addressed by the discipline. For Human Computation, in all of its incarnations, the central question distills to:

> How do we create new capabilities and derive knowledge through human participation in computational systems?

The pursuit of answers to this question leads to a "body of knowledge, principles, and practices" pertaining to the design and analysis of human computation systems.

Unique Worldview

A discipline manifests a unique perspective from which its constituents view the world. This perspective determines the framework of practice and is sufficiently complex to be divided into sub-disciplines. In HC, the uniqueness of this perspective arises in part from the unusual combination of five assumptions:

- **Behavioral**—Human Computation employs and studies human interaction.
- **Complex**—Human Computation necessarily involves a system of humans, which are themselves complex dynamic systems. It is within the structure of this complexity that new capabilities or intelligence may emerge.
- **Ecological**—Human Computation presumes an ecological perspective because participation is situated. Individual cognition and agency are part of an interactive system within which they exhibit reciprocal influence with other agents, both machine and human, as well as the environment.
- **Purposeful**—Human Computation is purposeful at the agent level, system level, or both, whether the goals are imposed overtly or manifest simply as a tendency toward some equilibrium state.
- **Engineered**—Human Computation is the product of engineering, whether information processing architecture, mechanism design, or simply a technosocial infrastructure that gives rise to new patterns of behavior. The engineer may be a person, a system, or even a process, such as evolution.

These five assumptions give rise to a multitude of sub-disciplines that derive from existing parent disciplines but constitute sub-disciplines by their *specific and unique application to human computation*. Among these are:

- *Theory of Computation*—the formal analysis and performance characterization of algorithmic behavior that involves human computational elements (see Crouser et al. this volume, for a ground-breaking foray into this sub-discipline)
- *Computer Engineering*—the development of a scalable and reliable computational infrastructure to support computation that combines machine and human processing elements (see chapter on "Crowd Agents", Lasecki and Bigham, this volume)
- *Distributed Computing* (particularly multi-agent architecture)—the theory and design of multi-agent computing systems in which some agents are humans (see Castelli et al., Durfee this volume)
- *Software Engineering*—a systemic approach to the design, development and testing of software that runs on human computational infrastructure (see Morishima's HC development platform, this volume)
- *Human-Computer Interaction*—the study, planning, and design of human interaction specific to the provision of information processing support to an HC system (see Reeves, this volume)
- *Artificial Intelligence*—the design of HC systems that exhibit intelligence (see Heylighen, this volume)
- *Machine Learning*—the design of machine-based algorithms that incorporate humans as either a source of learning bias or as dynamic resources for augmenting machine capabilities (e.g., human-based genetic algorithms—see Nickerson, this volume)

- *Cybernetics*—the control systems analysis of the constraints and possibilities of closed-loop human computation systems (see Nechansky, this volume)
- *Motivation Theory*—the theory of human participation behavior in HC systems and engineering incentive structures that maximize participation volume and quality (Mason, Ghosh, Reed et al., all this volume), and creating systems that themselves exhibit goal-directed behavior (see chapter "Organismic Computing," this volume)
- *Evolutionary Biology*—the study of evolution as an algorithmic approach to human computation (see Nickerson, this volume)
- *Cognitive Science*—the study and architecture of HC systems that think and the analysis of their thought processes (see Blumberg's chapter "Patterns of Connection"); also the comparative analysis of information processing capabilities between machines and humans (see Crouser et al., this volume)
- *Entomology*—the study of eusocial insect behavior as both an explanatory and generative model of superorganismic behavior in humans (see Moses, Pavlic & Pratt, both this volume)
- *Organizational Science*—that study of organizational workflow models as candidate HC architectures (see Brambilla & Fraternali)
- *Social Informatics*—the analysis of human social behavior in HC systems via quantitative modeling (see Lerman's introduction to the Analysis section in this volume)
- *Knowledge Engineering*—the encoding and locus of knowledge in HC systems (see Gil, Witbrock, both this volume)
- *Cultural Anthropology*—the use of culture as a model of transcendent state space for collective knowledge (see Smaldino & Richerson, this volume) and cultural evolution as a model for collective problem-solving (see Gabora, this volume)
- *Psychopathology*—the study of mental illness applied to the classification, diagnosis, and treatment of behavioral pathology in societies and superorganisms (see Blumberg & Michelucci, this volume)
- *Social Psychology*—the role of group dynamics and social cognition in collective intelligence and group efficacy (see Woolley & Hashmi, this volume)
- *Information Theory*—the ability to characterize the transformation of information (see Gershenson, this volume) by humans to inform the design and understanding of HC systems
- *Epistemology*—the interplay and representation of belief and truth in human-based computation (see Nechansky, this volume)
- *Cognitive Neuroscience*—the use of biological models of cognition (e.g., brains) to inform the design of distributed thinking systems in which networked nodes are human (see Larson-Prior, this volume)

The existence of such numerous and diverse sub-disciplines suggests that the underlying worldview is sufficiently substantive to support a discipline (Keen 1980). However, the most telltale sign of Human Computation's disciplinary maturity may be the active research referenced within these sub-disciplines.

Fig. 3 Five reference disciplines of human computation

To be clear, this list of sub-disciplines does not implicate Human Computation as a transdisciplinary field. The intended direction of applicability is from each of the parent disciplines to HC—not the reverse. That is not to neglect the potential applicability of HC to these or other disciplines, but that is not the relationship being conveyed here.

Reference Disciplines

Though new disciplines may emerge to solve problems not addressed by existing disciplines, they critically rely upon the knowledge, methods, and tools of the primary disciplines from which they borrow—their "reference disciplines". As indicated above, numerous disciplines contribute to Human Computation; however, only five of these (see Fig. 3), seem truly foundational to HC. In the absence of these reference disciplines the pursuit of human computation would seem untenable.

While the existence of these reference disciplines enables the pursuit of HC, it is their formal acknowledgement that supports the broad acceptance of Human Computation in the scientific community by anchoring its conceptual framework in established bodies of work.

Principles and Practices

Human Computation borrows, perhaps most directly, from Software Engineering in terms of theory, abstraction, design, and implementation. While the unique characteristics of HC will likely cause these principles and practices to evolve in new directions, this sub-discipline of the Computer Science reference discipline, serves as a reasonable starting point.

Research Agenda

An active research agenda with diverse lines of inquiry is fundamental to a thriving discipline. As evidenced by the rich and diverse body of work conveyed in this volume, HC seems to meet this criterion. However, it constitutes such an expansive and fertile space of research that designating sub-agendas easily becomes an arbitrary exercise in framing the dimensionality of the problem space. Nonetheless, consideration of the foundational assumptions (see above) of HC helps cluster active research into sensible sub-agendas. Taking this approach, we end up with these key research areas within Human Computation:

- Participation—incentivizing participation and modeling interactions in HC
- Application—architecting purposeful HC systems
- Efficacy—engineering circumstances conducive to synergy
- Security—creating HC systems robust to surreptitious participation
- Platform—creating tools and infrastructure to support HC development
- Analysis—the study of HC system behavior

Subdividing the research space in this way helps us locate our own research efforts among related work, and engage within interested sub-communities.

Education and Professionalism

The emergence of a discipline is reflected as much by the community structures in place to support sharing and learning as by its conceptual distinctiveness and technical maturity. Though in the near term this handbook may serve as a catalyst for new research, in the longer term it will persist as a community knowledge base of general principles, key ideas, and emerging research. Its diverse interdisciplinary authorship will serve to draw out latent HC community members from related disciplines. Other community structures such as a forthcoming interdisciplinary journal of human computation and a new professional society within IEEE will serve as a home to those expatriates.

Though HC has a rich history of workshops (e.g., HComp, SocialCom, CI, SoHuman, etc.), they have been historically hosted by conferences from reference disciplines, and populated primarily by participants originating from those disciplines. However, coinciding with the publication of this first edition handbook, the First AAAI Conference on Human Computation and Crowdsourcing will be held in November of 2013.

In my personal experience, using the term “human computation” in public produces blank stares and confusion. Though not part of Liles’ exposition on education, it might be worth considering that general acceptance of a discipline requires not just the formal education of scientists and practitioners, but also dissemination to the general public of a broad-based popular understanding about what human computation does. Proactive public engagement on this would serve to reduce potential misunderstandings about human computation, whether at the definitional level (e.g., "is this about humans using computers?") or the implementation level (e.g., "is this dehumanizing?").

The Birth of a Discipline

Most, if not all, of the foregoing indicators seem consistent with the present emergence of Human Computation as its own discipline. It is worth mentioning though, that while formal recognition as a discipline may seem beneficial to HC, distancing prematurely from parent disciplines carries its own potential liabilities,[5] such as becoming disconnected from communities and related work that have helped sustain HC up until now. Thus, it is critical that we proceed gingerly. Indeed, HC may be best served by preserving strong connections to related disciplines and tempering certain canonical aspects of disciplinary maturation. Perhaps we can borrow from the successes of Cognitive Science, a notable success story among "interdisciplinary disciplines".

This cautionary note notwithstanding, it is still of interest to consider the implications of the apparent disciplinary trajectory of HC based on the above analysis. The primary effect we might expect to result from this is an inflection point in the rate of advancement of HC research and development. Even over the relatively brief 9-month course of this handbook's development, I have borne witness to numerous interdisciplinary "epiphanies" within the book community microcosm. These revelations seemed to result from author exposures to HC-related work across communities that rarely interact. Even if one cannot reasonably extrapolate to the broader community from such anecdotal evidence, it is difficult to ignore the growing interest in this field. So what's next? Perhaps academic programs.

A Department of Human Computation?

Today, formal studies in human computation tend to occur as seminar courses within human-centered systems or distributed information systems programs in computer science departments. However, as the HC community coalesces around new research, tools, and community structures, we might expect to witness the emergence of formal programs in HC as we did with Cognitive Science in the late 1980s and early 1990s. These new programs could be driven by dedicated leaders and faculty who would, through their participation in such programs, begin to identify themselves more formally with the HC discipline, perhaps even referring to themselves as "human computation scientists".

---

[5] I owe special thanks to Mary Catherine Bateson for providing a valued counterpoint to the potential benefits of disciplinary identity, as well as for pointing out the relevance of public education for a new discipline.

Is there sufficient interest and activity in the field to support a dedicated department of human computation? There's certainly an industry demand for people who are capable of leveraging the power of the crowd. This suggests a commensurate demand for vocational degrees in crowdsourcing. Such a pull from industry may be enough to compel the right visionary dean to back a new department.

Better Yet, an Interdisciplinary Program

On the other hand, perhaps it would make more sense to advance formal education through interdisciplinary degree programs. This would obviate the risk of departmental isolation, which could be fatal for such a conceptually distributed field as HC. Indeed forcing people to choose between an established related reference discipline and a new speculative discipline could reduce the population of both. Furthermore, the barrier to entry for interdisciplinary programs is much lower than for new departments. Among other things, interdisciplinary programs often have minimal requirements for office space and new hires as resources are often shared across existing departments.

The evolution of such programs might begin with concentrations, for which university certificates are awarded. This might be followed by the emergence first of graduate degrees and eventually by undergraduate degrees as the field gradually migrates from specialized to mainstream status. The promotion of such interdisciplinary programs in HC could arise through the efforts of an HC professional society (such as the aforementioned one), by developing academic program requirements for different degree award levels. These core requirements would specify which reference disciplines should have representation among the program faculty, and include guidelines for ensuring a suitable core curriculum as well as recommended course materials. The requirements would be sufficiently flexible to strike a balance between ensuring core competencies in graduates and allowing universities to differentiate their programs. The resultant program specifications could thus be used in turnkey fashion by universities to implement unique programs, taking comfort in the standardization and broad acceptance that would derive from such a community-driven approach. Of further interest to such academic programs might be the conscientious oversight of HC technologies and their development, which is considered next.

A Conscience Committee

Human computation represents a promising means for solving extant and future problems. However, like any new technology it bears its own risks—through vulnerabilities, misuse, and outright abuse. Several contributors to this volume have begun to explore such categories of risk. Dan Thomsen considers human computation systems in the context of cyber security principles to anticipate susceptibility to coordinated attacks as well as vulnerabilities to subtler, though perhaps more insidious, surreptitious participant behaviors. James Caverlee, in his policy chapter on labor standards (this volume; see also Witbrock's introduction to the book section on Infrastructure and Architecture), raises the specter of exploitation and other abuses arising from the commoditization of human computational labor. And from a moral perspective, Juan Pablo Hourcade and Lisa Nathan (this volume) warn against the possibility that human computation could be used as a coercive force. These thought-provoking analyses serve to bootstrap a discussion that will endure for the foreseeable future.

