EDITED BY
LUIZ MOUTINHO
MLADEN SOKELE

INNOVATIVE RESEARCH METHODOLOGIES IN MANAGEMENT

VOLUME I: Philosophy, Measurement and Modelling

Innovative Research Methodologies in Management

Luiz Moutinho • Mladen Sokele
Editors

Innovative Research Methodologies in Management

Volume I: Philosophy, Measurement and Modelling

palgrave
macmillan

Editors
Luiz Moutinho
University of Suffolk, Suffolk, England, UK
The University of the South Pacific, Suva, Fiji

Mladen Sokele
Faculty of Electrical Engineering and Computing
University of Zagreb
Zagreb, Croatia

ISBN 978-3-319-64393-9 ISBN 978-3-319-64394-6 (eBook)
DOI 10.1007/978-3-319-64394-6

Library of Congress Control Number: 2017954805

Printed on acid-free paper

This Palgrave Macmillan imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To my VESF Lieutenant Colonel Dr. Hudek and to my late incredible and unique cat Jakey
Luiz Moutinho

To my lovely daughter Dunja
Mladen Sokele

Preface

I am very grateful to Palgrave for the fact that they were enthusiastic about this project. It has been one of my scholarly "pet" projects for some time. I was also extremely happy that I was able to secure the incredible collaboration of my friend Mladen Sokele as my co-editor of this book.

Methodologies are the basis for scientific research. However, scholars often focus on a very limited set of methodologies, partly due to a lack of knowledge about innovative methodologies outside their area. This is especially true in the area of management science. Providing management scholars with an education about methodologies outside their primary area of expertise is the goal of the proposal made to Palgrave Macmillan, global academic publisher, for a book providing a comprehensive presentation of innovative research methodologies: Palgrave's *Innovative Research Methodologies in Management*.

This book is to be positioned as a seminal collection of mind-stretching and thought-provoking research methodology essays. Hopefully, these research methods and topics will greatly enhance the research methodology "armoury" of management scholars and alter the research "modus operandi" of academic research output and published work.

One of the aims of the *Innovative Research Methodologies in Management* text is to identify and foster methodological research innovation in the academic management field. This book project seeks out research practices that are not highlighted through the majority of academic research outlets and journals, that are not highlighted typical research method courses or and to have an impact on the research process in novel ways. These innovative methodologies usually entail a combination of (i) technological innovation, (ii) the use of existing theory and methods in new ways, and (iii) interdisciplinary approaches. The project's focus on innovative research practices will range across new philosophical insights into academic research, new experimental designs to new technological research contexts, and new analytical techniques. Departing from the somewhat perennial situation of academic research methodolatry and scientism, this tome will be focusing on a series of rigorous methodological advances in many areas. The primary focus relies on emerging and bleeding-edge areas of academic research methodologies in management, making the contents of this book an authoritative source on the applications of these methodologies to management science.

Volume 1 is dedicated to the coverage of innovative research methodologies within the realms of research philosophy, research measurement, and research modeling.

Volume 2 is focused on chapters dealing with Futures Research, Biometrics Research, and Neuroscience Research in Management.

Chapter 1 (Zarkada, Panigyrakis, and Tsoumaka) introduces a panoply of metamodern mixed methods in management research dealing with Web 2.0+. It discusses metamodern socioeconomic phenomena, mixed methods designs, and sampling integrity, among many other topics. It follows an interesting light analogy and is very robust in terms of theoretical content. Interesting reflections are also included.

In Chap. 2, Hackett and Foxall tackle the topic of neurophilosophy. This is a chapter with a rich theoretical background. There is an interesting section on psychological constructs and neurophysiological events. There is also a challenging exploration into mereological understandings.

Kip Jones (Chap. 3) deals with emotivity and ephemera research. The content is focused on arts-led and biographical research as well as relational aesthetics. There are some interesting insights into neo emotivism. There is a challenging section on performative social science. There are also interesting comments on experimentation and the experimental redux.

Chapter 4 presents the novel approach—Abductive Thematic Network Analysis (ATNA) using ATLAS-ti written by Komalsingh Rambaree. This chapter introduces ATNA as a methodological approach for qualitative data analysis. It starts by providing a brief description on abductive theory of method and thematic analysis method. Then, it highlights how the two methods are combined to create ATNA. Using a qualitative dataset, this chapter demonstrates the steps in undertaking ATNA with a computer-aided qualitative data analysis software—ATLAS-ti v.7.5. The chapter concludes that ATNA provides to researchers a much-needed pragmatic and logical way of reasoning, organizing, and presenting qualitative data analysis.

Sullivan, Lao, and Templin (Chap. 5) deal with diagnostic measurement. With diagnostic measurement, the aim is to identify causes or underlying properties of a problem or characteristic for the purposes of making classification-based decisions. The decisions are based on a nuanced profile of attributes or skills obtained from observable characteristics of an individual. In this chapter, the authors discuss psychometric methodologies involved in engaging in diagnostic measurement. They define basic terms in measurement, describe diagnostic classification models in the context of latent variable models, demonstrate an empirical example, and express the broad purpose of how diagnostic assessment can be useful in management and related fields.

Yang and Fong (Chap. 6) explore the issues of incremental optimization mechanism for constructing a balanced, very fast decision tree for big data. Big data is a popular topic that highly attracts attentions of researchers from all over the world. How to mine valuable information from such huge volumes of data remains an open problem. As the most widely used technology of decision tree, imperfect data stream leads to tree size explosion and detrimental accuracy problems. Over-fitting problem and the imbalanced class distribution reduce the performance of the original decision tree algorithm for stream mining. In this chapter, the authors propose an Optimized Very Fast Decision Tree (OVFDT) that possesses an optimized node-splitting control mechanism using Hoeffding bound. Accuracy, tree size, and the learning time are the significant factors influencing the algorithm's performance. Naturally, a bigger tree size takes longer computation time. OVFDT is a pioneer model equipped with an incremental optimization mechanism that seeks for a balance between accuracy and tree size for data stream mining. OVFDT operates incrementally by a test-then-train approach. Two new methods of functional tree leaves are proposed to improve the accuracy with which the tree model makes a prediction for a new data stream in the testing phase. The optimized node-splitting mechanism controls the tree model growth in the training phase. The experiment shows that OVFDT obtains an optimal tree structure in numeric and nominal datasets.

Sokele and Moutinho (Chap. 7) introduce the Bass model with explanatory parameters. Over the 45 years, the Bass model is widely used in the forecasting of new technology diffusion and growth of new products/services. The Bass model has four parameters: market capacity; time when product/service is introduced; coefficient of innovation; and coefficient of imitation. Although values of coefficient of innovation and coefficient of imitation describe the process of how new product/service gets adopted as an interaction between users and potential users, their explanatory meaning is not perceptible. These authors explore this important gap.

Chapter 8 by Volker Nissen is titled "A Brief Introduction to Evolutionary Algorithms from the Perspective of Management Science." Summarizing, it is useful to differentiate between several perspectives. From a methodological point of view, the myriad of nature-inspired heuristics is rather confusing for casual users and definitely not helpful in creating more acceptance for metaheuristics in practical applications. Moreover, there is evidence (e.g., Weyland 2015; Sörensen 2015) that at least some of the nature-inspired concepts published recently (such as Harmony Search) are rather old wine in new skins. One should better look for over-arching and bearable concepts within evolutionary algorithms (EA) and related metaheuristics, putting together a common framework (or toolbox) that integrates different options of solution representation, search operators, selection mechanisms, constraint-handling techniques, termination criteria, and so on. Fortunately, such frameworks are available today. Then, properly choosing the components for a hybrid heuristic from such a framework requires a deep understanding of which components actually fit together and work well for certain classes of problems or search spaces.

Moreover, following the No-Free-Lunch Theorem (Wolpert and Macready 1997), problem-specific fine-tuning of heuristics remains important to achieve truly good results. Today, much of this is still more an art than a science, despite helpful textbooks like Rothlauf (2011). As a consequence, there are lots of interesting research issues to be solved along these lines. This process is indeed ongoing for several years now, but, as Sörensen et al. (2016) point out, also bypassed by useless initiatives to invent ever "new" nature-inspired heuristics.

For more than 20 years now, EA have become integrated in business software products (e.g., for production planning), so that, as a result, the end user is often unaware that an evolutionary approach to problem solving is employed (Nissen 1995). Today, large software companies like SAP use EA in their enterprise software. However, since customizing options are limited in these systems, it appears fair to say that the full power of EA is frequently not unleashed by such standardized approaches. This confronts us with a dilemma. Simple forms of EA that can be fairly easily understood and applied widely are of only limited power. If we want to use metaheuristics like EA to full extent, then this requires knowledge and experience in their design and application. Most users have neither the qualification nor the time to dive that deep into the matter. According to my own observations as an IT consultant, this unfortunately also holds for most consultants that could potentially help customers in applying modern heuristics. Thus, creating really powerful applications of EA today is frequently an issue for only a small number of highly specialized IT companies. This situation is unsatisfactory and could only be changed if the design and application of modern heuristics becomes an important topic in general management studies at universities and related higher learning institutions. The author argues that this should indeed be the case, because in today's digital era, there is strong evidence that we are entering an age of knowledge-based competition where a qualified workforce that is able to creatively use modern tools for data mining, ad-hoc reporting, heuristic optimization, artificial intelligence, and so on will make the difference in many branches of industry.

Volker Nissen also kindly contributes to Chap. 9 on applications of EA to management problems. EA (or Evolutionary Computation) represent as nature-inspired metaheuristics a branch of computational intelligence or soft computing. Their working is based on a rough abstraction of the mechanisms of natural evolution. They imitate biological principles, such as a population-based approach, the inheritance of information, the variation of solutions through crossover and mutation, and the selection of individual solutions for reproduction based on fitness. Different variants of EA exist, such as genetic algorithms, evolution strategies, and genetic programming.

This chapter reviews EA from the perspective of management applications where "management" indicates that predominantly economic targets are pursued. In general terms, the preferred area of application for EA, and other metaheuristics as well, is optimization problems that cannot be solved analytically or with efficient algorithms, such as linear programming, in a reasonable time or without making strong simplifying assumptions on the problem. Many of these problems are of a combinatorial nature, such as job shop scheduling, timetabling, nurse rostering, and vehicle routing, to name just a few. In practical settings, often the issue of "robustness" of a solution is equally important as "optimality," because the optimization context is characterized by uncertainty and changing conditions.

While the first variants of EA were already invented in the 1960s, it is in the last 15–20 years that these powerful methods of heuristic optimization have attracted broader attention also outside the scientific community. Providers of optimization software such as MathWorks with MATLAB, but also large players in enterprise resource planning (ERP), such as SAP AG with SAP APO, integrated EA in their canon of software-based optimization methods.

However, following the "No-Free-Lunch Theorem," metaheuristics such as EA always require a certain degree of adaptation to the individual problem at hand to provide good solutions. In practice, this can pose a problem for the inexperienced user of EA-based software products. On the scientific part, lately there appears to be an inflation of nature-inspired metaheuristic approaches, some of them related to EA and partly criticized for a lack of novelty.

The last chapter of Volume 1 is written by Beynon, Moutinho, and Veloutsou (Chap. 10). It is an exposition of the role of consideration sets in a Dempster-Shafer/Analytic Hierarchy Process (DS/AHP) analysis of consumer choice. Consumer behavior is often perceived through the notion of consideration sets. However, realistic modeling of consumer choice processes identifies impeding factors, including ignorance and non-specificity. In this chapter, the appeasement of these factors and the role of consideration sets are considered through the utilization of the nascent DS/AHP method of choice analysis. The central element in the DS/AHP analysis is the body of evidence (BOE), with certain BOE constructed at different stages in the analysis, then a number of different sets of results can be found. The research content is attempting to convey a more realistic approach for the individual consumer to undertake the required judgment-making process. The investigation is based on a group of consumers and their preferences on a number of cars over different criteria. The notion of consideration sets is shown to be fundamental within DS/AHP, and a novel approach to the aggregation of the preferences from the consumers is utilized. A notional approach to the identification of awareness, consideration, and choice sets is described, based on the levels of belief and plausibility in the best car existing in a group of cars, which could be compared with the algorithm developed by Gensch and Soofi (1995).

Volume 2 of *Innovative Research Methodologies in Management* starts with Chap. 1 by Simone Di Zio on the "Convergence of Experts' Opinions on the Territory: The Spatial Delphi and the Spatial Shang." The judgments of a panel of experts are of extreme usefulness when, in front of a decision-making problem, quantitative data are insufficient or completely absent. Experts' opinions are helpful in forecasting contexts, for the detection of innovative solutions or for the verification and refinement of consensus on objectives or alternative scenarios.

The way the views are collected is crucial, and without a rigorous methodology, any consultation process may become vain. In literature, there are many methods, but some are used for the ease of application rather than that for their scientific properties. Methods such as focus group, face-to-face interview, or online questionnaire are very popular but have quite important drawbacks.

Many of those disadvantages are overcome by the methods of the "Delphi family," whose prototype is the Delphi method, which involves the repeated administration of questionnaires, narrowing the range of assessment uncertainty without generating errors that result from face-to-face interactions. To date, the Delphi technique has a very high number of applications, and its success has produced a wide range of methods that are its variants.

In this chapter, we present two recent variants, called Spatial Delphi and Spatial Shang, applicable when consultations, and consequent decisions, concern matters of spatial location. The judgments of the experts are collected by means of points placed on a map, and the process of the convergence of opinions is built up through the use of simple geometric shapes (circles or rectangles). During the subsequent iterations of the procedure, the shapes become smaller and smaller, until to circumscribe a very small portion of territory that is the final solution to the research/decision problem.

After the discussion of the methods and the presentation of some practical applications, we propose some possible evolutions that most likely will produce a future increase in the application of these techniques.

Chapter 2, titled "Interactive Scenarios," is written by Theodore J. Gordon and Jerome Glenn. A scenario is a rich and detailed portrait of a plausible route to a future world, including issues and decisions that might have to be faced as the future unfolds. In this chapter, we are most interested in scenarios creation processes that involve inputs from more than a single person, involve feedback, and may use quantitative models to help establish a foundation for the scenario story. Several interactive methods and applications are described including (1) use of Delphi questionnaires to collect suggestions for scenario content and scenario axes, that is the major dimension that defines the domain of interest; (2) redrafting a scenario story based on feedback from reviewers; (3) integrating a cross-impact matrix, Futures Wheel, or other modeling system to help assure internal self-consistency and quantitative rigor; and (4) allowing the audience to determine the course of the scenario at key decision points. Also described is a clustering approach in which a multitude of scenarios are constructed, each of which differs in input assumptions resulting from uncertainties associated with the variables and the assumed policies. The chapter concludes with a brief speculation about the future of interactive scenarios: multi-mode presentations so realistic that users feel they are in the scenario world and make simulated decisions accordingly.

Chapter 3 by Raymond R. Burke introduces “Virtual Reality for Marketing Research.” Computer graphic simulations of retail shopping environments have become a popular tool for conducting marketing research, allowing manufacturers and retailers to test innovative marketing concepts with shoppers in realistic, competitive contexts. Virtual store tests can deliver more detailed behavioral data than traditional methods of consumer research and are faster and less expensive than in-store field tests. This chapter outlines the benefits and limitations of virtual reality simulations, describes the steps involved in creating and running a simulated shopping study, discusses the validity of the simulation technique, provides examples of several commercial and academic research applications, and summarizes the future prospects for using the virtual store for marketing research and other business applications.

Chapter 4 by Pestana, Wang, and Moutinho is titled “The Knowledge Domain of Affective Computing: A Scientometric Review.” The aim of this study is to investigate the bibliographical information about Affective Computing identifying advances, trends, major papers, connections, and areas of research. A scientometric analysis was applied using CiteSpace, of 5078 references about Affective Computing imported from the Web-of-Science Core Collection, covering the period of 1991–2016. The most cited, creative, bursts, and central references are displayed by areas of research, using metrics and throughout-time visualization. Interpretation is limited to references retrieved from the Web-of-Science Core Collection in the fields of management, psychology, and marketing. Nevertheless, the richness of bibliographical data obtained largely compensates this limitation. The study provides managers with a sound body of knowledge on Affective Computing, with which they can capture general public emotion in respect of their products and services, and on which they can base their marketing intelligence gathering and strategic planning. The chapter provides new opportunities for companies to enhance their capabilities in terms of customer relationships. The effect of emotions on brand recall mediated by gender using voice emotion response with optimal data analysis.

Chapter 5 is written by the same previous authors—Wang, Pestana, and Moutinho—and is titled "The Effect of Emotions on Brand Recall by Gender Using Voice Emotion Response with Optimal Data Analysis." Its purpose is to analyze the effect of emotions obtained by oral reproduction of advertising slogans established via Voice Emotion Response software on brand recall by gender, and to show the relevance for marketing communication of combining "Human-Computer-Interaction (HCI)" with "affective computing (AC)" as part of their mission. A qualitative data analysis did the review of the scientific literature retrieved from Web-of-Science Core Collection (WoSCC), using the CiteSpace's scientometric technique; the quantitative data analysis did the analysis of brand recall over a sample of Taiwanese participants by "optimal data analysis." Advertising effectiveness has a positive association with emotions; brand recall varies with gender; and "HCI" connected with "AC" is an emerging area of research. The selection of articles obtained depends on the terms used in WoSCC, and this study used only five emotions. Still the richness of the data gives some compensation. Marketers involved with brands need a body of knowledge on which to base their marketing communication intelligence gathering and strategic planning. The research provides exploratory research findings related to the use of automatic tools capable of mining emotions by gender in real time, which could enhance the feedback of customers toward their brands.

Chapter 6 by Jyrki Suomala provides an overview of "The Neuroscience Research Methods in Management." Enormous progress in understanding fundamental brain processes by using neuroscientific methods underlying management, marketing, and consumers' choice has been achieved. All thoughts and ideas of people are constituted by neural circuits. However, people have only limited conscious access to these neural circuits. As a result, an estimated 2 percent of thoughts are conscious, and the weakness of traditional research methods—like surveys and focus group interviews—is that they concentrate mainly for people's conscious part of mind. On the contrary, the main benefit by using neuroscientific methods is that there are many possibilities to uncover the unconscious brain processes, which are critical for human choice in management contexts. The chapter divides neuroscientific methods in biometrics and neuroimaging. The main biometric methods include eye tracking, face reading, skin conductance, and heart rate measurements. The main neuroimaging methods include EEG and fMRI. The description of each method is presented with examples in the management and marketing contexts. The analysis of each method includes benefits and drawbacks in these contexts. And finally, how much a method can predict the human behavior in the real context. This chapter introduces neuroscientific methods at an elementary level to the management science community. It gives basic concepts and ideas on the application of neuroscience in order to solve scientific and practical management problems through the use of specific neuroscientific methodologies.

Jyrki Suomala also introduces the topic on the "Benefits of Neuromarketing in the Product/Service Innovation Process and Creative Marketing Campaign" (Chap. 7). Most of the neuroscientific studies try to find a neural correlation between specific stimuli and brain circuits' activation patterns. However, the brain-as-prediction approach tries to find brain circuits' activation patterns during persuasive messages, which predict human behavior in future situations. The chapter describes the brain valuation circuits, whose activation pattern has been found critical for the prediction of human behavior in the future. In addition, the most common examples about the brain's valuation system will be presented from management's point of view. Whereas the academic community has produced a lot of research with neuroscientific tools, there is still a lot of room for applications of neuroscience to the innovation and marketing campaign processes. Thus, the chapter describes how different stakeholders can cooperate in order to find optimal products and optimal marketing campaign by using neuromarketing. Because the innovation process includes much uncertainty, it is very critical to encourage all participants to collaborate in the process. The chapter describes the benefits of this collaboration from management's point of view. In addition, the concrete examples of how researchers and business innovators together can apply neuromarketing in order to solve concrete innovation and management problems are presented. Finally, the future's opportunities of neuromarketing in innovation processes are also presented.

Chapter 8 of Volume 2 and of the whole book itself is by Robin Chark. It also deals with the nascent theme of neuromarketing. This chapter surveys methodological advancements in consumer research as a result of recent developments in neuroscience and behavioral genetics. This new approach is highly interdisciplinary and is termed "neuromarketing," paralleling similar developments in related fields such as neuropsychology and neuroeconomics. Researchers of this approach consider our behaviors to be the result of psychological processes embodied physiologically in the brain and nervous system. Thus, the biological influences may be rooted in our genes and shape the activities in our brain, and thus behaviors through actions on hormones and neurotransmitters. In this paradigm, our genotypes may be interpreted as a measure of individual differences, while brain activities may be observed and taken as more direct measures of the underlying psychological process. Such a neurobiological approach may take consumer research to another level and help answer some historically difficult questions. These questions are not confined to marketing scholars' interest in theory development; they encompass real-world marketing implications of concern to practitioners. To illustrate this, we review some recent findings in neuromarketing that make use of neuroimaging, twin study, and molecular genetics. We then discuss some recent trends in neighboring fields and their implications for the future of neuromarketing

In Chap. 9, Pacinelli presents the Futures Polygon. This futures research method is based on subjective probabilities and evaluations. It is a very interesting approach to the study of the future with facets that include subjective impacts. With varied degrees of probabilities, the Futures Polygon is designed for the expansion of scenarios. The methodology of the Futures Wheel is also incorporated. Methods of integration and important illustrations are introduced in the chapter.

We both hope that you like the challenging innovative research methodologies in management and that they will assist researchers in the enhancement of their scholarly research capabilities!

Our last word is a word of thanks to Liz Barlow, our editor, and Lucy Kidwell from Palgrave for all their help and understanding. Moreover, we wish to thank Liz for believing in this publishing project. We are both very grateful. We would also like to thank all the excellent contributors to the two volumes. Their reputation, mind-stretching methodologies, and collaborative stance were really very impressive. We are very indebted to all of them.

Enjoy the content of the two volumes, and we sincerely hope to have the readers' curation of their content and co-creation of future value in this domain that could help us prepare new writings……

May, 2017

Luiz Moutinho
Mladen Sokele

References

Nissen, V. (1995). An Overview of Evolutionary Algorithms in Management Applications. In J. Biethahn & V. Nissen (Eds.), *Evolutionary Algorithms in Management Applications* (pp. 44–97). Berlin: Springer.

Rothlauf, F. (2011). *Design of Modern Heuristics. Principles and Application*. Berlin: Springer.

Sörensen, K. (2015). Metaheuristics—The Metaphor Exposed. *International Transaction in Operational Research, 22*, 3–18.

Sörensen, K., Sevaux, M., & Glover, F. (2016). A History of Metaheuristics. In R. Marti, P. Pardalos, & M. Resende (Eds.), *Handbook of Heuristics*. Berlin: Springer. (to appear).

Weyland, D. (2015). A Critical Analysis of the Harmony Search Algorithm—How Not to Solve Sudoku. *Operations Research Perspectives, 2*, 97–105.

Wolpert, D. H., & Macready, W. G. (1997). No Free Lunch Theorems for Optimisation. *IEEE Transactions on Evolutionary Computation, 1*, 67–82.

Contents

Notes on Contributors

Malcolm J. Beynon is Professor of Uncertain Reasoning at Cardiff Business School, Cardiff, UK. His mathematical work centres around the understanding of uncertainty in data analytics, with particular emphasis on the employment of the methodologies of fuzzy set theory and Dempster-Shafer theory. He has written over 200 peer-reviewed articles across journal book chapters and for conferences. His work ranges from the development of novel mathematical approach to data analysis as well as the novel application of nascent techniques in marketing, business, and human resource management.

Simon Fong graduated from La Trobe University, Australia, with a BEng degree with honours in computer systems and a PhD degree in computer science in 1993 and 1998, respectively. Fong is now working as an associate professor at the Computer and Information Science Department of the University of Macau. Fong has written over 377 international conference and peer-reviewed journal papers, mostly in the areas of data mining, big data analytics, and optimization algorithms. He serves on the editorial boards of the *Journal of Network and Computer Applications*, IEEE's (Institute of Electrical and Electronics Engineers) *IT Professional* magazine, and various special issues of SCIE (Science Citation Index Expanded) journals.

Gordon Foxall is Distinguished Research Professor at Cardiff University. He received a PhD degree in industrial economics and business studies from the University of Birmingham, a PhD degree in psychology from the University of Strathclyde), and a higher doctorate (DSocSc) also from Birmingham. The author of over 300 refereed papers and chapters and over 30 books, he held visiting appointments at the Universities of Michigan and Oxford, and is a fellow of the Academy of Social Sciences, the British Psychological Society, and the British Academy of Management. His research interests include consumer behaviour analysis, philosophical implications of marketing neuroscience, and the theory of the marketing firm.

Paul M.W. Hackett is a psychologist with an interest in structured ontological accounts of behaviour in human and non-human animals. His research has addressed several applied domains including consumer behaviour, and his writing has appeared in prestigious journals. He has also published ten books including texts on research methods in consumer psychology. Hackett is a professor of ethnography in the School of Communications at Emerson College, Boston. He is also a visiting professor in the psychology department at Gloucestershire University and a visiting researcher in the psychology department at the University of Cambridge.

Kip Jones is Director of the Centre for Qualitative Research and Reader in Performative Social Science (PSS)–which uses tools from the Arts & Humanities in researching and/or disseminating Social Science research—in the Faculties of Media & Communication and Health & Social Sciences at Bournemouth University in the United Kingdom. Jones' accomplishments have been reported widely in the media, including BBC Radio 4, BBC TV news, Times Higher Education, LSE Impact Blog, *Sunday New York Times*, *International Herald-Tribune*, and *The Independent*.

Hongling Lao is a doctoral candidate in the Department of Educational Psychology at the University of Kansas. Lao has received a master's degree in educational psychology and research with honour from the University of Kansas in 2016. As a measurement scholar, Lao is a strong advocate for using measurement as a scientific tool to inform and improve our decision making broadly in academia, in industry, or in the contexts of daily life. Lao's research interest lies in the intersection of psychology, education, and measurement.

Luiz Moutinho is Full Professor of BioMarketing and Futures Research at DCU Business School, Dublin City University, Ireland. Moutinho's current research interests are mainly bio-marketing, neuroscience in marketing, evolutionary algorithms, human-computer interaction, the use of artificial neural networks in marketing, modelling consumer behaviour, Futurecast marketing, and tourism and marketing. He has written over 140 refereed academic journal articles and 29 books.

Volker Nissen holds the Chair of Information Systems Engineering in Services at Ilmenau University of Technology, Germany, since 2005. Prior to this, he pursued a consulting career, including positions as manager at IDS Scheer AG and director at DHC GmbH, Germany. In 1994, he received a PhD degree in economic sciences with distinction from the University of Göttingen, Germany. His current research interests include metaheuristic optimization, the digital transformation of the consulting industry, the management of IT agility, and process acceptance research. He is the author and editor of 19 books and some 200 other publications, including papers in IEEE's *Transactions on Evolutionary Computation*, IEEE's *Transactions on Neural Networks*, and *Annals of Operations Research*.

George G. Panigyrakis is Professor of Marketing at Athens University of Economics and Business since 2002. He completed his PhD degree at the University College of Wales. His research in the areas of marketing, integrated and marketing communications, brand management, international marketing, and services marketing has appeared in a number of books and journals. He has written 12 books and contributed over 80 papers to international conferences, edited volumes, and academic journals. He has served as a member of the organizing and scientific committees of several international scientific conferences as well as a member of the editorial boards of various scientific journals. He has extensive consultancy experience in companies and organizations of the private and public sector.

Komalsingh Rambaree is currently Associate Professor of Social Work at the University of Gävle in Sweden. He has a PhD degree in social work and social policy from the University of Manchester, UK. His research and teaching areas include computer-aided qualitative data analysis, international social work, eco-social work, and youth development.

Mladen Sokele is a lecturer at Zagreb University of Applied Sciences in the courses of Communication and Computer Techniques Chair and at Faculty of Electrical Engineering and Computing of the University of Zagreb as a guest lecturer in the course of Forecasting and Marketing of Telecommunication Services. His scientific fields of interest are techno-economic modelling and forecasting. Sokele is the author of more than 50 published scientific/professional papers and book chapters in the areas of telecommunications and informatics (modelling, analytical methods, simulation, forecasting, and expert systems). From 2002, he has regularly contributed as a speaker, moderator, or reviewer at telecoms and forecasting conferences and tutorials.

Meghan Sullivan is a doctoral candidate focusing in psychometrics, research methods, and statistics in the Department of Educational Psychology at the University of Kansas. She received her undergraduate degree specializing in psychology and mathematics from Bridgewater State University in 2013. Her research interests lie in the advancement of psychometric models and their applications in the educational and social sciences.

Jonathan Templin is a professor in the Department of Educational Psychology at the University of Kansas. He is co-coordinator of the Research, Evaluation, Measurement, and Statistics (REMS) Program. The main focus of Templin's research is in the field of diagnostic classification models—psychometric models that seek to provide multiple reliable scores from educational and psychological assessments. Templin's research program has been funded by the National Science Foundation and the Institute of Education Sciences. His research work has been published in journals such as *Psychometrika*, *Psychological Methods*, *Applied Psychological Measurement*, and the *Journal of Educational Measurement*.

Eugenia Tzoumaka is Lecturer in Sports Marketing at the American College of Greece. She holds a PhD degree in international marketing from the Athens University of Economics and Business, for which she has been awarded the international João Havelange Research Scholarship by the International Federation of Association Football (FIFA). She has written articles on celebrity brands in marketing journals and international conferences. Her current research interest focuses on celebrity brands, consumer-based brand equity, sports marketing, and social identification effects on consumer behaviour.

Cleopatra Veloutsou is Professor of Brand Management in the Adam Smith Business School of the University of Glasgow, Visiting Professor at the University of Bari, and Head of Marketing Research Unit of Athens Institute of Education and Research (ATINER). Her primary research interest is on brand management and has published about 45 articles in international academic journals. She is the co-editor of the *Journal of Product and Brand Management* since 2014 and is on the editorial boards of various journals. She has been the conference chair and a member of the organizing committee for a number of international academics conferences.