It is impossible to anticipate all of the technical risks and ethical dilemmas that will arise as human computation and spinoff technologies (e.g., artificial generalized intelligence, animal computation, etc.) evolve. Thus, to sustainably address these issues, it is imperative that a body is formed in perpetuity that is geographically and culturally diverse, and composed of human computation cognoscenti, scholars of ethics and morality, and representatives of policy. Such a "conscience committee" would engage regularly in technical risk and ethics analysis, perhaps employing formal methods, such as systemic risk analysis (see Renn and Klinke 2004), to ensure a multi-view perspective. The resultant findings would be disseminated to the public via a societal journal and inform new policies, that would be regularly revisited in the context of observed effects and new findings. Crucially, the existence and maintenance of this body would be built into the charter of a human computation society.

Conclusion

Regardless of how one envisions the applications or implications of human computation, its increasingly prevalent and complex role in society is indisputable. In this chapter, we have considered the technological plight of our species, the potential risks and rewards of human computation, the maturity of this evolving discipline, and a proposal for its scientific and practical advancement. Each of us contributing to this handbook has, in one form or another, encountered the transformative effects of human computation. Perhaps you have too. This book represents the beginning of a collective effort to shape tomorrow. Please join us in seizing our destiny to empower hope.

Acknowledgments The author would like to express his enduring gratitude to each of the 117 contributors to this volume, who are collectively catalyzing the emergence of human computation as a discipline. The author also wishes to gratefully acknowledge useful feedback from Mary Catherine Bateson, Kevin Crowston, Kshanti Greene, Antonio Sanfilippo, and Michael Witbrock.

References

Breiman RF (1996) Impact of technology on the emergence of infectious diseases. Epidemiol Rev 18(1):4–9

Bunn M, Maslin EP (2011) All stocks of weapons-usable nuclear materials worldwide must be protected against global terrorist threats. J Nucl Mater Manage 39(2):21–27

Friedman TL (1999) The Lexus and the olive tree/by Thomas L Friedman. Farrar, Straus and Giroux, New York

Horrock N (2001) FBI focusing on portable nuke threat. UPI. Retrieved 29 June 2013, from http://www.upi.com/Top_News/2001/12/21/FBI-focusing-on-portable-nuke-threat/UPI-90071008968550/

IAEA (2007) IAEA information system on illicit trafficking and other unauthorized activities involving nuclear and radioactive materials. International atomic energy agency. Retrieved from http://www.iaea.org/newscenter/features/radsources/pdf/fact_figures2007.pdf

Joy B (2000) Why the future doesn't need us. Wired. 8(4). Retrieved from http://www.wired.com/wired/archive/8.04/joy_pr.html

Joy B (2006) Bill Joy: what I'm worried about, what I'm excited about. [Video file]. Retrieved from http://www.ted.com/talks/bill_joy_muses_on_what_s_next.html

Keen P (1980) MIS research: reference disciplines and cumulative tradition. In: Proceedings of the first international conference on information systems, Philadelphia, Pennsylvania, USA pp 9–18

Knutson TR, McBride JL, Chan J, Emanuel K, Holland G, Landsea C, … Sugi M (2010) Tropical cyclones and climate change. Nat Geosci 3(3):157–163. doi:10.1038/ngeo779

Kurzweil R (2001) Raymond Kurzweil: question and answers. Presented at the are we becoming an endangered species? Technology and ethics in the twenty first century, Washington National Cathedral. Washington, DC, USA. Retrieved from http://www.kurzweilai.net/are-we-becoming-an-endangered-species-technology-and-ethics-in-the-twenty-first-century

Kurzweil R (2006) The singularity is near: when humans transcend biology. Penguin, New York

Liles DH, Johnson ME, Meade L (1996) The enterprise engineering discipline. In: Proceedings of the fifth annual industrial engineering research conference

Raymond ES (1997) The cathedral and the bazaar. Retrieved from http://www.catb.org/esr/writings/homesteading/cathedral-bazaar/

Renn O, Klinke A (2004) Systemic risks: a new challenge for risk management. EMBO Rep 5(Suppl 1):S41–S46. doi:10.1038/sj.embor.7400227

Sanfilippo A, Riensche R, Haack J, Butner S (2013) Psychosocial and cultural modeling in human computation systems: a gamification approach. In: Michelucci P (ed) The handbook of human computation. Springer, New York

Silvert W (2002) A bifurcation model of speciation due to environmental change. In: Proceedings of first workshop on information technologies application to problems of biodiversity and dynamics of ecosystems in North Eurasia - WITA'2001. Novosibirsk, Russia. Retrieved from http://academia.edu/197962/A_bifurcation_model_of_speciation_due_to_environmental_change

Solomon S, Plattner G-K, Knutti R, Friedlingstein P (2009) Irreversible climate change due to carbon dioxide emissions. Proc Natl Acad Sci 106:1704–1709. doi:10.1073/pnas.0812721106

Woolf AF (2010) Non-strategic nuclear weapons. DIANE Publishing, Darby, Pennsylvania, USA

Index

A
Accelerometer, 123, 164, 305
Accuracy, 102, 124, 125, 144, 147, 159, 175, 184, 193, 294, 325, 369–371, 374, 375, 381, 383, 385, 386, 401, 407, 526, 528, 547, 549, 598, 605, 607, 610, 621, 622, 628, 630, 655, 658, 661, 663–669, 732, 733, 771, 776
AceWiki, 290
ACO. *See* Ant Colony Optimization (ACO)
Action-coordination algorithms, 635
Activation cascades, 779–788
Active learning, 44
Activism, 98, 425, 996, 1016
Adaptation, 199, 554, 575, 720, 902, 911, 962, 982–985
Adaptive agents, 281, 367–376
Additive tasks, 468
Adrenaline, 512, 546
Adversarial advantage, 893
Advocacy groups, 995, 997
Agency, 54, 96, 97, 102, 200, 222, 226, 230, 256, 263, 339, 416, 418, 433, 434, 513, 572, 835, 1022, 1031
Agent-based framework, 762
Agent-based modeling, 746, 761–766, 812
Agents, 7, 25, 44, 57, 84, 139, 216, 281, 367, 379, 397, 429, 449, 467, 476, 505, 509, 527, 532, 551, 598, 610, 617, 627, 633, 677, 679, 725, 746, 761, 803, 895, 897, 940, 962, 1026
Aggregate, 43, 196, 229, 241, 286, 339, 349, 369, 383, 394, 396, 422, 487, 514, 577, 597, 603, 606, 608, 609, 611, 627, 677, 680, 710, 732, 735, 741, 746, 769–771, 774, 777, 843, 861, 874, 880, 928, 937, 939, 940, 954, 962, 966, 967, 969, 970, 973, 975, 995
Aggregating responses, 386, 387
Aggregation, 40, 154, 193, 229, 240, 247, 309, 367, 378, 382, 383, 387, 394, 412, 421, 423, 431, 441, 468, 492, 499, 568, 575, 577, 583, 598–612, 629, 651, 653, 655, 656, 669, 729, 903, 904, 1003
AI. *See* Artificial intelligence (AI)
Algorithm, 19, 26, 58, 62, 84, 89, 102, 122, 133, 156, 187, 206, 216, 279, 288, 309, 325, 333, 369, 397, 412, 438, 447, 464, 509, 545, 585, 597, 600, 617, 625, 633, 641, 649, 675, 682, 745, 753, 786, 792, 803, 870, 902, 937, 980, 996, 1028
Almanac, 6, 15–17
ALS. *See* Amyotrophic lateral sclerosis (ALS)
Altruism, 173, 272–274, 806
Amazon Mechanical Turk (AMT), 90, 101, 134, 144–147, 187, 210, 333, 335–339, 341, 342, 377, 423, 548, 551, 553, 562, 680, 684, 725, 732, 741, 832, 837–839, 841, 843, 885, 931
America Online, 863
AMT. *See* Amazon Mechanical Turk (AMT)
Amyotrophic lateral sclerosis (ALS), 105, 112–115, 117, 118, 120–122, 126
Analogical reasoning, 288
Analytic gaming, 749
Anaphoric co-reference, 681
Anchoring effect, 547, 549, 607
Android, 125, 286, 299, 328, 533, 537, 970, 1027
Angry Birds, 335

P. Michelucci (ed.), *Handbook of Human Computation*,
DOI 10.1007/978-1-4614-8806-4,

Anomalies, 92, 175, 176, 178, 183, 184
Anonymization, 852–853, 871
Ant(s), 25–34, 280, 349, 636, 708, 797, 798, 911, 912, 917–923, 926–929, 931–933, 936–954, 961, 962, 966, 967, 976, 1027
Ant Colony Optimization (ACO), 26, 27, 636, 941
Anthropology, 3, 108, 416, 805, 1032
Anti-scrip laws, 829
Anxiety, 55–57, 976
Apache, 793, 798
Apache web server project, 793
Aphasia, 987
API, 260, 341, 507, 554, 570, 576, 1028
Applause, 791
AR. *See* Augmented Reality (AR)
Archaeological, 92, 174–175, 178, 184, 985, 987, 1016
Archaeology, 174, 1018
Archimedes, 271, 272
Architecture, 4, 9–11, 26, 43, 45, 51, 57, 85, 132, 146, 174, 188–190, 196, 200, 201, 257, 260, 279, 319, 323, 324, 326, 327, 438, 451, 457, 481–482, 505–508, 514, 528, 568, 570–571, 575–578, 681, 795–797, 808, 854, 857, 865, 866, 873, 874, 879, 944, 1028, 1029, 1031, 1032, 1037
Argument culture, vii
Armed conflict, 894, 993–996, 998, 1000, 1001, 1004–1007
Army ants, 926–928, 936, 937, 946
ARPANET, 863
Art, 3, 5, 16, 92, 132, 147, 149, 187, 188, 190, 192–194, 197, 200, 201, 403, 404, 411, 416, 448, 546, 584, 637, 641, 644, 826, 915, 1012, 1017, 1028
Artificial artificial intelligence, 95
Artificial intelligence (AI), 84, 95, 102, 132, 279, 283, 287, 352, 398, 507, 509, 519, 531–535, 543, 546, 549, 597, 598, 625, 630, 633, 827, 898, 1031
Artwork(s), 190–193
Assistive systems, 511
Associative thought, 451–452
Astronomy, 14, 90, 91, 153, 157, 164, 697, 698
Astroturf, 840, 843
Asynchronous, 63, 107, 145, 499, 510, 529, 554, 558, 565, 571, 709, 710, 712, 1028
Atomic micro-tasks, 431, 433
Attention economy, 961–976
Audience, 134, 137, 310, 311, 313, 403, 404, 534, 682, 684, 685, 862, 864, 868, 873, 917, 939, 952, 1000, 1018
Audio annotation, 397, 680
Augmented, 46, 86, 280, 282, 292, 301, 317–330, 334, 487–490, 492, 583, 873, 894, 924, 930, 935, 939–941, 970
Augmented reality (AR), 280, 282, 301, 317–330, 334, 335, 487–490, 492, 495, 516–517, 519, 785, 873, 894, 939–941, 970
Autocatalytic, 449
Automated workflows, 626–628
Autonomy, 617
Autopoietic, 449, 457
Axon, 41, 42

B

Baby tooth collection, vi
Bayesian Networks, 368, 369
BBS. *See* Bulletin board services (BBS)
Bee(s), 26, 918, 924–926, 933, 934, 962
Behavioral dysfunction, 53, 54
Behavioral pathology, 58, 1032
Berners-Lee, T., 131
Biased assimilation, 805
Bifurcation, 951, 1024
Big crisis data, 102, 104
Big data, 66, 86, 104, 127, 153, 248, 669, 751–758, 847–855, 870
Biodiversity, 294
Biology, 3, 4, 42, 279, 416, 515, 762, 770, 849, 912, 953, 981, 1032
Biosphere 2 Evapotranspiration Experiment, 92, 165
Bird counting, 163
Blendr, 299, 300
Blog(s), 97–99, 101, 172, 177, 187, 196, 197, 285, 649, 651, 655, 656, 658, 659, 662, 665, 667, 668, 685, 695, 696, 762, 824, 838, 841, 860, 864, 1000
BOINC, 153, 964
BOSSA, 11, 154, 569
BOTTARI, 327–329
Bottom-up, 954, 963, 1027
Brain, 4, 26, 39–45, 51, 55, 56, 72, 117, 268, 271, 477, 479–482, 615, 848, 849, 851, 883, 894, 897–907, 936, 939–941, 961, 979, 983, 987, 994, 1025, 1026, 1032
Brainstorm(ing), 247, 403, 463, 468, 608, 644, 705, 712

Braintalk, 106–109, 112
Breadth-first, 448
Brier score, 370, 372–375
Budget ReACTion Project (BRP), 996
Bulletin boards, 106, 139, 863
Bulletin board services (BBS), 863
Bureaucracy, 285, 836, 907
Business process, 92, 255–263, 282, 424, 506

C
CAL. *See* Cultural algorithm (CAL)
Calculate, 62, 263, 271, 370, 399, 404, 406, 464, 557, 705, 783, 794, 799
Calculation(s), 15–20, 22, 84, 157, 167, 209, 416, 464, 898, 903, 987, 1025
Call to action, 96
Cancer, 106, 107, 109–111, 118, 120, 270, 287, 292, 394, 805, 887
Captcha, 597
Cascade, 42, 534, 629, 631, 753–757, 762, 779–788, 942, 943
Case studies, 92, 96, 165–166, 174–182, 201, 561–572, 582, 587, 610, 749, 763–766, 793, 798–799, 894, 1024
Catastrophic loss, 990
Cathedral and the Bazaar, 1027
Cellular computation, 25
Challenge propagation, 903, 904
Chaotic, 45, 63, 456
Charles, B., 4, 13, 14, 18, 23
ChatNet, 863
Cheating, 307, 308, 401, 412, 690, 692
Chemistry, 14, 51, 849, 851
Chorus, 506, 517–518
Citizen computation, 280, 297–315
Citizen computing, 280, 297–315
Citizen journalism, 1002–1003, 1005
Citizen science, 46, 84, 90–92, 103, 118, 153–161, 163–165, 167, 298, 306, 315, 339, 412, 424, 600, 676, 677, 680, 686, 692, 695–700, 725, 726, 732
CityExplorer, 307, 308
CitySourced, 306
Classifier, 103, 154–156, 158, 159, 161, 621, 947
Click manipulation, 842
Clickworkers, 6, 10, 154, 551, 824
Climate change, 57, 58, 80, 163, 165–166, 424, 838, 994, 1026
ClimatePrediction.net, 153
Cloud(s), 72, 119, 140, 155, 261, 349, 350, 524–528, 665, 1015
Cloud computing, 72, 665, 1015
CMC. *See* Computer mediated communication (CMC)
Coevolution, 1024
Cognitive architecture, 4, 9–11, 51, 57, 85, 481–482
Cognitive bias(es), 387, 507, 749, 803, 804, 880, 886
Cognitive constraints, 479
Cognitive limitations, 382, 549, 751
Cognitive load, 476, 478, 487, 499, 547, 684, 746, 755, 757
Cognitive overload, 936–937, 939, 943, 945, 946, 950
Cognitive risks, 838
Cognitive science, 512, 762, 1032, 1035
Cognitive surplus, 103, 127, 1011–1013, 1015–1018
Collaboration, 3, 86, 117, 120, 126, 155, 157, 161, 164, 174, 178, 180, 182, 185, 189, 201, 247, 248, 252, 258, 280–282, 314, 340, 347, 378, 401–403, 413–415, 422–424, 428–429, 431, 433–437, 440–442, 475–477, 482, 488, 489, 493–495, 498, 499, 505–508, 543, 598, 615–619, 621, 623, 631, 636–638, 676, 680, 690, 696, 703, 705–708, 711, 712, 715–722, 741, 762, 791, 819, 820, 824, 835, 842, 857–874, 882, 883, 887, 899, 907, 930, 980, 986, 999, 1000, 1006, 1016, 1029
Collaboration efficacy, 476–477, 482, 703
Collaborative knowledge production, 422, 424
Collaborative ontology development, 292–293
Collaborative problem solving, 347, 348, 422, 424, 425, 618, 709, 806
Collapse informatics, 993, 1017
Collection design pattern, 90
Collective action, 85, 281, 421–442, 763, 769, 932
Collective behavior(s), 26, 27, 84, 477, 745, 747, 748, 761–766, 770
Collective computation, 26, 34, 936, 939–946
Collective intelligence, 85, 333, 338, 424, 427, 477, 513, 616, 676, 679, 680, 682, 703–712, 797, 899, 904–906, 980, 1017, 1032
Collective intelligence factor, 705, 706, 797
Collective journalism, 188
Collective judgement, 602, 604–607, 612
Collective memory, 499, 514, 990, 1023

Collective narrative(s), 188–190, 200, 201
Collective patterns, 764, 766
Collective problem solving, 280, 347–364, 463, 464, 618, 894, 1032
Collective reasoning, 86, 176, 282, 482–484, 486, 499
Collective search, 282, 463–473
Collective sensing, 492
Collective wisdom, 247, 599, 939
Collective writing, 201
Colony, 26–34, 636, 797, 911, 912, 915–933, 936–950, 952–954, 961, 962, 966–968, 971, 972, 975
Colony computation, 27–33
Combinatorial markets, 368, 373, 374
Combinatorial optimization, 397
Combinatorics, 478–480
Commodities, 44, 810, 811, 883, 994, 997
Communication, 3, 4, 13, 17, 26–32, 34, 39–46, 51–54, 56, 57, 59, 61, 64, 75, 85, 86, 91, 92, 108, 109, 111, 117, 163, 187, 194, 198, 201, 202, 249, 252, 261, 282, 297, 320, 326, 347, 382, 422, 423, 431, 433, 437, 442, 463, 465, 467–471, 473, 478, 479, 485–489, 499, 507, 552, 573, 579, 597, 634, 636, 637, 676, 677, 695, 696, 703, 708, 709, 711, 715, 716, 719, 720, 746, 751, 793, 798, 805, 810, 811, 833–835, 857, 861–864, 866, 868, 870, 897, 899–901, 906, 907, 916, 920, 934, 944, 951, 979, 983, 984, 986, 987, 994, 995, 1002, 1014–1018, 1025
Communities
 community data, 867
 community-driven, 137, 1036
 community structure, 44, 748, 779, 780, 786, 788, 1034, 1035
Commuting pattern, 587
Companion systems, 532, 543
Compensation, 210, 380, 388, 430, 823–825, 827–831, 834, 839
Compensatory tasks, 468
Competition, 89, 90, 125, 137, 157, 272, 273, 299, 309, 311, 463, 468, 471, 697, 706, 736, 741, 752, 763–766, 806, 880, 893, 907, 917, 921, 925, 926, 928, 954
Competition design pattern, 89
Competitive disruption, 839
Complex
 complexity theory, 546, 549, 597
 complex system(s), 3, 4, 27, 33, 39, 45, 58, 161, 792, 941, 1024, 1025
 complex task(s), 144, 147, 384, 398, 401, 422, 423, 434, 532, 625, 626, 792, 797, 799
 complex technologies, 894, 984–987, 990, 1016
Composition, 40, 46, 63, 288, 421, 438, 540, 554, 558, 575, 656, 1026
CompuServe, 106, 863
Computational agents, 85, 382, 483, 505, 617, 633, 637–639, 806, 1026
Computational analysis, 438, 746–747, 761–766
Computational nodes, 412, 416, 417
Computational resources, 26, 134, 159, 412, 507, 621, 639, 944, 945, 1011, 1015, 1029
Computational system, 25, 34, 83, 161, 187, 189, 190, 197, 253, 279–281, 335, 340, 343, 394, 396–399, 402, 421, 431, 435, 506, 507, 509, 617, 633, 637–639, 763, 840, 962, 980, 1030
Computer-mediated communication (CMC), 676, 677, 695, 715, 720, 863
Computer-supported cooperative work (CSCW), 411, 899
Concept interaction, 452–454
Conceptnet, 525, 680
Conceptual modeling, 135, 138–141, 144–146
Conditional probability, 348, 349
Conflict, 21, 53, 54, 79, 132, 161, 227, 273, 347, 395, 436, 437, 708, 721, 831, 888, 894, 907, 912, 917, 936, 993–1008
Conformity, 135, 288, 563, 805
Conjunctive tasks, 468
Connectedness, 984–985, 987, 994, 1016
Connection, 4–12, 41, 43, 45, 55, 63, 116, 195, 248, 249, 260, 288, 297, 303, 314, 325, 329, 333, 336, 342, 351, 401, 439, 464, 466, 472, 477, 479–482, 509, 526, 576, 651–653, 655, 659, 665, 692, 740, 751, 762, 763, 796, 804, 806, 827, 830, 850, 862–864, 866, 867, 870, 873, 903, 904, 930, 937, 950, 954, 987, 998–1000, 1006, 1032, 1035
Connectivity, 43, 112, 172, 173, 248, 329, 341, 477, 481, 506, 590, 748, 762, 782–786, 894, 980, 990, 999, 1015, 1017, 1018
Connectome, 479
Consensus opinion, 608

Consistent output, 514
Constraint network, 352, 360–362
Consumer awareness, 997–998
Contamination, 1014
Content-plus-Link, 653, 654
Context
collapse, 866, 868–869
collision, 866, 868, 869
Continuous crowdsourcing, 510–513
Continuous learning, 266–270
Contractual obligation, 834
Convergent thinking, 706
Convergent thought, 451, 712
Conversational computation, 531–543
Cooperative, 26, 34, 258, 266, 411, 424, 426, 431, 434, 435, 442, 477, 511, 633, 634, 636–638, 708, 712, 795, 797, 899, 911
Cooperative behavior, 34, 477
Cooperative Intelligent Agents, 633–635
Cooperative problem-solving (CPS), 634, 636
Coordinated action, 86, 282, 483–486, 499, 898, 906
Coordinated attacks, 842, 1032
Coordination, 27, 85, 96, 97, 100, 285, 286, 293, 310, 413, 414, 422, 488, 495, 496, 498, 552, 569, 576, 578, 631, 635–638, 709, 731, 748, 791, 835, 861, 900, 901, 903–905, 916, 936, 940, 964, 974, 976, 989, 1013
Copyleft, 913, 915
Counting, 163, 482
CplusL, 653
CPS. *See* Cooperative problem-solving (CPS)
Craigslist, 172, 971
Creative, 92, 112, 172, 253, 286, 388, 423, 442, 447, 451, 452, 500, 546, 598, 608, 646, 707, 711, 912, 954, 999, 1022
Creativity, 44, 84, 191, 546–547, 619, 706, 710–712, 803, 806, 813, 893, 894, 975, 1013
Credit risk, 92, 215–241
Crime, 102, 516, 600, 606, 819, 842
Crisis informatics, 993, 1003–104, 1017
Crisis, M., 95, 101–102
Crisis map, 96–98, 101, 102
Crowd agents, 499, 509–520, 1028, 1031
Crowd algorithms, 509
Crowd-based mitigation, 844
CrowdCrafting, 95, 100, 101, 154
CrowdDB, 562
Crowdflower, 95, 96, 100, 101, 144, 146, 407, 423, 508, 548, 551, 625, 824, 829, 837
CrowdForge, 562
Crowdfund(ing), 20, 21
Crowding-out effect, 547
Crowd innovation, 248, 250–252
CrowdMAP, 146
Crowd-powered, 510, 511, 513, 516, 517, 844, 1026
Crowd-powered systems, 510, 511, 513, 844
Crowdsource, 96, 148, 182, 380, 489, 531, 553, 565, 568, 631, 732, 879, 989
Crowdsourced data collection, 680
Crowdsourced human computation, 184, 191, 335, 337, 377, 692
Crowdsourcing
API, 562
marketplace(s), 216, 238, 239, 281, 377–388, 839
CrowdSPARQL, 144–147
Crowdturf(ing), 840–842, 844
Crowd4U, 508, 561–572
Crowd-workers, 421–442
Cryptography, 411
CSCW. *See* Computer-supported cooperative work (CSCW)
CUbRIK, 438, 442
Cultural algorithm (CAL), 448, 456, 458
Cultural anthropology, 1032
Cultural evolution, 281, 447–458, 979–990, 1032
Cultural modeling, 803–814
Culture, 3, 54, 59, 188, 196, 202, 247, 394, 447–450, 453, 455, 457, 863, 893, 894, 907, 948, 979, 980, 982–984, 987–990, 999, 1001, 1012, 1014, 1023, 1032
Cumulative cultural evolution, 979–990
Cumulative culture, 894, 980, 982–984, 987–990, 1023
Curation, 132, 506, 524–527, 697, 969
Curious Cat, 532–538
Currency, 103, 220, 230, 234, 236, 237, 485, 741, 824, 828–831, 894, 974
Cyber-crime, 819
Cybernetic(s), 3, 72–73, 561–572, 1032
Cybernetic dataspaces, 561–572
Cyc, 288, 506, 533–534, 537, 539
Cyc FACTory, 288
CyLog, 507, 508, 561–572

D

DAGGRE, 367–370
DAI. *See* Distributed Artificial Intelligence (DAI)