Hang Yang is working as a senior engineer and researcher at the Central Research Institute of China Southern Power Grid. He worked for a business intelligence company in Hong Kong and a telecommunication company in Macau. Yang has written over 40 international conference and peer-reviewed journal papers, mostly in the areas of data mining, big data analytics, artificial intelligence, and information security.

Anna K. Zarkada is Associate Professor of Marketing in the Department of Business Administration of the Athens University of Economics and Business, Greece. She has worked in various universities in the UK, Japan, Australia, and Russia. Her research in the areas of B2B and B2G marketing ethics, consumer behaviour on the Internet, international marketing, cross-cultural negotiations, and services marketing has appeared in a number of books and journals and has won various awards. Her current work focuses on virtual brand communities, the effect of Web 2.0 technologies on services branding and pricing, personal branding, and the philosophy of marketing.

List of Figures

List of Tables

1

Hosting a Successful Metamodern Party: Mixed Methods Management Research on the Web 2.0+

Anna K. Zarkada, George G. Panigyrakis, and Eugenia Tzoumaka

Introduction

Three decades ago McGrath 1981 stressed that "a single observation is not science" (p. 191) and amply demonstrated that all research methods are inherently incomplete and fraught with often fatal imperfections. He offered methodological pluralism—"bowling [dilemmas] over with multiple methods … embedded in multiple designs, using multiple strategies" (p. 209) "selected from different classes of methods with different vulnerabilities" (p. 207)—as the solution for transcending methodological vulnerabilities, maximising the theoretical and practical desiderata and capturing the nuances of rich data.

Mixed methods designs, in which qualitative and quantitative techniques complement and enhance each other, can help overcome the inherent limitations of quantitative and qualitative methods because they simultaneously provide data depth and breadth whilst safeguarding

A.K. Zarkada (✉) • G.G. Panigyrakis • E. Tzoumaka
Department of Business Administration, Athens University of Economics and Business, Athens, Greece

L. Moutinho, M. Sokele (eds.), *Innovative Research Methodologies in Management*,
https://doi.org/10.1007/978-3-319-64394-6_1

generalisability and transferability of results (Hesse-Biber and Leavy 2008; Johnson et al. 2007). Moreover, multiple methods and sources of data minimise the danger of common method variance (the "variance that is attributable to the measurement method rather than to the construct of interest") which is a concern in approximately 41% of attitude measures (Podsakoff et al. 2003, p. 879). Finally, they enhance triangulation as they allow for findings to be cross-checked (Bryman and Bell 2007).

Despite their having been found to work well for many disciplines (Hewson 2008), especially those that are naturally "multifaceted [and] crossing national, cultural, organizational, and personal boundaries" (Sedoglavich et al. 2015, p. 257), and their increasing popularity (Bryman and Bell 2007; Cui et al. 2015; Hesse-Biber and Leavy 2008; Hewson 2008), they account for only 4–9% of the total business literature (Harrison 2013). One of the main reasons for their limited use is cost (Kemper et al. 2003). In this chapter, we propose that a large part of the resource restrictions (such as limited time and funds or scarcity of equipment and competent data collectors) plaguing traditional mixed methods research designs can be reduced by using the internet for qualitative research sampling and quantitative data collection. We demonstrate that online research increases rather than sacrifices reliability, validity and generalisability on the altar of cost-efficiency.

Practicalities, however, are but a small and rather mundane part of the necessity for management scholars to explore novel methodological approaches. It is the emergent reality of the twenty-first-century world that forces us to re-examine both our tools, attitudes and identities. The understanding that complex interactions of multiple stakeholders over boundary spanning networks cause time-delayed effects that cannot be solved analytically by applying deterministic linear models is not new (Lutha and Virtanen 1996). From applications of chaos theory (Arnaboldi et al. 2015; Murphy 1996) to various management problems we also know that the qualitative properties of dynamic systems cannot be captured by the cross-sectional data collection techniques of modernity. Everyday universal experiences, such as consumption, for example, have been demonstrated to be "so diverse, variable,… esoteric, … and dependent on the specific nexus of the person, the object and the context as to be rendered totally immaterial and thus, incommensurable to modelling … by the tools [used in] the modern and post-modern milieu" (Panigyrakis and Zarkada 2014a, p. 18).

Mixed methods designs—by virtue of their inherent dynamism, decentralisation, multiplicity and multifacetedness—not only control for the context of management practice of the twenty-first century identified in Panigyrakis and Zarkada (2014a, b), namely the remnants of the hyperreality, fragmentation and juxtaposition of opposites that characterised postmodernity, but also transcend the fluidity of personal and communal identities and the brutal sociocultural restructuring that comes with the transition to metamodernity. Quantitative methods alone cannot detect Baudrillard's (1988) "fantastic cages" of consumption or what Lacan described as the powerful images that reside between language and the unconscious, feed desire for the sake of desire (Sharpe 2005) and form the bases of Sternberg's (1995) "iconic capitalism". At the same time, qualitative methods alone can only capture valuable but largely ungeneralisable subjectivities thus limiting the resulting theories' practical applicability in a globalised economy consisting of billions of interconnected consumers, entrepreneurs and employees and millions of interacting organisations, institutions and markets.

Metamodern socioeconomic phenomena, however, take place in the yet largely uncharted territories of the Web 2.0+ as much, if not more, than they do in the physical world. Cyberspace, Augmented and Virtual Reality experiences are as real as the chairs they sit on to their partakers. Space and time are reconfigured and the loci and nature of communications between people and organisations are shifting. The one-way controlled transmissions of information over broadcast media of modernity have evolved into incontrollable multi-party conversations over Online Social Media Networks. Content, experiences and emotions are Posted, Shared, Liked and Commented upon alongside organisational communications. It is thus obvious that the internet is fast becoming the single largest and most readily accessible repository of "digital life stories, an invaluable database of socio-demographics, opinions, needs, desires, values, grievances and hates" (Zarkada and Polydorou 2013,

p. 108). New collective identities emerge, old ones are renegotiated, reputations and brands that had been carefully crafted over decades are deconstructed, virtual teams replace hierarchies and remote work is becoming the norm, e-commerce volumes increase and even governments invite online bids to public auctions. It follows that management researchers need to be where their subjects are: in the largest ever village square.

Indeed, the 3.7 billion most affluent, educated and influential people in the world (i.e. 89% of the North American, 74% of the European and 73% of the Australian and Oceanian population) meet daily online (Internet World Stats 2016) and interact freely over physical, national, linguistic and psychic barriers. What is more important, they spend on average 6.6 hours per day living the World Wide Web (web) experience (Kemp 2016). Even the Japanese and South Koreans, the laggards in internet usage, are online for about 3 hours every day (Kemp 2016). Internet-based research has for over a decade now been quite popular (Hewson 2008; Wang and Doong 2010), mainly because of its time and cost benefits and despite concerns over the quality of the data it yields and its generalisability (Fricker 2008; Hewson 2008). Since these concerns were voiced, however, the frame bias concern has been practically eliminated by the rapid adjustment of internet users' demographics to include the over 65-year-olds, the poor and the uneducated (Deutskens et al. 2004; Wang and Doong 2010) as well as people living in remote areas of the developing world (Dahir 2016). The opposite is actually fast becoming the case: it is the use of traditional media that excludes whole generations who shun print and increasingly switch off broadcast media (Luck and Mathews 2010), that is, most of GenY, the Millennials and all those whose birth pictures were posted on Facebook and are now old enough to own tablets. It is obvious that internet-based research is a sounder methodological approach to reaching large, dispersed or interest-based populations than pen-and-paper or telephone surveys (Hewson 2008). After all, the world average of internet users has tripled during the past decade (it reached 43.9% at the end of 2015) whilst the fixed telephone line subscription rate has declined to the level of 1998 (14.34%) (International Telecommunication Union 2016a) and mail volume has been halved (United States Postal Service 2006). What is more important is that almost 60% of the world's population (Statista 2016) carry the web in their smart mobile devices with them wherever they go.

In this chapter, we propose that, to understand the "click and mortar" world that twenty-first-century people and organisations inhabit, and to be able to study, not only the unstructured and multifaceted emergent problems but also the traditional research themes which are being reconstituted by Web 2.0+ technologies and mentalities, management scholars need to be able to reach their subjects both in their physical and their avatar forms using new and exciting methods. We add to the voices (c.f. Hewson 2008) that call for internet-mediated mixed methods research as a solution to overcoming resource constraints. We also argue that these methods serve the purpose of addressing current and future social circumstance efficiently whilst safeguarding data quality by using freely available technologies such as web analytics and e-marketing techniques. We aim to assist the management research community in overcoming the well-documented (Harrison 2013; Harrison and Reilly 2011) limited familiarity with both mixed methods designs and IT functionalities.

As an example, we offer our study of consumer-based brand equity of celebrity footballers (Tzoumaka and Zarkada 2013, 2016; Zarkada and Tzoumaka 2014, 2015; Zarkada et al. 2014). The inability of traditional methodological approaches to serve consumer culture theory and the organisation-stakeholder meaning cocreation process has been well documented (Panigyrakis and Zarkada 2014a) so we needed to develop novel approaches for decomplexifying and organising an emergent research area. We applied an exploratory sequential mixed methods design. This type of research, despite its advantages, is actually quite rare (Abeza et al. 2015; Creswell and Plano Clark 2011; Harrison and Reilly 2011). Our study comprised (a) a short online survey to identify focus group participants, (b) five traditional focus groups meetings in two cities and (c) a complex, quasi-experimental, self-selected web-based questionnaire utilising a one-group post-test protocol (Gaines et al. 2007), the closest to pure experimentation (Fricker 2008). Whilst designing the study, we found very little practical advice regarding applications of state-of-the-art technologies that are popular in business but still rarely used for academic purposes. The following sections present the procedures we developed to cover the gaps in the literature on efficient and effective online:

(i) Sampling
(ii) Participant recruitment
(iii) Data collection

We provide checklists with criteria for selecting appropriate techniques and practical tips on how to apply them. Our suggestions are derived as much from our experiences as from our mistakes. Finally, we reflect on how technological innovations affect not only the tools but also their users.

Managing Twenty-first-century Management Research: The Party Planners' Checklist

The internet is the vastest meeting place the world has ever experienced—a distinct but also deeply enmeshed in the collective global everyday experience social milieu where friends, foes and strangers alike engage in multi-party user-controlled meaningful but also silly conversations, traces of which are for ever hosted on millions of networked computers all over the world. Management researchers cannot stay out of these conversations that evolve uncontrollably *ad infinitum* thus generating valuable data. Being present, however, is not enough to make sense of the evolving phenomena. Neither is being a wallflower in this "endless party where people invite themselves" (Zarkada and Polydorou 2013, p. 93).

Management researchers need to become hosts of their own carefully orchestrated data collection parties. Imaginative use of technology can make these parties highly visible and so exciting that people will want to massively attend and bring the host a nice present: their experiences, feelings and opinions.

Online Sampling: Drawing Up the Party Guest List

The foremost concern of online sampling is that the research population is actually present and adequately active on the web. Commercial and official sources need to be carefully examined in tandem to avoid errors of coverage. For example, in our study of footballer brand equity the research population was defined as sports fans with an interest in professional soccer. From industry reports we established the percentage of the country's population that have an interest in soccer. To ensure that the internet population characteristics did not differ from the general and the research population we used a combination of data sources such as national statistics, databases compiled by international agencies (International Telecommunication Union 2016b; Internet World Stats 2016) and commercial reports (Internet Live Stats 2014; MediaScope Europe 2012).

Sampling integrity is maintained by considering the nature of the study in relation to triangulated documentation of habits of the target population. We were seeking the opinions of people on celebrity footballers so we met our subjects in the milieu in which they meet their idols. Greeks spend more time on the web than on any other medium such as television, radio, newspapers or magazines (MediaScope Europe 2012) and sports fans are the heaviest of all internet users (European Interactive Advertising Association 2008). Also, sports-related search terms consistently top the national popularity lists (Google 2016) and sports sites are always amongst the most visited ones (Alexa 2016). Finally, sports fans are found all over the country so the internet is the fastest and most reliable way of reaching people in remote areas as well as ensuring that they are all approached at the same time and in exactly the same way—something that could not have been guaranteed had we, for example, used pen-and-paper surveys administered by the research team at football stadia and club refectories at immense cost.

Participant Recruitment: Inviting the Guests

Twenty-first-century people are playful, easily distracted online community members who multi-task and media mesh at exponentially increasing rates (Luck and Mathews 2010). Institution-generated messages are generally perceived as "pathetic… not funny… not interesting… not know[ing] who we are or car[ing]" (Hanna et al. 2011, p. 267). The good party hosts' main objectives are to shine through the media clutter, attract the attention of potential research participants and engage them long enough to collect their valuable data. Marketing practice and advertising theory provide valuable tools for promoting the research.

Promotional Media Selection: Choosing How to Send The Invitations

In theory, unrestricted self-selected web sampling gives researchers the opportunity to access individuals who are difficult or very costly to locate and reach (Fricker 2008). In practice, however, the degree to which the opportunity will be taken full advantage of depends on the dissemination medium. The key medium selection criteria and our insights on how to apply them are summarised in Table 1.1. They are (i) relevance (there is little to be gained by placing a call to participate in a focus group on sports in a cooking magazine), (ii) appeal (as indicated by circulation data), (iii) audience profile (to check for representativeness of the sample), (iv) partisanship (to control for sampling and response biases), data sharing (i.e. full access to medium usage patterns) and (vi) the cost of placing an advertisement or advertorial (or the effort required to set up a sponsorship agreement).

Whilst for print or broadcast media circulation and audience demographic data can be hard to find or unreliable (as they are provided by the department selling advertising space on the medium), for the online ones, statistics are readily accessible through free (or low-cost) website analytics tools such as Alexa.com, siteworthtraffic.com and Similarweb.com. Multiple sources are, again, required to triangulate the medium evaluation data. SiteWorthTraffic, for example, shows unique visitors and page views per day as well as trends whilst Alexa provides country-wide and global rankings as well as detailed visitor demographics. Further to reliable hard data, qualitative examination of each medium is also required. Open partisanship is a main concern as it automatically excludes

Table 1.1 Criteria for promotional media selection

Criteria	Tips
Relevance	Make sure the topics covered by the medium are closely aligned to the research theme
Appeal	Use multiple sources to cross-check metrics such as • Ranking of the medium in relation to its competitors in the region of interest • Average session duration (time spent on site at each visit) • Number of pages clicked • Bounce rate (percentage of visitors who enter the site but leave it without interacting with the site instead of continuing to other pages within the same medium) • Overall web traffic statistics (an aggregate metric comprising the number of visitors and the number of pages they visit) • Content curation (the content they pick from other sources and upload to the medium) • Content sharing (the content other media pick and reproduce from the medium that is being evaluated) • Quantity, quality and relevance of the user-generated content (such as comments and discussions) posted under relevant topics
Audience profile	• Make sure the demographics of the medium are representative of the population • Check medium access patterns in terms of time (e.g. early morning or late evening), place (e.g. home, school, train or office) and technology (e.g. pc, smartphone or tablet) and use them as survey design parameters
Partisanship	• Make sure that there is no conflict of interest between the medium and the topic of research • Check whether the medium is being perceived as biased in any way related to the study
Cost	Negotiate a media sponsorship or content sharing agreement and/or a price that includes multiple promotional opportunities
Data sharing	Request full access to web analytics for all content and promotional activities and failing that ensure that you get regular reports

users that belong to rival groups or hold different opinions. Public perceptions of the medium's lack of independence, whether based on fact or not, are an even bigger threat as they can introduce uncontrollable response bias. Finally, the data needed for medium evaluation are also needed for the *a posteriori* sample quality assessment. So, full access to site traffic analytics needs to be granted to the research team by the site owners throughout the project.

Research Promotion Tools: Designing the Invitations

Recruitment for internet-based self-selected surveys has usually been carried out using banner advertisements on web pages (Fricker 2008). Advances in IT and digital marketing practices, however, have since provided additional alternatives, namely Content and Social Media Marketing, which have yet to be evaluated in the context of academic research. We here present a comparison of the costs and results of all three techniques which we applied sequentially on the most popular internet-based sports medium in the country. First, we run banner ads and then we applied Content and Social Media Marketing techniques in tandem. To evaluate the effect of each promotional technique we used data provided by Google Analytics.

Banner Ads

The exposure of internet users to banners is usually measured by counting impressions, that is, how many times the banner was displayed on users' screens provided by the website that displays the banner. There are several problems with the direct placement of banner ads. Impressions count even when the screen is automatically refreshed periodically by the site thus counting the additional impressions on the same user's screen as new impressions. It also cannot differentiate between above and below the fold placement (i.e. how much of the screen the user sees without scrolling as the algorithm cannot factor for screen size and resolution). Moreover, impressions, as a measure, cannot account for the visitors' using ad blocking software. "Active desktop ad blocker usage has quadrupled globally since 2013, with around 220 million users employing ad-blocking technology today. Consequently, 32% of all page views worldwide are now impacted by ad blocking" (Hancock 2016, p. 1). For example, in the USA 45 million active users do not see website advertisements whilst in the UK the number of ad block users grew by 82% in 12 months (PageFair Team 2015, p. 1). At the time of data collection, Greece was "leading the way with an average of 24.5% of [the] online populations using adblocking software" (PageFair and Adobe 2014, p. 7). Finally, there is no way to account for "banner blindness" (Stec 2015), the fact that over 70% of internet users ignore banner advertising (eMarketer 2014) and certain age groups, such as the 18- to 34-year-olds, pay them even less attention than they do to TV, radio and print advertisements (Stec 2015).

The clickthrough rate (CTR), the ratio of clicks on the banner to the number of total impressions, is another measure of the conversion rate of a banner ad. On a global scale, the average CTR across all formats and placements is 0.06% (Stec 2015) and researchers can benchmark their placements against the performance of similar advertisements by industry, country, formats, placement and size using free internet tools such as the Benchmark tool on richmediagallery.com. For our project, we had a three-frame Flash animated medium rectangle (300*250) skyscraper banner placed to the left sidebar on both the home and the dedicated (football) page. The cost of the ad placement at the time the research was carried out was approximately €10,000 per week. During the one-week period that the banner ad was left in place, it yielded 1,648,000 impressions, 329 attempts to respond to the survey (a minuscule 0.02% interest rate, below the country average of 6% for same type and size advertising but comparable to the average 3‰ CTR achieved by web banner promotions of academic research) but only 41 fully completed questionnaires (i.e. a rather small 12% response rate). Thus, the ROI of our banner advertising was unacceptably low as the cost per participant, had we paid for the ad, would have been €243.90.

Content Marketing

Content marketing is based on a *quid pro quo* logic: instead of yelling to attract attention, like you do when advertising, you give something valuable (informative or entertaining content) to get something valuable (attention, clicks, conversions or, in our case, data) in return. Instead of being the irritating commercial you become the exciting show (O'Brien 2012). The website that first hosted the advertisement, later, featured an interview with a research team member written by a sports journalist. The article discussed a topic that is important to football fans but also included information about the research, stressing its academic nature, the university affiliation (the premier Business School in the country) and the FIFA funding (the top football institution in the world) to increase its perception of credibility, seriousness and relevance. It also mentioned that all focus groups participants would enter a draw for a season ticket for their favourite team. We used the same egoistic and altruistic appeals of the banner advertisement as calls to action in the sidelines and also inserted multiple hyperlinks to our online questionnaire in the text.

In approximately 48 hours the article web page yielded 8,425 unique page views, 1,351 clicks to the questionnaire (a satisfactory 16% interest rate) and 1,274 completed questionnaires (an amazing 94% response rate). What is most impressive here is the commitment to the research and the level of trust the respondents to the screening survey demonstrated. They completed the questionnaire after having read the instructions and accepted the terms of the survey which were (a) to participate in the focus group meetings and (b) to provide full personal data (name, surname, email address and mobile telephone number).

We also used content marketing to promote the survey for the quantitative phase of the research. There were no incentives for filling in the quite long and complex questionnaire, but again, the response rate was a very satisfactory 33.4% (much higher than the 10–25% typically reported in Deutskens et al. 2004; Manzo and Burke 2012; Sánchez-Fernández et al. 2012; Sauermann and Roach 2013). Overall, 87.37% of the traffic on the survey website came from clicks on the links incorporated in the article. Interestingly, 14.45% of those clicks came from the mobile version of the site hosting the article. Finally, about 10% of the traffic came from sites that reproduced the content. Traffic from the article and its reproductions had a very low bounce rate (17.83%) thus further strengthening the argument for using content marketing to promote research. Having taken the egoistic motive away, we believe that this result strengthens the relevance argument: when people care about the topic they happily give their time and opinions. Upon comparing the results with those of the banner ad we also believe that the article played the role of reducing participants' perceived risks.

Social Media Marketing

The online articles were also pushed through the media group's relevant social media platforms. The leverage for both surveys was poor as only about 2.1‰ of the clicks to the focus group screening survey and 2.11% of the clicks on the quantitative study page came from media sponsor's Facebook posts. Moreover, the bounce rate of the Social Media-generated traffic was a quite high 48.33%. In our study, Social Media Marketing was supportive of the content marketing effort and controlled by the sponsor's marketing personnel so our data is insufficient to fully evaluate its appropriateness for academic purposes.

In Table 1.2 we summarise our experiences and provide guidelines for putting IT and digital marketing practices to work for academic research based on what we learnt through creative trial and error.

Managing Incentives: Getting the Invitation Accepted

Offering monetary and quasi-monetary incentives for participation has long been common practice in qualitative research for which extra effort and commitment is required of the participants (Deutskens et al. 2004; Fricker 2008; Morgan 1997; Wang and Doong 2010). Moreover, during the past decade, internet-based data collection has increased, so response rates have decreased, thus increasing the need to offer incentives for participation (Teitcher et al. 2015). There is evidence that incentives increase online survey participation by about 27% but they also have the potential to encourage multiple submissions (Manzo and Burke 2012; Teitcher et al. 2015). There is conflicting evidence on the effect of incentives on response rates (Sánchez-Fernández et al. 2012; Sauermann and Roach 2013). We attribute the success of our screening survey to (a) the

Table 1.2 Promotional techniques

Alternatives	Tips
Banner advertising	• Check o CTR for similar type of ads in the region and in sites related to the industry of interest through multiple sources o Ad blocking software penetration in the region of interest through multiple sources • Carefully negotiate the placement • Consider using Google Display Network (instead of negotiating with sites for ad placement you specify the audience segmentation parameters and Google does the placement of the ad) • Have the ad professionally designed and produced • Do not use the outdated flash technology as this does not display properly on all screens. Use static images, GIFs and, if budget permits, videos
Content marketing	• Control for sampling and other biases introduced by the medium and/or the text • Include many different calls to action and hyperlinks in and around the main text • Carefully negotiate concurrent promotions, multiple articles and access to web analytics
Social media promotion through the medium's owned media	• Carefully consider o the reach, o style and o appeal of the medium's SM portfolio and those of its elements (Facebook, twitter, etc.) • Request access to detailed platform analytics (e.g. Facebook demographics and usage patterns during the week and day) • Study the comments and shares of the users to calculate the risks of message distortion • Check the content of the posts to ensure that response biases are not introduced by the wording of the text after it is condensed to comply with message length restrictions • Use the services of professional designers to produce visual content appropriate for SM • Negotiate bundle price for numerous carefully timed posts
Mixed mode	Check for systematic response variance across subsamples defined by the entry point to the survey which you track by creating custom links on each medium through Campaign URL Builder

relevance and value of the prize and (b) alleviating respondents' perceived risks of participation. The prize, a "lottery incentive with a high payoff and a low chance of winning" (Sauermann and Roach 2013, p. 273) was something they really wanted: a season ticket to their favourite football team. The prospective participants were presented with prudently crafted legal documents explaining both the prize draw process and the data protection safeguards.

To reduce participants' perceived risk and make sure they trusted that the prize would be awarded through a transparent and unimpeachable procedure, we employed the services of a notary public to write the Terms and Conditions document that preceded the online survey and to design and oversee the lottery process. Not only were the terms of the competition clearly explained, but also, details of the time and place of the draw as well as the ways by which the winners would claim their prizes were provided before they completed the survey.

Another problem with online surveys is that there really is no guarantee of respondent anonymity as the IP addresses of the visitors to the survey website can be recorded. In the case of our focus group recruiting and screening survey the problem was compounded by the need to collect the personal and contact details needed to arrange the focus group meetings. So, for both surveys—even for the quantitative one where no names and contact details were required—we employed the services of an academic specialising in online privacy issues who worked together with a lawyer to prepare a Privacy Policy Disclaimer. Both legal documents were presented as hyperlinks in the first and the last pages of the electronic surveys and respondents had to click a button to accept the terms and enter the survey and another one to submit their responses. Further to the conditions standard university ethics stipulate, we reassured potential respondents that (a) all safety measures were taken for the domains to be free of viruses and other threats to their computers, (b) no further communication would ever be attempted and (c) the contact details and IP files would be destroyed upon completion of the research.

Moreover, several filters were built into the focus group screening survey to ensure that the respondents' time was not wasted and that no personal data that was not absolutely necessary was collected. For example, residents of cities other than the ones where we intended to run focus groups were thanked for their attempt to complete the survey and the session was terminated at the third question, after about 20 seconds. The efficacy of the filters is evident in that out of the 1680 people that attempted to respond to the survey, 976 were eliminated, thus also reducing data handling and screening time and effort. Finally, we offer our insights in confidence of their efficacy as only 3% of the email addresses we collected were not valid.

Data Collection: Party Time!

The proof of the good host is in the superb guest party experience. After being allowed into subjects' computer-mediated private spaces and managing to generate high-quality data, the research also needs to be perceived as interesting, fun and amazing enough to motivate its subjects to Like and Share the research with their Friends and Followers, thus creating snowball effects.

The criteria for selecting the technology for designing, hosting and administering an online survey are (a) user interface and experience (UI/UE), (b) researcher interface (c) credibility, (d) hosting and (e) instrument self-promotion.

For the survey that recruited participants for the focus groups we used Google Docs which is free and very easy to use but has limited room for aesthetic adjustments. It comes with free hosting but offers no web analytics data. As respondents were to supply their details, the analytics were superfluous and aesthetics and advanced programming functionalities were considered of limited value for a short and simple screening survey.

For the lengthy and sophisticated quantitative survey of the model building and testing phase of the research, we employed a web developer and programmer to customise LimeSurvey (https://www.limesurvey.org/), an open source survey application. The modifications we made were (1) to create and attach an algorithm to randomly assign the questionnaire versions required for the quasi-experimental design (photos and bios of local, global, active and retired celebrity footballers),

(2) to install IP address authentication for filtering out returning users, (3) to track the page from which the user had been redirected, (4) to modify standard Likert scales to include a "no opinion" option needed for scale cleaning at the measure construction phase of the analysis, (5) to add the logos of the university and funding agency to increase credibility and alleviate perceived risks, (6) to customise and aesthetically improve the default templates and (7) to add "buttons" with the logos of various popular platforms through which the participants could invite members of their online social networks to participate in the study.

The survey was uploaded to a university server to increase the credibility of the research by clearly signalling the purely academic nature of the research and to allow us to collect via Google Analytics the page visit and visitor profile data needed to test for sample quality and representativeness.

We performed numerous ex-post quality controls to address potential online survey pitfalls (Schmidt 1997) such as contamination and skewing of results by accidental, fraudulent or malicious multiple submissions by the same individual—an increasingly common and serious problem in online research (Teitcher et al. 2015). We performed manual and visual checks for outliers and irregular patterns in questionnaire completion time, variables and cases with too many repetitive, outlying or missing values. We also checked for duplicate or irregular IP addresses (such as too many Chinese IPs on a survey written in Greek) using the tools freely available on NirSoft.net. Moreover, we used Google Analytics data to control for self-selection bias effects by comparing our data set demographics with those of the website and the pages through which the questionnaire was promoted.

Both survey samples were representative both of the internet and the football fan population in Greece and in line with similar European (Bauer et al. 2005) and Greek studies (Athanasopoulou et al. 2011) so we offer our insights in confidence that the strategies and tactics we employed were efficient, effective and efficacious. In Table 1.3, we summarise the techniques we used and found them to produce the desired results. We also list some suggestions derived both from our mistakes and from our experiences with subsequent virtual data collection parties we hosted.

Conclusions and Reflections: The Hosts' After Party

With the full benefit of hindsight, and after a lot of soul searching, we feel that we hosted an overall successful party in which academia and practice got better acquainted. First, we evaluated not only over three decades of mixed methods and almost two decades of online research reports published in academic journals and handbooks but also the experiences of professional e-marketers showcased in commercial websites and blogs. Then, we identified, evaluated and used multiple sources of information not traditionally accessed for academic research. Based on the secondary data, we set up and run a media collaboration for participant recruitment. We also identified, reviewed and tested different data collection instrument building and hosting platforms. In essence, we recruited, selected and managed a dynamic virtual team of graphic artists, media and IT practitioners. Finally, we critically examined the results of our e-adoptions and innovations.

The process of designing and executing the project was not always smooth. Neither were our understandings automatically self-evident to our media partners and tech-services suppliers. Explaining what we wanted and understanding what was technically possible often proved to be a struggle but, we are happy to report, we managed to work through our experiences to turn them into shareable insights. We here present a practical online mixed methods research guide and a set of tried and tested methodological tools for the twenty-first century. With this chapter, we firmly reconfirm the applicability and argue for the necessity—if not the inescapability—of on- and offline mixed methods management research. We contribute to knowledge by enriching academic practice with insights gained by businesses and by providing managers with academically sound testing of their practices. Thus, we offer a guide for bringing academically solid management research practices on and in line with the realities of the web2.0+ lived experience.