DARPA, 351, 500, 638, 728, 839
Darwin, C., 163, 457
Darwinian, 448–450, 457
Data annotation, 132, 526–527
Data avalanches, 172, 185
Database mining, 581, 590
Data collection, 86, 113, 125, 163–165, 167, 168, 172, 174, 309, 311, 424, 499, 582–585, 589, 680, 684, 689, 752, 853, 874, 940, 967, 970, 971, 974, 975
Data interlinking, 134, 136
Data mining, 102, 326, 339, 769, 866, 870, 871
Data philanthropy, 96, 104
Data quality, 132, 323, 377, 381, 382, 384–387, 394, 396, 398, 405, 506, 524, 525, 527, 528, 570, 688, 692
Data reduction, 154, 159–161
Data sources, 110, 143, 507, 561–567, 570–571, 581, 583–586, 589, 651, 669, 808, 940, 1004
DAWN, 4, 982
DBPedia, 142, 143, 147, 148, 524, 526
De-anonymizing data, 852
Decency, 831, 834
Decentralized, 27, 29, 135, 574, 791, 901, 912, 925, 926, 929, 932–937, 939–941, 946, 950, 952, 954
Decentralized harmony, 933–936
Decision field theory, 59
Decision-making, 4, 27, 31, 41, 55, 119, 164, 191, 247, 339, 424, 427, 467, 761, 803, 804, 806, 814, 835, 937, 946
Decisions, 4, 16, 27, 29, 41, 44, 55, 73–77, 79, 80, 91, 98, 99, 109, 116, 119, 124, 137, 139, 154, 155, 159, 164, 171, 184, 191, 193, 207, 216, 238, 247, 256, 257, 266, 339, 348, 350, 352, 356–359, 364, 368, 370, 380, 382, 396, 415, 424, 427, 437, 464, 467, 481, 514, 547, 564, 581, 586, 588, 608, 609, 625, 627–629, 631, 634, 635, 637, 638, 662, 664, 680, 683, 688, 690, 691, 706, 707, 711, 712, 749, 757, 761, 771, 776, 781, 791, 803–808, 814, 823, 824, 835, 847–852, 870, 880–882, 926, 937–943, 946, 965, 980, 995, 1006, 1007, 1015, 1018
Decision theory, 625
Declarative knowledge, 131
Deco, G., 43, 562
Decompose, 144, 465, 532, 584, 587, 625, 626, 631, 637, 675
Decomposing problems, 273
Deforestation, 1014
Deliberation groups, 600, 603, 607–609
Delicious, 769
Delphi technique, 710
Democracy, 862, 995–996, 1007, 1016
Democratization of technology, 1021
Demonization, 998, 999
Dendrite, 41, 42, 480
Depth-first, 448, 450
Derivatives, 215–226, 229, 230, 234, 236, 240, 241
Descriptive analysis, 466
Design pattern, 89–91, 257, 280, 425, 447, 1028
Diffusion, 42, 45, 470, 753–757, 762–765, 788, 870
Digg, 747, 748, 751–756, 769–777, 796, 843
Digital archives, 200
Digital art projects, 188
Digital breadcrumbs, 870
Digital footprints, 592
Digital humanities, 437, 439
Digital intelligence, 954
Digital natives, 297
Digital volunteers, 97
Dignitary interest, 825, 834–835
Dignity, 1017
Diminishing returns, 29, 282, 475, 498, 627, 786
Disaster, 6, 91, 92, 95–104, 174, 180, 184, 339, 424, 464, 569, 907, 985–987, 1002, 1003
Disaster relief, 6, 1002
Disaster response, 91, 95–104, 339
Disclosure, 126, 235, 833–834, 855, 866–869, 871, 873, 880, 883, 926
Discomfort, 1014
Disease, 21, 45, 102, 105–108, 111, 115, 117–127, 237, 283, 287, 292, 394, 465, 600, 753, 754, 770, 804, 853, 907, 1003, 1017, 1024, 1027
Disjunctive tasks, 468
Disruptive computation, 962
Disruptive human computation, 84
Disruptive technology, 893
Dissemination, 44, 45, 52, 857, 894, 976, 987–989, 1034
Distributed Artificial Intelligence (DAI), 598, 633, 634, 637, 638
Distributed cognition, 85, 414, 899, 900, 906
Distributed cognitive architecture, 85
Distributed computation (DC), 25, 63, 153, 281, 447–458, 483, 574, 633, 939, 964, 979–990, 1031
Distributed constraint optimization, 635

Distributed constraint satisfaction, 635
Distributed human sensors, 335
Distributed intelligence, 85, 283, 894, 897–907
Distributed intelligent agent algorithms, 598, 633–639
Distributed intelligent agents, 598, 633–639
Distributed problem solving, 57, 58, 85, 421, 996
Distributed thinking, 4, 5, 7–11, 85, 483, 1032
Distributed work platforms, 832, 835
Divergent thinking, 706–707
Divergent thought, 451, 706–708, 710–712
Division of labor, 14–16, 475, 618, 619, 903, 904, 917, 927, 946–950, 983
DNA, 6, 11, 25, 954, 986, 1022
DNA fingerprinting, 986
Domain models, 808
Do Not Track, 605, 853
Dopamine, 125, 847–855
Dopplr, 303, 304
Double-bind, 54
Do you know, 651
Drive theory, 55
Duolingo, 732, 733, 837
DYK widget, 651, 652, 656
Dynamical systems, 969, 1024
Dynamic curriculum, 269
Dynamic systems, 40–41, 45, 1031
Dystopia(n), 506

E
Earthquake, 95–97, 100, 103, 163, 164, 178–180, 339, 1002, 1004
EBay, 239, 431, 526, 527, 531, 739, 843
Eco-law, 574–579
Ecological, 237, 484–486, 495, 498, 500, 508, 762, 798, 918, 921, 962, 998, 1031
Ecological footprint, 998
Ecological perception, 485, 500
Ecological systems, 798
Ecology, 31, 574, 575, 747, 762, 1024
eCommerce, 131
Economic(s), 3, 14, 15, 21, 22, 54, 110, 215, 218–220, 223, 238, 283, 337, 380, 456, 477, 510, 604, 609, 610, 676, 727, 739, 741, 820, 827, 853, 861, 870, 879, 880, 894, 906, 965, 966, 971–976, 986, 994, 1000, 1011, 1012, 1014, 1017, 1018
Economic optimization problem, 966, 972–975
Economies of attention, 763
Economy, 14, 15, 21, 33, 34, 191, 272, 274, 347, 741, 763, 829, 883, 893, 906, 907, 961–976, 987, 1016
Ecosystem(s), 3, 27, 39, 166, 508, 574, 576–578, 584, 761, 844, 870, 907, 913–915, 928, 953, 994, 1014
Education(al), 44, 92, 156, 163–169, 266, 268–270, 313, 336, 385, 719, 725, 727, 732, 872, 884, 888, 907, 972, 986, 988, 994, 1000–1001, 1018, 1030, 1034–1036
Efficient Market Hypothesis, 238, 239, 610
Einstein@home, 153
Electronic literature, 92, 187–202
Electronic medical record (EMR), 109, 110, 126
Electronic voting, 711
Elicitation method, 602, 603
Elite, 912, 915, 918, 1016, 1018
Elitism, 643
EmailValet, 511
EMB. *See* Enterprise microblogging (EMB)
Embodied, 27, 28, 639, 860, 898, 902, 1017
Embodied cognition, 908
Emergence, 58, 68, 131, 176, 257, 281, 285, 329, 414, 433, 447, 449, 496, 498, 574, 646, 651, 703, 715, 747, 766, 792, 812, 840, 920, 939, 940, 983, 1030, 1034–1037
Emergent, 3, 4, 7, 9, 33, 39, 44, 84, 176, 178, 184, 188, 252, 253, 280, 281, 457, 637, 745, 748, 761, 771, 777, 792, 866, 894, 930, 933, 949, 954, 963–965, 976, 980, 983, 1021, 1025, 1027
Emergent behavior, 4, 39, 44, 280, 281, 748, 777, 792, 894
Emergent human computation, 963
Empathy, 281, 402–406, 993–996, 1006–1008
Empirical analysis, 745, 746
Employee, 17, 20, 248, 255, 388, 655, 656, 659, 679, 825–831, 833, 834, 998
Employment, 22, 770, 823–828, 831, 833, 836
Employment law(s), 823–825, 827, 828, 831, 833
EMR. *See* Electronic medical record (EMR)
Endangered species, 165, 1024
Enemies, 968
Enforced independence, 937–939
Engineer, 16, 17, 22, 30, 34, 131, 135, 188, 292, 482, 506, 577, 598, 745, 834, 895, 901, 970, 980, 986, 1031
Engineered, 30, 58, 59, 84, 475, 505, 827, 842, 851, 963–965, 969, 1022, 1031

Engineered human computation, 475, 505, 842, 963, 1031
Engineering, 30, 61, 131–150, 257, 279, 287, 290, 317, 354, 411, 458, 479, 578, 719, 961, 975, 976, 987, 1028–1034
Enterprise microblogging (EMB), 709
Entity linking, 147, 150, 525, 526
Environmental monitoring, 421
Environmental stress, 994
Environment design, 573–579
Epidemic(s), 45, 369, 754–756, 762, 770, 870
Epidemiology, 124, 339, 747, 762, 770
Epistemic, 78–80, 603, 605, 611
Epistemology, 3, 4, 71–73, 75–80, 1032
ePluribus Solver, 188
Equilibrium, 54, 55, 57, 236, 351, 448, 453, 455, 603, 609, 677, 727, 730, 731, 733, 734, 737, 738, 935, 981, 1024, 1031
ESP game, 6, 7, 10, 322, 397, 399, 411, 412, 417, 418, 510, 561, 565–567, 637, 726, 727, 729–731, 741
Ethics, 283, 412, 418, 805, 1024, 1037
Eureka, 271–272
Eureka moment, 271–272
Eusocial insect colonies, 912, 921, 931, 961
Eusocial insects, 477, 482, 798, 894, 911–913, 918, 921, 929–946, 953, 961, 1032
Evapotranspiration, 92, 165–168
Evil, 893, 994
EVOC. *See* Evolution of culture (EVOC)
Evolution, 11, 26, 58, 62, 131, 185, 238, 270, 281, 282, 293, 324, 447–458, 480, 508, 571, 574, 578, 641, 644, 762, 774, 775, 812, 820, 853, 858, 863–864, 871, 894, 902, 905, 912, 913, 915, 918, 921, 929, 939, 953, 979–990, 1013, 1031, 1032, 1036
Evolutionary, 16, 119, 282, 449, 457, 479–481, 598, 636, 641–646, 748, 901, 911, 912, 980, 981, 983, 989, 990, 1032
Evolutionary algorithms, 598, 641–646
Evolution of culture (EVOC), 449–451, 455, 456
Exception handling, 546
Existential, 72, 74–76, 80, 534, 1023, 1025, 1027
Existential risk, 1025, 1027
Experiment, 92, 119, 123, 134, 146, 156, 165, 167, 168, 177, 189, 190, 193, 194, 207, 213, 271, 306, 355, 357, 369, 374, 375, 379, 405, 407, 438, 480, 482, 485, 488–498, 508, 518, 527–529, 545, 548, 644–646, 653, 659, 660, 663, 746, 769, 770, 780, 784, 786, 787, 797, 804, 836, 885, 894, 913, 914, 938, 942, 943, 952, 976
Expert-based crowdsourcing, 431–434, 440
Expertise, 7, 23, 103, 118, 135, 164, 166, 273, 287, 292, 293, 322, 348, 417, 423, 424, 427, 440–442, 524, 558, 605, 619, 629, 634, 635, 644, 646, 691, 803, 809, 832, 849, 1004
Exploitation, 27–29, 34, 256, 257, 299, 380, 429, 467, 471–472, 629, 820, 832, 847–855, 944, 974, 1037
Exploits, 27, 132, 225, 257, 260, 262, 289, 297, 313, 429, 437, 470, 471, 630, 837–844, 885, 964, 974, 982, 984
Exposure function, 754, 755
External populations, 840, 843
Extrinsic motivation, 400, 429, 676, 677
Eyespy, 307, 308, 417

F

Facebook, 41, 44, 65, 66, 111, 139, 140, 172, 197, 257, 299, 300, 302–304, 320, 324, 400, 404, 438, 469, 651, 658, 664, 682, 684, 685, 687, 709, 710, 813, 828, 830, 841, 863–865, 868–874, 903, 986, 1015, 1016
Face2face, 299, 300
Face-to-face collaborations, 340, 716–718, 720, 722
Face-to-face groups, 706–708, 712
Fairness, 283, 603, 820, 825, 828, 831, 972
Fascism, 4
FDA. *See* Food and Drug Administration (FDA)
Feature extraction, 125, 183
Feature selection, 465, 472
Federal Emergency Management Agency (FEMA), 96, 97, 99, 100, 842, 843
Federal Reserve, 217, 236
Feedback, 28, 59, 75, 77, 86, 145, 146, 156, 168, 175, 177, 184, 188, 197, 205, 206, 208–213, 236, 256, 262, 272, 293, 355, 399, 400, 402, 404, 406, 435, 437, 485, 489, 500, 509, 513, 514, 564, 585, 586, 656, 669, 682, 710–712, 798, 849, 850, 898, 902, 935, 941, 943, 944, 964, 965, 1037
Feedback loop, 75, 77, 177, 236, 485, 849, 850, 941