Our team comprises a Gen Xer, a Baby Boomer and a Millennial, so we are an adequately representative sample of the business academic

Table 1.3 Technical aspects of designing and hosting online data collection instruments

Criteria	Tips
User interface and user experience (UI/UX)	• Invest in the services of experienced professional graphic artists and web developers • Have responsiveness checked thoroughly on all devices and Operating Systems used by the population • Use multiple filters in the survey design to *a priori* control sample characteristics instead of wasting respondents' time to collect data you will later discard
Researcher interface	• Pay close attention to the programming required for the delivery of a useable data file (e.g. make sure that responses to Likert scales are delivered ready-coded into numbers and not as the words that appear on the survey) • Specify questions as mandatory to collect only the responses of committed and interested participants, filter out internet lurkers and thus save on data cleaning time
Credibility	• Use the university and/or funding organisation logos on the cover page, at the bottom of survey pages and provide hyperlinks to the relevant pages of their websites • Provide valid contact details and hyperlinks to the profiles of the investigator(s) on the university website
Hosting	• Host the data collection instrument on a secure server to which you have access for maintenance and analytics. If you choose to use a survey creation and data collection application, do not host the survey on their server • Use web analytics tools for response rate calculations • Consider using IP Authentication to filter out malicious response attempts or fraud (especially when offering incentives for participation) but also check the data set visually and manually
Instrument self-promotion	• Add Social Media buttons at the entry and thank you pages for easy snowballing • Buy a carefully chosen domain name and invest in SEO (applying on- and off-page refinements so that the site will be indexed and ranked successfully by the search engines) to increase survey visibility and domain authority

community. We feel that the three aspects of twenty-first-century mixed methods research that are the most alien to contemporary researchers are (i) securing sampling integrity online, (ii) selecting appropriate media and cost-effective techniques for the promotion of the research and (iii) the technicalities of online data collection. In Tables 1.1, 1.2 and 1.3 we provide our hard-gained insights and suggestions based on what we did and worked, what we tried and found out does not work and what we now know we should have done.

The media sponsorship that made the application of the promotional strategies presented here possible was the result of mobilising pre-existing personal networks. This is not always possible, however. What researchers need to do is understand the roles, benefits and challenges of owned, paid and earned media so that they make sure they strike the right balance of effort, time and funds expended to achieve adequate promotion of the research and to ensure sampling adequacy and integrity. Owned media (web and mobile sites, blogs, etc. dedicated to the project) are controllable, versatile and cost efficient and, over time, they provide visibility and build relationships with potential respondents and journalists so they generate both data and earned media. Earned media (the publicity that is generated by people that have shown an interest in the research and they choose to promote it through their own media) might be hard to measure, impossible to control and slow to grow but they have the benefits of being transparent, long-lived and as credible as their source—at least to the source's audience. Paid media (the researcher-paid leverage of the power of other channels through advertising, paid searches and content marketing) can feed the owned and support the earned media but it is becoming increasingly more difficult for them to cut through media clutter, adblocking software and audience boredom. So, if we were to do it all over again, we would start by building a blog, website and relevant Social Media pages dedicated to the research for recruiting participants and media collaborators. We would upload carefully crafted articles, in plain language and lay terms, to highlight the broader context of the research without giving away hypotheses or findings that would introduce bias we could not later control for. We would also run a carefully planned email marketing campaign.

Technology changes fast, so what is now is not tomorrow. Hence, management researchers need not only to familiarise themselves with but to constantly stay in touch with developments in both IT and marketing regardless of their field of work. IT and marketing developments change the tools of the academic trade. The deeper issue that emerges from the discussion above, however, is how the tools change their users. Long gone are the days of the Ivory Tower, from which the university researcher descended gracefully to meet subjects that were eager to share their opinions over a cup of coffee, at the street corner or over the telephone. Twenty-first-century researchers are—whether we like it or not—entrepreneurs, fund-raisers and project managers as well as the mass marketers of their work and themselves.

In the UK, the Arts and Humanities Research Council funds academics that are "listenable", that is, those that have the mental flexibility to engagingly parry journalists, the ability to "dumb down" complex ideas, the right looks and a pleasant voice (Tickle 2012). Even though it has been argued that, no matter how famous, scholars cannot be classed as celebrities (Leslie 2011), the fact remains that twenty-first-century ones find it hard to resist the lure (Kurzman et al. 2007) of publicly displayed authority for entertainment purposes that makes them spend more time in studios than in studies thus commanding speaking engagement that look like a fortune to their off-the-limelight colleagues. Thus, academic careers become similar to those of fashion models—all about building and exploiting "field-specific social and cultural capital" (Parmentier et al. 2012). It seems that the new realities make managing the necessary "modifications in [the academics'] role identity" (Jain et al. 2009, p. 922) a prerequisite for attracting research funds, media sponsorships for their projects and even students to their universities (Joseph et al. 2012)—all in the course of serving science.

Acknowledgements The research project used as an example here was funded by FIFA under João Havelange Research Scholarship by CIES Research and sponsored by the digital publishing group 24MEDIA. The authors gratefully acknowledge the technical advice provided by Ms Katerina Fotiadi and the helpful comments of Dr Ilias Kapareliotis.

References

Abeza, G., O'Reilly, N., Dottori, M., Séguin, B., & Nzindukiyimana, O. (2015). Mixed Methods Research in Sport Marketing. *International Journal of Multiple Research Approaches, 9*(1), 40–56. https://doi.org/10.1080/18340806.2015.1076758.

Alexa. (2016, December 23). Top Sites in Greece November 2016. Retrieved from http://www.alexa.com/topsites/countries/GR

Arnaboldi, M., Lapsley, I., & Steccolini, I. (2015). Performance Management in the Public Sector: The Ultimate Challenge. *Financial Accountability & Management, 31*(1), 1–22.

Athanasopoulou, P., Zafeiropoulou, G., Siomkos, G. J., Assiouras, I., & Douvis, J. (2011). *Consumer Behaviour in the Arena: A Classification of Football Fans*. Paper presented at the 4th Annual EuroMed Conference of the EuroMed Academy of Business.

Baudrillard, J. (1988). *The Consumer Society: Myths and Structures (Société de Consommation: Ses Mythes, Ses Structures)* (C. Turner, Trans.). London: Sage.

Bauer, H. H., Sauer, N. E., & Schmitt, P. (2005). Customer-Based Brand Equity in the Team Sport Industry: Operationalization and Impact on the Economic Success of Sport Teams. *European Journal of Marketing, 39*(5–6), 496–513.

Bryman, A., & Bell, E. (2007). *Business Research Methods* (2nd ed.). Oxford: Oxford University Press.

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and Conducting Mixed Methods Research* (2nd ed.). Thousand Oaks: SAGE Publications.

Cui, X., Zhou, Q., & Liu, L. (2015). Using Online Field Surveys in e-business Research: Reflections on a Referent Study. *International Journal of Electronic Business, 12*(4), 345–363. https://doi.org/10.1504/ijeb.2015.074610.

Dahir, A. L. (2016). Smartphone Use Has Doubled in Africa in Two Years. *Quartz Africa*, (SMART FUTURE). Retrieved from https://qz.com/748354/smartphone-use-has-more-than-doubled-in-africa-in-two-years/

Deutskens, E., de Ruyter, K., Wetzels, M., & Oosterveld, P. (2004). Response Rate and Response Quality of Internet-Based Surveys: An Experimental Study. *Marketing Letters, 15*(1), 21–36. https://doi.org/10.1023/b:mark.0000021968.86465.00.

eMarketer. (2014). Traditional or Digital Ads? Millennials Show Mixed Feelings. *eMarketer.com*. Retrieved from eMarketer.com website: https://www.emarketer.com/Article/Traditional-Digital-Ads-Millennials-Show-Mixed-Feelings/1010747#sthash.m3HSKh07.dpuf

European Interactive Advertising Association. (2008). *Sport and the Shift to Interactive Media 2008, Pan-European Results*. Retrieved from http://www.sponsors.de/uploads/tx_svsstudiengaenge/EIAA_-_Sport_and_the_Shift_to_Interactive_Media_2008.pdf

Fricker, R. D. J. (2008). Sampling Methods for Web and E-mail Surveys. In N. G. Fielding, R. M. Lee, & G. Blank (Eds.), *The SAGE Handbook of Online Research Methods* (pp. 195–216). London: SAGE Publications.

Gaines, B. J., Kuklinski, J. H., & Quirk, P. J. (2007). The Logic of the Survey Experiment Reexamined. *Political Analysis, 15*(1), 1–20. https://doi.org/10.1093/pan/mpl008.

Google. (2016). Google Trends. Retrieved December 29, 2016, from Google https://www.google.com/trends/

Hancock, L. (2016). B2B Programmatic vs Ad Blockers: Who's Winning. *Perspectives, Marketing and Sales*. Retrieved from dun&bradstreet website: http://www.dnb.com/perspectives/marketing-sales/b2b-programmatic-vs-ad-blockers-who-is-winning.html

Hanna, R., Rohm, A., & Crittenden, V. L. (2011). We're All Connected: The Power of the Social Media Ecosystem. *Business Horizons, 54*(3), 265–273. https://doi.org/10.1016/j.bushor.2011.01.007.

Harrison, R. L., III. (2013). Using Mixed Methods Designs in the Journal of Business Research, 1990–2010. *Journal of Business Research, 66*(11), 2153–2162. https://doi.org/10.1016/j.jbusres.2012.01.006.

Harrison, R. L., III, & Reilly, T. M. (2011). Mixed Methods Designs in Marketing Research. *Qualitative Market Research: An International Journal, 14*(1)), 7–26. https://doi.org/10.1108/13522751111099300.

Hesse-Biber, S. N., & Leavy, P. (2008). *Handbook of Emergent Methods*. New York: Guilford Press.

Hewson, C. (2008). Internet-Mediated Research as an Emergent Method and It's Potential Role in Facilitating Mixed Methods Research. In S. N. Hesse-Biber & P. Leavy (Eds.), *The Handbook of Emergent Methods* (pp. 543–570). New York: Guilford Press.

International Telecommunication Union. (2016a). Fixed Telephone Subscriptions (per 100 People). Retrieved December 21, 2016, from The World Bank Group http://data.worldbank.org/indicator/IT.MLT.MAIN.P2

International Telecommunication Union. (2016b). Individuals Using the Internet. Retrieved December 21, 2016, from The World Bank Group http://data.worldbank.org/indicator/IT.MLT.MAIN.P2

Internet Live Stats. (2014). Internet Users by Country. Retrieved from http://www.internetlivestats.com/internet-users-by-country/

Internet World Stats. (2016). World Internet Usage and Population Statistics, June 30, 2016—Update. Retrieved December 6, 2016, from Miniwatts Marketing Group http://www.internetworldstats.com/stats.htm

Jain, S., George, G., & Maltarich, M. (2009). Academics or Entrepreneurs? Investigating Role Identity Modification of University Scientists Involved in Commercialization Activity. *Research Policy, 38*(6), 922–935.

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a Definition of Mixed Methods Research. *Journal of Mixed Methods Research, 1*(2), 112–133. https://doi.org/10.1177/1558689806298224.

Joseph, M., Mullen, E. W., & Spake, D. (2012). University Branding: Understanding Students' Choice of an Educational Institution. *Journal of Brand Management, 20*(1), 1–12.

Kemp, S. (2016). *Digital in 2016*. Retrieved from http://wearesocial.com/uk/special-reports/digital-in-2016

Kemper, E. A., Stringfield, S., & Teddlie, C. (2003). Mixed Methods Sampling Strategies in Social Science Research. In A. Tashakkori & C. Teddie (Eds.), *Handbook of Mixed Methods in Social and Behavioral Research* (pp. 273–296). Thousand Oaks: Sage.

Kurzman, C., Anderson, C., Key, C., Lee, Y. O., Moloney, M., Silver, A., & Van Ryn, M. W. (2007). Celebrity Status*. *Sociological Theory, 25*(4), 347–367. https://doi.org/10.1111/j.1467-9558.2007.00313.x.

Leslie, L. Z. (2011). *Celebrity in the 21st Century: A Reference Handbook*. Santa Barbara: ABC-CLIO.

Luck, E. M., & Mathews, S. W. (2010). What Advertisers Need to Know About the iYGeneration: An Australian Perspective. *Journal of Promotion Management, 16*(1–2), 134–147. https://doi.org/10.1080/10496490903574559.

Lutha, I., & Virtanen, I. (1996). Analyzing the Behaviour of a Non-linear Advertising Campaign Model; an Application of Bifurcation Theory, Lyapunov Exponents and Correlation Dimension. In P. Walden, M. Brannback, B. Back, & H. Vanharanta (Eds.), *The Art and Science of Decision-Making* (pp. 138–151). Åbo: Åbo Akademi University Press.

Manzo, A. N., & Burke, J. M. (2012). Increasing Response Rate in Web-Based/Internet Survey. In L. Gideon (Ed.), *Handbook of Survey Methodology for the Social Sciences* (pp. 327–343). New York: Springer.

McGrath, J. E. (1981). Dilemmatics: The Study of Research Choices and Dilemmas. *American Behavioral Scientist, 25*(2), 179–210.

MediaScope Europe. (2012). *Greece Launch Presentation Summary*. Retrieved from http://www.iab.gr/files/1/research/Mediascope/mediascope_2012_greece_summary%20launch%20presentation.pdf

Morgan, D. L. (1997). *Focus Groups as Qualitative Research* (2nd ed.). Thousand Oaks: Sage Publications, Inc.

Murphy, P. (1996). Chaos Theory as a Model for Managing Issues and Crises. *Public Relations Review, 22*(2), 95–113. https://doi.org/10.1016/S0363-8111(96)90001-6.

O'Brien, J. (2012). How Red Bull Takes Content Marketing to the Extreme. *Mashable*. Retrieved from http://mashable.com/2016/12/22/robotic-mobilization-device/#YRcwESBSraqi

PageFair, & Adobe. (2014). *Ad Blocking Goes Mainstream*. Retrieved from https://downloads.pagefair.com/wp-content/uploads/2016/05/Adblocking-Goes-Mainstream.pdf

PageFair Team. (2015). The 2015 Ad Blocking Report. *PageFair*. Retrieved from PageFair website: https://pagefair.com/blog/2015/ad-blocking-report/

Panigyrakis, G., & Zarkada, A. (2014a). New Philosophical Paradigms in Marketing: From Amoral Consumerism to Axiological Societing. In L. Moutinho, E. Bigne, & A. K. Manrai (Eds.), *The Routledge Companion on the Future of Marketing* (pp. 25–50). Oxford: Routledge.

Panigyrakis, G., & Zarkada, A. (2014b). A Philosophical Investigation of the Transition from Integrated Marketing Communications to Metamodern Meaning Cocreation. *Journal of Global Scholars of Marketing Science, 24*(3), 262–278. https://doi.org/10.1080/21639159.2014.911494.

Parmentier, M.-A., Fischer, E., & Reuber, A. (2012). Positioning Person Brands in Established Organizational Fields. *Journal of the Academy of Marketing Science*, 373–387. doi:10.1007/s11747-012-0309-2

Podsakoff, P. M., MacKenzie, S. B., Jeong-Yeon, L., & Podsakoff, N. P. (2003). Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies. *Journal of Applied Psychology, 88*(5), 879.

Sánchez-Fernández, J., Muñoz-Leiva, F., & Montoro-Ríos, F. J. (2012). Improving Retention Rate and Response Quality in Web-Based Surveys. *Computers in Human Behavior, 28*(2), 507–514. https://doi.org/10.1016/j.chb.2011.10.023.

Sauermann, H., & Roach, M. (2013). Increasing Web Survey Response Rates in Innovation Research: An Experimental Study of Static and Dynamic Contact Design Features. *Research Policy, 42*(1), 273–286. https://doi.org/10.1016/j.respol.2012.05.003.

Schmidt, W. C. (1997). World-Wide Web Survey Research: Benefits, Potential Problems, and Solutions. *Behavior Research Methods, Instruments, & Computers, 29*(2), 274–279. https://doi.org/10.3758/bf03204826.

Sedoglavich, V., Akoorie, M. E. M., & Pavlovich, K. (2015). Measuring Absorptive Capacity in High-Tech Companies: Mixing Qualitative and Quantitative Methods. *Journal of Mixed Methods Research, 9*(3), 252–272. https://doi.org/10.1177/1558689814523677.

Sharpe, M. (2005). Jacques Lacan (1901–1981). In J. Fieser & B. Dowden (Eds.), *The Internet Encyclopedia of Philosophy (IEP): A Peer Reviewed Academic Resource*. Retrieved from http://www.iep.utm.edu/lacweb/

Statista. (2016). Mobile Phone Internet User Penetration Worldwide from 2014 to 2019. Retrieved December 28, 2016, from Statista https://www.statista.com/statistics/284202/mobile-phone-internet-user-penetration-worldwide/

Stec, C. (2015). 20 Display Advertising Stats That Demonstrate Digital Advertising's Evolution. Retrieved from https://blog.hubspot.com/marketing/horrifying-display-advertising-stats#sm.000whbz5ucmadwf10k71tymg9oh01

Sternberg, E. (1995). The Economy of Icons. In W. T. Anderson (Ed.), *The Truth About the Truth: De-confusing and Re-constructing the Postmodern World* (pp. 82–85). New York: Jeremy Tarcher/Putnam.

Teitcher, J. E. F., Bockting, W. O., Bauermeister, J. A., Hoefer, C. J., Miner, M. H., & Klitzman, R. L. (2015). Detecting, Preventing, and Responding to "Fraudsters" in Internet Research: Ethics and Tradeoffs. *The Journal of law, Medicine & Ethics : A Journal of the American Society of Law, Medicine & Ethics, 43*(1), 116–133. https://doi.org/10.1111/jlme.12200.

Tickle, L. (2012). So You Want to Be the New Brian Cox? … How to Become a Celebrity Academic. *The Guardian International*. Retrieved from https://www.theguardian.com/education/2012/may/14/celebrity-academic-radio-tv-funding

Tzoumaka, E., & Zarkada, A. (2013, September 25–27). *Towards a Model of Consumer Engagement with Celebrity Brands*. Paper presented at the 4th EMAC Regional Conference, St. Petersburg, Russia.

Tzoumaka, E., & Zarkada, A. (2016, June 22–24). *'He Had a Meaning in My Mind' Unpacking Celebrity Footballer Brands*. Paper presented at the 4th International Conference on Contemporary Marketing Issues, Heraklion, Greece.

United States Postal Service. (2006, January 13, 2016). A Decade of Facts and Figures. Retrieved from https://about.usps.com/who-we-are/postal-facts/decade-of-facts-and-figures.htm

Wang, H.-C., & Doong, H.-S. (2010). Nine Issues for Internet-Based Survey Research in Service Industries. *The Service Industries Journal, 30*(14), 2387–2399. https://doi.org/10.1080/02642060802644926.

Zarkada, A., & Polydorou, C. (2013). You Might Be Reputable but Are You 'Liked'? Orchestrating Corporate Reputation Co-creation on Facebook. In T. Bondarouk & M. R. Olivas-Lujan (Eds.), *Advanced Series in Management—Social Media in Strategic Management* (Vol. 11, pp. 87–113). London: Emerald.

Zarkada, A., & Tzoumaka, E. (2014, July 15–18). *Exploring Soccer Fans' Schemata Regarding Global VS Local Human Brands*. Paper presented at the 2014 Global Marketing Conference Singapore.

Zarkada, A., & Tzoumaka, E. (2015, May 26–29). *The Effect of Footballer Brand Characteristics on Fans' Ticket Purchase Intention*. Paper presented at the European Marketing ACademy 2015, Leuven, Belgium.

Zarkada, A., Tzoumaka, E., Siomkos, G., & Panigyrakis, G. (2014, December 1–3). *Achievement-Based Celebrities as Objects & Instruments of Consumption*. Paper presented at the Australia-New Zealand Marketing Academy (ANZMAC2014 Agents of Change), Brisbane, Australia.

2

Why Consumer Psychology Needs Neurophilosophy

Paul M.W. Hackett and Gordon R. Foxall

Summary

In this essay we will suggest that neuromarketing-based consumer behavior research (research that utilizes neurophysiological measures in its methods) will better understand the consumer through the incorporation in their approach of an interdisciplinary neurophilosophy. Neurophilosophy is the melding of neuroscience and philosophy in the investigation of the human mind. In this essay we will suggest that neurophilosophy possesses explanatory sprezzatura, as being endowed with a nonchalant effortlessness in its assertions: Informative and innovative insights often reward the researcher employing a neurophilosophical perspective. The question however may be asked regarding the applied veracity of these answers. Keeping in mind such approbations, we will advise a degree of vigilance

P.M.W. Hackett (✉)
University of Gloucestershire, Cheltenham, MA, UK

University of Cambridge, Cambridge, UK

G.R. Foxall
Cardiff University, Cardiff, UK

L. Moutinho, M. Sokele (eds.), *Innovative Research Methodologies in Management*,
https://doi.org/10.1007/978-3-319-64394-6_2

in interpreting the findings of this form of research. We issue this caveat due to the heterogeneous ontological status of the neurophysiological events, psychological constructs, and various forms of human behavior which may make meaningful interpretations of research that includes these different constructs and events, far from simple.

In order to explore these claims we will, in our essay, consider the existential nature of the varied components of ontologies of the types of variables that are frequently associated with neuromarketing/consumer behavior. We employ the term "ontology" in our writing in a sense that delimits our area of interest to being with the "basic elements" of consumer behavior. On this definition, "consumer behavior understanding" is ontological and is seen to be a whole with the following composite parts: neurophysiological events (we are specifically concerned with events as these are measured through appropriate technologies, such as functional magnetic resonance imaging (fMRI), electroencephalography (EEG), steady state topography (SST) and biometric measurement of heart and respiratory rates, galvanic skin responses, eye tracking, etc.) and psychological constructs (and again, we are particularly interested in these as they are assessed using, e.g., psychometric test batteries, attitude scales, and other tools for investigating affective, cognitive, and sensorimotor responses to marketing materials). As well as stating this rudimentary ontology for the investigation of neuromarketing and neuroscientific studies of consumer behavior, we propose that the interplay of parts to whole within the ontology is of vital importance in interpreting consumer behavior. We therefore investigate the status of the ontological components as a mereology and make claim to the existence of differences in the status of components and that these dissimilarities are significant when interpreting consumer behavior.

Introduction

The prospect of neuromarketing[1] fills many with excitement, nay joy, even as it fills others with horror. At its crudest, the possibility of manipulating consumer behavior through the modification of potential buyers' unconscious information processing and decision making could be expected to arouse the same dread and alarm as the feasibility of advertisers' exploiting subliminal perception did when it was first mooted.

Neuromarketing (see, e.g., Morin 2011 and Fisher et al. 2010) has seen enormous growth as a subdiscipline within marketing over the previous decade and a half,[2] and this escalation in neuromarketing as a discipline is demonstrated by the large number of books that have been published in just the closing few months of 2016 (examples include Gallucci 2016; Niemczyk 2016; Szymczak 2016; Zeidler 2016). As a discipline within marketing, and as an area of academic study, neuromarketing is rooted in the assessment of neurophysiological activity (the functioning of the central and peripheral nervous systems) as a basis for conducting research into consumer behavior within a marketing context (see Hammou et al. 2013 for a review of neuromarketing's position within marketing). From the controversies that marked its initial reception, neuromarketing is now widely employed in advertising and marketing. It is of particular interest to this essay to note the use, on the one hand, neuroscience's use of technological methods to measure neurophysiological events, and on the other hand, psychological research that has typically gathered data through self-reported psychological tools. Advocates for the adoption of a neurophysiological approach have the hope and aim that by employing neuroscience's techniques, neuromarketing will allow the brain of the consumer to be directly probed and will thus remove error associated with psychological methods that rely upon, and court error due to, consumers' reflecting upon and reporting their experiences.

Yet neuromarketing, somewhat in contrast, seems to evoke a neutral, even muted, reaction, at least among marketing managers and, perhaps more interestingly, some marketing academics who see their role as predominantly that of furthering managerial concerns. That such aspects of consumer choice as brand and retail preferences might come under greater executive control clearly motivates those for whom industrial objectives are paramount to seek out tools of neuromarketing on which to base market strategies that enhance corporate effectiveness and profitability. The hope that consumer cognition might turn out to be reducible to neurophysiological functioning, raises both expectations and concern over the marketing professionals' potential to influence consumers' thought and actions (Marques dos Santos and Moutinho 2016). However, it is time for market academia take stock, not simply because of the ethical concerns such matters evoke, but also because the place of marketing in the academy requires that it rise to meet the intellectual concerns and opportunities that flow from the marriage of consumer research with neuroscience.

One of the reasons for holding to the notion regarding the exemplary usefulness of neurophysiological measures may be stated as follows. There is a strand to the understanding consumer behavior, via a deeper comprehension of the neural underpinnings that are correlated with choice behaviors, that extends far beyond neuromarketing. We claim that that form of enquiry ought to enthuse marketing academics in particular since it raises fundamental intellectual questions about the nature of consumption in a very broad sense, and with marketing practice and its relationship to consumers' choices. It is the prospect of *consumer neuroscience,* the possibility that the neural correlates of consumer behavior can be identified is intriguing for consumer theorists and marketing analysts who conceive their intellectual endeavors within the broader milieu of the social, behavioral, and biological sciences. The putative explanation of consumer choice in terms of neural events promises to link consumer psychology with biology and to relate some of the most ubiquitous aspects of human behavior to their evolutionary origins. As Bickle (2003, p. xiii) puts it:

> We move closer every day to actually having something that human beings have speculated about for centuries, a purely physical account of behavioral causes.

The result could be a genuine consilience of the social and biological sciences. In particular, these prospects call for an understanding of the implications for consumer and marketing research of the philosophical ramifications of the emerging science and technology of neural processing. Furthermore, we call for a neurophilosophy that mediates the understanding that arises from marketing neuroscience. Later in this essay, we will however note that while there may be correlational evidence that associates consumer behavior with both the processes of the central nervous system and ultimately consumer behavior, caution is needed in interpreting these associations due to inherent differences in the ontological status of the components of consumer behavior understanding built upon neurophysiology.

As an area of study and understanding, neurophilosophy is centrally concerned with the ontological and methodological bases of psychology: first, with what existential claims can be made for the components of psychological explanation such as beliefs and attitudes (see, for instance, Foxall 2008, 2014a, b, 2016a, b); second, with the ways in which psychological concepts can legitimately enter explanations of behavior (see, inter alia, Bennett and Hacker 2003; Churchland 1986). Can we claim a distinct kind of existence for cognitive variables or does their deployment in behavioral explanations simply mark an alternative mode of speaking about neurophysiological entities which, under some understandings of causation, are deemed the actual causes of movements and actions? Hardcastle and Stewart (2009) prompt us to ask questions such as: Does fMRI scanning genuinely provide a "cerebrascope" which permits the tracking of cognition? Do the firing patterns of neurons constitute actual thoughts, or the physical representations of thoughts, just correlates of thoughts, or what? Two questions we ought to pose before we give way to speculation are: How far can neurophysiological events shed light on cognition? And what is the role of behavior, especially verbal behavior, in interpreting neural events? (Bickle 2009).

The answers to these two questions are of vital importance in determining the appropriateness of employing neurophysiological explanations of consumer behavior and behavioral predictions that are based upon these. Through the adoption of the conceptual rigor and procedures associated with the philosophy of science, the philosophy of neuroscience makes efforts to elucidate the methods and the ensuing results that come about from neuroscientific research. In this chapter we argue that consumer psychology is in need of a neurophilosophy that relates neuroscience to the explanation of behavior via the construction of cognitive processing mediated through an appreciation of the differential status of the neurophysiological, psychological, and behavioral variables. More specifically, we claim that while neurophysiology may offer a correlative guide to understanding of consumer cognition and ultimately behavior, caution is needed when interpreting the status of such predictive claims.

On such an understanding, neurophilosophy relates to and addresses the implications of neurophysiological research for the explanation of behavior. We contend that the potential role of neuroscience in consumer psychology is in need of discussion and in-depth interrogation, and later in this chapter we go on to consider some of the difficulties involved in interdisciplinary research of the required kind. We conclude by considering both the applied utility and the acceptability of neuroscientific findings regarding consumer behavior. However, in the next section of this essay we undertake an exposition of the ontological status of neuroscientific explanations of consumer behavior. We continue by asking what is the most appropriate level of analysis for this work if a convincing case is to be made for a consumer neurophilosophy that genuinely contributes to the incorporation of neuroscience in the explanation of consumer choice.

Psychological Constructs and Neurophysiological Events

When considering the role and status of psychological constructs, neuronal activity and neuroscientific knowledge, in relation to consumer behavior, of cardinal importance to our attempts to provide answers to such enquiries is the ontological and mereological[3] status of these entities. More explicitly, questions that address the mereological and ontological status may be asked which include the following: What is the ontological status of psychological constructs and are they simply neurophysiological entities? Are psychological constructs better thought of as behavioral entities, or do they have a status peculiar to themselves? We contend that the answers to these questions are both complex and intricate. Of even more importance is an appreciation of the ontological and mereological nature of these constructs viewed as a complex of items with differential status. These requirements, we contend, are of importance in attempts to proffer neuroscientific-based answers to behavioral questions. Therefore, we will briefly consider the meaning of the terms "ontology" and "mereology."

Several different definitions of the term "ontology" may be identified, each of which has emerged from a different academic or professional discipline. As a result, despite these definitions' seeking ostensibly to increase the precision and specificity of discussions of the substantive contents of consumer psychology, each has developed its own particular and idiosyncratic ontological forms of understanding. Examples of the disciplines that have put forth a specific ontological understanding include philosophy; logic; and technology and information and computer sciences. Respectively, these disciplines have proffered the following definitions of ontology as a branch of metaphysics with a specific concern with the nature of being; an a priori set of hypothesized entities assumed by a theory; explanations of existence within a systematic format; and the specification in a thorough manner of extant elements in a system and their interrelationships. If the above definitions are inspected for commonalities, what emerges is the notion that ontology is in some way associated with an interest in studying and taking into consideration the rudimentary components of existence. Furthermore, the specific fundamental aspects of an ontology are defined by the discipline using the term and the context of its use.

In this essay we use the term "ontology" in the sense of basic or fundamental parts of our experiences as these apply to different forms of explanation of consumer behavior. More specifically still, we employ the term "ontology" to allow the formal explication of more basic or fundamental psychological and neurophysiological constructs as these are understood in the context of consumer behavior. Ultimately, we are calling for a neurophilosophical mediation to allow the appropriate interpretation of psychological and neurophysiological research. In attempting to achieve these aims, we assume an ontology of psychological and neurophysiological events may be better understood, and the meaning of the ontology better revealed, by considering the ontological nature of these components.[4]

As we are specifying that multiple fundamental psychological and neurophysiological events (an ontology of fundamental aspects) can be associated with consumer behavior, there is an immediate requirement for us to consider any interaction of constructs that may be present within our ontology. Thus, we are concerned with a "… theory of part-hood or composition" (Harte 2002, p. 7). Harte and others have investigated theories of part-hood relations. Such associations may be between parts to whole; alternately they may address the relationships of part to part within a specific whole (see, e.g., Kleinschmidt 2014; Casati and Varzi 1999; Churchland 2007; Simons 1987; Surma et al. 1992). Mereology has been used in reference to one of the formal languages that have been developed to describe part-to-whole relationships. However, we use the term "mereology" in a broader, more universal sense as

> the theoretical study (formal or informal) of parts, wholes, and the relations (logical or metaphysical) that obtain between them. (Arlig 2015)

Mereological understandings may focus upon general principles in an attempt to establish how the relationships that underlie an entity and its constituent parts and related to the whole of the entity. In a mereology, both the whole event that is being investigated along with its parts may be concrete or abstract (Varzi 2016).

Another characteristic of mereological accounts is that they have a purposeful focus to either explain or describe. A variety of philosophers have considered part and whole understandings in their investigations and have developed their own understanding of mereology (e.g., Aristotle, Brentano, Husserl, Kant, Leibniz, Lesniewski, and Whitehead).

As our interest is in the relationships between the psychological and neurophysiological events (parts) and their relationship to consumer behavior (whole or total) mereological principles are of importance to us in our exposition. It is important to understand that the first stage in the development of our consumer behavior mereology entails the initial specification of the fundamental aspects of consumer behavior (an ontological evaluation). It is also important that we take great care in assembling a mereology for assisting the understanding of consumer behavior that we systematically and explicitly reference the relationships embodied between the psychological constructs and neurophysiological events as these come together to assist understanding of consumer behavior. Finally, such a typology should be focused to specifically address either consumer behavior in general or particular aspects of this (such as consumer choice, browsing behavior, the effects on behavior of price reductions, etc.). The complexity of the interactions within a mereology that incorporates psychological constructs and neurophysiological events may be based upon any specified number of appropriate constructs or events, where the precise number is determined by the consumer behavior under investigation.

In conducting our ontological/mereological analysis we attempt to understand the relationships between psychological and neurophysiological events, and consumer behavior and the implications of these relationships within a categorial system or ontology. Psychological constructs and neurophysiological events can be thought of as being entities that are functionally or materialistically distinct aspects of consumer behavior. However, they may conversely be conceived as non-separable qualities or quantities in terms of their relationship to a consumer functional or a consumer behavioral totality.

We have briefly provided some explanations and definitions of our terms of reference and it is now apposite to consider some of the structural relationships that may be present in a mereological account of such an intricate behavioral domain. Mereological, part/whole accounts have a series of common formal properties (ingredient, overlap, disjoint, identicalness, asymmetry, transitivity, supplementivity) which enable part/whole relationships to be described and differentiated. Part-to-whole relationships may therefore be characterized as being:

– Ingredient: (a) is the ingredient of (b) if (a) is part of (b) or (a) is (b).
– Overlap: (a) overlaps (b), and vice versa, if something has a common ingredient.
– Disjoint: (a) and (b) are disjoint if they do not overlap.
– Identicalness: if (a) and (b) have the same parts, they are identical.
– Asymmetric: if (a) is part of (b), (b) is not a part of (a).
– Transitive: if (a) is a part of (b) and (b) is a part of c, then (a) is a part of c.
– Supplementive: if (a) is part of (b), then there is a part of (b) that has no common part with (a).

The above are somewhat formal theoretical terms, which we now consider in relation to our analyses of psychological constructs and neurophysiological events associated with consumer behavior.

Consumer and marketing research, and the analysis of marketing management, rely heavily upon psychological (cognitive) constructs such as beliefs and desires, emotions and perceptions, attitudes and intentions: theoretical entities that are usually inferred in some way on the basis of behavioral or neurophysiological observations or both. However, this inference may be of questionable status if the psychological constructs being inferred are not independent qualities to the behavioral or neurophysiological observations from which they are insinuated.[5]

In an attempt to shed light upon the veracity of such an inference, and the nature of psychological constructs and neurophysiological events, we now consider the qualities a neurophysiological and/or psychological events/construct may possess in relation to the explication of a specific consumer behavior. This is achieved by applying the above seven characteristics of part-hood to consumer behavior-related neurophysiological and psychological events and constructs.

Ingredient

When we conduct consumer behavior research, a specific psychological construct (a) may be identified as being the ingredient of neurophysiological event (b) if the psychological construct (a) is part of the neurophysiological event (b) or if the psychological construct (a) is neurophysiological event (b). In an effort to better understand this statement, we can consider this rule in relation to a specific consumer behavior, for example, the act of browsing sale items in an online store. In this situation the psychological constructs that may be associated with this online form of consumer behavior include perception, attention, evaluation, and so on. It seems unreasonable to state that any of these psychological constructs (a), for example, perception is the neurophysiological event (b) itself. This unreasonableness is supported by the complexity of psychological construct (a), say in this instance, attention, where the intricacy of attention obviates claims as to the identical nature of construct (a), attention, with event (b) browsing. Browsing (b) may be highly complex in itself but for a complex construct (a), attention, to be identical with event (b), browsing, (b) browsing must be identical in its complexity to the complexity of attention (a) and the complexity of the interrelationship itself between (a) and (b) attention and browsing must similarly be identical.

However, the other characteristic of being an ingredient is that the psychological construct or constructs listed above (a) is or are a part or parts of the neurophysiological event (b), seems to be inevitable. This statement may be made with respect to the complexity of constructs (a) and (b), as they are both defined as addressing consumer behavior, have at least this definitional quality in common and on this understanding both constructs and events may constitute ingredients. Another characteristic that may typify the constructs we are investigating is that of overlap.

Overlap

For a psychological construct (a) to possess the quality of overlap with neurophysiological event (b), and/or vice versa, the psychological construct must have a common ingredient with the neurophysiological event (b). This seems to be the most easily met criterion under a definition that utilizes the seven relationships of part-hood. This is due to the fact that we have already established, that in the context of consumer behavior psychological and neurophysiological events have common ingredients which are foci of concern in consumer behavior.

Disjoint

A disjoint characteristic is, in some ways, the opposite relationship to that of overlap, as under the condition of disjoint a psychological construct (a) and neurophysiological event (b) are said to be disjointed if they do not overlap. As we have already determined that overlap does occur between these two forms of constructs and events, at least in the sense of their common subject matter of consumer behavior, and as both underpin consumer behavior, a disjoint characteristic would not appear to be present.

Identicalness

Identicality is of particular interest to the two forms of entities we are considering, as identicalness can be said to exist if psychological construct (a) and neurophysiological event (b) have the same parts. This statement is meant in an exclusive as well as an inclusive sense and as the ontological status of psychological constructs and neurophysiological constructs may be conceived in many conceptually functional and descriptively different ways, identicalness would not seem an appropriate characteristic to attribute to these constructs: (a) and (b). For example, the neural activity associated with a consumer behavior, while being intimately associated with the cognitions themselves, are not the cognitions if only in terms of the mode of expression. Consequently, there are necessarily parts of constructs (a) and (b) that cannot be identical.

Asymmetric Relationship

When considering consumer behavior, if a specific psychological construct (a) is part of neurophysiological event (b), then, the two constructs possess an asymmetric relationship if neurophysiological event (b) is however not part of psychological construct (a). To illustrate this, asymmetry implies that while the psychological constructs associated with this online browsing behavior (perception, attention, evaluation, etc.) are part of the neurophysiological event (e.g., the output of an fMRI), the fMRI output is not part of perception, attention, evaluation, and so on. On this definition, it seems probable that in a specific consumer behavioral context, if an action potential (a neurophysiological entity) was associated with perceiving a certain product (the perception itself being a specific psychological construct) then the reverse relationship would be true. Consequently, asymmetry does not appear to be applicable, at least under the conditions of our example.

Transitive Relationship

In the context of consumer behavior, a transitive relationship requires that: if psychological construct (a) is part of neurophysiological event (b), and neurophysiological event (b) is a part of psychological or neurophysiological event (c), then psychological construct (a) is a part of psychological or neurophysiological event (c). A contextualized rendering of this relationship is as follows. If a product choice behavior (a) is part of an fMRI output event (b), and the fMRI output event (b) is also part of perceptual behaviors related to product (c), then choice behavior (a) is also a part of perceptual behavior (c). In this example the claim may be made that this relationship does indeed hold and transitive part-hood is present. However, transitive conditions appear to again be heavily dependent upon the specific consumer behavior, neurophysiological event and psychological construct that is being investigated.

Supplementive Relationship

For a supplementive part-to-whole relationship to exist between neurophysiological and psychological components of a specified consumer behavior, the state must exist where a psychological construct (a) is part of neurophysiological event (b), while a part of neurophysiological event (b) exists that has no common part with psychological construct (a). Thus, for a supplementive relationship to be present among psychological constructs and neurophysiological events, in order to form explanations of consumer behavior, the following must hold true. A consumer behavior (a), for example, a product choice, is part of a neurophysiological event (b), such as an fMRI output, while part of this fMRI output (b) has no common part with the choice behavior (a). As with all of the examples we have provided the exact nature of the psychological construct and the neurophysiological event are of importance in determining the veracity of the part-hood statement. However, in the present example, due to the imprecise nature of the fMRI process, there are parts of an fMRI output that are not associated with or part of the choice behavior. If therefore we take another example and we change the fMRI to be an evoked potential this part-hood characteristic may also be refuted as there are components of the evoked potential, for example, the dying off of the potential, that are not part of the choice behavior but are rather a consequence of it. It would therefore appear that a supplementive nature may typify the part-whole, part-part relationship.

To summarize the characteristics that appear to be present in a part-hood understanding of psychological constructs and neurophysiological events associated with consumer behavior, we are able to state that these constructs and events possess the following:

- They have common ingredients.
- They overlap.
- They are transitive.
- They are supplementive.
- They are not disjointed.
- They are not identical.
- They are not asymmetric.

The consequences of the above tentatively offered seven part-hood relationships have implications for our understanding of neurophysiologically driven consumer behavior research. An example of these can be seen if we consider what are usually identified as being distinct, integrated, and complete forms of consumer behavior. On our account these may be divided into sub-constituents while still constituting a complete systemic whole. For example, the parts of a purchasing action are behaviorally coherent functional units in respect of the buying behavior taken as a whole. Consumers' communications and interactions may similarly be identified as possessing functionally or behaviorally distinct parts that together form entire interactions or communications. Furthermore, within the context of consumer behavior, the relationships between parts to whole may be temporary and/or may vary across time, such that a particular construct may not always be present in a particular behavioral whole: In this case, the mereological definition of psychological constructs and neurophysiological events relationships will change. The question may therefore be asked as to how temporal differences that exist between psychological and neurophysiological events may best be understood when applied to our attempts to understand consumer behavior. The answer to this would seem to be that research must be designed that explicitly takes into account the differences that must exist between ostensibly similar consumer behaviors due to (1) the ontological differences between neurophysiological events and psychological constructs, in conjunction with (2) the acknowledgment of a wide variety of internal conditions or states of affairs of individual consumers along with a consideration of consumer behavior settings and contexts.

So, when we ask again the questions that motivated this essay we are in a position to proffer tentative answers. We asked, what is the ontological status of psychological constructs? Are they simply neurophysiological entities? Or behavioral? Or do they have a status peculiar to themselves? From the above analyses we suggest that the ontological status of neurophysiological events and psychological constructs are distinctly different but that they possess similarities and that these differences and similarities vary in accordance with their part-whole relationship to specific consumer behaviors. Arising from these differences it seems apparent that psychological constructs cannot be thought of as simply neurophysiological or behavioral events. Instead, psychological constructs possess a unique status within mereological accounts of consumer behavior.

In the preceding pages of this essay we have claimed that an ontologically distinct status exists between consumer behavior and neurological and psychological correlates of this behavior. The claims we have made are rooted in our adoption of a neurophilosophical perspective. However, the "neuro-" prefix has been found to cause some who read research accounts that originate from a "neuro-" discipline with enthusiasm, while others are less enthusiastic. As we have proffered the utility of neurophilosophical accounts of consumer behavior, in this final section we consider selected research that has evaluated the acceptability of neuromarketing.

We have claimed neurophysiological events and psychological constructs offer strands that offer the potential to develop a more complete understanding of consumer behavior. This assertion rests upon the distinctive ontological and mereological status that we have demonstrate to exist between the knowledge components associated with neurophilosophical and psychological accounts. However, in order for neurophilosophy, neuroscientific and neuromarketing accounts to become influential aspects of consumer behavior knowledge, it has been claimed that researchers need to address practical challenges. For example, Fisher et al. (2010) investigated the professional issues associated with neuromarketing. The researchers reviewed neuromarketing's history and conducted an exploratory study of the content neuromarketing websites. They identified the heterogeneity of neuromarketing practice, which was discovered to employ a variety of technological approaches. They identified that neuromarketing companies made use of academics in their practices and that the media coverage of neuromarketing is disproportionately enlarged in comparison to peer-review articles on the subject. They also cautioned over claims regarding the ability to predict consumer behavior made in neuromarketing practice. This last point is of particular relevance to this essay as we have stated that care must be taken when using psychological constructs and neurophysiological events to typify and understand consumer behavior. Claims of behavioral prediction require an even greater degree of certainty than statements of association (against which we cautioned, although were optimistic).

Fisher et al. (2010) engaged in what they called the search for quantification and for the attribution of certainty within realms previously seen as indefinite aspects of human behavior. They concluded that important implications are associated with neuromarketing in terms of academic-industrial partnerships. Moreover, these implications are associated with conducting responsible research and with the general public's understanding of the brain. These findings are related to the ontological status of neuroscientific research in general and to the need to present the findings of consumer behavior research at an appropriate level. The ontological and mereological understanding of consumer behavior incorporates notions of inequality between the constituent parts, whole, part-to-part, and part-to-whole relationships, and it is a challenge to those working in neuromarketing to make widely accessible the outcomes from their research.

As well as contributing to understanding within consumer behavior, the advent of neuromarketing science has resulted in a multitude of varied questions that are associated with consumer perceptions of neuromarketing. If neuromarketing is to be influential it needs to develop a means to educate the public about its subject, as the involvement of consumers as participants in their research is fundamental to neuromarketing enquiry. The research of Elitza Bakardjieva and Allan Kimmel (2016) describes the findings of questionnaire surveys that have investigated the role of personal constructs associated with perceptions of neuromarketing research. They claim that "neuromarketing research knowledge, attitudes toward science, attitudes toward technology, and ethical ideology" are seminal factors associated with perceptions of the ethicality of, and attitudes toward, neuromarketing research. Moreover, these are factors in participants expressed willingness to participate in neuromarketing research studies. These findings warn against the embodiment of neuroscience or neurophysiological research results in a manner that is either imprecise or inaccessible to the intended audience. Perhaps, developing an ontological understanding of the differential status of neuroscience, psychology, and behavioral accounts, and understanding the mereological interplay of these elements, may assist the researcher to present their findings in a consistent and clear format.

Recently, mereological studies have been conducted that have taken location into their accounts of part-hood relations (see Gilmore 2014). The physical location or region is obviously an important part of many accounts of what constitutes the identity of an event. Some areas of research within consumer behavior are location specific or location related and the employment of a mereological understanding in these cases could yield useful findings. For example, if we are interested in the development of shopping areas, or in other forms of situated consumer behavior, the incorporation of location within a mereology may reveal interesting understanding.

Conclusions

We conclude by asking two overarching questions that arise from the above discourse. The first question is:

> In what ways, if any, must the mereological structure of a specified consumer behavior mirror the mereological structure of the psychological constructs and neurophysiological events used to explain this behavior?
>
> The second question is:
>
> In what ways, if any, must the mereological relationships between some consumer behaviors mirror the mereological relationships between the neurophysiological events and psychological constructs used to explain these behaviors?

Neither of these questions may be answered from our research, but both questions suggest useful lines of investigation of the ontological and mereological status of a neuroscience of consumer behavior. Furthermore, we contend that neurophilosophy offers an approach to answering these questions and also has the potential to avail greater understanding of consumer behavior.

Notes

1. The attitudinal position that one assumes in terms of appraising neuromarketing as a joyous or fearful occurrence may be related to a more general orientation an individual may assume toward the incorporation of neuroscience in tandem with a host of other disciplines. This way of developing novel fields of study has been termed "neuroculture" and includes neuroaesthetics, neurotheology, neuroeducation, and several other interdisciplinary amalgams (Frazzetto and Anker 2009; Rolls 2012).
2. One of the first mentions or documented activities associated with the development of neuromarketing is evidenced in Zaltman and Kosslyn's 2000 patent application for "Neuroimaging as a marketing tool."
3. Mereology is the study of interrelationships between parts and wholes. For further information regarding mereology, the interested reader is guided toward the following texts: Calosi and Graziani (2014); Hackett (2014, 2016, 2017); Kleinschmidt (2014); Simmons (2000).
4. Thus, in this essay we propose a framework or structure for the ontological units of consumer behavior and claim the utility of a neurophilosophically based ontological understanding of consumer behavior.
5. If independence cannot be established between behavioral or neurophysiological observations and the psychological constructs they are seen to be indicative of, then the inference of the latter based upon the former may be self-fulfilling.

References

Arlig, A. (2015). *Medieval Mereology*. The Stanford Encyclopedia of Philosophy (Fall 2015 Edition), Edward N. Zalta (Ed.). URL = https://plato.stanford.edu/archives/fall2015/entries/mereology-medieval/.

Bakardjieva, E., & Kimmel, A. J. (2016). Neuromarketing Research Practices: Attitudes, Ethics, and Behavioral Intentions. *Ethics & Behavior*. https://doi.org/10.1080/10508422.2016.1162719

Bennett, M. R., & Hacker, P. M. S. (2003). *Philosophical Foundations of Neuroscience*. Oxford: Wiley-Blackwell.

Bickle, J. (2003). *Philosophy and Neuroscience: A Ruthlessly Reductive Account*. Dordrecht: Kluwer.

Bickle, J. (Ed.). (2009). *The Oxford Handbook of Philosophy and Neuroscience*. Oxford: Oxford University Press.

Calosi, C., & Graziani, P. (2014). *Mereology and the Sciences: Parts and Wholes in the Contemporary Scientific Context (Synthese Library)*. New York: Springer.

Casati, R., & Varzi, A. C. (1999). *Parts and Places: The Structures of Spatial Representation*. Cambridge, MA: The MIT Press.

Churchland, P[aul]. (2007). *Neurophilosophy at Work*. Cambridge: Cambridge University Press.

Churchland, P[atricia], S[mith]. (1986). *Neurophilosophy: Toward a Unified Science of the Mind/Brain*. Cambridge, MA: MIT Press.

Fisher, C. E., Chin, L., & Klitzman, R. (2010). Defining Neuromarketing: Practices and Professional Challenges. *Harvard Review of Psychiatry, 18*(4), 230–237. https://doi.org/10.3109/10673229.2010.496623

Foxall, G. R. (2008). Reward, Emotion and Consumer Choice: From Neuroeconomics to Neurophilosophy. *Journal of Consumer Behaviour, 7*, 368–396.

Foxall, G. R. (2014a). Neurophilosophy of Explanation in Economic Psychology: An Exposition in Terms of Neuro-Behavioral Decision Systems. In L. Moutinho, E. Bigné, & A. K. Manrai (Eds.), *Routledge Companion to the Future of Marketing* (pp. 134–150). London/New York: Routledge.

Foxall, G. R. (2014b). Cognitive Requirements of Competing Neuro-Behavioral Decision Systems: Some Implications of Temporal Horizon for Managerial Behavior in Organizations. *Frontiers in Human Neuroscience*, *8*, Article 184, 1–17. https://doi.org/10.3389/fnhum.2014.00184.

Foxall, G. R. (2016a). *Addiction as Consumer Choice: Exploring the Cognitive Dimension*. London/New York: Routledge.

Foxall, G. R. (2016b). Metacognitive Control of Categorial Neurobehavioral Decision Systems. *Frontiers in Psychology (Theoretical and Philosophical Psychology), 7*(170), 1–18. https://doi.org/10.3389/fpsyg.2016.00170.

Frazzetto, G., & Anker, S. (2009). Neuroculture. *Nature Reviews Neuroscience, 10*, 815–821.

Gallucci, F. (2016). *Neuromarketing*. Milano: Egea.

Gilmore, C. (2014). *Location and Mereology.* The Stanford Encyclopedia of Philosophy (Fall 2014 Edition), Edward N. Zalta (Ed.). URL = https://plato.stanford.edu/archives/fall2014/entries/location-mereology/.

Hackett, P. M. W. (2014). *Facet Theory and the Mapping Sentence: Evolving Philosophy, Use and Application*. Basingstoke: Palgrave Macmillan.

Hackett, P. M. W. (2016). *Psychology and Philosophy of Abstract Art: Neuroaesthetics, Perception and Comprehension*. Basingstoke: Palgrave.

Hackett, P. M. W. (2017). *The Perceptual Structure of Three-Dimensional Art: A Mapping Sentence Mereology, Springer Briefs in Philosophy*. Heidelberg: Springer.

Hammou, K. A., Galib, M. S., & Melloul, J. (2013). The Contributions of Neuromarketing in Marketing Research. *Journal of Management Research, 5*(4), 20–33.

Hardcastle, V. G., & Stewart, C. M. (2009). fMRI: A Modern Cerebrascope? The Case of Pain. In J. Bickle (Ed.), *The Oxford Handbook of Philosophy and Neuroscience* (pp. 179–199). Oxford: Oxford University Press.

Harte, V. (2002). *Plato on Parts and Wholes: The Metaphysics of Structure*. Oxford: Oxford University Press.

Kleinschmidt, S. (2014). *Mereology and Location*. Oxford: Oxford University Press.

Marques dos Santos, J. P., & Moutinho, L. A. (2016). Decision-"Making" or How Decisions Emerge in a Cyclic Automatic Process Parsimoniously Regulated by Reason. In G. R. Foxall (Ed.), *The Routledge Companion to Consumer Behavior Analysis*. London/New York: Routledge.

Morin, C. (2011). Neuromarketing: The New Science of Consumer Behavior. *Society*, *48*(2), 131–135: 61–62.

Niemczyk, C. (2016). *Neuromarketing. Kundenkommunikation Und Markenfuhrung Fur Die Unternehmenszukunft*. Munich: Grin Verlag.

Rolls, E. T. (2012). *Neuroculture: The Implications of Brain Science*. Oxford: Oxford University Press.

Simons, P. (1987). *Parts. A Study in Ontology*. Oxford: Oxford University Press.

Simmons, P. (2000). *Parts: A Study in Ontology*. Oxford: Clarendon Press.

Surma, S. J., Srzednicki, J. T. J., Barnett, J. D., & Rickey, V. F. (1992). *Collected Works*, 2 vols. Dordrecht/Warszawa: Kluwer/Polish Scientific Publishers.

Szymczak, A. (2016). *Neuromarketing: Implications for Dove Digital Campaign*. Saarbrücken: LAP Lambert Academic Publishing GmbH & Co. KG.

Varzi, A. (2016). *Mereology*. The Stanford Encyclopedia of Philosophy (Winter 2016 Edition), Edward N. Zalta (Ed.), forthcoming. URL = https://plato.stanford.edu/archives/win2016/entries/mereology/.

Zaltman, G., & Kosslyn, S. M. (2000, August 8). *Neuroimaging as a Marketing Tool*. 6,099,319. U.S. Patent.

Zeidler, N. (2016). *Neuromarketing Im Internet*. Munich: Grin Verlag.

3

Emotivity and Ephemera Research

Kip Jones

The early waves of renewed interest in the narrative paradigm (or the narrative "turn" in qualitative research as it developed in the late 1970s, 1980s, and early 1990s) and the onset of the "post-modern era" in qualitative approaches established protocols, procedures, and language that, by now, are repeated habitually. By 2000 Denzin could tell us:

> We live in narrative's moment. The narrative turn in the social sciences has been taken. The linguistic and textual basis of knowledge about society is now privileged. Culture is seen as performance.
>
> Everything we study in contained within a storied, or narrative, representation. Indeed, as scholars we are storytellers, telling stories about other people's stories. We call our stories theories. (Denzin 2000, p. xi)

This chapter originally appeared as "A Report on an Arts-Led, Emotive Experiment in Interviewing and Storytelling," *The Qualitative Report* 2015 Volume 20, Number 2, Article 6, 86–92, http://www.nova.edu/ssss/QR/QR20/2/jones6.pdf

K. Jones (✉)
Bournemouth University, Poole, UK

L. Moutinho, M. Sokele (eds.), *Innovative Research Methodologies in Management*,
https://doi.org/10.1007/978-3-319-64394-6_3

As our skills at in-depth interviewing continued to develop, we became better and better at acting as but "silent witnesses" to the lives of others. Ethical considerations and sensitivities became ethical procedures and limitations over time. As the subtleties of the interview environment became more familiar, at the same time, our encounters with strangers became more constrained by committees and the management culture pervading academia. These drove narrative researchers further into taking the position of the "neutral observer" and the disengaged participant.

In addition, we began to routinely repeat what are by now shop-worn words in our academic outputs such as *rigour, robust, thick, embodied,* and *evocative* to support (or deny?) our emotive tendencies. Most of those words have been repeated *ad infinitum* for more than 20 years now, degenerating into no more than code words signalling membership in a particular scholarly community. They have become words without force. Perhaps it time now to look both inward and elsewhere (to culture, to the arts, to literature, and so forth) to find fresh inspiration and vocabulary to support a new "emotive" participatory approach to our encounters with others.

There is a New Emotivity emergent in academia worth exploring

- **Time and time again, when given the opportunity, scholars long to connect emotionally with the people about whom they are writing.**
- **The difficulty encountered for academics wishing to write creatively is that we are programmed to repeat (endlessly) what we've read to establish "validity."**
- **When you write to provoke (arouse) readers emotionally, don't mimic words you've read to do it. Instead, chose unique words that equal your experience.**
- **Scholars realising the soundness of their emotional connectivity need to find their own language to express feeling—a new language not simply justified by the idiom preceding them.**

What is this new "emotivity"? In spite of constraints and time and time again when given the opportunity, scholars long to connect emotionally with the people whom they are investigating. Indeed, scholars realising the soundness of their emotional connectivity yearn for a language to express these feelings—a new language not simply justified by the idiom preceding them. The difficulty encountered for academics wishing to write emotively and creatively is that we are programmed to repeat (endlessly) what we've read to establish "validity." Rather than repeat words that have preceded us in the literature, Neo Emotivism (Jones 2014) asks us to choose unique words that bring to life our unique interactions with others by beginning with ourselves. Perhaps if we return to C. Wright Mills and *The Sociological Imagination*:

> The emotions of fear and hatred and love and rage, in all their varieties, must be understood in close and continual reference to the social biography and the social context in which they are experienced and expressed. …The biography and the character of the individual cannot be understood merely in terms of milieux, and certainly not entirely in terms of the early environments … When we understand social structure and structural changes as they bear upon more intimate scenes and experiences wear able to understand he causes of individual conduct and feelings of which men (*sic*) in specific milieux are themselves unaware. (Mills 1959, 2000, pp. 161–162)

Performative Social Science: A Methodological and Theoretical Base

The workshop under discussion is an example of the continuing and ongoing development of a *Performative Social Science* (Jones 2012), or the use of tools from the arts in researching and/or disseminating social science studies. Many social scientists have begun to turn to the Arts for both inspiration and practical assistance in answer to frustrations with more standard ways of carrying out and/or diffusing research. What "performative" refers and relates to in these contributions and elsewhere is the communicative powers of research and the natural involvement of an "audience," whether that be connecting with groups of citizens, peers, or students, a physical audience or a cyber-audience, or even a solitary reader of a journal or a book.

Relational Aesthetics (Bourriaud 2002) offers a theoretical ground for the complexities of connections across seemingly disparate disciplines such as the Arts and Sciences and for further exploration of the synergies between both disciplines as well as communities beyond the academy. Nicolas Bourriaud's *Relational Aesthetics* is suggested as a starting point because it offers a post-modern, contemporary philosophy that allows academics to think about aesthetics and the use of platforms from the Arts across disciplines in refreshing ways.

Relational Art is located in human interactions and their social contexts. Central to it are inter-subjectivity, being-together, the encounter and the collective elaboration of meaning, based in models of sociability, meetings, events, collaborations, games, festivals, and places of conviviality. By using the word "conviviality," the emphasis is placed on commonality, equal status, and relationship (Hewitt and Jordan 2004, p. 1). *Relational Aesthetics* or "socialising art" often comprises elements of interactivity, but its most noticeable characteristic is its socialising effect. Through such efforts, it aims to bring people together and to increase understanding.

The goals of a *Performative Social Science* based in *Relational Aesthetics* are:

1. To dramatically demonstrate through meaningful impact, the value and worth of in-depth Social Science research carried out, interpreted, and/or disseminated through use of tools from the Arts and Humanities.
2. To further substantiate the methods of *Performative Social Science* in which community is central to (re)discovering meaning and utility through a Relational Art (Bourriaud 2002), located in human interactions and their social contexts. Central to Relational Art are inter-subjectivity, being-together, the encounter and the social construction of meaning.
3. Through relational artistic activity, to strive "to achieve modest connections, open up (one or two) obstructed passages, and connect levels of reality kept apart from one another" (Bourriaud 2002, p. 8).

Perhaps then, there is a "New Emotivity" emergent in academia worth exploring.