FEMA. *See* Federal Emergency Management Agency (FEMA)
Filter bubble, 862, 866, 870–871
Financial derivatives, 218
Financial incentives, 110, 687–688, 725, 741
Find-fix-verify, 510, 626
Fireflies, 791, 792, 797
Fission, 921–928
Flax spinning machine, 89
Flickr, 111, 143, 172, 320, 323, 734, 769, 865
Flocking, 791, 797
Fold-it, 886
Food and Drug Administration (FDA), 119, 805
Forage, 27–32, 911, 932, 936, 937, 942, 943, 951, 952
Foraging, 27–29, 31, 32, 466, 636, 708, 911, 922, 933, 936, 937, 941–946, 948, 949, 951, 952, 985
Forced disclosure, 866, 869
Forecasting, 367–372, 374, 375, 602, 603, 902
Forums, 64, 105–108, 112, 113, 115, 117, 155, 266, 347, 380, 383, 388, 426, 428, 658, 685, 696, 697, 725, 726, 728, 729, 736, 826, 835, 860, 882, 884, 901, 904, 907, 912, 915, 916, 920, 996, 1017
Fossil fuels, 1011–1018
Foundation, 3–12, 14, 21, 34, 100, 118, 125, 149, 185, 241, 248, 286, 290, 294, 359, 360, 362, 411, 425, 574, 584, 587, 599, 619, 633, 636, 646, 700, 722, 740, 741, 798, 800, 859, 901, 918, 921–926, 928, 930, 931, 938, 941, 946, 951, 952, 954, 976
Foursquare, 301–304, 583, 842, 940
Frame problem, 454
Fraud, 274, 380, 404–406, 830, 834, 1001
Freebase, 141, 142, 149
Freelance, 15–17
FriendCompass, 326, 327, 329
Friend-of-Friend, 653, 654
Friendster, 863, 864
FullCircle, 299, 300
Functional fixedness, 267, 271
FusionCOMP, 561, 572

G

GalazyZoo, 6, 10, 91, 154–156, 158, 159, 164, 378, 412, 418, 696–698, 700
Game-based human computation, 813
Game mechanics, 248, 313, 393, 398, 400, 491, 526, 806, 966, 968–970, 972
Games with a Purpose (GWAP), 91, 92, 137–144, 149, 173, 205–207, 212, 216, 238, 280, 281, 298, 307–315, 402, 411, 412, 423, 430, 432, 551, 558, 561, 565, 676, 677, 680, 681, 683, 687, 690, 725–727, 729–731, 824
Game theory, 416, 564, 565, 571, 633, 725–741, 972
Game the system, 430, 974
Gamification, 96, 103, 143, 173, 301, 728, 736–740, 803–814, 962, 965
Gaming, 154, 156–159, 174, 252, 253, 298–302, 305, 311, 317, 322, 398, 400, 401, 404–406, 408, 412, 432, 490, 682, 685, 695, 749, 803, 806, 808–809, 823, 829, 864, 953, 961–976
General Problem Solver (GPS), 97, 123, 127, 297, 299, 305, 314, 318, 322, 352, 417, 488, 492, 538, 583, 585, 966, 970, 973, 986
Genes, 118, 287, 823, 918, 983, 986
Genetic, 26, 31, 32, 39, 55, 56, 118, 123, 127, 257, 274, 394, 448, 617, 636, 644, 902, 923, 928, 929, 931, 980, 983, 1022, 1023, 1031
Genetic algorithm(s), 26, 31, 32, 448, 617, 636, 644, 902, 980, 1031
Genetic programming, 448, 1022
Genghis Khan, 174
Genome, 515
Genomic, 105, 118, 290, 294
Genotype, 118–119, 641
Geocaching, 299, 301
Geographic information system (GIS), 179
Geolocated, 492
Geo-spatial boundary, 310
Gibson, W., 102
GIS. *See* Geographic information system (GIS)
Global, 25, 33, 34, 40, 41, 43–46, 65, 66, 97, 101, 123–125, 154, 163, 164, 178, 188, 191, 193, 194, 205, 216, 247, 283, 293, 299, 333, 452, 466–472, 477, 523, 529, 634, 762, 766, 782, 783, 785, 787, 796, 820, 825, 835, 842, 860, 862, 868, 894, 897–907, 935, 936, 939–941, 945, 954, 972, 984, 986–990, 1000, 1002, 1011, 1014–1017, 1024
Global brain, 477, 894, 897–907, 939–941
Global knowledge repositories, 523
Global network, 44, 46, 193, 299, 825, 860, 894, 905, 1017

Global synchrony, 796
Goal state, 55, 71, 75, 80, 352, 353
Goal-values, 72–77, 79, 80
GoblinsNGold, 300, 302
Gold standard, 119, 156, 159, 208, 210, 386, 405, 604, 683, 689, 691
Google, 10, 110, 121, 124, 172, 179, 188, 197–199, 235, 258, 299–302, 329, 339, 405, 533, 600, 680, 685, 691, 863, 864, 870, 874, 915, 939, 940, 966, 970, 988, 995, 1016
Google+, 66, 710, 864, 874
Google Flu Trends, 124, 339
Google Glass, 124, 329, 486, 487, 966, 974
Google ingress, 301, 302
Governance, 792, 835, 1001, 1012, 1015
Government, 16, 17, 20, 21, 54, 65, 101, 105, 123, 127, 221, 222, 225, 229, 231, 258, 347, 376, 524, 586, 824, 842, 866, 883, 988, 994, 996, 1005, 1006, 1015, 1016
GPS. *See* General Problem Solver (GPS)
Grexit, 369–375
Grey market, 829
Ground truth, 175, 178, 179, 184, 386, 627, 732, 733
Group collaboration, 429, 712
Group data, 867
Group effectiveness, 703
Group efficacy, 475, 477–479, 482, 486, 488, 1032
Group intelligence, 477
Group interaction, 422, 423, 428, 431, 442, 608
Group IQ, 477
Groupon, 304
Group synchrony, 794, 798, 799
Groupthink, 803
Guerrilla tactics, 996
GuessIt, 405–407
Guess What?!, 140, 141
GWAP. *See* Games with a Purpose (GWAP)

H
Haiku, 190, 193–197, 200
Haiti Earthquake, 95–97, 100
Handwriting recognition, 509
Hazard, 224–226, 228, 241, 430, 431, 433, 434, 437
HCI. *See* Human computer interfaces (HCI)
Health, 42, 91, 101–103, 105–127, 223–225, 229, 249, 285, 304–306, 310, 479, 524, 586, 804, 826, 852, 853, 907, 948, 979, 994, 1013–1015
Health and safety risks, 1014
Health Information Technology for Economic and Clinical Health (HITECH), 110
Healthnet, 285
Heartbeat, 305
Herding, 791
HEU. *See* Highly enriched uranium (HEU)
Hierarchical, 40, 41, 43, 44, 52, 143, 350, 491, 530, 717, 974
Hierarchical leadership, 717
Hierarchy, 41, 74, 75, 77, 135, 138, 250, 254, 482, 492, 584, 587, 590, 629, 793, 916, 920
Highly enriched uranium (HEU), 1022
HIT, 145, 146, 307, 317, 384, 828, 850, 851, 1005
HITECH. *See* Health Information Technology for Economic and Clinical Health (HITECH)
HIV. *See* Human immuno virus (HIV)
Holistic, 58, 414, 898
Homeostasis, 26, 54, 55, 68
Homo ludens, 281, 393–408
Homo sapiens, 912, 1013
Honeybees, 911, 925, 933, 934, 939
Hotlist, 304, 305
Howe, J., 13, 424, 551, 679
HPS. *See* Human-Provided Services (HPS)
HPV. *See* Human-papillomavirus (HPV)
Hub Culture, 863
Human affordances, 619
Human agents, 27, 34, 84, 507, 532, 598, 617, 638, 803, 900, 903, 941, 1028, 1029
Human-as-a-service, 333
Human Assisted Turing Machines, 549
Human-based evolutionary algorithm, 642–644
Human computation algorithms, 416, 417, 549, 597
Human computational work, 823
Human-computer collaboration, 505, 506, 598, 615–617, 623
Human-computer environments, 712
Human-computer interaction, 281, 334, 411–418, 509, 720, 721, 1018
Human computer interfaces (HCI), 279, 281, 411–418, 993, 996
Human-computer symbiosis, 897, 899
Human connections, 1006
Human connectome, 479
Human data, 176, 454, 562, 563, 581–591
Human experts, 436, 441, 629, 996
Human Genome, 515
Human immuno virus (HIV), 111, 115, 119, 394

Human innovation, 893, 980
Human intelligence, 102, 194, 206, 436, 509, 519, 561, 562, 569, 680, 803, 813, 953, 1011, 1025, 1026
Human Intelligence task (HIT), 569, 680
Humanitarian, 91, 95,–98, 100, 101, 103, 104, 842, 1003, 1004
Humanitarian aid, 91
Humanitarian response, 96, 97, 101, 104
Humanizing connections, 998–1000
Human judgement, 599–612
Human-papillomavirus (HPV), 805
Human-Provided Services (HPS), 552–557
Human sensors, 335, 339, 581, 582, 584, 728
Human society, 89, 899, 911
Human specialists, 638
Human tasks, 145–147, 338, 552–555
Hunter-gatherer, 980
Hurricane Sandy, 95, 99–100, 102, 104, 842, 1026
Hybrid human machine computing, 137
Hybridized work flow, 625, 628–629
Hybrid systems, 132, 156, 378

I

ICT. *See* Information and communication technologies (ICT)
IDE. *See* Integrated Development Environment (IDE)
Idea, 4, 11, 13, 20, 21, 52, 62, 90, 119, 131, 132, 167, 184, 190, 192, 248, 249, 251, 252, 256, 267, 269, 270, 282, 314, 323, 338, 341, 347, 383, 394, 401–403, 405, 421, 424, 450, 451, 453, 463, 477, 478, 505, 511, 513, 514, 549, 563, 575, 581, 584, 597, 599, 605, 626, 636, 643, 644, 646, 687, 707, 709, 712, 733, 763, 788, 834, 848, 850, 868, 870, 886, 893, 984, 1006, 1025
Ignorance, 907, 936–939, 943–946, 950
Illicit nuclear trafficking, 804, 809–813
Image classification, 732
Image labeling, 397, 509, 617, 621, 741, 838
Immune system, 25, 26
Inbreeding, 922–924
Incentive, 34, 154, 216, 217, 225–226, 236, 240, 299, 301, 309, 400–402, 421, 423, 430, 432, 517, 518, 524, 553, 561, 564, 565, 567, 569, 601, 609, 675, 686, 725, 726, 728, 729, 734, 736, 739, 741, 832, 833, 851, 885, 894, 950, 969, 974, 989, 1032
Incentive structure, 34, 561, 564, 567, 569
Incentivizing participation, 728, 736, 1034
Increasing returns, 475
Independent cascade model, 753–756
Individual autonomy, 617
Individual learners, 981, 982
Industrial psychology, 704
Inference engine, 533, 534
Information
  diffusion, 753–757, 764
  processing, 13, 17, 42, 51–59, 84–86, 97, 172, 333, 343, 415, 477, 479, 481–483, 487, 488, 617, 716, 721, 722, 852, 897, 899, 906, 954, 969, 1029, 1031, 1032
  retrieval, 92, 131, 205–213, 320, 421, 584
  spread, 746, 748, 751, 753, 755
  theory, 4, 61, 62, 65, 66, 68, 1032
  visualization, 622
Information and communication technologies, 897, 899, 1016, 1017
Information and communication technologies (ICT), 67, 297, 573–579, 715, 717, 870, 897, 899, 900, 902, 903, 1017
Infrastratego, 808
Infrastructure, 30, 91, 100, 101, 103, 132, 235, 294, 318, 322, 326, 328, 336, 341, 478, 479, 489, 505–508, 533, 552, 554, 555, 573, 574, 576–579, 638, 675, 719, 864, 887, 894, 906, 944, 979, 987, 989, 1006, 1015, 1018, 1023, 1026, 1028, 1029, 1031, 1034, 1037
In-game mechanics, 969
Innocentive, 90, 378, 463, 988, 989
Innovation, 20, 34, 44, 46, 92, 247–253, 351, 431, 449, 644, 703, 706, 893–895, 907, 912, 967, 972, 980, 984, 986, 988–989, 1021, 707, 762
Innovation network, 247–253
InPhO system, 145–146
Insight problems, 378, 450
Instrumental privacy, 866
INT, 810, 812
Integrated Development Environment (IDE), 260, 579, 1029
Intellectual property, 831, 882–883, 888, 922–924, 929, 988, 989
Intelligent, 7, 9–10, 131, 247, 248, 251, 253, 298, 424, 505–508, 510, 511, 513, 514, 519, 520, 598, 616, 633–639, 679, 680, 705, 706, 792, 897, 900, 903, 907
Intellipedia, 600