Experimentation

The penny began to drop for me when the Bournemouth University ARTS in Research (AiR) Collaborative met up recently for two days of experimentation. I am familiar with health and social care academics having a proclivity towards sensitivity to the often-emotional stories of others gleaned through their investigative encounters. What surprised and encouraged me were faculty and students attending the workshop from Media, Design, Engineering and Computing and Tourism with the same ache to connect emotionally with their subjects and to acknowledge the "first person" in their dialogues.

There was a sweet nostalgia present in my informal biographic encounters with fellow academics on this occasion, wistful for the days of the likes of David Bowie and Kate Bush. Their recollections were often about how we used to be before we were led to believe that we needed to behave (differently). It was life in the British academy pre-RAE and REF—the "pre-REFaelites," to coin a phrase. It was often dialogue reminiscent of a time in our shared lives of both emotional conflict and emotional connect.

What does this tell us? Indeed, scholars often find their own narratives in the stories that participants tell them for their research. A big part of Neo Emotivism is embracing this phenomenon instead of backing away from it. The relationships that can be established through such connections are potent and ripe with possibilities for innovation and change in academic connectivity.

The Experiment Redux

The newly formed and loosely organised ARTS in Research (AiR) Collaborative at Bournemouth University was called together for two days for the purpose of a workshop on biography, narrative, and arts-based approaches to collecting and disseminating the personal stories of others by using our own.

The instructions were deceptively simple: The ARTS in Research (AiR) Collaborative would like you to contribute to an experiment. Please bring your past as a present to a workshop. You will give it to someone else. They get to keep it.

Look through that box at the back of the wardrobe or in the loft—the one with bits and pieces that you have been unable to throw away because they represent you and your past. You are going to give some of them away now.

Find some of those precious objects to include in a small packet. Objects might include a paperback novel, pamphlets, railroad tickets, stamps, old letters or photographs (from when photographs used to be physical things), a mix-tape cassette (still have any?) or a 45 record, a food-stained recipe card, a small piece of clothing, an accessory like a ribbon or a badge, sheet music, keys, post cards, used concert or theatre tickets, a self-penned poem or a song, or a drawing. Select a few of the objects that tell a particular story from a particular time in your life. Finally, find a box or something else to put them in or wrap them in. Wrap them lovingly, using beautiful materials, perhaps ones that you also have collected. No more than could fill a cigar box or a shoebox at most.

Bring your gift to the workshop. You will agree to exchange presents with one person, a stranger, someone chosen for you by random. You will talk to each other, telling each other stories about the contents. You might make some notes, but be a good listener/observer.

After eating lunch with your partner, we will gather to begin to create individual projects around the earlier exchanges.

Day Two: will be "Show & Tell"—more show than tell. You will present your partner's story in five minutes using any media of your choosing that is convenient. You may want to have your phone, your iPad, or your laptop with you. You will be creating "narrative postcards" of the stories that you have experienced on Day One.

Participants were then reminded of advice from one of my favourite characters, Little Edie, from *Grey Gardens* (Maysles and Maysles 1975): "It's difficult to draw a line between the past and the present—awfully difficult."

Potential participants were asked to consider:

- Other than listening, how do we gather life stories?
- How do we involve participants in "gifting" us with their stories?
- Other than dry academic reports, how can we retell these stories in sensitive and ethical way to wider audiences?
- How do the stories themselves inspire creativity in retelling them?
- How can we involve participants in the retelling of their stories?
- How much of their story is also our story?
- When is the gathering of the story itself, itself the story?
- How willing are we to let go of ourselves?

Promised benefits of participating included:

- Form new relationships with colleagues across disciplines and Schools.
- Experiment with arts-based methods to gather data and represent/disseminate research findings.
- Develop more participatory relationships/collaborations with research participants.
- Explore visual and tactile methods of gathering data using all of the senses.

Ephemera

Ephemera can be defined as things that exist or are used or enjoyed for only a short time. These are sometimes collectable items that were originally expected to have only short-term usefulness or popularity. Such objects can have added value for the researcher willing to move into interpretation of the visual, the physical, the auditory, and the sensual. In my audio/visual production for *The One about Princess Margaret* (Jones 2007), for example, the presentation relied upon auto-ethnography/auto-biography/auto-ephemera to describe its author as a member of a culture at a specific time and place. As my first foray as the "Reluctant Auto-ethnographer," the production used tools from the arts as a powerful way to recover yet interrogate the meanings of lived experiences.

As auto-ephemera, it documented minor transient personal moments of everyday life: something transitory, lasting a day.

I mention this early experiment because they sometimes stay with us and take on a life of their own. I am now developing the treatment and script for a full-length coming-of-age, gay rom-com feature film based on this earlier short piece's story. There are files full of ephemera facilitating the development of this new project. Indeed, objects resuscitate memories and enrich the telling of them.

The Results

There was initial trepidation from the gathered workshop participants, particularly around sharing with a "stranger" and with the requirement to "give away" their objects to that stranger. Once they were informed on the day that the "gift" requirement would be up for negotiation between the partners, participants felt somewhat easier. Still, it was very much an adventure into the unknown, full of excitement, but also some nervousness. From the very start, both partners were aware that they would be taking in turns identical roles with each other, first as the listener, then as the storyteller.

One of the most intense realisations from this experiment was recalling that in many of our previous more usual interview experiences of asking a stranger to reveal intimate details about her/his life, we assumed our own neutrality and distance from the story and the storyteller her/himself. The experiment allowed those who usually would be on the receiving end of a stranger's tale, to reverse roles. By doing this, we learned a great deal about what it feels like to reveal one's often most private self to an unfamiliar person. I doubt any of us will go back to interviewing without having been profoundly changed by this experiment.

The assignment of producing a five-minute presentation for the second day also was not without some concerns for participants. It was a case of a quick turnaround, only having late afternoon and the evening of the first day to come up with the presentations for the next day. This, however, encouraged participants to think creatively about the task and use ingenuity.

After the pairs had shared their stories, I offered a few examples of working with data creatively to kick off a discussion and brainstorming for the five-minute presentations the following day. It was an exploration of finding ways and means of responding creatively to detailed data as well as dealing with time and material constraints. Copies of a few chapters from *Michael Kimball Writes Your Life Story (On a Postcard)* (Kimball 2013) were shared. His compilation of the book's life stories started in 2008 as a performance piece at an arts festival. I am a great fan of Kimball's writing (I refer to his and the work of some of the other conceptual novelists as "the new writing"). I often recommend his books to fellow academics as a kind of intellectual colonic irrigation to improve their scholarly literary outpourings. Kimball is someone at whom Proust would have smiled. He constructs, through simple sentences, complex situations and ideas. He is particularly skilful at describing innermost thoughts and feelings and the meniscus that both separates and joins those two intertwining elements in lives. In *Postcards*, Kimball reduces a life to a *soupçon* of a story, usually entailing no more than a few hundred words.

I then turned to a story I recently had heard on *BBC Radio 3* during the Intermission discussion of a Met broadcast of Puccini's *Manon Lescaut*. The opera is based on an early eighteenth-century French novel by Abbé Prévost, controversial at its time and banned in Paris. Puccini's opera followed on the success of Massenet's French opera of the story a few years earlier. The story of Manon was a popular one in late nineteenth-century Europe. The tale of Manon's life, however, is a complicated one to tell. Puccini decided to portray Manon's life in four acts, each act representing a discrete time and occurrence in her life. Although cautioned that the audience would never grasp the plot of the whole story, Puccini insisted that each act was like ***a postcard sent from a life story*** and should be thought of in those creative terms.

Finally, I showed a short three-minute video, *I Can Remember the Night* (Jones 2001), which I made years ago now that encapsulates the driving event in one woman's life story. It was produced from just one paragraph from her three-hour interview, transformed into an audio/visual production. In this clip "Polly," a 65-year-old woman from the Midlands in the UK, recalls the time as a child when her parents sat her down and asked her which of them she wanted to be with. Her story, re-narrated by three players, represents how this traumatic event became an enduring memory throughout the various stages of her life. This memory not only provided clues to her identity, but also represented an event that played a significant role in the way that she made decisions for the rest of her life.

Other than these few examples offered as inspiration rather than as instruction, participants engaged in a learning process through participation in the two days of workshop activities. Suffice it to say that the resulting presentations were far more illuminating, exciting, and inventive than most PowerPoint lectures could be.

I will not go into the details of the first day's storytelling or the second day's resulting performances here. The group decided that they would like to present their emotive responses to the two days fully, perhaps in an article or for a conference presentation that they will produce as a collaborative effort.

A few of the comments that followed the workshop, however, may whet your appetite for staging an experiment of your own and the development(s) from the AiR workshop to follow:

- *Thank you all for the incredible willingness to be inventive, creative and think/be outside "the box."*
- *An illuminating two days of deep sharing. I was honoured to be there and look forward to more creative adventures together.*
- *Inspiring. An artful and generative suspension of "normal" activity.*
- *I can't remember ever attending such an inspiring event "in house."*

Indeed, scholars often find their own narratives in the stories that people tell them for their research. A big part of Neo Emotivism is embracing this phenomenon instead of backing away from it. The relationships that can be established through such connections are potent and comprised of more than the sum of their parts.

The early waves of renewed interest in qualitative and narrative approaches (or the qualitative and narrative "turns" in research as they were called in the early 1990s) established protocols, procedures, and a language that, by now, is repeated habitually. Perhaps it time now to look elsewhere, (to culture, to the arts, to literature, etc. both past and present), to find fresh inspiration and vocabulary to support our new emotive efforts. For example, I often recommend that academics read the contemporary fiction of conceptual novelists such as Michael Kimball in order to unleash creativity and a new, uncluttered way of using language in their academic writing. Should we continue to routinely repeat what are by now shop-worn words in our academic outpourings such as *rigour, robust, thick, embodied,* and *evocative* to support (or deny?) our emotive tendencies? Most of those words have been repeated *ad infinitum* for more than 20 years now, degenerating into no more than code words signalling membership in a particular scholarly community. They have become words without force.

The first step in reporting emotive encounters in research, therefore, is moving away from concepts that have evolved from measurement—terms like "empathic validity," "reliability," and so on. Rejecting the use of statistical language to describe the emotional components of our labours is key to communicating an understanding of the Hows and Whys of the human condition. The second step is to find our own individual language (a descriptive and poetic one?) that does not mimic the status quo language of a specific scholarship simply because of our insecurities or longing to fit in with a particular club or movement.

Acknowledging the emotive connections in our work doesn't mean simply producing wishy-washy, touchy-feely texts either. In fact, Neo Emotivism insists upon tougher, more resilient, profoundly compassionate yet hard-hitting productions. This is accomplished through the creative use of language—textural/visual/physical—or some new mode of communication that we haven't even attempted yet.

Feel free to use this outline of the workshop for your own purposes. I will end by saying that it was also an experiment in organising and facilitating a two-day workshop without funding or charging for participation. The Media School's Narrative Group generously covered costs for coffees and teas for each day. I hustled room bookings, having to locate each day's activities in a different location. We paid for our own lunches and on the second day ate in a local Italian restaurant—all at one long table, having by then shared all of our stories as a group. There was no "teaching," no lecturing, and no "expert," no flipchart and no PowerPoint.

We met up as a group of mostly strangers for our arts-led adventure and, after two days, left as friends. Bourriaud would be proud.

References

Bourriaud, N. (2002; English version). *Relational Aesthetics.* Dijon: Les Presses du Reel.

Denzin, N. (2000). Narrative's Moment. In M. Andrews (Ed.), *Explorations in Sociology, Psychology, and Cultural Studies* (pp. xi–xii). New Brunswick: Transaction Publishers.

Hewitt, A., & Jordan, M. (2004). Talking Up the Social. *Press Corps 2004 Liverpool Biennial.* Retrieved from http://www.everydayarchive.org/press-corps/writing/hewitt-jordan.pdf

Jones, K. (2001). I Can Remember the Night (Video). *The Sociological Cinema.* Retrieved from http://www.thesociologicalcinema.com/videos/i-can-remember-the-night

Jones, K. (2007). How Did I Get to Princess Margaret? (And How Did I Get Her to the World Wide Web?) *Forum: Qualitative Social Research Special Issue on Virtual Ethnography, 8*(3). Retrieved from http://www.qualitative research.net/index.php/fqs/rt/suppFiles/281/617/

Jones, K. (2012). Connecting Research with Communities Through Performative Social Science. *The Qualitative Report, 17*(18), 1–8. Retrieved from http://www.nova.edu/ssss/QR/QR17/jones.pdf

Jones, K. (2014). (The Grand Theory of) Neo Emotivism. *Social Science Space*. Retrieved from http://www.socialsciencespace.com/2014/10/the-grand-theory-of-neo-emotivism/

Kimball, M. (2013). *Michael Kimball Writes Your Life Story (on a Postcard)*. Atlanta: Publishing Genius Press.

Maysles, A., & Maysles, D. (1975). *Grey Gardens* (Documentary Film). (Producer). New York: Portrait Films, (Directors) Albert and David Maysles.

Mills, C. W. (1959). *The Sociological Imagination*. Oxford: Oxford University Press.

Mills, C. W. (2000). *The Sociological Imagination*. Oxford: Oxford University Press.

4

Abductive Thematic Network Analysis (ATNA) Using ATLAS-ti

Komalsingh Rambaree

Introduction

Qualitative research methods rest on three types of reasoning—inductivism, deductivism, and abductivism. Through inductivism, qualitative data are organised and structured for theorisation based on gathered evidence. This type of reasoning is commonly used in the Glaserian grounded theory approach. In deductivism, theory precedes observation. Through deductivism, a theoretical framework is developed and used by researchers to gather evidence from the field. Deduction is therefore a logical form of reasoning which is based on making deduction regarding a selected theory by gathering qualitative evidence. The deductive approach within qualitative research is more commonly used when researchers undertake, for example, content analysis or discourse analysis.

This chapter focuses on the third type of logical reasoning which is referred as abductivism. The use of abductive reasoning in research was introduced and advocated by Charles Sanders Peirce in the 1950s as

K. Rambaree (✉)
University of Gävle, Gävle, Sweden

L. Moutinho, M. Sokele (eds.), *Innovative Research Methodologies in Management*,
https://doi.org/10.1007/978-3-319-64394-6_4

discovery, relating the term to the process of providing scientific explanation based on the newly found facts (as referred in Levin-Rozalis 2004). For Pierce, abduction is part of a broader pragmatic methodological process of inquiry for forming hypotheses or suggestions, through the use of back-and-forth reasoning between theory and empirical evidence (Dubois and Gadde 2002; Morgan 2007; Feilzer 2010). In particular, abduction is about discovering new concepts, ideas, or explanations by finding events, which lack theorisation within current discourses (Thornberg 2012).

Abductive reasoning is commonly used within mixed-methods research. Researchers using mixed-methods reject the 'purist' notion that positivist and constructivist ontologies are irreconcilable (Cupchik 2001). Instead, they promote the combination of qualitative and quantitative methodologies as a pragmatic and an efficient way of getting the benefits from both approaches. For a pragmatist researcher, the most important question is to find out what s/he wants to know (Hanson, as quoted in Feilzer, p. 8). Pragmatism is based on abductive reasoning which acknowledges 'uncertainty' in what is found as evidence. Thus, it is important to note that any knowledge produced through pragmatic research using abductivism is considered as being relative and not absolute. In other words, abductivism is firmly rooted on the belief that a theory (or whatever has been theorised) is provisional, tentative, and in need of confirmation (Cooper and Meadows 2016).

Abductive Theory of Method

Over the years, scientists have developed different strategies of using abductive reasoning in research, such as backward reasoning, probabilistic evaluation of explanations, eliminations of implausible explanations, testing abducted hypotheses by further empirical investigations, introduction of new concepts or theoretical models, and analogical reasoning based on conceptual abstraction (Thornberg 2012). Based on Pierce's idea on abduction, Gilbert Harman therefore introduced *Inference to the Best Explanation* model within abductive reasoning, which made abduction an appealing topic for philosophy (as referred in Paavola 2004, 2015). The governing idea of *Inference to the Best Explanation* is that explanatory considerations are a guide to inference, that scientists infer from the available evidence to the hypothesis which would, if correct, best explain that evidence (Lipton 2000).

In a similar way, Haig (2005a, b, 2008a, b) and Haig and Evers (2016) argue that scientific knowledge in social and behavioural research is also based on abduction as a way of reasoning from factual premises to explanatory inferences. Using such a type of reasoning, social and behavioural researchers try to associate gathered data with ideas for logical explanation. Abduction is therefore used not as means of drawing conclusion but, rather, as a logical means of inferencing (Reichertz 2009). In other words, abduction is used as a process towards reaching conclusions through the use of analogical reasoning between existing knowledge and the discovery that needs to be explained. Haig (2008a) provides a sound description of *Inference to the Best Explanation* as given below:

> F1, F2, … are surprising empirical facts.
> Hypothesis H explains F1, F2, …
> No other hypothesis can explain F1, F2, … as well as H does.
> Therefore, H is accepted as the best explanation. (p. 1015)

Using such pragmatic way of reasoning, Haig (2005a) proposes Abductive Theory of Method (ATOM) which can be considered to be broader than both the inductive and hypothetico-deductive accounts of scientific method. As mentioned earlier, pragmatism in research offers scientists the flexibility of providing understanding on social phenomena through an integrated methodology with back-and-forth movement between theory and evidence (Morgan 2007; Feilzer 2010). This flexibility is sometime very important in research, as it provides scientists with the freedom to approach research questions in a non-coercive manner as it is in deductive and inductive approach (Haig 2005a). Haig (2008b: 1020) therefore proposes the following steps with regard to ATOM:

1. *Detection of Phenomena*: Sets of data are analysed in order to detect robust empirical regularities, or phenomena.
2. *Theory Generation*: Once detected, these phenomena are explained by abductively inferring the existence of underlying causal mechanisms.

(Abductive inference involves reasoning from phenomena, understood as presumed effects, to their theoretical explanation in terms of underlying causal mechanisms.)
3. *Theory Development*: Upon positive judgements of the initial plausibility of these explanatory theories, attempts are made to elaborate on the nature of the causal mechanisms in question. (This is done by constructing plausible models of those mechanisms by analogy with relevant ideas in domains that are well understood.)
4. *Theory Appraisal*: When the theories are well developed, they are assessed against their rivals with respect to their explanatory goodness. (This assessment involves making judgements of the best of competing explanations.)

Haig's (2008b) description is somewhat similar to some form of Grounded Theory Analysis (GTA). However, Haig (2005a: 386) argues that GTA can be regarded as an abductive method in the sense that it explains the qualitative data patterns from which theories are derived; however, 'it does not confine itself to existential abduction, and it imposes weaker constraints on the abductive reasoning permitted by the researcher than does exploratory factor analysis'. For instance, Strauss and Corbin's GTA has been criticised for being overly prescriptive, lacking explanatory power, and minimising the influence of existing theories and researchers' biases (Hodkinson 2016). Moreover, Haig and Evers (2016) argue that Glaser and Strauss' formulation of GTA does not make systematic use of the philosophy of science. While grounded theory still offers useful tools for the organisation of qualitative research, it is only in relation to abduction with exploratory factor analysis that theory construction becomes meaningful (Timmermans and Tavory 2012: 169). However, Haig and Evers (2016) conclude that ATOM does not replace grounded theory method but becomes an additional option, as the demand for methodological pluralism ensures a place for both in the scientists' toolkit.

Furthermore, Haig's (2005b, 2015) ATOM goes beyond theory generation by using exploratory factor analysis for making the appraisal of the developed theory providing innovative ideas by providing guides on using a generated theory as an analytical framework. The theory appraisal phase therefore becomes a continuation within the research process where developed hypothesis through abductive reasoning are tested (deductive manner). For instance, Haig (2005b: 326) argues that, 'exploratory factor analysis functions as a data analytic method that contributes to the detection of empirical regularities', whereas 'confirmatory factory analysis can contribute to the goal of empirical adequacy in the subsequent hypothetico-deductive appraisal of common causal theories'.

Abductive Thematic Network Analysis

Thematic analysis is the process of identifying patterns in seemingly random information found in the collected data (Boyatzis 1998). In thematic analysis researchers organise segments of gathered data into 'themes', a process which is facilitated by coding. Braun and Clarke (2013) define a code as a word or a brief phrase that captures the essence of why a researcher think that a particular bit of data may be useful. Seal (2016: 452) distinguishes that, 'a code is a descriptor of a data segment that assigns meaning, whereas a theme is a theoretical construct that explains similarities or variations across codes'. Codes and themes are essential components of the data reduction process within thematic data analysis.

Braun and Clarke (2013) describe thematic analysis as a process of identifying and reporting patterns from the gathered evidence in a descriptive manner using back-and-forth movement between gathered evidence and the thematic description. According to Bazeley (2013), effective thematic analysis requires using gathered qualitative data to build a comprehensive, contextualised, and integrated understanding or theoretical model of what has been found, with an argument drawn from empirical evidence based across the data. Over the years, researchers using thematic analysis have come with creative and innovative techniques in the way to identify, organise, and present themes (Morse 2011; Vaismoradi et al. 2016).

For instance, Attride-Stirling (2001) presents Thematic Network Analysis (TNA) as a creative and systematic way of identifying and reporting themes in qualitative research. Within TNA, researchers study the data to identify themes and then develop graphical representation/s of the linkages between the themes. According to Attride-Stirling (2001), the networks between the themes are merely a graphical tool to organise themes and show the interconnectivity between them in order to facilitate the subsequent analysis. TNA is a flexible method of qualitative data analysis which can be data driven or theory driven or even a combination of both (Rambaree and Faxelid 2013). By having a combination of both data- and theory-driven TNA, researchers can enhance the rigour in the analysis process by dealing with biases of self and others. TNA therefore requires researchers to, in priori, think about their qualitative data in a critical and analytical way (Seal 2016). However, the central part of TNA is where researchers relate the principal themes and patterns that emerged in the analysis to the original questions and then propose explanations to the questions (Attride-Stirling 2001).

Innovation and creative in qualitative research has often come in the form of a new analytical tool (Taylor and Coffey 2009; Morse 2011; Vaismoradi et al. 2016). Given that TNA requires back-and-forth movement and it is possible to apply an abductive theory of method within such an analysis process, it becomes appealing to combine the technique of thematic network analysis with the abductive reasoning as an innovative way of analysing qualitative data (Rambaree and Faxelid 2013). ATNA becomes an innovative and creative way of undertaking qualitative data analysis with a combination of ideas borrowed from Haig's (2005a) Abductive Theory of Method (ATOM) and Attride-Stirling's (2001) Thematic Network Analysis (TNA). ATNA can be broadly defined as an abductive way of reasoning in looking at and explaining the linkages between the emerging themes from the analysis of the gathered qualitative data (Rambaree and Faxelid 2013). However, it is important to point out that not all recommendations of Haig's (2005a) ATOM and Attride-Stirling's TNA have been strictly followed. The steps proposed in this chapter are the author's own recommendations for undertaking ATNA as a pragmatic way of analysing qualitative data using some of the recommendations made by Haig (2005a) and Attride-Stirling (2001) (Fig. 4.1).

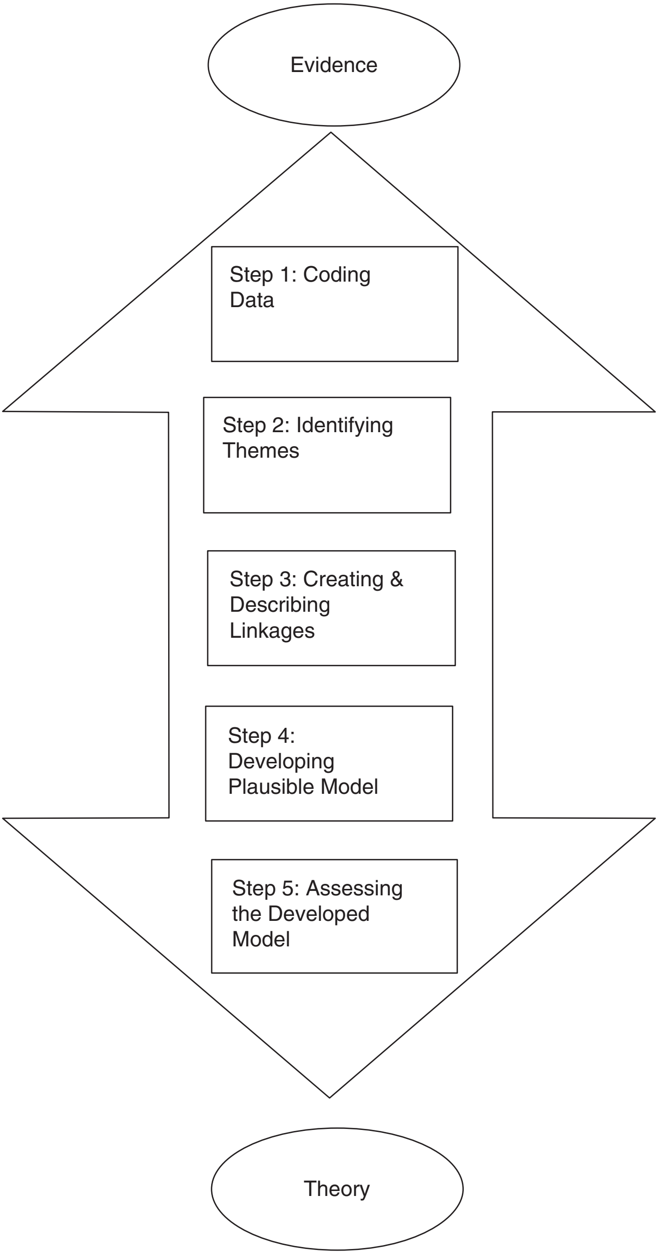

Fig. 4.1 Process and steps in ATNA (Source: Adapted from Rambaree and Faxelid (2013: 176))

ATNA in Researching International Social Work Practicum: An Exemplar

In this part of the chapter, a data set is used as an example to demonstrate, in a simplified manner, how ATNA can be applied in qualitative research.[1] The aim of this exercise is to develop a model that can help in theorising how social work students describe their cross-cultural experiences during their field practice in a foreign country.[2] The research questions are: (a) What kind of cross-cultural observations social work students make? (b) How do the students act/react to cross-cultural differences?

The data set considered here is the practicum placement reports of some Swedish social work students[3] from a particular Swedish University. The selected social work students, as cases, have been on field practice within a social work organisation for about two months outside Sweden, during the period 2010–2014. In their practicum placement reports, students write about their cross-cultural experiences. A total number of 22 students' reports are selected and analysed, for this chapter, using ATLAS-ti v.7.5.13 software.[4]

Steps for setting up a project (New Hermeneutic Unit) in ATLAS-ti and adding your materials as Primary Documents (P-Docs) for analysis[5] (refer to Fig. 4.2 as an example).

- Right click on Project (far top-left corner).
- Go to Save As, and then give a name to the project (example: Cross-Cultural).
- Click on Project. Then move curser to Add Document(s) and then after click the option, Add Documents to select the source where your materials (P-Docs) are located. Select all materials that you want to assign in ATLAS-ti as P-Docs for analysis.

ATNA Step 1: Coding Data

Coding is much more than just labelling a segment of data. It is a vital building block in qualitative data analysis and a fundamental skill that

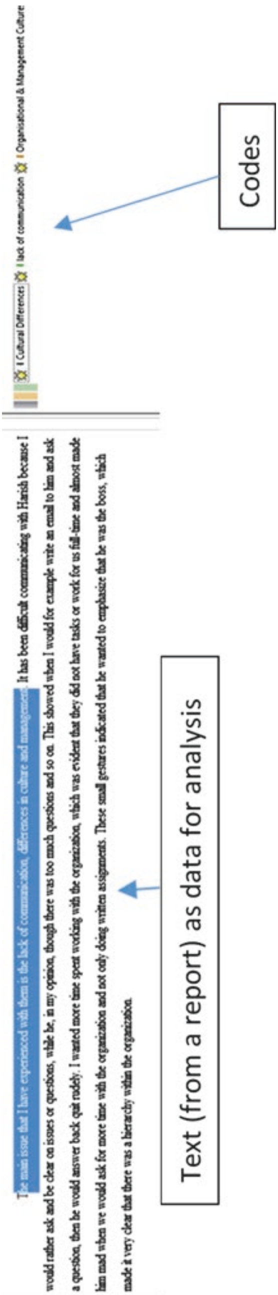

Fig. 4.2 Coding in ATLAS-ti

researchers apply in accessing evidence for testing assumptions and making conclusions (Bazeley 2013; Braun and Clarke 2013). Open coding starts by systematically and meticulously examining the data and identifying appropriate codes from selected segments of empirical data. A 'good code' tries to capture and represent the qualitative richness of the phenomenon under study (Boyatzis 1998; Fereday and Muir-Cochrane 2006). Haig (2015) argues that phenomena are the appropriate source of evidence for the explanatory theories. Braun and Clarke (2013) remind researchers that coding is an organic and evolving process that provides researchers deeper understanding of phenomena being studied through the gathered empirical data.

During the data analysis process, researchers therefore need to revisit the codes and modify, if required, according to latest reflexivity captured within analytical memo. Birks et al. (2008: 69) argue that, 'the very nature of qualitative research requires the researcher to assume a reflexive stance in relation to the research situation, participants and data under study'. Reflexivity basically means making reflections through memos (notes) on the knowledge construction within the whole research process that allows researchers to be aware about biases in analytical process. Analytical reflexive memos assist researchers in making conceptual leaps from raw data to those abstractions, such as codes and themes that explain research phenomena in the context in which it is examined (Birks et al. 2008).

In order to start with the open coding process, follow the below given steps in ATLAS-ti (refer to Fig. 4.2).

- Under P-Docs, select the document that you want to start coding.
- Once the document is open, you can select the segment of data (quotation) that you want to code.
- Then, right click to go on Coding to choose 'Enter Code Name'.
- Give a name to your code (example: 'Cultural Differences' is chosen as a code).

As mentioned earlier in this chapter, codes provide the basis for developing themes within the process of thematic data analysis. While deciding a name for the code it is important to bear in mind the research questions and the overarching goal of the research project. The code name needs to direct the data analysis towards organising themes that could help in reaching the overarching goal of the research project. If required, researchers can use the 'Edit Comment' function in ATLAS-ti to make reflexive notes on a code (as shown in Fig. 4.2). The steps are:

- Right click on the code.
- Then click on 'Edit Comment'.
- In the panel, make notes.
- Finally, go to 'Comment' section within the panel, and click Save.

ATNA Step 2: Identifying Themes

Once the gathered data are coded,[6] the following step is to identify themes. A theme is a fuzzy concept based on organisation of codes that qualitative researchers use to characterise the phenomena being studied (Ryan and Bernard 2003; Fereday and Muir-Cochrane 2006). It emerges from a segment of data through coding, categorising, and analytical reflection (Saldana, as referred in Bazeley 2013). A theme has therefore a high degree of generality that pull ideas together regarding the subject of inquiry (Vaismoradi et al. 2016). To identify a theme, researchers need to study the codes in relation to the respective associated quotations and context, and try to make analytical reflective memos on selection of codes that can be pulled together as a concept in providing explanation towards answering the set research objectives/question(s). It is through a systematic study of the codes that researchers extract the salient, common, or significant themes in the coded part of the empirical data (Attride-Stirling 2001).

Within ATLAS-ti, a theme can be identified by regrouping codes that show pattern towards answering the research question. Under ATLAS-ti there is no such thing labelled as theme, but only codes that could be assigned as a 'Family' or using 'Merged Codes' as 'Theme'. Therefore, the 'The Family Manager' function can be used for this particular purpose. It is worth noting that a theme (referred as Family under ATLAS-ti) can share codes with other theme/s. Therefore, a researcher can use a code under different themes, if required, to start mapping pattern in the gathered data. The steps to follow in ATLAS-ti for creating a theme are as given below (Fig. 4.3):

- Click on 'Codes'.
- Bring curser to 'Families', and then click on 'Open Family Manager'.
- When the panel is open, click on 'Families' and then click on 'New Family'.
- Assign a name—which will become a 'Theme'.

Step 3: Creating and Describing Linkages

At this stage, researchers create and describe linkages between the themes using abductive reasoning. Here, the emphasis is on considering how the different themes intersect to create a constellation (network) in expanding the observed pattern/s towards answering the research question/s (Bazeley 2013). It is important to note that themes can be organised in different rank order. Attride-Stirling (2001: 388) presents three levels of themes, which are: (i) Basic Themes, lowest-order premises that are evident in the text, but on their own they say very little about the text or group of texts as a whole; (ii) Organising Themes, categories of basic themes grouped together to summarise more abstract principles and simultaneously group the main ideas proposed by several basic themes to dissect the main assumptions underlying a broader theme that is especially significant in the texts as a whole; and (iii) Global Themes, superordinate themes encapsulating the principal metaphors in the text as a whole and indicate what the texts as a whole are about within the context of a given analysis.

The central aspect in TNA is the analytical reflexivity use in linking the themes. The linkages create flow in describing the observe phenomena within the gathered data. However, Attride-Stirling (2001: 393) points out that, 'the networks are only a tool in analysis, not the analysis itself'. Researchers therefore need to go deeper in deconstructing the gathered data by further exploring the themes, the linkages, and the emerging patterns to provide interpretative explanation on the phenomena being

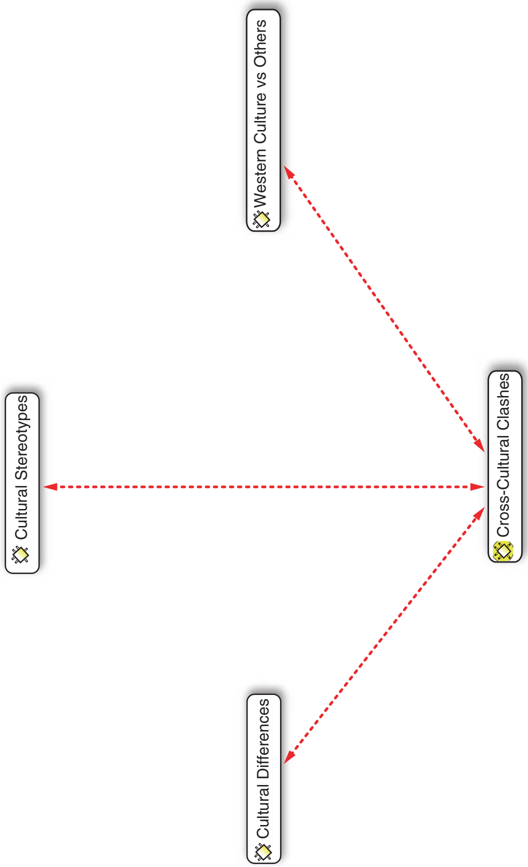

Fig. 4.3 Identifying theme in ATLAS-ti

observed with reference to the gathered evidence. This particular task requires the beginning of abductive reasoning (through making inferences) with a back-and-forth movement between themes and gathered evidence (data) to identify what is emerging as knowledge in answering the research question/s. This task is mainly done through creating and saving analytical reflective memos that are specifically dedicated to the linkages between the 'Themes'.

Within ATLAS-ti, the following steps can be followed for creating and describing linkages between the themes:

- Click on 'Networks'.
- Then go to 'Network View Manager'.
- Click on 'Create a New Item' (far left corner, there is a folder Icon).
- Give a name to your network (in the exemplar, it is labelled as 'Cross-Cultural Network').
- Click on it to open the networking panel.
- Then, click on Codes (not on the top row, but on row level three from the top next to P-Docs).
- Drag the 'Families' (Themes) that you want to explore the link in between in the network panel (from the left-hand side of the panel, not from the list of codes).
- Right click on each of the 'Themes' that have been dragged into the 'Network' panel; select 'Import Neighbours'.
- Then after, click on the 'Colour' Icon (round with multiple colours) on the 'Network' panel to select 'Colour by Density & Groundedness'.

The colour differences shown density (number of links to other codes and memos) and groundedness (number of links to quotations). In addition to pre-existing links (red ones, from the creation of themes from codes), new linkages in between the 'Themes' and 'Codes' can be drawn, if required (black one). Within the new links relationship between two 'Themes' and in between a 'Theme' and a 'code' can be assigned.[7] In addition, analytical memos on each of the themes that are being analysed can be created by clicking on them and selecting 'Edit Comment' to write up memos (notes), as explained and shown previously (in Figs. 4.2 and 4.4).

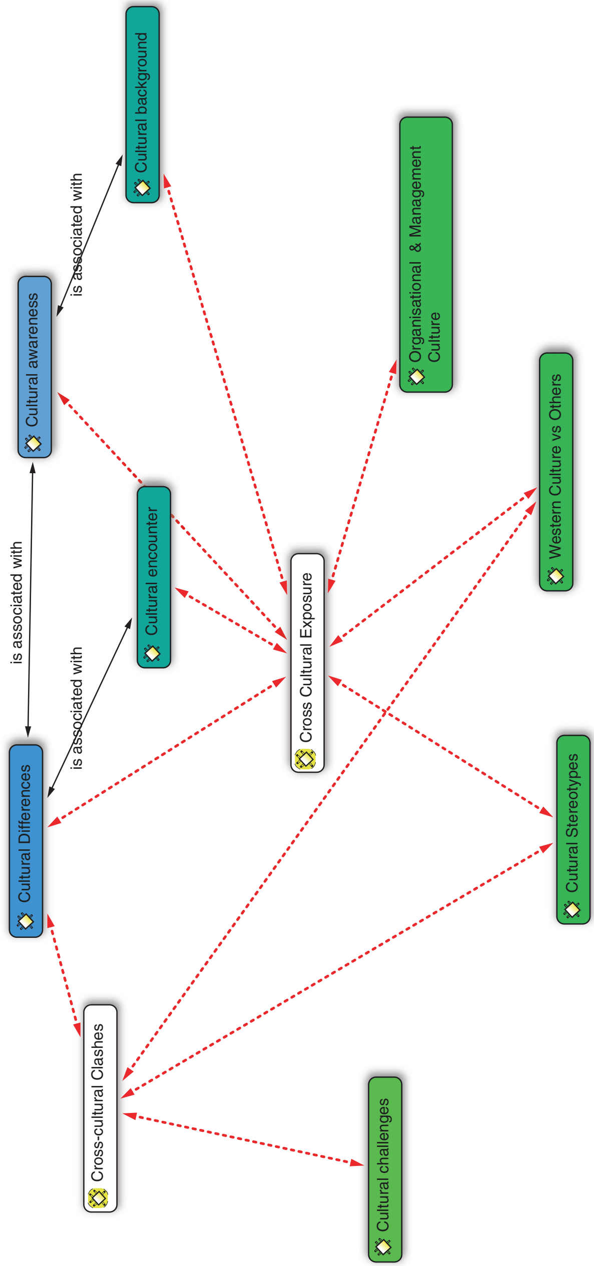

Fig. 4.4 Creating and describing linkages between themes with ATLAS-ti

Step 4: Developing Plausible Model

At this stage of analysis, the themes within a network are arranged in the form of a plausible model that could be used to facilitate explanation on the patterns being observed in reaching answer/s to the set research questions. Here, a model can be broadly defined as a graphical presentation that facilitates explanation on inter-linkages (e.g. cause-effect types) between themes as a system. In essence, modelling becomes a crucial step towards the theorisation of an observed phenomena. In a similar vein, Haig (2005a) posits that the construction of appropriate analogical models serves to assess the plausibility of our expanded understanding regarding the phenomena being studied.

For this particular task, Haig's (2005a, b) guidelines on analogical modelling, which is central in abduction theory of method, are followed. An analogical model of an unknown subject or causal mechanism is based on the pragmatic strategy of conceiving it in terms of what is already known, for instance from information available in discourses (Haig and Evers 2016). In the exemplar research, an extensive literature review of existing discourses (mostly journal articles and book chapters) related to cross-cultural learning during field practice was carried out using ATLAS-ti, to identify what is already known. For undertaking analogical modelling, Kolb's (1984) model of 'Experiential Learning' (refer to Fig. 4.5) was identified and used as an analogical model for theorising the known with respect to what is already known (extensive from literature review).

Kolb (1984) presents a theoretical model on experiential learning with four stages linked in a cycle as shown in Fig. 4.5. Kolb's (1984) model describes on how different people learn by integrating their practical experiences with reflection. According to his model, learning process from practice exposure takes place through four distinct phases: (1) feeling, through being involved in concrete experience; (2) watching, making reflective observation; (3) thinking, abstract construction of concepts; and (4) doing, making active experiment. Kolb's (1984) model has therefore relevant ideas in field of experiential learning that are well understood.

According to this model, it is therefore argued that social work cross-cultural learning takes place by direct experience of the social work

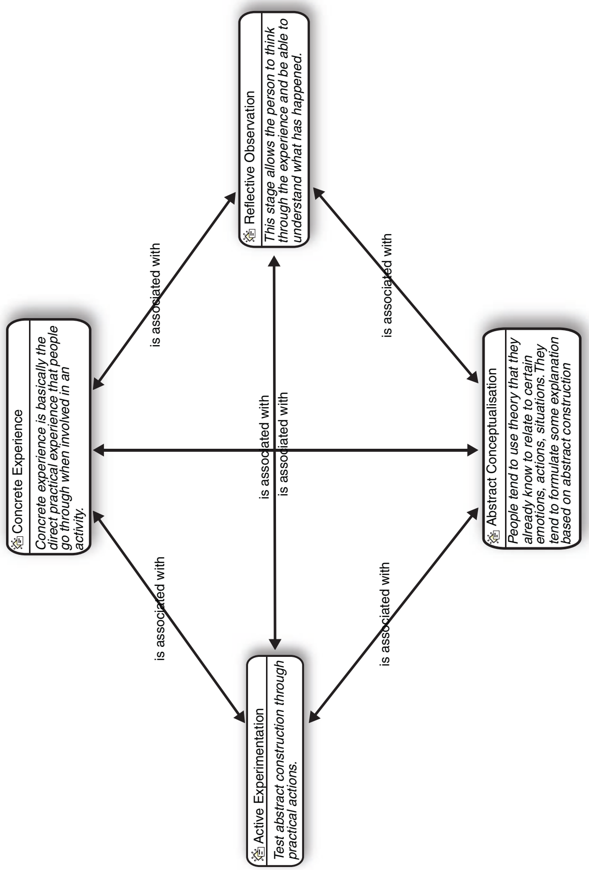

Fig. 4.5 Experiential learning model (for analogical use) (Source: Adapted from Kolb (1984))

practice in different cultural contexts, by reflecting on the cultural experiences during the practice, by conceptualising and thinking abstractly about different cultures, and by being active in putting cross-cultural learning in social work practice (Koob and Funk 2002). But, Kolb's model is limited in the sense that it does answer the set research questions: (a) What kind of cross-cultural observations social work students make? (b) How do the students act/react to cross-cultural differences? For this particular reason, Kolb's model is as an analogical model to develop a plausible model that can provide the basis for theorisation about how social work students describe their cross-cultural experiences during their field practice in a foreign country.

In developing the plausible models, the analogical reasoning used can be written and saved using 'Edit Comment' feature in ATLAS-ti.

- Download and organise literatures relevant to the subject matter for reviewing the analogical model (example: Kolb's Experiential Learning) in a folder in your computer.
- Upload the literatures from your computer to ATLAS-ti and start coding and organising thematic linkages (follow same steps 1–3 using literatures instead of data empirical data from the field).
- Using the developed linkages from both the literature review (including a selected analogical model) and the analysis of the empirical data from the field to create models that could support theorisation of new findings. Steps in ATLAS-ti are:
- Click on 'Networks'.
- Then click on 'New Network View'.
- Give a name (example: Theoretical Model on Cross-Cultural Experiences).
- When the network panel is open, drag Themes & Codes (from both the analogical model and the empirical data) to develop a nascent (plausible) model to support theorisation, related to the research aim.
- Organise the themes in an explanatory pattern, for example, by renaming central themes in an orderly manner, so as to facilitate the theorisation in a structured way (using analogical modelling). For instance, Stage 1: Cross-Cultural Experience, Stage 2: Cultural Observations, and so on.

See an example of analogical modelling in Fig. 4.6.

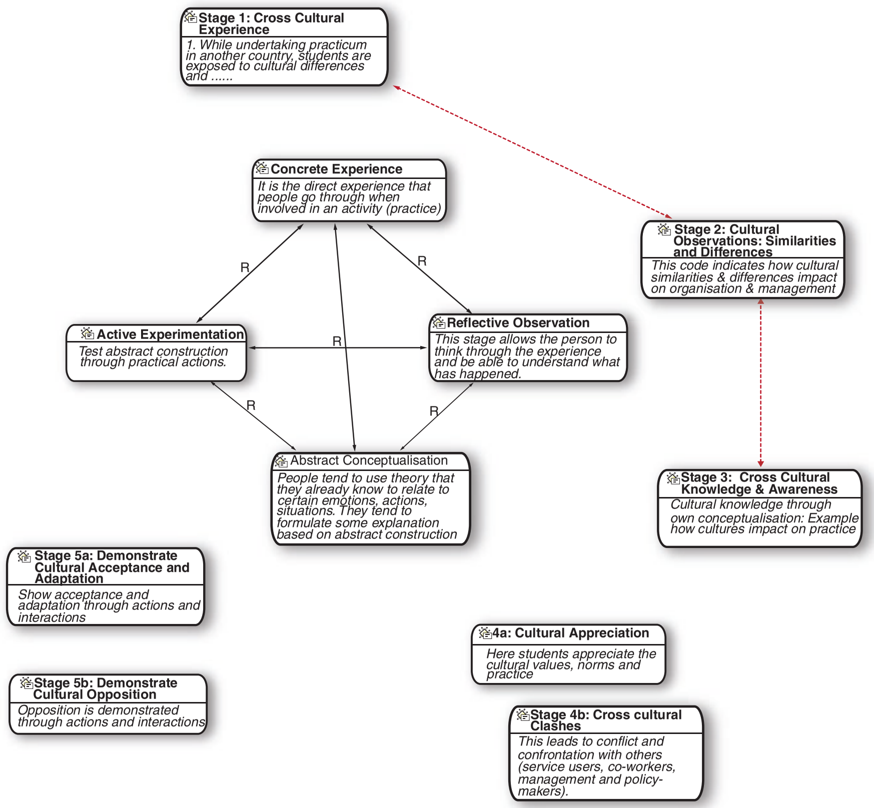

Fig. 4.6 Developing a nascent model using analogical reasoning (Mapped on Kolb's (1984) Experiential Learning)

Step 5: Assessing the Developed Model

When the plausible model based on the empirical data is completed (as shown in Fig. 4.7), researchers need to assess it for 'explanatory goodness' in comparison with existing model/s or explanation/s (from literature review) (Haig 2008a, 2015. The plausible model is therefore assessed for being theoretically elegant, coherent, and scientific (Lipton 2000). At this particular stage, two related techniques advised by Haig (2005a) become central within the data analysis process, which are termed as 'inference to the best explanation' and 'theory explanation coherence'.

Haig and Evers (2016: 85) argue:

> Inference to the best explanation is founded on the belief that much of what we know about the world is based on considerations of explanatory merit. Being concerned with explanatory reasoning, inference to best explanation is a form of abduction. It involves accepting a theory when it is judged to provide a better explanation of the evidence than its rivals do.

Haig and Evers (2016: 85) also point that the determination of the explanatory coherence of a theory is made in terms of three criteria, which are:

(i) *Consilience* (explanatory breadth): by explaining a greater range of facts
(ii) *Simplicity*: by making fewer special or ad hoc assumptions
(iii) *Analogy*: by supporting itself through analogy to theory/ies that scientists already find credible

Within ATLAS-ti the memo functions can be used to capture the 'inference to the best explanation' and 'theory explanation coherence'. Researchers can open new memo and write up qualitative inferential analytical and theoretical explanatory memos. Each created memos can be saved and linked with the themes or the relations between the themes in

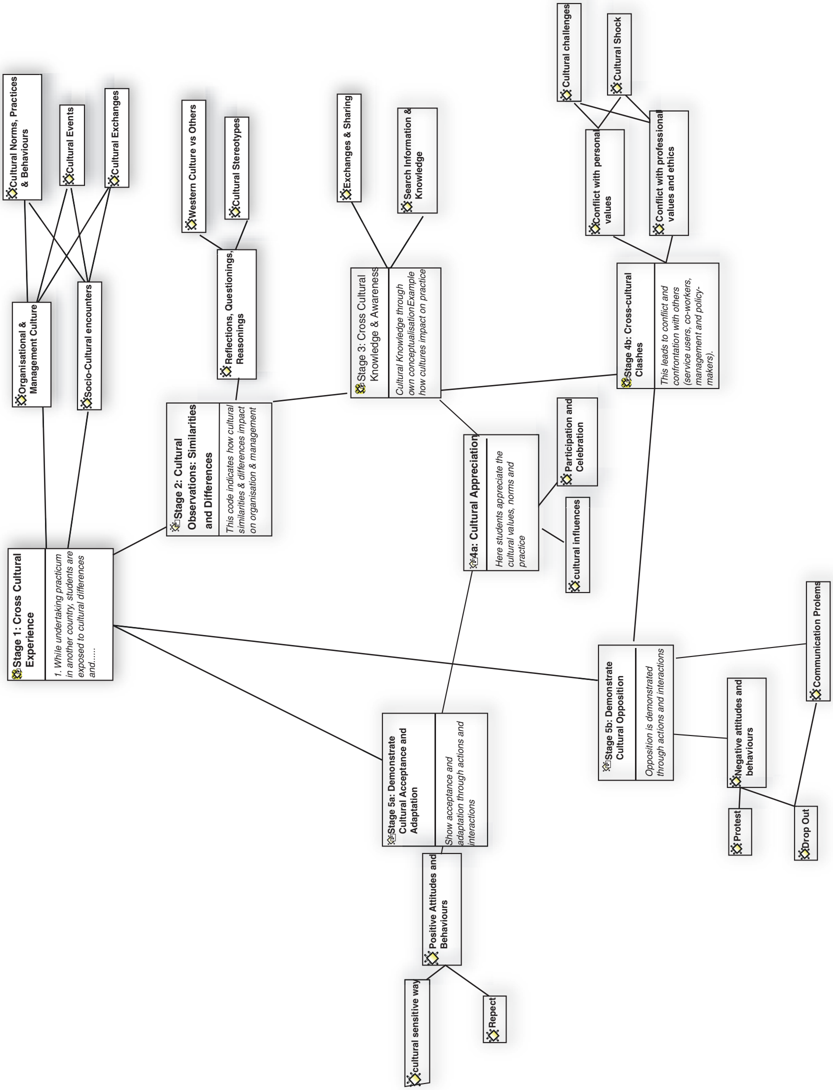

Fig. 4.7 Developed model on cross-cultural experiences

the network view panel of ATLAS-ti. For the final stage of ATNA, the steps within ATLAS-ti are:

- Click on 'Network' and then 'Network View Manager'.
- Select the network that has been developed as a model for theoretical explanation (in the exemplar study, it is labelled as 'Theoretical Model Cross-Cultural').
- Once the model (network of themes) has opened, select the themes/codes that have been used for analogical modelling by 'Right Clicking' on each of them and then selecting 'Remove from View'. This will leave only the nascent (plausible) theoretical model.
- Expand the nascent theoretical model with memos and additional codes or themes (if necessary).
- To create memos, click on 'Memo', and then click on 'Create Free Memo'. Give a name to the memo (example is the exemplar study: Analytical Memo: Cultural Competence Explanation).

In this way, a plausible model is developed to theorise how social work students describe their cross-cultural experiences during their field practice in a foreign country. The plausible model is based on analogical reasoning from a well-known theory (Kolb's 1984 model) on experiential learning. This nascent plausible model develops through the analysis of gathered empirical evidence. It helps to answer the set research questions through inference to best explanation as compared to the analogical theory. The theorisation is done in a logical and simple manner but with coherence and consilience.

Limitations and Conclusion

This chapter needs to be considered with some limitations. Firstly, it is limited to analysis in textual format. However, materials for analysis in ATLAS-ti can also be in audio and video format. Somewhat similar steps, as shown in this chapter, can be followed while using materials in other formats—such as audio/video. Secondly, the whole explanatory process has been simplified in this chapter, so that readers can have a better understanding of the application of a theoretical model of qualitative analysis. Finally, the chapter has focused on the qualitative aspects of abductive theory of methods. In particular, ATNA, as a qualitative data analysis methodological approach, can be used as a basis for quantitative analysis, such as exploratory factor analysis.[8]

To conclude, ATNA can be considered as a core component within mixed-methodology research. It provides a pragmatic and logical way of reasoning, organising, and presenting qualitative data analysis. ATNA helps researchers to structure qualitative data analysis through stepwise application of abductive theory of method. Such approach allows researchers to go into deep details in exploring and working with qualitative data. It therefore allows researchers to theorise their findings through the development of conceptual/thematic model/s, which can be tested and validated through further research. It therefore brings rigour to qualitative data analysis.

Notes

1. For ethical reason some parts, that are not relevant to this chapter, are hidden in some of the visual demonstrations.
2. In a country other than where they are born and/or studying.
3. Swedish students—born and raised in Sweden.
4. For more detailed explanation on using the software, refer to Friese (2013, 2014).
5. Read more about setting up project from ATLAS-ti free manual available at http://atlasti.com/wp-content/uploads/2014/05/atlasti_v7_manual_201312.pdf?q=/uploads/media/atlasti_v7_manual_201312.pdf.
6. Coding from list is also possible, after open coding has been done. Refer to ATLAS-ti manual for more details on coding.
7. Read more about assigning relationship in network from the ATLAS-ti manual.
8. Refer to Haig and Evers (2016), for more information on exploratory factor analysis.

References

Attride-Stirling, J. (2001). Thematic Networks: An Analytic Tool for Qualitative Research. *Qualitative Research, 1*(3), 385–405.

Bazeley, P. (2013). *Qualitative Data Analysis*. London: SAGE.

Birks, M., Chapman, Y., & Francis, K. (2008). Memoing In Qualitative Research: Probing Data and Processes. *Journal of Research in Nursing, 13*(1), 68–75. https://doi.org/10.1177/1744987107081254.

Boyatzis, R. (1998). *Transforming Qualitative Information: Thematic Analysis and Code Development*. Thousand Oaks: SAGE.

Braun, V., & Clarke, V. (2013). *Successful Qualitative Research: A Practical Guide for Beginners*. London: SAGE.

Cooper, G., & Meadows, R. (2016). Conceptualising Social Life. In N. Gilbert & P. Stoneman (Eds.), *Researching Social Life* (pp. 10–24). London: SAGE.

Cupchik, G. (2001). Constructivist Realism: An Ontology that Encompasses Positivist and Constructivist Approaches to the Social Sciences. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research, 2*(1), Article 7. Retrieved February 12, 2011, from http://qualitative-research.net/fqs-texte/1-01/1- 01cupchik-e.htm

Dubois, A., & Gadde, L.-E. (2002). Systematic Combining: An Abductive Approach to Case Research. *Journal of Business Research, 55*(2), 553–560.

Feilzer, M. Y. (2010). Doing Mixed Methods Research Pragmatically—Implications for the Rediscovery of Pragmatism as a Research Paradigm. *Journal of Mixed Methods Research, 4*(1), 6–16.

Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development. *International Journal of Qualitative Methods, 5*(1), Article 7. Retrieved February 12, 2011 from http://www.ualberta.ca/~iiqm/backissues/5_1/html/fereday.htm

Friese, S. (2013). *ATLAS.ti* 7 User Guide and *Reference*: *ATLAS.ti*. Berlin: Scientific Software Development GmbH. Retrieved December 15, 2016, from http://atlasti.com/wpcontent/uploads/2014/05/atlasti_v7_manual_201312.pdf?q=/uploads/media/atlasti_v7_manual_201312.pdf

Friese, S. (2014). *Qualitative DATA analysis with ATLAS.ti* (2nd ed.). Los Angeles: SAGE.

Haig, B. D. (2005a). An Abductive Theory of Scientific Method. *Psychological Methods, 10*, 371–388.

Haig, B. D. (2005b). Exploratory Factor Analysis, Theory Generation, and Scientific Method. *Multivariate Behavioral Research, 40*(3), 303–329.

Haig, B. D. (2008a). Scientific Method, Abduction, and Clinical Reasoning. *Journal of Clinical Psychology, 64*, 1013–1018.

Haig, B. D. (2008b). Précis of "An Abductive Theory of Scientific Method". *Journal of Clinical Psychology, 64*, 1019–1022.

Haig, B. D. (2015). Commentary: Exploratory Data Analysis. *Frontiers in Psychology, 6*(1), 1247.

Haig, B. D., & Evers, C. W. (2016). *Realist Inquiry in Social Science*. New York: SAGE.

Hodkinson, P. (2016). Ground Theory and Inductive Research. In N. Gilbert & P. Stoneman (Eds.), *Researching Social Life* (pp. 98–115). London: SAGE.

Kolb, D. A. (1984). *Experiential Learning: Experience as the Source of Learning and Development* (Vol. 1). Englewood Cliffs: Prentice-Hall.

Koob, J. J., & Funk, J. (2002). Kolb's Learning Style Inventory: Issues of Reliability and Validity. *Research on Social Work Practice, 12*(2), 293–308.

Levin-Rozalis, M. (2004). Searching for the Unknowable: A Process of Detection—Abductive Research Generated by Projective Techniques. *International Journal of Qualitative Methods, 3*(2), Article 1. Retrieved February 12, 2011, from http://www.ualberta.ca/~iiqm/backissues/3_2/pdf/rozalis.pdf

Lipton, P. (2000). Inference to the Best Explanation. In W. H. Newton-Smith (Ed.), *A Companion to the Philosophy of Science* (pp. 184–193). London: Blackwell.

Morgan, D. L. (2007). Paradigms Lost and Pragmatism Regained. *Journal of Mixed Methods Research, 1*, 48–76.

Morse, J. (2011). Molding Qualitative Health Research. *Qualitative Health Research, 21*(8), 1019–1021.

Paavola, S. (2004). Abduction as a Logic and Methodology of Discovery: The Importance of Strategies. *Foundations of Science, 9*(3), 267–283.

Paavola, S. (2015). Deweyan Approaches to Abduction? In U. Zackariasson (Ed.), *Action, Belief and Inquiry—Pragmatist Perspectives on Science, Society and Religion* (pp. 230–249). Nordic Studies in Pragmatism 3. Helsinki: Nordic Pragmatism Network.

Rambaree, K., & Faxelid, E. (2013). Considering Abductive Thematic Network Analysis with ATLAS.ti 6.2. In N. Sappleton (Ed.), *Advancing Research Methods with New Media Technologies* (pp. 170–186). Hershey: IGI Global.

Reichertz, J. (2009). Abduction: The Logic of Discovery of Grounded Theory. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research, 11*(1), Art. 13. Retrieved February 12, 2011, from http://nbn-resolving.de/urn:nbn:de:0114-fqs1001135

Ryan, G., & Bernard, H. R. (2003). Techniques to Identify Themes. *Field Methods, 15*(1), 85–109.

Seal, A. (2016). Thematic Analysis. In N. Gilbert & P. Stoneman (Eds.), *Researching Social Life* (pp. 444–458). London: SAGE.

Taylor, C., & Coffey, A. (2009). Editorial—Special Issue: Qualitative Research and Methodological Innovation. *Qualitative Research, 9*(5), 523–526.

Thornberg, R. (2012). Informed Grounded Theory. *Scandinavian Journal of Educational Research, 56*(3), 243–259.

Timmermans, S., & Tavory, I. (2012). Theory Construction in Qualitative Research: From Grounded Theory to Abductive Analysis. *Sociological Theory, 30*(3), 167–186.

Vaismoradi, M., Jones, J., Turunen, H., & Snelgrove, S. (2016). Theme Development in Qualitative Content Analysis and Thematic Analysis. *Journal of Nursing Education and Practice, 6*(5), 100–110.