Interaction paradigm, 334, 488
Interactive crowds, 509–520
Interactive genetic algorithm, 644
Internal representation, 287
Internet, 4, 5, 13, 23, 71, 106–109, 124, 126, 137, 172, 173, 175, 179, 187, 190, 196–200, 297, 314, 378, 382, 384, 393, 423, 464, 599, 679, 680, 684, 685, 735, 813, 839, 842, 857, 860, 862–864, 870, 873, 897, 901–903, 905–907, 912, 915–918, 920, 931, 939, 952, 953, 975, 986–988, 999, 1015, 1018, 111120
Internet of Things, 975
Intrinsic motivation, 273, 281, 400, 402, 429
Intuitiveness, 258, 507, 548, 549
Invasive, 29, 30, 165, 174, 853, 941, 949, 974, 1014
Invisible hand, 238, 901
Invisible Human Computation, 341–342
iPhone, 299
iPod, 335
iPoPP, 100
Iterative workflow, 626

J

Junaio, 318, 319
Justice, 327, 805, 825, 976, 1004–1005

K

Kepler, 156
Khan, Genghis, 174
Knowledge
  acquisition bottleneck, 898
  base, 131, 134, 137, 141, 147–149, 199, 288, 328, 533–535, 540, 1034
  engineering, 131–150, 287, 290
Knowledge retrieval, 92, 131, 205–213, 421, 584
Kongregate, 404, 406

L

Labor, 7, 13–23, 76, 89–92, 144, 164, 187, 190, 191, 239, 268, 335, 341, 342, 379–381, 384, 387, 388, 412, 418, 475, 509, 618, 619, 820, 823–836, 903, 904, 917, 927, 946–950, 971, 972, 976, 983, 989, 1013, 1014, 1037
Laboratory, 22, 153, 156, 163, 165–166, 705, 746, 941
Labor market, 92, 144, 239, 379, 384, 388, 400, 828, 829, 835, 836
Labor standards, 191, 412, 418, 823–836, 1037
Labour, 13–16, 154, 507, 679
Language translation, 421
LarKC, 326
Layar, 318
Leaderlessness, 930–933, 939, 946
Leadership maintenance, 915–917
League of Legends, 103
Legion, 510, 514–517, 519
Leonhard, E., 14
Leveraging diversity, 926–929
Libya Revolution, 95, 97–98
Life in a day, 999
Life long learning, 861
Limited attention, 746, 747, 752, 756–757, 763, 764, 766
Limited capacity systems, 478
Linear threshold model, 780, 781, 786, 788
Linguistics, 416
Linked Data, 140, 145–149, 523–528
LinkedDataSail, 528
LinkedIn, 41, 66, 258, 261, 651, 658, 863, 903
Linked Open Data, 140, 290, 506, 523–528
Linked Open Data Cloud, 524–528
Link prediction, 866, 871
Linux, 793, 1027
LISTSERV, 863
Literary projects, 188, 190
Live Semantic Annotations (LSA), 575–579
Location Aware Services, 326
Location-based applications, 280, 297, 298, 310–312
Location-based deception, 842
Logical type, 52, 53, 58
LSA, Live Semantic Annotations (LSA)
Ludens, 281, 393–407
Ludic elements, 281, 395–398, 400–402, 407
Lurkers, 107, 109, 428

M

Machine
  affordances, 619–622
  agents, 57, 509, 1029
  intelligence, 182, 424, 513, 517, 803, 810, 813
  learning, 77, 84, 92, 102, 103, 124, 125, 136, 153, 156, 182, 183, 196, 379, 472, 585, 589, 627, 629, 645, 692, 769, 843
Machinery, 14–18, 457, 634, 936

Makeoutclub, 863
Malicious, 26, 175, 689–691, 819, 820, 838, 840, 842, 880, 881, 884, 887
Malware, 838
Manifesto, 108, 894, 1021–1037
Manufacture serendipity, 266, 267
MapHook, 304, 305
Market
  of lemons, 388
  place, 380, 382, 553, 839, 949, 980
Markov decision process, 627
Markov random field (MRF), 348
Mars, 6, 154, 265, 703, 824, 1014
MAS, 574–576
Masively multiplayer online role-playing games (MMORPG), 489, 807
Massively distributed, 57, 85, 999
Massively distributed problem solving, 57
Massively multiplayer online games, 96, 103
Massively Open Online Courses (MOOCs), 732, 1000, 1001
Mass self-communication, 862, 863, 866
Mathematical aggregation, 600, 603–608
Mathematical analysis, 747–748
Mathematical Tables Project, 17–22, 90
Mathematician, 13, 14, 16, 19, 22
Mathematics, 14, 19, 22, 993
mClerk, 337, 338, 341
MEA. *See* Means-ends analysis (MEA)
Mean field analysis, 780–786
Means-ends analysis (MEA), 352–354, 362
Me-centric, 428, 429
Mechanism design, 472, 1031, 734940
MediaWiki, 286, 290
Membership in a larger whole, 483
Meme, 303, 763–766, 841
Memetic algorithms, 282, 643
Memory, 26, 30–33, 43, 56, 63, 171, 196, 414, 451, 496, 506, 507, 512, 514, 547–549, 615, 618, 621, 710, 747, 764–767, 797, 803, 898, 899, 944, 972–974, 976, 979, 987, 1005, 1023, 1028
Mental illness, 4, 51, 52, 1007, 1032
Meta-rules, 819
Metric for synchrony, 794
Microtask(s), 22, 96, 98–103, 137, 144–149, 281, 334, 342, 379, 388, 423, 568, 569, 570, 572, 677, 680, 827, 829
Microtask crowdsourcing, 137, 144–145, 147
Microworking, 680, 682, 687, 692
Middle East peace process, 265
Middleware, 11, 554, 555, 557, 579, 1028
Minimum wage, 191, 381, 824, 826–828, 830, 843
Misrepresentation, 429, 838, 998
Mixed-initiative system(s), 617
MMORPG. *See* Masively multiplayer online role-playing games (MMORPG)
Mobi, 511, 630, 631
Mobile apps, 298, 300, 303, 307, 312, 315, 995
Mobile AR, 318–322, 325, 326, 329, 330
Modality, 90, 279–283, 431, 438, 441, 488, 491, 493–495, 498, 548, 583, 635, 939, 940, 1028
Modular graphs, 781
Modularity, 779, 788, 804, 1029
Monetary reward, 553, 886
MOOCs. *See* Massively Open Online Courses (MOOCs)
Moral(s), 18, 224–226, 228, 241, 342, 430, 431, 433, 434, 437, 705, 983, 1016, 1037
Moral hazard, 224–226, 228, 241, 430, 431, 433, 434, 437
Morphology, 41, 42, 700, 932
Motivating users, 682, 686–687
Motivation, 79, 156, 158, 159, 161, 206–207, 256, 267, 272–274, 281, 282, 303, 305, 311, 338, 377, 382, 385, 388, 398, 400–402, 407, 429, 432, 484, 549, 615–616, 633, 653, 656, 675–677, 680–681, 686–688, 696–699, 711, 884, 885, 887, 964, 994, 1001, 1032
Mozilla web browser, 793
MRF. *See* Markov random field (MRF)
MSA. *See* Multiple sequence alignment (MSA)
Multi-agent systems, 369
Multi cellular, 25
Multi-channel chat, 710
Multiple choice questions, 385, 386
Multiple instance learning, 183, 184
Multiple sequence alignment (MSA), 394, 515
Mutation, 641–643, 645, 646, 954
MySpace, 863, 864, 868

N

NASA, 6, 1014
Nash equilibrium, 730, 731, 733, 734
National Bureau of Standards, 21
Natural computation, 25, 961
Natural human computation, 961–964, 969

Natural human computation (NHC), 961–966, 970–974, 976
Natural language, 108, 131, 136, 137, 141, 288, 393, 397, 485, 517, 533–535, 540–542, 548, 625, 631, 686, 694, 884
Natural language processing (NLP), 108, 131, 136, 137, 141, 288, 517, 625, 629, 686, 884
Natural language technique (NLT), 108, 136, 288
Natural systems, 26, 39, 482, 488, 748, 762
Netflix, 90, 531, 649
Network analysis, 46, 86, 215–241, 428, 745, 748, 862, 871
Network architecture, 457, 795–797
Network density, 480, 481
Network effects, 387
Network evolution, 762
Network theory, 40, 44–46
Neural, 4, 26, 39–46, 63, 156, 449, 634, 791, 898, 899, 903, 947
Neural network, 63, 449, 899
Neurobiology, 477, 820, 848–851, 853
Neurochemistry, 848–850
Neuron, 26, 40, 42, 63, 480, 481
Neuronal, 40–42, 481
Neuroplasticity, 477
Newgrounds, 404, 406
News item recommendation, 664–668
Newton, 93
NHC. *See* Natural human computation (NHC)
Nike+, 305
NLP. *See* Natural language processing (NLP)
NLT. *See* Natural language technique (NLT)
Nodes, 45, 63–66, 147, 228, 231, 248, 249, 251, 348–351, 354, 357–360, 363, 368–374, 412, 413, 416, 417, 469, 477, 480, 481, 524, 545, 566, 576, 579, 748, 754, 762, 779–788, 795, 796, 803, 809, 902, 971, 1032
Nominal group technique, 710
Nonverbal signals, 707
NP hard problem, 546
Null hypothesis, 355, 356
Nurotransmitter(s), 42, 43, 56, 58, 849

O

Obama, B., 995
Objective, 84, 105, 109–111, 117, 123, 124, 207, 208, 228, 256, 303, 315, 354, 457, 458, 464–469, 498, 612, 634, 642, 644, 717, 719, 810, 862, 1018
Objective function, 464–469
Object-oriented, 132, 155, 280
Object-oriented programming, 132
Object recognition, 325
Obsessive-compulsive disorder (OCD), 56–58
OCR. *See* Optical character recognition (OCR)
oDesk, 15, 210, 216, 384, 431, 551, 625
Offline transcription, 509
Off-shore, 1014
Old Weather project, 160
Oligogyny, 930–933, 939, 946
Online collaboration, 676, 703, 820
Online collective action, 248, 424–426, 433
Online economies, 825
Online knowledge exchanges, 426
Online marketplaces, 377, 380, 387, 739
Online social space, 858
Online teams, 703, 704, 707–709, 711
OntoBay, 526
OnToGalaxy, 393, 401
Ontogeny, 1025
Ontological classification, 144–145, 150
Ontology
  alignment, 134, 142, 146, 150
  management, 134
OntoPronto, 138–139, 142, 525, 526
OntoTube, 526
OntoWiki, 290
OODA, 257
Open access publishing, 119
Open Science Collaboration, 378
Open Source Software (OSS), 426, 749, 791, 792, 794, 798–799, 894, 912, 913, 917, 918, 921, 929, 1027, 1028
OpenStreetMap (OSM), 96, 100, 101, 143, 297, 309, 323, 324, 1003, 1004
Optical character recognition (OCR), 337, 515, 546, 625
Organismic, 54, 86, 282, 475–500, 1026, 1032
Organismic computing, 54, 86, 282, 475–500, 1026, 1032
Organizational behavior, 704
Organizational management, 704
Organizational structures, 912, 930
Ornithology, 163
OSM. *See* OpenStreetMap (OSM)
OSS. *See* Open Source Software (OSS)
Oxford English Dictionary, 90, 378

P

PageRank, 234, 600, 691
Pandora, 531, 893
Paradox, 426, 606, 980–982
Parallel distributed processing, 909
Parkinson's disease, 115, 118, 124–126