5

Diagnostic Measurement

Meghan Sullivan, Hongling Lao, and Jonathan Templin

Analyzing information and resolving problems often start with a diagnosis. But to obtain a diagnosis, data must first be obtained to infer the different mechanisms giving rise to the problem or characteristic. In management research, as indeed in other forms of research, it is not uncommon to confront studies in which classification judgments are the primary concern. With diagnostic measurement, the aim is to identify causes or underlying properties of a problem for the purposes of making classification-based decisions. The decisions are based on a nuanced profile of attributes or skills obtained from observable characteristics of an individual: for example, diagnosing whether a job applicant has the skills required of the job, an employee is meeting satisfactory progress on a task, or whether a firm is using effective strategies to retain clientele. Another common example of diagnoses is in medical fields, in which certain symptoms that have a high probability of occurring together indicate the presence of a disorder. Classifying respondents in these domains based on the commonalities of their observed characteristics serves the

M. Sullivan (✉) • H. Lao • J. Templin
University of Kansas, Lawrence, KS, USA

L. Moutinho, M. Sokele (eds.), *Innovative Research Methodologies in Management*,
https://doi.org/10.1007/978-3-319-64394-6_5

practical aim of having ease in communication by assigning categories instead of scores and retaining continuity between analysis and subsequent decision making.

In this chapter, we discuss psychometric methodologies involved in engaging in diagnostic measurement. We define basic terms in measurement, describe diagnostic classification models in the context of latent variable models, demonstrate an empirical example, and express the broad purpose of how diagnostic assessment can be useful in management and related fields.

Before delving into the complexities of diagnostic modeling and measurement, we will first introduce basic terms and concepts of testing and measurement. The primary focus of modern psychometric methods is to measure unobservable phenomena indirectly from manifest indicators or observable properties of objects or individuals. The hypothetical concepts to be measured from an assessment, referred to as *constructs,* are derived from researchers' abstractions or theories aimed to explain human behavior in psychology, or a state of knowledge in education. The manifest indicators that compose the construct are observed responses to items (e.g., questions or tasks). By definition, constructs are more general than specific as they relate the common cause of multiple item responses to an elaborate theory regarding complex explanations of behavior or phenomena. In other words, the nature of the relationships between multiple observable events can be explained by a general underlying construct.

The outcome of measuring constructs may be quantifying scores or numerical values for respondents or identifying whether an individual falls into different categories of the construct (e.g., classifying students as masters or non-masters of a skill). But before this can be done, empirical data must be obtained and analyzed with a proper psychometric model according to the construct definitions. Such psychometric models that researchers can use for hypothetical constructs as we have defined here fall under the heading of latent variable models, which are designed to relate the observed response data to the construct and obtain a score or profile of sorts per respondent. This is done to quantify the magnitudes of the construct and as a basis to test hypotheses or refine theories as researchers see fit. We will briefly introduce latent variable models before exploring the distinctions of models designed for diagnostic measurement.

Latent Variable Models

The common assumption of latent variable models is that the covariance (i.e., shared variance) among observed item responses is accounted for by the latent construct of the respondent, which is represented mathematically as a latent variable. Compared to traditional regression and analysis of variance approaches, the observed predictors (i.e., independent variables) are instead replaced with latent variables, where the item responses are regressed on the latent variable. In statistical terminology, the latent variable is considered a random variable that is described by a probability distribution (e.g., Gaussian distribution, Bernoulli distribution) with a desired function and shape. The primary goal of describing the latent variables is to capture the main features of the construct and the empirical processes underlying it, while also inferring respondents' amount or class based on their estimated values of the latent variable.

After conditioning on the latent variable or *attributes* in the case of diagnostic classification models (DCMs), the observed responses are assumed to be independent. This condition is referred to as *conditional* or *local independence* (e.g., Hambleton et al. 1991; Rupp et al. 2010). That is, item responses are only a function of the set of latent variables they measure, in that responses are only related due to their common relationship to the latent variable. The local independence assumption holds that the joint probability of observing two or more patterns of item responses (across all items) can be calculated as a product of their conditional response probabilities, given the latent variables and covariates. For a more comprehensive synthesis of latent variable modeling, readers are referred to the works of Skrondal and Rabe-Hesketh (2004), McDonald (1999), and Bartholomew et al. (2011).

Examples of latent variable models used to obtain the scale values (e.g., latent trait or latent factor) or profiles (e.g., latent class) for a respondent include *item response theory* (e.g., de Ayala 2009), *confirmatory factor analysis* (e.g., McDonald 1999), or *latent class models* (e.g., Hagenaars and McCutcheon 2002; Lazersfield and Henry 1968), to name a few. Although each of these modeling paradigms have different names, they are all close relatives with distinguishing features dependent on the assumed statistical distribution of the latent and observed variables (i.e., continuous or categorical; see Bartholomew et al. 2011 for a review). Recently, more general modeling frameworks have been introduced that can involve both continuous and categorical latent and observed variables, such as *generalized linear mixed models* (e.g., McCulloch and Searle 2001; Stroup 2012; Skrondal and Rabe-Hesketh 2004). In effect, the response format of the data, the assumed statistical distribution of the latent variable, and the corresponding psychometric model chosen should all conform to the requirements of the construct and the pragmatic decisions to be made. For our purposes, we will focus on the diagnostic modeling of latent variables demanding classification of objects or individuals to aid in validating theoretical models of discrete-valued constructs and provide specialists with tools to give diagnostic feedback for making diagnostic decisions.

Diagnostic Measurement

Measurement for the purpose of making classification-based decisions or identifying typologies consists of determining whether an object or individual falls in the same or different categories with respect to the construct. Researchers must collect data and choose a psychometric model that involves a latent variable that reflects the nature of the construct to obtain a statistically driven classification of respondents. A common class of psychometric models used for the purposes of diagnostic measurement are called *diagnostic classification models* (DCMs; e.g., Rupp and Templin 2008b; Rupp et al. 2010), or in the field of education, *cognitive diagnosis models* (e.g., Leighton and Gierl 2007; Nichols et al. 1995). These models are founded on the statistical framework of latent class models (e.g., Lazersfield and Henry 1968; Macready and Dayton 1977) but are restricted in a sense that they classify respondents based on a collection of predetermined discrete latent variables, commonly called *attributes* in the DCM literature.

DCMs can provide management scholars with tools to make classification-based decisions, such as identifying whether employees or firms are meeting satisfactory performance to diagnose where deficits may be occurring. This is done by classifying the unit of analysis, whether at the micro (e.g., employee-level) or macro (e.g., firms or organizations) level, into groups or *latent classes* according to specified rules as defined by the construct and in the assessment construction process. For example, individuals or organizations may be classified to a latent class or *attribute profile* according to their patterns of ownership of different attributes, which is based on their observed responses to items. In business and management fields, such attributes assessed may be aptitude, personality, or essential job-related skills to inform hiring decisions.

A distinguishing feature of tests designed for use with diagnostic methods is that the model assesses multiple discrete attributes at a time, a term referred to as multidimensionality. The constructs measured by these methods are therefore somewhat narrow in focus, where multiple specific-yet-interrelated components covering the domain of interest are defined with a fine granularity. With DCMs, the multidimensional latent variable is categorically valued, where the ordered categories of the assessed attributes could represent mastery or non-mastery of a skill or presence or absence of a disorder. The categorical attributes provide targeted differential feedback on a set of multiple skills in the form of a multivariate attribute profile for a respondent, which is synonymous with a latent class. This is distinct from common unidimensional applications of latent variable modeling, where the researchers' goal is to measure a single, broadly defined concept with a continuous, interval-level scale of measurement (e.g., item response theory), to obtain a unidimensional score. Inferences to be made about the construct are therefore broad in scope as these methods intend to define where respondents lie on the continuum relative to others. As such, a primary benefit of using DCMs is differential feedback because of their ability to define whether a respondent is close to possessing the discrete set of attributes. As such, a diagnosis can be directly accomplished from the model rather than subjectively determined at a cut point value of a continuum.

In practice, it is common for researchers to use models involving continuously distributed latent variables, such as multidimensional item response theory (M-IRT; e.g., Ackerman et al. 2003) or multidimensional confirmatory factor analysis (M-CFA; e.g., McDonald 1999), and partition the interval scale(s) into ordered, adjacent categories to be used for making classification-based decisions. This involves performing secondary analyses to determine thresholds that define group membership, or making subjective judgments based on expert opinions. Examples of such judgments include passing or failing a licensure test or identifying advanced, proficient, needs improvement, or failing students in state-wide summative exams. In these cases, the choice of the statistical distribution underlying the latent variable is continuous; however, the decision to be made from the score obtained is categorical. The process of determining where categories should be assigned according to a particular score may be done by arbitrarily splitting the score by half or assigning a cut point based on experts' consensus of what is minimally sufficient in a process called standard setting (see McClarty et al. 2013, for review). DCMs offer an advantage over subjective judgments because classifications are operationalized directly by the statistical model of the data in a single analysis rather than by content-based standards or percentiles. Furthermore, DCMs are able to reliably estimate respondent attribute profiles with fewer items per attribute than its continuous counterparts (Templin and Bradshaw 2013b). However, the large-scale educational testing industry has not implemented recent psychometric developments into operational testing programs, particularly those in categorical data analysis and diagnostic measurement, which may explain their the limited use in empirical research involving fields that make use of psychometric methods (Templin et al. 2016).

Diagnostic Classification Models

DCMs have been developed on the statistical framework of latent class models for the purposes of making diagnostic decisions. DCMs are constrained versions of latent class models, such that positive item responses are evidence for the presence of specific unobserved attributes. The characteristics of each class and number of latent classes, represented by unique attribute profiles, are specified beforehand based on the construct definitions. Thus, the approach of DCMs is similar to confirmatory factor analytic techniques that are theoretically driven as opposed to exploratory procedures, such as exploratory factor analysis or clustering techniques (Chiu et al. 2009). For multiple attributes, DCMs become "multiple classification" latent class models where each respondent is characterized by their membership in multiple latent classifications (Maris 1995, 1999). DCMs meet the goal of obtaining a diagnosis because of these distinct latent classes, in which having or not having a specific trait is sought. Those in the distinct classes are grouped based on their similar response patterns to others, rather than from assigned cut points or secondary analyses.

DCMs can be used for multiple scale types of observable variables, including binary data (e.g., Henson et al. 2009), ordinal data, nominal data (e.g., Templin and Bradshaw 2013a), and continuous data (e.g., Bozard 2010). Furthermore, the number of discrete categories of the latent variables need not be binary, but can also have multiple categories (i.e., polytomous; Rupp and Templin 2008b). To illustrate typical analyses using DCMs, we will detail a general modeling framework, including modeling the latent variable structure and item responses, estimating model parameters, and evaluating model fit. Although a number of DCMs have been proposed with varying degrees of generalization, we introduce the log-linear cognitive diagnosis model (LCDM; Henson et al. 2009) because of its general formulation that is tractable and flexible compared to other more specific models commonly found in the literature.

Formulation of the Model

The LCDM defines a measurement model to characterize the relationship between observable items and the set of latent variable(s). In the case of DCMs in general, the latent variables are represented by categorical attributes which provide statistically driven classifications of respondents into predetermined groups. The measurement model uses multivariate techniques to relate multiple observed categorical item responses, often binary, using a Bernoulli distribution with a logistic link function, to the attribute(s) within an item response function. Once the latent variable is defined in its measurement model and evaluated in terms of its fit with the data, structural relationships between other observed variables or latent variables can be examined using a similar lens of structural equation modeling literature.

The *logistic* or *logit* link function takes a bounded binary outcome, such as the domain of probabilities between 0 and 1, and uses its natural logarithmic function to map the model-predicted outcomes onto a continuous, unbounded space. This is done to accommodate explanatory variables (e.g., covariates or attribute effects) that could draw fitted values outside of this interval (Agresti 2002). Though it appears to transform the outcome, it leaves the data intact and only the conditional mean probability of the outcome is transformed. Because of this, the item response function for binary data following a Bernoulli distribution does not specify measurement error because it is automatically incorporated through the models parameterization. However, measurement error is incorporated in model via the item parameters and is reflected by probabilities of responses that are not zero or one. Other link functions can also be used depending on the specific type of model chosen, such as when the data is polytomous (e.g., multiple categories).

A Q-matrix (Tatsuoka 1983) is used to represent whether an attribute is measured by an item, where an element may be either 0 or 1 to indicate if the item measures the attribute or not. Items can have multiple elements of 1 if they measure more than one attribute, often referred to as a complex item structure. The specification of the Q-matrix should be guided by theory since it represents the operationalization of the construct and defines the design of the diagnostic assessment. This is a crucial first step as any inaccuracy can largely impact the quality of parameter estimates and correct classification of respondents' attribute profiles (Kunina-Habenicht et al. 2012; Rupp and Templin 2008a). Building the Q-matrix is a complex process that requires expert knowledge of the construct and may demand qualitative research to verify that the response processes of respondents are those intended and to establish content validity (Boorsboom and Mellenbergh 2007).

A saturated measurement model for items, in which the all possible model parameters are estimated, predicts the log-odds of a correct response based on the items' measured attributes and the respondents' latent class. For an item measuring a single attribute, the probability of a correct response is a function of the respondents' predicted status on the attribute measured (e.g., presence or absence of a diagnosis; mastery or non-mastery of a skill). For an item measuring more than one attribute, the LCDM provides the log-odds of a correct response predicted by the respondents' status on all of the items' measured attributes, along with attribute interactions to accommodate differences in the effect of having positive values of both, one, or none of the measured attributes.

To illustrate for a single item measuring two attributes taken by a respondent r, the probability of a positive response, π_{ic}, for a given latent class or attribute profile $\boldsymbol{\alpha}_r = \boldsymbol{\alpha c}$, is given as

$$\pi_{ic} = P\left(Y_{ri} = 1 | , \alpha_{r1} | , \alpha_{r2}\right) = \frac{\exp\left(\lambda_{i,0} + \lambda_{i,1,(1)}\alpha_{r1} + \lambda_{i,1,(2)}\alpha_{r2} + \lambda_{i,2,\left(1^*2\right)}\alpha_{r1}\alpha_{r2}\right)}{1 + \exp\left(\lambda_{i,0} + \lambda_{i,1,(1)}\alpha_{r1} + \lambda_{i,1,(2)}\alpha_{r2} + \lambda_{i,2,\left(1^*2\right)}\alpha_{r1}\alpha_{r2}\right)} \quad (5.1)$$

where $Y_i \in \{0, 1\}$ is the observed response to a dichotomously scored item $i, i = 1, 2, \ldots, I$. The two attributes measured, α_1 and α_2, represent the latent variables measured and are contained within a vector $\boldsymbol{\alpha c} = [\alpha_{r1}, \alpha_{r2}]$. The intercept parameter, $\lambda_{i,0}$, represents the log-odds of a correct response ($Y_i = 1$) for a respondent who has a predicted latent class membership of $\boldsymbol{\alpha c} = [0, 0]$, where $a_1 = 0$ and $a_2 = 0$. The simple main effect parameters for the two attributes, $\lambda_{i,1,(1)}$ and $\lambda_{i,1,(2)}$, characterize the change in the log-odds of a positive response when the corresponding attribute is mastered but not the other attribute. The subscripts for these parameters are separated by a comma to illustrate the (1) item for which the parameter is associated, (2) the level of the effect (i.e., 0 for the intercept, 1 for the simple main effect, 2 for the two-way interaction, etc.), and (3) the attributes in parentheses that are involved in the effect. The interaction parameter, $\lambda_{i,2,\left(1^*2\right)}$, represents the additional change in the logit of a correct response when both attributes are positive beyond the sum of the main effect parameters ($\boldsymbol{\alpha}_r = [1, 1]$). When greater than two attributes are measured by an item, higher-order interaction effects can and should be included to accommodate changes in the predicted log-odds when different attribute status constellations are present. Statistical constraints are imposed on main effect and interaction parameters such that they are always positive to ensure that having a greater number of positive attribute profile statuses corresponds to having a higher conditional probability of a positive observed response.

As you may have noticed, the LCDM measurement model for dichotomous attributes is comparable to a fully crossed factorial ANOVA with two levels per each design effect. The difference is that the latent attribute statuses correspond to the dummy-coded design effects for categorical variables, where the number of levels relates to the number of statuses the attributes can take (Rupp et al. 2010). In a similar fashion, the LCDM has similar parameter interpretations and can include all possible combinations of factors or attributes up to a final K-way interaction. Furthermore, nonsignificant highest-order parameters can be dropped from the model sequentially. By using this top-down approach to model building, a saturated model is the baseline and the final model retained generally has fewer parameters.

The general form of the saturated LCDM predicting the probability of a positive response to an item, π_{ic}, for a respondent r with attribute profile $\boldsymbol{\alpha}_r = \boldsymbol{\alpha}_c$, is given by:

$$\pi_{ic} = P\left(Y_{ri} = 1 | \alpha_{\boldsymbol{c}}\right) = \frac{\exp\left(\lambda_{i,0} + \lambda_{\boldsymbol{i}}^{\boldsymbol{T}} \mathbf{h}\left(\alpha_{\boldsymbol{c}}, \mathbf{q}_{\boldsymbol{i}}\right)\right)}{1 + \exp\left(\lambda_{i,0} + \lambda_{\boldsymbol{i}}^{\boldsymbol{T}} \mathbf{h}\left(\alpha_{\boldsymbol{c}}, \mathbf{q}_{\boldsymbol{i}}\right)\right)}. \tag{5.2}$$

The parameter $\lambda_{i,0}$ is the same intercept parameter defined above in Eq. (5.1), indicating the logit of a correct or positive response given no positive statuses on the attributes measured by item i as indicated in the Q-matrix. The simple main effects and interaction parameters are encompassed in a vector, $\boldsymbol{\lambda i}$, of size $(2^A - 1) \times 1$, for A attributes. The function $\mathbf{h}(\boldsymbol{\alpha c}, \mathbf{q}\boldsymbol{i})$ is also a vector the same size as $\boldsymbol{\lambda i}$, pertaining to the linear combinations of $\boldsymbol{\alpha c}$ and $\mathbf{q}\boldsymbol{i}$, and is expressed as:

$$\lambda_{\boldsymbol{i}}^{\boldsymbol{T}} \mathbf{h}\left(\alpha_{\boldsymbol{c}}, \mathbf{q}_i\right) = \sum_{a=1}^{A} \lambda_{i,1,(a)} \alpha_{ca} \mathrm{q}_{ia} + \sum_{a=1}^{A} \sum_{k>1} \lambda_{i,2,\left(a^* k\right)} \alpha_{ca} \alpha_{ck} \mathrm{q}_{ia} \mathrm{q}_{ik} + \ldots \tag{5.3}$$

This is dependent on the number of attributes items measure and the number of ways attribute profiles and items can be combined. The items are indicated to measure attributes in the Q-matrix, as defined here for an item i by $\boldsymbol{qi}$. Main effects are represented by $\lambda_{i,1,(a)}$ for attribute a, and second-order interaction parameters are represented by $\lambda_{i,2,\left(a^{*}k\right)}$ between attributes a and k. These are formed depending on whether both the item-by-attribute Q-matrix elements are 1 for q_{ia} and q_{ik} and the attribute statuses of the respondent are 1, α_{ca} and α_{ck}. When a respondent does not possess an attribute, all main effect and interaction item parameters indicated by the related attribute are dropped from the item response function. Because item parameters are fixed, respondents with the same attribute profile will have the same probability of a positive response.

Common special cases of the LCDM in which different parameterizations of π_{ic} are defined include the *deterministic inputs, noisy and-gate* (DINA; Haertel 1989; Junker and Sijtsma 2001); the *deterministic inputs, noisy or-gate* (DINO; Macready and Mitchell Dayton 1977); or the *noisy input, deterministic output and-gate* (NIDA; Maris 1999; Junker and Sijtsma 2001) models. These models are common in the literature and reflect different assumptions of underlying processes of the latent attributes. However, they inflict strict assumptions that may, for example, constrain equality across items or attributes. Although placing these types of constraints can ease interpretations at the test level, we recommend using a more general model at the item level with a flexible parameterization such as the LCDM.

By starting with a saturated LCDM baseline at the onset of model specification, parameters can be removed based on their empirical contributions and Q-matrices can be updated by changing corresponding elements from ones to zeroes. LCDM refinement by removing nonsignificant item parameters can be done to statistically validate and find the best fitting (yet still interpretable) model to represent the relationships between items and attributes within the data. In a similar vein, methods used in confirmatory factor analytic traditions can also be used with DCMs to add item parameters if they significantly improve model-data fit (Bradshaw et al. 2014; Jöreskog 1993; Brown 2013), such as with omitted Q-matrix entries that may critique theory and help refine constructs. The LCDM has the added benefit in that it does not require additional identification constraints, particularly when there is additional complexity due to complex items that measure more than one attribute.

Beyond the specification of the measurement model that relates items to attributes, the LCDM also includes a structural component to indicate the proportion of respondents with each attribute profile and, additionally, how attributes are related to each other. This is represented by a parameter, ν_c, to denote the base-rate probability of being in a particular class (or having a particular attribute profile), c. Since the LCDM is a confirmatory model and specifies the latent classes *a priori,* the number of estimated structural parameters is equal to $2^A - 1$ for a saturated model with A dichotomous attributes. With increasing attributes measured as indicated in the Q-matrix, the number of estimated structural parameters increases exponentially.

The respondent's attribute profile follows a multivariate Bernoulli distribution (or MVB; Maydeu and Joe 2005) for binary attributes or a multivariate multinomial distribution for DCMs with more than two category attributes. The structural parameter, ν_c, represents the probability that a respondent r is in a particular class c ($c = 1, \ldots, 2^A$), for respondents of the same attribute profile (Rupp et al. 2010). The structural portion defines the associations between attributes and combines with the measurement component of the model to give the marginal LCDM likelihood function for binary item responses with two-category attributes for a respondent r:

$$P\left(Y_r = y_r\right) = \sum_{c=1}^{2A} \nu_c \prod_{i=1}^{I} \pi_{ic}^{Y_{ir}} \left(1 - \pi_{ic}\right)^{1-Y_{ir}} \tag{5.4}$$

Assuming independent item responses after conditioning on the respondents' latent class, the dichotomous item responses are modeled as a product of Bernoulli trials for each item. For dichotomous items modeled using a general LCDM with fully crossed ANOVA-like factors with all possible parameters specified, there are $2^A - 1$ estimated ν_c parameters (the last parameter is determined by the other parameters because the ν_c must sum to 1).

The marginal distribution of the attributes provides information about the population-level proportions of respondents that possess the attribute. For a given attribute a , the sum of the ν_c parameters over all classes in which the attribute is possessed provides the marginal proportion of respondents having the attribute. That is, $\dot{v}_a = \sum_{c=1}^{C} \pi_c \alpha_{ca}$. A summary of the joint distributions containing the 2^A ν_c parameters that form all pairwise associations between two selected attributes is contained in 2 × 2 contingency tables of positive attribute statuses. The cells in the table contain the proportion of respondents who have positive or negative statuses for the two attributes and can be used to quantify the strength of association between the two attributes using tetrachoric correlations (Rupp et al. 2010).

Common models that place a structure on attribute associations are outlined in Rupp et al. (2010). For ease of exposition, our next section illustrating an example data analysis of personality traits will use an unstructured structural model in which all possible structural parameters are estimated and no constraints are placed on attribute profile membership probabilities.

Application of the Diagnostic Classification Model

Now that we have defined a general model to obtain latent attribute profiles from our data, we can demonstrate its use with an empirical example. This section illustrates the implementation of DCMs by starting from the purpose of data analysis to the conclusions drawn. We intend this section as a modeling guide rather than an analysis result summary and hope to make the concepts described above more concrete.

A public personality dataset is selected as our example, with two considerations. First, a public dataset is selected so that interested readers can freely access the data online and replicate the example for practice. Second, personality tests might be used in personnel selection, a relevant topic in management research (e.g., Morgeson et al. 2007; Schmidt and Hunter 1998; Thomas and Scroggins 2006).

The purpose of examining this data is to screen potential employees for their predisposition to certain personality traits that may inform whether they would make quality candidates for future employment. By using a

DCM, the data can be used to determine the proportion of respondents with certain endorsements of traits and the relationships between the traits in the sample. The data is the "bfi" dataset from the "psych" package of *R* software (*R* Core Team 2016) and contains 2800 respondents answering 25 items. There are five personality traits that characterize our latent attributes, resulting in $2^5 = 32$ latent classes. Each trait is measured by five items with a simple structure, so there are no interactions between attributes within our item response functions of our measurement model. Sample items and their Q-matrix entries are depicted in Table 5.1.

Several items were recoded based on the keys provided in the description webpage. For our example, the original 6-point Likert responses are dichotomized into yes/no responses, a similar step taken by Finch and Bronk (2011). In practice, compressing rich information contained in continuous data is not recommended because fine-grained distinctions between adjacent response categories are lost. When the information is beyond the research scope, data compression is a compromise, for example, using principle components analysis to reduce data dimensionality in big data. However, in the context of our illustration, dichotomous response data is desired to map onto common data response types found in the literature. The aim here is to exemplify model implementation while avoiding extra complexity involved in modeling continuous responses with DCMs.

Table 5.1 Example Q-matrix for sample items for five personality attributes

	Attribute				
Sample item	Agreeable	Conscientious	Extraversion	Neuroticism	Openness
Am indifferent to the feelings of others	1	0	0	0	0
Do things in a half-way manner	0	1	0	0	0
Make friends easily	0	0	1	0	0
Panic easily	0	0	0	1	0
Spend time reflecting on things	0	0	0	0	1

The purpose of our analysis is to select an agreeable, conscientious, extraverted, non-neurotic, and open employee from a sample of job candidates because we believe that people with these traits are best qualified to perform the job tasks. This attribute profile is represented by $\boldsymbol{\alpha c}$= [1, 1, 1, 0, 1]. Thus, in order to increase hiring efficiency and reduce cost, we use this personality test as a screening criterion before proceeding to an in-person interview.

Based on the DCM we have selected, there is an underlying belief that our personality traits, as unobservable latent variables, are categorically distributed instead of continuous. This is arguable, but for our purpose of demonstration, we specify two categories that are operationalized as possession of personality attributes. For example, having a "0" within a latent class for extraversion indicates introverts, whereas a "1" indicates extraverts. This is dependent on how the items are scored, where having a positive response indicates greater probability of endorsing the item and possessing the attribute. Without too much prior constraints on the model parameters, a saturated LCDM measurement model and unstructured structural model is preferred.

After data has been collected, prepared for analysis, and the details of our model are specified (e.g., Q-matrix), we must select a software package that can implement our choice of model. Some considerations of which software to use may be its statistical soundness, supporting documentation, and its frequency of reference in peer-reviewed papers. Users' prior software experience and price are additional considerations. We recommend using the commercial software *M*plus (Muthén and Muthén 2012) because of the available step-by-step guides for DCMs (Rupp et al. 2010; Templin and Hoffman 2013). The development of *M*plus syntax could be laborious; however, automation is possible using an online SAS macro for creating *M*plus syntax via SAS® software (SAS 2011), or via *R* software (R Core Team 2016).

*M*plus software is also selected because it supports the addition of model constraints required by our confirmatory models. Two potential issues that can occur when using confirmatory model estimations without proper model constraints include (1) having a locally optimal solution and (2) label switching.

Local optimal solution is a universal concern to latent class analysis, either confirmatory or exploratory, and either with or without model constraints. It refers to the case when a model solution is optimal in a local sample space, but not necessarily optimal in a global sample space. Various methods have been proposed to solve this problem (e.g., Pintér 2002, 2006).

On the other hand, label switching is specific to confirmatory related models (e.g., Redner and Walker 1984; Stephens 2000). It refers to the meaning of the latent class shuffles after estimation. It likely happens when not specifying any constraints in confirmatory related model, allowing reallocating meaning to latent class labels during estimation. Without constraints, the confirmatory models are treated as exploratory models. With intractable latent class meanings, the results are misleading and confusing.

After running *M*plus, a series of model outputs based on the syntax specification are provided. The first set of results is general model fit information and estimated parameters for our measurement model. The second set is the attribute classifications for respondents, providing model-predicted probabilities per latent class membership and the final classifications for respondents based on their highest probabilities per attribute.

The first thing to check is whether the model converged from the general output. If the model fails to converge, no estimated parameters are provided and other solutions to obtain convergence should be considered (e.g., increasing iterations). If the model converges, the next step is to evaluate model fit, such as the log likelihood (LL = −35221.33), Akaike information criteria (AIC = 70604.66), and Bayesian information criteria (BIC = 71085.58). However, although useful for model comparison and selection (log likelihood for nested models, AIC and BIC for non-nested models), none of these outputs are appropriate to decide the absolute fit of a single model. An appropriate measure of absolute model fit comes from the bivariate fit information using *M*plus option TECH 10 (see Rupp et al. 2010 for review). Once we have determined that our model fits the data, we can proceed to interpreting our results.

Among the estimated parameters, those from the measurement model are of little interest here to our selection purpose as we are primarily interested in respondent attribute classifications. This information can be found from the structural model as shown in Table 5.2 which illustrates the proportion of respondents classified in each latent class. These proportions provide a description of the personality distribution of our candidates. Based on the results, 30 percent of job candidates are in the latent 30th class, α_{30}= [1, 1, 1, 0, 1], which corresponds to respondents who are agreeable, conscientious, extravert, non-neurotic, and open. The next step is to select from candidates who are classified into class α_{30} according to the last column from the person classification output for further consideration of employment.

As with any inference, precautions should be taken when using any test for personnel selection. First, comprehensive research to evaluate the appropriateness and legitimacy of using a test to select the candidates should be completed. Test development or existing tests that correspond to the context and purpose of implementation should be selected and evaluated for their content validity and construct representation (Kane 2013). This includes checking (1) the relevance and representativeness of the test content to the job (i.e., content representativeness), (2) the consistency of test scores over repeated administrations (i.e., reliability), and (3) the accuracy of using the test scores to make the selection decisions (i.e., validity). Some resources that outline the test development process for use with DCMs are given by Bradshaw et al. (2014) and by Rupp et al. (2010).