Participation, 83, 92, 109, 111, 118, 153, 163–165, 169, 177, 187–190, 194, 197, 200, 238, 247, 255, 256, 272, 274, 282, 306–307, 310, 311, 320, 339, 387, 417, 422, 423, 428–430, 441, 445, 489, 499, 589, 610, 617, 675–677, 679, 685, 686, 688, 692, 697–700, 706–708, 710–712, 715, 725–729, 736–738, 740, 823, 836, 862, 864, 869, 884, 887, 907, 917, 929, 940, 975, 996, 997, 1011, 1017, 1018, 1028, 1030–1032, 1034, 1035
Participatory sensing, 86, 338, 581–583, 586, 590
PatientsLikeMe, 111, 112, 113, 115–119, 121–123, 852
PatientsLikeMe.com, 852
Pattern, 6, 9, 10, 41, 54, 62, 74, 82, 84, 90, 160, 192, 282, 447, 469, 470, 481, 510, 577, 587, 700, 898, 948, 951
Pattern recognition, 6, 74, 84, 700
Patterns of connection, 4–12, 479
Payoff matrix, 565
PCR. *See* Polymerase chain reaction (PCR)
PD, Phrase Detective (PD)
Peer-review(d), 167, 695
People recommendation, 650, 651, 653–655, 668, 669
Performance metrics, 814
Periodic, 62
Personal incentives, 686
Personalisation techniques, 588
Perspective hierarchy, 584
Persuasive computing, 993
Pervasive computing, 86, 280, 334–338, 340–343, 573, 574
Pervasive gaming, 338
Pervasive human computing, 280, 333–343
Pharmaceutical, 54, 56
Phase transition(s), 40, 41, 781, 786, 930, 950, 953, 1026
Phenotype, 118–119
Pheromones, 26–28, 31, 32, 936
Pheromone trails, 28–30, 32, 280, 349, 708, 941, 943, 944, 963
Philanthropy, 95, 104
Philosophy, 71, 72, 145, 524, 611
Phishing, 838
PhotoCity, 339
Phrase Detective (PD), 681–684, 690, 691
Phylo, 6, 11, 12, 394, 397–399, 402
Phylogeny, 55
Physics, 416
Physiology, 34, 848, 849
PicBreeder, 644
Piece-rate compensation, 828
Pieplue, 808
Pipelined tasks, 625
Planet, 156, 159, 160, 458, 907, 989, 992, 1014, 696905
Planetary, 894, 897, 905, 906, 1024
PLASH. *See* Platform for Location Aware Services with Human Computation (PLASH)
Platform, 91, 92, 96, 99–103, 106, 117, 118, 121, 145, 146, 153–155, 175, 179, 189, 193, 197, 238, 248, 258, 260, 266, 290, 304, 306, 322, 326, 328, 329, 338, 339, 367, 377, 379, 384, 387, 405, 424, 491, 499, 508, 525–529, 532–534, 537, 552–555, 561–571, 684, 685, 696, 725, 732, 763, 777, 798, 823, 831–835, 862, 868, 903, 924, 971, 975, 976, 1000, 1028, 1029, 1034
Platform for Location Aware Services with Human Computation (PLASH), 326, 327
Pocess in parallel, 25, 980
Poduct traceability, 998
Poetry engines, 188
POI, 300, 307, 308, 310, 318
Poi Friend, 328
Points of interest, 102, 143, 144, 304, 307, 308, 318, 323, 326, 526
Policy makers, 241, 824
Polistine, 912, 913, 929
Polistine wasps, 912, 929
Political, 4, 41, 44, 54, 76, 104, 191, 241, 379, 437, 438, 751, 779, 804, 805, 838, 868, 871, 907, 976, 986, 994–997, 999, 1002, 1003, 1007
Political campaigns, 779, 995, 996
Political engagement, 994–997
Politics, 599, 972, 995–997
Pollution, 339, 586, 907
Polydomy, 952, 953
Polymerase chain reaction (PCR), 1022
Ponerine ants, 912, 917–918, 920, 922, 926, 929
Population, 45, 101, 123, 124, 134, 138, 172, 265, 268–270, 273, 297, 348, 351, 356, 379, 393, 449, 463, 477, 485, 489, 492, 493, 498, 589, 609, 617, 641, 643, 645, 647, 738, 739, 745–747, 753–755, 757, 770, 774, 777, 784, 787, 788, 794, 842, 843, 880, 881, 933, 963–965, 980–982, 984–987, 990, 994, 1014, 1036

Poverty reduction, 1001–1002
PPSR. *See* public participation in scientific research (PPSR)
Prediction markets, 10, 216, 238, 367–375, 600, 603, 607, 609–611, 635, 732
Preventing conflict, 995
Primitive functions, 8, 562, 975
Principal-agent, 429–431, 433, 435
Privacy, 126, 516, 589, 651, 658, 668, 820, 825, 831, 833, 834, 838–840, 847–855, 857–874, 967, 1003, 1015
Privacy By Design, 847–855
Privacy leakage, 838–840
Privacy settings, 868, 869, 871–873
Private data, 867
Prizes, 687, 726, 885, 954
Problem solving, 7, 46, 57, 58, 85, 131, 273, 280, 333, 343, 347–365, 378, 393, 395, 421, 422, 424, 425, 430, 437, 458, 463, 464, 472, 618, 633, 634, 636–638, 709, 806, 880, 882, 886, 888, 894, 905, 989, 996, 1026, 1032
Problem solving pavers (PSP), 348, 351, 364
Problem solving template, 364
Prodigy, 86
Programming language, 507, 536, 563–564, 1028–1029
Progressive social agendas, 1012, 1015, 1018
Project Noah, 306
Propaganda, 76, 838, 842, 870
Proscriptive analysis, 466
Protein folding, 157, 158, 173, 510, 695, 837, 965
Protocol, 64, 167, 527, 552, 554–558, 634
Proton ontology, 138, 143, 525
Prototype, 118, 257, 281, 403, 810
Prototyping, 256, 257, 260
Provenance, 290, 293, 294, 524, 525, 528, 1005
PSP. *See* Problem solving pavers (PSP)
Psychological manipulation, 400, 547
Psychology, 3, 79, 413, 414, 416, 463, 512, 605, 611, 676, 679, 704, 719, 726, 804, 805, 859
Psychometric, 627, 804, 805, 813
Psychopathology, 51–58
Psychosis, 124, 749, 803–814
Psychosocial, 124, 749, 803–814
Public data, 867
Public health, 123, 124, 907
Public participation in scientific research (PPSR), 153
Puzzle(s), 3, 6, 157, 360, 476, 477, 482, 695, 752, 755, 756, 874, 962, 965, 976

Q
Q&A services, 561
QR codes, 971
Quantified self, 105–127
Quantum mechanics, 452
Query, 92, 132, 144, 145, 147, 205–213, 286, 290, 325, 339, 379, 464, 524, 528, 534, 541, 542, 562, 570, 586, 709, 903
Query refinement, 92, 205–207, 209, 210, 212, 213
Question and answer, 424, 433, 799
Question and answer system, 424
Questionnaire, 110, 111, 118, 120, 189
QuikTurkit, 510
Quora, 725, 728
Quorum, 926, 945, 946
Qurk, 562

R
Racism, 994
Random, 21, 26, 31, 45, 62, 67, 138, 142, 143, 175, 178, 183, 195, 271, 303, 313, 348, 350, 365, 386, 404, 406, 430, 464–467, 469, 470, 621, 636, 642, 643, 645, 657, 735, 754, 762, 765, 780, 782–787, 794, 796, 798, 812, 871, 885, 935, 939, 941, 945, 951, 954, 981, 982
Rapport game, 139–140
Rational datasources, 565–567
RCA. *See* Root cause analysis (RCA)
Real-time, 6, 96, 98, 124, 138, 142, 189, 310, 387, 389, 401, 433, 509–519, 631, 649, 650, 711, 712, 843, 898, 940
Real-time crowdsourcing, 509–520
Real-time feedback, 711, 712
Real-time social streams, 649, 650
Reasoning, 73, 84, 86, 131, 176, 282, 286–290, 328, 352, 368, 397, 414, 427, 433, 441, 442, 482–484, 486, 499, 506, 534–536, 582, 606, 618, 619, 637, 638, 705, 736, 803, 805, 807, 813, 935, 1013
Recall, 102, 144, 208, 209, 211–213, 450, 451, 618, 621, 700, 898, 972–974
Recaptcha, 379, 962
reCAPTCHA, 200, 601, 939, 964
Recombination, 194, 641–643, 645, 646
Recommender systems, 585, 588, 649–651, 656, 658, 668, 669
Reconciliation, 993, 1004–1006

Recruit, 28, 31, 109, 117, 120, 226, 293, 315, 379–381, 384, 437, 512, 578, 601, 650, 688, 728, 769, 880, 943
Recruitment, 29–33, 117, 120, 177, 273, 380, 381, 421, 424, 433, 601, 602, 650, 942–944, 946, 951
Recursive recall (RR), 450–451, 456
Red balloon challenge, 22, 839
Reengineer, 188
Refactor, 18
Refactoring, 14
Registry, 553, 554, 556
Regressive social trends, 1011
Regulatory
 authorities, 825
 capital, 218, 234, 236, 240
Relational database, 563
Reliability of data, 386
Remote person call, 333
Repression, 838, 1016
Reproduction, 192, 636, 911, 912, 914, 919, 921, 922, 925–929, 952
Reputation
 market, 239
 systems, 431, 739, 832, 833, 843
Requester portal, 553
Resource Description Framework (RDF)
 agents, 527
 triples, 290
Response aggregation, 607
Restorative restructuring, 282, 450–451, 453–455, 457
Result-sharing algorithm, 634
Retrieval, 92, 131, 136, 205–212, 421, 581, 584, 586, 587, 589, 590, 862, 953
Reward, 158, 173, 185, 205, 216, 217, 225–226, 233, 240, 272–274, 313, 337, 338, 382, 400, 401, 407, 506, 507, 533. 553, 564, 675, 682, 686, 687, 725–728, 730–734, 739–741, 819, 835, 849, 850, 882, 884–888, 968
RFID, 587
Risk, 20, 22, 92, 126, 215–241, 270, 299, 350, 380, 388, 609, 633, 637, 687, 720, 749, 804–806, 813, 830, 837, 838, 870, 871, 881, 883, 888, 928, 981, 994, 998, 1007, 1024, 1025, 1036, 1037
 perception, 804–806, 813
Roamler, 306, 307
Robot, 161, 289, 512, 515, 631, 708, 934, 935
 collaboration, 708
Robotic(s), 26, 618, 792, 898, 993, 1007, 1023
Robust computing, 26
Roger's Paradox, 980–982
Role-playing games, 806, 807, 1000
Role-playing games (RPG), 100, 806, 807
Root cause analysis (RCA), 352, 354–355
ROSETTA@home, 153
Routing task, 26
RPG. *See* Role-playing games (RPG)
RR. *See* Recursive recall (RR)
Runtastic, 305

S

SAPERE, 508, 573–579
Satellite imagery, 98–101, 174–178, 182, 184, 834, 1003
SBS. *See* Software-Based Services (SBS)
Scalable, 25–27, 29, 30, 33, 102, 108, 172, 182, 489, 528, 549, 619, 646, 835, 1031
Scales, 39, 41, 43, 45, 46, 53, 63, 65, 110, 173, 253, 728, 737, 738, 983
Schema, 131, 133, 134, 155, 288, 290, 524, 556, 563, 564, 568
Schizophrenia, 52–54
Schooling, 791
Scientific crowdsourcing, 153, 154, 157
Scientific inquiry, 91, 167, 696, 849
Scrip, 829
SCVNGR, 302, 303
Search and discovery, 92, 171–185
Search and rescue, 92, 96, 174
Search engine, 199, 235, 587, 841, 843, 902
Security(s), 20, 54, 98, 174, 221, 222, 229, 236, 597, 718, 809, 810, 819–821, 833, 854, 869, 879–888, 1011, 1012, 1034, 1037
Seizure, 107, 481
Sekai camera, 320, 321
Self-censorship, 869
Self-disclosure, 866–869
Self-interested agents, 633, 726
Self-interested individuals, 422–423, 425, 428, 442
Self-organization, 68, 428, 452, 574, 792, 894, 897–907
Self-organized, 22, 449, 791, 792, 967, 969, 971
Self-regulation, 239, 849
Self-report, 105, 111, 114
Semantic, 85, 91, 131–136, 139, 140, 144–146, 148, 280, 286–290, 292, 294, 322, 348, 351, 352, 359–360, 379, 404, 421, 432, 433, 508, 523–529, 543, 574, 575, 578, 579, 653, 669, 865

Semantic (*cont.*)
annotation, 132, 134, 135, 144, 525, 526
mediawiki, 290
network, 139, 140, 286, 288, 351, 352, 359–360, 525
web, 91, 131–133, 290, 294, 523–529
wiki, 290, 292
Semantic Automated Discovery and Integration (SADI), 527
Semantic query, 432
Semiotics, 62
Semi-structured knowledge, 286–289
Sensate reality, 487, 873
Serendipitous discovery, 155, 156, 495
Serious games, 803, 806
Service oriented protocols, 551–558, 1028
SETI@home, 153, 159
Shareable resources, 924–926
Shared goals, 76, 77, 80, 476, 617
Shared sensing, 82, 282, 483, 486, 499
Short-term memory, 547
Signaling, 42, 251, 430, 437, 708, 941
Simple Knowledge Organization System (SKOS), 526
Simulated, 31–33, 157, 159, 451, 469, 489, 492, 510, 754, 756, 762, 764, 765, 798, 799, 941–943, 1029
augmented reality, 280, 282, 301, 317–330, 487–490, 492, 873, 894, 939–941, 970
Simulation(s), 31, 44, 369, 487, 489, 745–748, 754–755, 761, 763, 764, 766, 783, 785, 806, 808, 812, 813, 849, 905, 916
Singularity, 953, 954, 1025–1026
Sir Tim Berners-Lee, 131
Situated, 187, 339, 340, 416–418, 486, 864, 898, 1031
SixDegrees.com, 863
Size weight and power, 480
Skierarchy, 629
Skilled workers, 555, 646, 924
SKOS. *See* Simple Knowledge Organization System (SKOS)
Skout, 299, 300
Skype, 98, 903
Smart cards, 587–589
Smart Cities, 143
Smart phone, 103
Smartphone, 123–125, 299, 526, 581, 583, 586, 589, 966
SmartSheet, 551
SME, 256, 263
SMS, 96, 187
SNA. *See* Social Network Analysis (SNA)
Snowmaggedon, 1002
Social aggregation system, 651
Social behavior, 86, 256, 616, 745, 747, 749, 751–758, 769–777, 858, 964, 969, 970, 979, 1027, 1032
Social business process management, 255, 424
Social capacity, 262–263
Social capital, 428, 433, 437, 650, 793, 861–862, 864, 865
Social collectives, 421–442
Social community, 858
Social computing, 4, 33, 86, 137, 315, 616, 746, 748, 749, 779, 862, 899
Social dilemma, 471, 472
Social distance, 999
Social dynamics, 608, 720, 762, 770–771, 774
Social engagement, 676, 698, 850, 859
Social epistemology, 71, 73, 75–80
Social graph, 248, 438–440
Social incentives, 685, 687
Social influence, 762, 774, 981
Social informatics, 86, 751–758, 1032
Social information, 30–32, 752–753, 865
Social intelligence, 803, 813
Social interactions, 44, 45, 65, 92, 124, 255–258, 261, 263, 340, 427, 477–479, 484, 687, 719, 720, 745–748, 757, 762, 850–851, 857
Social knowledge, 86, 285–294
Social learners, 981, 982
Socially useful goals, 256–257
Social media, 44, 95–98, 101, 107, 256, 260, 308, 327, 339, 428, 439, 589, 599, 649, 650, 655, 658–659, 662, 665, 668, 669, 709, 710, 736, 746, 747, 751, 752, 755, 757, 763, 766, 769–771, 775, 777, 813, 838, 840–844, 857, 865, 907, 986, 995, 999, 1006
Social media manipulation, 841
Social network
applications, 940
sites, 196, 650, 653, 658, 668
Social Network Analysis (SNA), 86, 428, 862, 871
Social norms, 428, 739, 858, 974
Social perceptiveness, 477, 676, 705–712
Social policy, 265
Social privacy, 866, 868, 871
Social progress, 894, 1011–1018
Social psychology, 611, 704, 726, 1032
Social recognition, 248, 250–253
Social recommendation, 598, 649–669

Social recommender systems, 649, 650, 668
Social sciences, 414, 762
Social stability, 994
Social surveillance, 869
Social synchrony, 748–749, 792–794, 797–799
Social trends, 1011
Social web, 874
Social wisdom, 899
Society, 7, 9, 14, 21, 39, 41, 44, 72, 89, 90, 163, 174, 185, 197, 230, 236, 388, 762, 792, 855, 857–860, 862, 870, 893, 899–901, 906, 907, 911–913, 929–930, 953, 976, 986, 990, 1012, 1014, 1016, 1017, 1034, 1036, 1037
Socio-cultural, 226, 227, 619, 745, 749, 804, 805, 976, 998
Software-Based Services, 552
Software-Based Services (SBS), 552, 553, 555
Solar energy, 1014
Solidarity, 835
Somali Crisis, 95, 98–99
Somatic, 111
SONAR, 651, 653–655, 660
Source control, 918–920
Soylent, 510–512, 546
Spam, 145, 146, 157, 200, 838, 840
SPARQL, 145, 147, 528
Sparse network, 861
SpeechInk, 551
SpotTheLink, 142–143, 526
SRS, 668
Stable computation, 84, 902, 962, 964, 969
Stable human computation, 84
Stack Overflow, 273, 799
Stakeholder(s), 135, 832, 833, 835, 854, 880, 1006
Standards, 8, 21, 165, 191, 290, 412, 418, 428, 552, 554, 555, 558, 602, 607, 737, 738, 820, 823–836, 987, 1037
Standby Volunteer Task Force, 97, 98, 100
Stardust@home, 6, 10, 154
State diagram, 770, 772, 773
State space, 486, 492, 529, 536, 990, 1028
Static, 62, 106, 111, 320, 328, 450, 457, 491, 590, 684, 949
Statistical hypothesis testing, 352, 355–357
STEM, 165, 168, 411, 917
Stigmergy, 26, 708, 712, 903, 904, 908
Stochastic, 28, 63, 216, 219, 224, 225, 228, 232, 233, 237, 240, 469, 636, 747, 748, 769–777, 787
  modeling, 747, 748, 769–777
Stock market, 791, 953
StrangerRec, 655
Stranger recommendation, 655–658
Stress, 55–58, 781, 994
Structured populations, 779–788
Subclass, 146, 287, 292, 537
Subjective, 110, 123, 237, 382, 397, 415, 417, 464, 562, 603, 699, 862, 942
SubvertAndProfit, 840
Superclass, 146
Super-classifiers, 154
Supercomputers, 980
Superorganisms, 26, 57–58, 636, 941, 1032
Supervised machine learning, 588
Surreptitious behavior, 1037
SWaP, 480, 481
Swarm
  behavior, 467, 636, 637, 935, 963, 967, 969–971
  intelligence, 26, 636
  robotics, 26
  theory, 967
Symbiosis, 897, 899, 954
Symbiotic, 188, 201, 954
Synchronization, 485, 498, 748, 749, 791–797, 916
Synchronous, 63, 312, 333, 481, 489, 510, 512, 513, 518, 519, 715
  workers, 519
Synchrony, 29, 745, 748–749, 791–800
  patterns, 799
Synergetically, 897
Synergistic, 32, 477, 495, 954
Synergy, 32, 34, 185, 476–478, 482, 484, 488, 489, 498, 499, 748, 797–799, 1026, 1034
Synthesis, 4, 83–86, 513, 515, 519, 533, 541, 849
Systems theory, 85, 969

T
Tacit knowledge, 989
TalkTo, 107, 479, 601, 1018
Task completion, 378, 388, 510, 683, 687, 716, 720, 839
Task decomposition, 388, 414, 631
Task marketplace, 553
Task-sharing algorithm, 634
Tasmania, 985, 986
Taxonomy, 4, 83–86, 145, 148, 212, 290, 551, 629, 837
Technique, 21, 22, 289, 352, 412, 487, 490, 616, 618, 658, 710, 799, 843, 962, 980, 1022

Technological singularity, 953, 1025, 1026
Technology mediated, 57, 86, 475, 479, 489, 617, 715–722, 912, 917–918, 926, 1026
Technology-mediated collaboration, 475, 489, 715–722
Technology-mediated leadership, 717–718, 917–918, 926
Technosocial predictive analysis, 815
TEDx, 95, 102, 110
Telecommunications, 61, 171
Temporal discounting, 848–850
Temporal workflows, 630
Text editing, 421
Text message, 414
The Middle East, 265
Thermocycler, 1022, 1023
The WELL, 106
Thick community, 860–861
Thin community, 860–861
Tim Berners-Lee, 131
Top-down, 916, 954, 961, 963, 974, 996, 1027
Topology, 64, 282, 762, 844, 901, 950
Tournament selection, 642
Tour Route Recommendation, 326
Training, 91, 110, 123, 135–137, 156–158, 161, 164, 165, 174, 175, 184, 196, 292, 293, 305, 398, 431, 507, 515, 516, 519, 534, 543, 629, 646, 688, 689, 806, 832, 835, 929, 996, 1001
Transaction cost effectiveness, 379
Transformation, 62, 63, 234, 260, 1032
Transparency, 256, 380, 387, 388, 833, 854, 917, 998
Trial-and-error, 901, 906, 980, 981
Trippy, 303, 304
Trust, 111, 252, 253, 274, 294, 380, 399, 404, 407, 428, 430, 433, 437, 651, 664, 668, 676, 716, 718–720, 832, 833, 843, 847, 852, 854, 861, 868, 870, 874, 883, 885, 901, 998, 1003, 1004
Trust & Tracing, 101, 111, 252, 253, 273, 294, 380, 399, 404, 407, 428, 430, 433, 437, 651, 654, 668, 676, 716, 718–720, 832, 833, 843, 847, 852, 854, 861, 868, 870, 874, 883, 885, 901, 998, 1003, 1004
Truth and Reconciliation Committees, 1006
T-test, 212, 794
Turing, A., 62
Turkopticon, 380, 832, 843, 844
Tweet(s), 189, 747, 751, 753, 864, 1018
Tweets per day, 104, 863
23andMe, 118, 119, 123
Twitter, 41, 44, 65, 66, 68, 96, 100, 102, 104, 187, 189, 190, 196, 200, 257, 320, 321, 323, 325, 328, 339, 378, 438, 439, 479, 589, 601, 664, 685, 687, 709, 747, 751–753, 755, 756, 763, 764, 769, 813, 841, 842, 844, 863, 864, 964, 986, 994, 1002
TxtEagle, 336–378
Typhoon Pablo, 95, 100–101, 1026

U

UAV, 178
UbiAsk, 323
Ubiquitous collective computation, 939–941
Ubiquitous Human Computing, 339, 832
UML, 132
Unified output, 513
United Nations, 97, 98, 994, 997
Unsupervised machine learning, 617
Unwitting participation, 190, 197
UrbanMatch, 143–144, 308, 309, 323, 526
Urbanopoly, 309, 323
URL, 147, 199, 209, 261, 320, 752, 753, 842
USENET, 106, 863
User interface design, 293, 414
User weighting, 159–160
Use-value, 970, 973, 974
Ushahidi, 6, 10, 95, 96, 837, 1002, 1026
Utest, 489
Utopia, 506, 888

V

VCS. *See* Virtual citizen science (VCS)
Verbosity, 729
Viral marketing, 779, 780
Virtual
  assets, 828, 829, 831
  community, 99, 106, 298, 645, 860–862, 864
  currency, 103, 824, 828–831
  economy, 829
  environment, 489, 823
  pet, 139–140
  reality, 487, 490, 860
  sweatshop, 3
  teams, 638, 716
  workforce, 839
Virtual citizen science (VCS), 695–700
Virtuous circle, 105–127
Virus, 111, 157, 757, 1024
Visual analytics, 180, 617, 998

Visualization, 113, 115, 173, 177, 273, 293, 431, 437, 440, 441, 524, 622, 631, 799, 966, 1003
Visual language, 258, 282
VizWiz, 336, 378, 510–512, 518
Volunteer attrition, 686
Volunteer computing, 162
Volunteered Geography, 297
Vote(ing), 141, 143–145, 159, 184, 249, 252, 258, 260, 266, 370, 371, 422, 431, 436, 440, 441, 515, 517, 518, 626, 628, 629, 635, 711, 727, 729, 734–735, 751, 752, 755, 771–777, 808, 840, 841, 845, 964, 997

W
Wage protection, 836
War, 18–22, 56, 207, 836, 842, 907, 993–995, 997–999, 1002, 1003, 1006, 1007
Warfare, 996, 1006–1007
Wasp(s), 911–918, 921, 929, 937
Waste-management, 911
W3C. *See* World Wide Web Consortium (W3C)
Weak ties, 256, 257, 861, 865, 873
Wearable computing, 124, 334
Web
  of Data, 290, 292, 524
  of Documents, 524, 660
  ratio, 257, 261
  of Services, 523, 527, 528, 574, 577, 863
  of Things, 975
  of Thoughts, 581, 582
Web 2.0, 111, 134, 320, 322, 328, 599, 600, 861
Weblogs, 68, 294, 296, 392, 701, 758, 766, 767
WebML, 260, 261
WhoKnows?, 527
Whole Earth Catalog, 106, 108, 111
Wicked problem, 265, 266, 268, 270, 272, 1018
Wiki, 142, 148, 189, 190, 249, 250, 285, 286, 290, 292, 297, 307, 309, 359, 655, 659, 665, 667, 1003, 1022
Wikimedia, 143, 149, 285, 286, 290
  commons, 143, 149
Wikipedia, 65–66, 111, 138, 144, 147–149, 172, 249, 250, 273, 285, 286, 289–292, 294, 297, 304, 307, 309, 359, 385, 423, 426, 524, 525, 582, 584, 600, 603, 638, 676, 677, 679, 680, 682, 686, 690, 697, 752, 799, 844, 880, 904, 988, 995, 1022
Wikis, 64, 187, 285, 287–292, 426, 651, 658, 659, 662, 665, 667, 668
Wikitude, 318
William, G., 102
Wisdom, 134, 247, 251–253, 279, 333, 367, 369, 375, 381, 383, 599–612, 689, 699, 937–939
Wisdom of crowds, 383, 599–612, 899, 937, 938
WITNESS, 1005, 1006, 1035
WordHunger, 141–142
WordNet, 141, 404
Workflow, 11, 57, 136, 183, 249, 310–312, 388, 421, 439, 545, 598, 625–629, 631, 901, 903, 904
Working memory, 615, 621, 898
WorldBoard, 320
World Cellular Model, 808
WorldHunger, 141–142, 283
World population, 986
Worldview, 180, 265–267, 449, 835, 1030–1033
Worldview Evolution, 449
World Wide Web, 91, 132, 523, 524, 551, 863
World Wide Web Consortium (W3C), 523

Y
Yahoo! Answers, 561, 736, 841
Yelp, 318, 320, 328, 531, 734, 963–965, 969, 973
YouTube, 111, 172, 261, 320, 526, 527, 649, 769, 813, 1000

Z
ZenCrowd, 147, 526
Zooniverse, 91, 103, 154, 155, 159, 160, 676, 695–700