Diagnostic classification models are useful when the primary goals are classifying respondents into multivariate attribute profiles and examining relationships between latent variables that assume discrete distributions, such as a Bernoulli distribution for binary attributes as outlined here or a Multinomial distribution for polytomous categories (e.g., positive, neutral, negative). In our example above, we could have used polytomous categories of attributes to represent more than two categories for each personality trait observed in the sample. Those having a probability range of between .60 and 1 as indicated by endorsing more items that indicate extraverts could be classified as such, whereas those having smaller probabilities between 0 and 0.40 could indicate introverts. The range between these extremes, 0.40 and 0.60, could indicate respondents who are indifferent in their magnitudes of extraversion and introversion or

Table 5.2 Estimated structural model parameters for five personality attributes

	Attributes					
Latent class	Agreeableness	Conscientiousness	Extraversion	Neuroticism	Openness	*νc*
1	0	0	0	0	0	.02
2	0	0	0	0	1	.02
3	0	0	0	1	0	.05
4	0	0	0	1	1	.03
5	0	0	1	0	0	.00
6	0	0	1	0	1	.00
7	0	0	1	1	0	.00
8	0	0	1	1	1	.01
9	0	1	0	0	0	.02
10	0	1	0	0	1	.02
11	0	1	0	1	0	.01
12	0	1	0	1	1	.03
13	0	1	1	0	0	.00
14	0	1	1	0	1	.01
15	0	1	1	1	0	.00
16	0	1	1	1	1	.01
17	1	0	0	0	0	.01
18	1	0	0	0	1	.02
19	1	0	0	1	0	.02
20	1	0	0	1	1	.02
21	1	0	1	0	0	.01
22	1	0	1	0	1	.05
23	1	0	1	1	0	.02
24	1	0	1	1	1	.06
25	1	1	0	0	0	.02
26	1	1	0	0	1	.02
27	1	1	0	1	0	.03
28	1	1	0	1	1	.02
29	1	1	1	0	0	.02
30	***1***	***1***	***1***	***0***	***1***	***.30***
31	1	1	1	1	0	.02
32	1	1	1	1	1	.13

Note: *νc* represents the proportion of candidates in each latent class, to the second decimal. The underlined 30th class is our selection criteria, representing agreeable, conscientious, extravert, non-neurotic, and open

there is insufficient information to correctly classify them as extraverts or introverts. The widths of these bands are determined by the construct definition and can be construed as a function of the utility in their interpretations (Rupp et al. 2010; Jang 2005). For our dichotomous attributes, if the probability of being in a class is 0.50, this could indicate a

50/50 chance of being an introvert or extravert. The uncertainty involved in classifying the respondent in this instance would necessitate more information, likely from additional item responses or external sources.

Chapter Conclusion

In this chapter we described how multidimensional latent variable models involving fine-grained attributes are represented by diagnostic classification models. We demonstrated a saturated measurement model and an unstructured structural model for personality traits that could be used to provide diagnostic feedback so that test administrators can determine the classifications of respondents among these traits. We hope our demonstration of diagnostic measurement techniques has introduced the framework with which a novice user of DCMs can implement such methods and expand their research repertoire.

References

Ackerman, T. A., Gierl, M. J., & Walker, C. M. (2003). Using Multidimensional Item Response Theory to Evaluate Educational and Psychological Tests. *Educational Measurement: Issues and Practice, 22*, 37–53. https://doi.org/10.1111/j.1745-3992.2003.tb00136.x.

Agresti, A. (2002). *Categorical Data Analysis*. Hoboken: Wiley.

Bartholomew, D. J., Knott, M., & Moustaki, I. (2011). *Latent Variable Models and Factor Analysis: A Unified Approach*. Hoboken: Wiley.

Boorsboom, D., & Mellenbergh, G. J. (2007). Test Validity in Cognitive Assessment. In J. P. Leighton & M. J. Gierl (Eds.), *Cognitive Diagnostic Assessment for Education: Theory and Applications*. Cambridge: Cambridge University Press.

Bozard, J. L. (2010). *Invariance Testing in Diagnostic Classification Models* (Masters Thesis). University of Georgia.

Bradshaw, L., Izsák, A., Templin, J., & Jacobson, E. (2014). Diagnosing Teachers' Understandings of Rational Numbers: Building a Multidimensional Test Within the Diagnostic Classification Framework. *Educational Measurement: Issues and Practice, 33*, 2–14. https://doi.org/10.1111/emip.12020.

Brown, C. (2013). *Modification Indices for Diagnostic Classification Models* (Unpublished Doctoral Dissertation). University of Georgia, Athens.

Chiu, C.-Y., Douglas, J., & Li, X. (2009). Cluster Analysis for Cognitive Diagnosis: Theory and Applications. *Psychometrika, 74*, 633–665. https://doi.org/10.1007/s11336-009-9125-0.

de Ayala, R. J. (2009). *The Theory and Practice of Item Response Theory*. New York: Guilford Press.

Finch, W. H., & Bronk, K. C. (2011). Conducting Confirmatory Latent Class Analysis Using *M*plus. *Structural Equation Modeling: A Multidisciplinary Journal, 18*(1), 132–151. https://doi.org/10.1080/10705511.2011.532732.

Haertel, E. H. (1989). Using Restricted Latent Class Models to Map the Skill Structure of Achievement Items. *Journal of Educational Measurement, 26*, 301–321.

Hagenaars, J. A., & McCutcheon, A. L. (2002). *Applied Latent Class Analysis*. Cambridge: Cambridge University Press.

Hambleton, R. K., Swaminathan, H., & Jane Rogers, H. (1991). *Fundamentals of Item Response Theory*. Newbury Park: Sage.

Henson, R. A., Templin, J. L., & Willse, J. T. (2009). Defining a Family of Cognitive Diagnosis Models Using Log-Linear Models with Latent Variables. *Psychometrika, 74*, 191–210. https://doi.org/10.1007/s11336-008-9089-5.

Jang, E. E. (2005). *A Validity Narrative: Effects of Reading Skills Diagnosis on Teaching and Learning in the Context of NG TOEFL* (Unpublished Doctoral Dissertation). University of Illinois, Urbana-Champaign.

Jöreskog, K. G. (1993). *Testing Structural Equation Models*. In K. A. Bollen & J. Scott Long (Eds.), *Testing Structural Equation Models*. Newbury Park: Sage.

Junker, B. W., & Sijtsma, K. (2001). Cognitive Assessment Models with Few Assumptions, and Connections with Nonparametric Item Response Theory. *Applied Psychological Measurement, 25*, 258–272. https://doi.org/10.1177/01466210122032064.

Kane, M. T. (2013). Validating the Interpretations and Uses of Test Scores. *Journal of Educational Measurement, 50*, 1–73. https://doi.org/10.1111/jedm.12000.

Kunina-Habenicht, O., Rupp, A. A., & Wilhelm, O. (2012). The Impact of Model Misspecification on Parameter Estimation and Item-Fit Assessment in Log-Linear Diagnostic Classification Models. *Journal of Educational Measurement, 49*(1), 59–81. https://doi.org/10.1111/j.1745-3984.2011.00160.x.

Lazersfield, P. F., & Henry, N. W. (1968). *Latent Structure Analysis*. Boston: Houghton Mifflin Company.

Leighton, J. P., & Gierl, M. J. (2007). *Cognitive Diagnostic Assessment for Education: Theory and Applications*. Cambridge: Cambridge University Press.

Macready, G. B., & Mitchell Dayton, C. (1977). The Use of Probabilistic Models in the Assessment of Mastery. *Journal of Educational Statistics, 2*, 99–120. https://doi.org/10.3102/10769986002002099.

Maris, E. (1995). Psychometric Latent Response Models. *Psychometrika, 60*(4), 523–547. https://doi.org/10.1007/BF02294327.

Maris, E. (1999). Estimating Multiple Classification Latent Class Models. *Psychometrika, 64*(2), 187–212. https://doi.org/10.1007/BF02294535.

Maydeu, A., & Joe, H. (2005). Limited- and Full-Information Estimation and Goodness-of-Fit Testing in 2n Contingency Tables. *Journal of the American Statistical Association, 100*, 1009–1020. https://doi.org/10.1198/016214504000002069.

McClarty, K. L., Way, W. D., Porter, A. C., Beimers, J. N., & Miles, J. A. (2013). Evidence Based Standard Setting: Establishing a Validity Framework for Cut Scores. *Educational Researcher, 42*, 78–88. https://doi.org/10.3102/0013189X12470855.

McCulloch, C. E., & Searle, S. R. (2001). *Generalized, Linear, and Mixed Models*. New York: Wiley.

McDonald, R. P. (1999). *Test Theory: A Unified Approach*. Mahwah: Erlbaum.

Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007). Reconsidering the Use of Personality Tests in Personnel Selection Contexts. *Personnel Psychology, 60*(3), 683–729. https://doi.org/10.1111/j.1744-6570.2007.00089.x.

Muthén, L. K., & Muthén, B. (2012). *Mplus Statistical Modeling Software: Release 7.0*. Los Angeles: Muthén & Muthén.

Nichols, P. D., Chipman, S. F., & Brennan, R. L. (1995). *Cognitively Diagnostic Assessment*. Mahwah: Erlbaum.

Pintér, J. D. (2002). Global Optimization: Software, Test Problems, and Applications. In P. M. Pardalos & H. Edwin Romeijn (Eds.), *Handbook of Global Optimization* (Vol. 2, pp. 515–569). Boston: Springer.

Pintér, J. D. (2006). Global Optimization: Scientific and Engineering Case Studies. In J. D. Pintér (Ed.), *Nonconvex Optimization and Its Applications* (Vol. 85). New York: Springer Science & Business Media.

R Core Team. (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna. https://www.R-project.org/

Redner, R. A., & Walker, H. F. (1984). Mixture Densities, Maximum Likelihood and the EM Algorithm. *SIAM Review, 26*(2), 195–239.

Rupp, A. A., & Templin, J. (2008a). The Effects of Q-Matrix Misspecification on Parameter Estimates and Classification Accuracy in the DINA Model. *Educational and Psychological Measurement, 68*, 78–96. https://doi.org/10.1177/0013164407301545.

Rupp, A. A., & Templin, J. (2008b). Unique Characteristics of Diagnostic Classification Models: A Comprehensive Review of the Current State-of-the-Art. *Measurement, 6*(4), 219–262. https://doi.org/10.1080/15366360802490866.

Rupp, A. A., Templin, J., & Henson, R. A. (2010). *Diagnostic Measurement: Theory, Methods, and Applications*. New York: Guilford Press.

SAS Institute. (2011). *The SAS System for Windows*. Cary: SAS Institute.

Schmidt, F. L., & Hunter, J. E. (1998). The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 85 Years of Research Findings. *Psychological Bulletin, 124*(2), 262–274. https://doi.org/10.1037/0033-2909.124.2.262.

Skrondal, A., & Rabe-Hesketh, S. (2004). *Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models*. New York: Chapman & Hall/CRC.

Stephens, M. (2000). Dealing with Label Switching in Mixture Models. *Journal of the Royal Statistical Society, Series B: Statistical Methodology, 62*(4), 795–809. https://doi.org/10.1111/1467-9868.00265.

Stroup, W. W. (2012). *Generalized Linear Mixed Models: Modern Concepts, Methods, and Applications*. Boca Raton: CRC Press.

Tatsuoka, K. K. (1983). Rule-space: An Approach for Dealing with Misconceptions Based on Item Response Theory. *Journal of Educational Measurement, 20*, 345–354. https://doi.org/10.1111/j.1745-3984.1983.tb00212.x.

Templin, J., & Bradshaw, L. (2013a). *Nominal Response Diagnostic Classification Models* (Manuscript Under Revision).

Templin, J., & Bradshaw, L. (2013b). The Comparative Reliability of Diagnostic Model Examinee Estimates. *Journal of Classification, 10*, 251–275. https://doi.org/10.1007/s00357-013-9129-4.

Templin, J., & Hoffman, L. (2013). Obtaining Diagnostic Classification Model Estimates Using *M*plus. *Educational Measurement: Issues and Practice, 32*, 37–50. https://doi.org/10.1111/emip.12010.

Templin, J., Bradshaw, L., & Paek, P. (2016). A Comprehensive Framework for Integrating Innovative Psychometric Methodology into Educational Research. In A. Izsák, J. T. Remmillard, & J. Templin (Eds.), *Psychometric Methods in Mathematics Education: Opportunities, Challenges, and*

Interdisciplinary Collaborations, *Journal for Research in Mathematics Education Monograph Series*. Reston: National Council of Teachers of Mathematics.
Thomas, S. L., & Scroggins, W. A. (2006). Psychological Testing in Personnel Selection: Contemporary Issues in Cognitive Ability and Personality Testing. *Journal of Business Inquiry, 5*, 28–38.

6

Incremental Optimization Mechanism for Constructing a Balanced Very Fast Decision Tree for Big Data

Hang Yang and Simon Fong

Introduction

Big data is a popular topic that highly attracts attentions of researchers from all over the world. How to mine valuable information from such huge volumes of data remains an open problem. Although fast development of hardware is capable of handling much larger volumes of data than ever before, in the author's opinion, a well-designed algorithm is crucial in solving the problems associated with big data. Data stream mining methodologies propose one-pass algorithms that are capable of discovering knowledge hidden behind massive and continuously moving data. Stream mining provides a good solution for such big data problems, even for potentially infinite volumes of data.

H. Yang (✉)
China Southern Power Grid,
Guangzhou, Guangdong, China

S. Fong
Department of Computer and Information Science,
University of Macau, Macau SAR, Zhuhai Shi, China

L. Moutinho, M. Sokele (eds.), *Innovative Research Methodologies in Management*,
https://doi.org/10.1007/978-3-319-64394-6_6

Decision tree learning is one of the most significant classifying techniques in data mining and has been applied to many areas, including business intelligence, health-care and biomedicine, and so forth. The traditional approach to building a decision tree, designed by greedy search, loads a full set of data into memory and partitions the data into a hierarchy of nodes and leaves. The tree cannot be changed when new data are acquired, unless the whole model is rebuilt by reloading the complete set of historical data together with the new data. This approach is unsuitable for unbounded input data such as data streams, in which new data continuously flow in at high speed. A new generation of algorithms has been developed for incremental decision tree, a pioneer of which using a Hoeffding bound (*HB*) in node-splitting is so called Very Fast Decision Tree (VFDT) (Pedro and Geoff 2000), which can build a decision tree simply by keeping track of the statistics of the attributes of the incoming data. When sufficient statistics have accumulated at each leaf, a node-splitting algorithm determines whether there is enough statistical evidence in favor of a node-split, which expands the tree by replacing the leaf with a new decision node. This decision tree learns by incrementally updating the model while scanning the data stream on the fly. In the past decade, VFDT has been extended to some improved algorithms, inheriting the use of HB (see in section "Background"). This powerful concept is in contrast to a traditional decision tree that requires the reading of a full dataset for tree induction. The obvious advantage is its real-time mining capability, which frees it from the need to store up all of the data to retrain the decision tree because the moving data streams are infinite.

On one hand, the challenge for data stream mining is associated with the imbalanced class distribution. The term "imbalanced data" refers to irregular class distributions in a dataset. For example, a large percentage of training samples may be biased toward class *A*, leaving few samples that describe class *B*. Both noise and imbalanced class distribution significantly impair the accuracy of a decision tree classifier through confusion and misclassification prompted by the inappropriate data. The size of the decision tree will also grow excessively large under noisy data. To tackle these problems, some researchers applied data manipulation techniques to handle the imbalanced class distribution problems, including under-sampling, resampling, a recognition-based induction scheme

(Nitesh et al. 2004), and a feature subset selection approach (Mladenic and Grobelnik 1999).

On the other hand, despite the difference in their tree-building processes, both traditional and incremental decision trees suffer from a phenomenon called over-fitting when the input data are infected with noise. The noise confuses the tree-building process with conflicting instances. Consequently, the tree size becomes very large and eventually describes noise rather than the underlying relationship. With traditional decision trees, the under-performing branches created by noise and biases are commonly pruned by cross-validating them with separate sets of training and testing data. Pruning algorithms (Elomaa 1999) help keep the size of the decision tree in check; however, the majority are post-pruning techniques that remove relevant tree paths after a whole model has been built from a stationary dataset. Post-pruning of a decision tree in high-speed data stream mining, however, may not be possible (or desirable) because of the nature of incremental access to the constantly incoming data streams. Incremental optimization seeks for solutions that evolve over time in response to environmental changes. In general, there are three performance metrics for incremental problems: *ratio*, *sum*, and *demand* (Hartline 2008). *Ratio metric* uses a worst-case measurement to determine the distance between the optimal solution and the solution made at each time step. *Sum metric* is the expected value metric over all time steps. A weight function while summing solution values can easily settle the problem of natural bias for late-stage solution. *Demand metric* is a decision metric measuring the degree of specific quantitative requirements satisfaction.

VFDT handles streaming data that tree structure keeps on updating when new data arrive. It only requires reading some samples satisfying the statistical bound (referring to the HB) to construct a decision tree. Since it cannot analyze over the whole training dataset in one time, normal optimization methods using full dataset to search for an optima between the accuracy and tree size do not work well here.

Our previous work has provided a solution for sustainable prediction accuracy and regulates the growth of the decision tree to a reasonable extent, even in the presence of noise. Moderated Very Fast Decision Tree (MVFDT) is a novel extension of the VFDT model (Yang and Fong 2011) that includes optimizing the tree-growing process via *adaptive tie-breaking threshold* instead of a user pre-defined value in VFDT.

In this chapter, for optimizing VFDT, we devise a new version, so called optimized VFDT (OVFDT), which can provide an incremental optimization on prediction accuracy, tree size, and learning time. The contributions of OVFDT are: (1) it contains four types of functional tree leaf that improve the classification accuracy; (2) it inherits the mechanism of MVFDT that uses an adaptive tie-breaking threshold instead of a user pre-defined. To this end, it may suit for the aforementioned real applications; (3) it contains an incremental optimization mechanism in the node-splitting test that obtains an optimal tree structure as a result. By running simulation experiments, the optimized value of adaptive tie is proved to be ideal for constraining the optimal tree growth.

Background

Decision Tree in Data Stream Mining

A decision tree classification problem is defined as follows: N is a set of examples of the form (X, y), where X is a vector of d attributes and y is a discrete class label. k is the index of class label. Suppose a class label with the k_{th} discrete value is $\mathrm{y_k}$. Attribute X_i is the i_{th} attribute in X, and is assigned a value of x_{i1}, x_{i2}… x_{iJ}, where $1 \leq i \leq d$ and J is the number of different values X_i. The classification goal is to produce a decision tree model from N examples, which predicts the classes of y in future examples with high accuracy. In data stream mining, the example size is very large or unlimited, $N \to \infty$.

VFDT algorithm (Pedro and Geoff 2000) constructs an incremental decision tree by using constant memory and constant time-per-sample. VFDT is a pioneering predictive technique that utilizes the Hoeffding bound. The tree is built by recursively replacing leaves with decision nodes. Sufficient statistics n_{ijk} of attribute X_i with a value of x_{ij} are stored in each leaf with a class label assigning to a value y_k. A heuristic evaluation function $H(\cdot)$ is used to determine split attributes for converting leaves to nodes. Nodes contain the split attributes and leaves contain only the class labels. The leaf represents a class according to the sample label. When a sample enters, it traverses the tree from the root to a leaf, evaluating the relevant attributes at every node. Once the sample reaches a leaf, the sufficient statistics are updated. At this time, the system evaluates each possible condition based on the attribute values; if the statistics are sufficient to support one test over the others, then a leaf is converted to a decision node. The decision node contains the number of possible values for the chosen attribute according to the installed split test. The main elements of VFDT include, first, a tree-initializing process that initially contains only a single leaf and, second, a tree-growing process that contains a splitting check using a heuristic function $H(\cdot)$ and Hoeffding bound (HB). VFDT uses information gain as $H(\cdot)$.

The formula of HB is shown in (6.1). HB controls over errors in the attribute-splitting distribution selection, where R is the range of classes' distribution and n is the number of instances that have fallen into a leaf. To evaluate a splitting-value for attribute X_i, it chooses the best two values. Suppose x_{ia} is the best value of $H(\cdot)$ where $x_{ia} = \arg\max H(x_{ij})$; suppose x_{ib} is the second best value where $x_{ib} = \arg\max H(x_{ij}), \forall j \neq a$; suppose $\Delta H(X_i)$ is the difference of the best two values for attribute X_i, where $\Delta H(X_i) = \Delta H(x_{ia}) - \Delta H(x_{ib})$. Let n be the observed number of instances, HB is used to compute high confidence intervals for the true mean r_{true} of attribute x_{ij} to class y_k that $r - \text{HB} \leq r_{\text{true}} < r + \text{HB}$ where $r = (1/n)\sum_i^n r_i$. If after observing $n_{\min}$ examples, the inequality $r + \text{HB} < 1$ holds, then $r_{\text{true}} < 1$, meaning that the best attribute x_{ia} observed over a portion of the stream is truly the best attribute over entire stream. Thus, a splitting-value x_{ij} of attribute X_i can be found without full attribute values even when we don't know all values of X_i. In other words, it does not train a model from full data and the tree is growing incrementally when more and more data come.

$$HB = \sqrt{\frac{R^2 \ln\left(\frac{1}{\delta}\right)}{2n}}$$

In the past decade, several research papers have proposed different methodologies to improve the accuracy of VFDT. Such incremental decision tree algorithms using HB in node-splitting test are so called Hoeffding Tree (HT). HOT (Pfahringer et al. 2007) proposes an algorithm producing some optional tree branches at the same time, replacing those rules with lower accuracy by optional ones. The classification accuracy has been improved significantly while learning speed is slowed because of the construction of optional tree branches. Some of options are inactive branches consuming computer resource. Functional tree leaf is originally proposed to integrate to incremental decision tree in VFDTc (Gama et al. 2003). Consequently, Naïve Bayes classifier on the tree leaf has improved classification accuracy. The functional tree leaf is able to handle both continuous and discrete values in data streams, but no direct evidence shows it can handle such imperfections like noise and bias in data streams. FlexDT (Hashemi and Yang 2009) proposes a Sigmoid function to handle noisy data and missing values. Sigmoid function is used to decide what true node-splitting value, but sacrificing algorithm speed. For this reason, the lightweight algorithm with fast learning speed is favored for data streams environment. CBDT (Stefan et al. 2009) is a forest of trees algorithm that maintains a number of trees, each of which is rooted on a different attributes and grows independently. It is sensitive to the concept-drift in data streams according to the sliding-window mechanism. VFDR (Gama and Kosina 2011) is a decision rule learner using HB. The same as VFDT, VFDR proposes a rule-expending mechanism that constructs the decision rules (ordered or unordered) from data stream on the fly.

There are two popular platforms for implementing stream-mining decision tree algorithms. Very Fast Machine Learning (VFML) (Hulten and Domingos 2003) is a C-based tool for mining time-changing high-speed data streams. Massive Online Analysis (MOA) (Bifet et al. 2001) is Java-based software for massive data analysis, which is a well-known open source project extended from WEKA data mining. In both platforms, the parameters of VFDT must be pre-configured. For different tree induction tasks, the parameter setup is distinguished.

MOA is an open source project with a user-friendly graphic interface. It also provides several ways to evaluate algorithm's performance. Hence, some VFDT-extended algorithms have been built-in this platform. For example, the VFDT algorithms embedded in MOA (released on Nov. 2011) are: Ensemble Hoeffding Tree (Oza and Russell 2001) is an online bagging method with some ensemble VFDT classifiers. Adaptive Size Hoeffding Tree (ASHT) (Bifet et al. 2009) is derived from VFDT adding a maximum number of split nodes. ASHT has a maximum number of split nodes. After one node splits, if the number of split nodes is higher than the maximum value, then it deletes some nodes to reduce its size. Besides, it is designed for handling concept-drift data streams AdaHOT (Bifet et al. 2009) is also derived from HOT. Each leaf stored an estimation of current error. The weight of node in voting process was proportional to the square of inverse of error. AdaHOT combines HOT with a voting mechanism on each node. It also extends the advantages using optional trees to replace the tree branches of bad performance. Based on an assumption "there has been no change in the average value inside the window", ADWIN (Bifet and Gavalda 2007) proposes a solution to detect changes by a variable-length window of recently seen instances. In this chapter, the OVFDT algorithm is developed on the fundamental of MOA platform. All experiments are also run on MOA platform.

Relationship Among Accuracy, Tree Size, and Time

When data contains noisy values, it may confuse the result of heuristic function. The difference of the best two heuristic evaluation for attribute X_i, where $\Delta\bar{H}\left(X_i\right)=H\left(x_{ia}\right)-H\left(x_{ib}\right)$, may be negligible. To solve this problem, a fixed tie-breaking τ, which is a user pre-defined threshold for incremental learning decision tree, is proposed as pre-pruning mechanism to control the tree growth speed (Hulten et al. 2001). This threshold constrains the node-splitting condition that $\Delta\bar{H}\left(X_i\right)\leq \mathrm{HB}<\tau$. An efficient τ guarantees a minimum tree growth in case of tree size explosion problem. τ must be set before a new learning starts; however, so far there has not been a unique τ suitable for all problems. In other words, there is not a single default value that works well in all tasks so far. The choice of τ hence depends on the data and their nature. It is said that the excessive invocation of tie breaking brings the performance of decision tree learning declining significantly on complex and noise data, even with the additional condition by the parameter τ.

A proposed solution (Geoffrey et al. 2005) to overcome this detrimental effect is an improved tie-breaking mechanism, which not only considers the best (x_{ia}) and the second best (x_{ib}) splitting candidates in terms of heuristic function but also uses the worst candidate (x_{ic}). At the same time, an extra parameter is imported, α, which determines how many times smaller the gap should be before it is considered as a tie. The attribute-splitting condition becomes: when $\alpha \times (H(x_{ia}) - H(x_{ib})) < (H(x_{ib}) - H(x_{ic}))$, the attribution x_{ia} shall be split as a node. Obviously, this approach uses two extra elements, α and x_{ic}, which bring extra computation to the original algorithm.

In addition to the tie-breaking threshold τ, $n_{\min}$ is the number of instances a leaf should observe between split attempts. In other words, τ is a user-defined value to control the tree-growing speed, and $n_{\min}$ is a user-defined value to control the interval time to check node-splitting. The former is used to constrain tree size and the latter is used to constrain the learning speed. In order to optimize accuracy, tree size, and speed for decision tree learning, first of all, an example is given for demonstrating the relationship among these three factors for data streams.

In this example, the testing datasets are synthetic added bias class. We use MOA to generate the tested datasets. *LED24* is the nominal data structure and *Waveform21* is the numeric data structure. Both datasets share the origins with the sample generators donated by UCI machine learning repository. *LED24* problem uses 24 binary attributes to classify 10 different classes. The goal of *Waveform21* task is to differentiate between three different classes of waveform, each of which is generated from a combination of two or three base waves. It has 21 numeric attributes. The data stream problem is simulated by large number of instances, which are as many as one million for both datasets. The accuracy, tree size, and time are recorded with changing the pre-defined values of τ and $n_{\min}$. From Table 6.1, we can see that:

- In general, the bigger tree size brings a higher accuracy, even caused by the over-fitting problem, but taking more learning time.
- τ is proposed to control the tree size growing. A bigger τ brings a faster tree size growth, but longer computation time. But because the memory is limited, the tree size does not increase while τ reaches a threshold ($\tau = 0.7$ for LED24; $\tau = 0.4$ for Waveform21).
- $n_{\min}$ is proposed to control the learning time. A bigger $n_{\min}$ brings a faster learning speed, but smaller tree size and lower accuracy.

Table 6.1 Comparison of VFDT using different τ and n_{min}

τ	**0.10**	**0.20**	**0.30**	**0.40**	**0.50**	**0.60**	**0.70**	**0.80**	**0.90**	**1.00**
LED24 (n_{min} = 200)										
Accuracy (%)	75.88	76.97	77.14	77.42	77.47	77.50	77.56	77.56	77.56	77.56
#Leaf	143.00	522.00	1124.00	1857.00	2618.00	3723.00	3743.00	3743.00	3743.00	3743.00
Time (sec)	8.70	9.91	10.92	11.51	11.92	12.78	12.32	12.32	12.43	12.45
Waveform21 (n_{min} = 200)										
Accuracy	85.00	86.53	86.61	86.72	86.72	86.72	86.72	86.72	86.72	86.72
#Leaf	506.00	1565.00	2492.00	2623.00	2623.00	2623.00	2623.00	2623.00	2623.00	2623.00
Time	17.80	18.50	18.89	18.69	18.77	18.72	18.53	18.72	18.74	18.72
n_{min}	**200**	**300**	**400**	**500**	**600**	**700**	**800**		**900**	**1000**
LED24 (τ = 0.7)										
Accuracy	77.5611	77.5867	77.4565	77.3472	77.2557	77.1417	77.1412		77.0847	76.9887
#Leaf	3743	2405	1826	1383	1244	1057	935		804	689
Time	11.66887	10.93567	10.40527	10.07766	9.469261	9.42246	9.032458		9.172859	8.642455
Waveform21 (τ = 0.4)										
Accuracy	86.7218	86.5226	86.3028	85.9499	85.9119	85.6378	85.6707		85.7318	85.2165
#Leaf	2623	1800	1363	1103	940	806	703		644	572
Time	18.31452	17.70611	17.34731	17.03531	16.84811	16.58291	16.61411		16.61411	16.2709

However, the only way to detect the best tie-breaking threshold for a certain task is trying all the possibilities in VFDT. It is impractical for real-world applications. In this chapter, we propose the adaptive tie-breaking threshold using the incremental optimization methodology. The breakthrough of our work is the optimized node-splitting control, which will be specified in the following sections.

Incrementally Optimized Decision Tree

Motivation and Overview

OVFDT, which is based on the original VFDT design, is implemented on a test-then-train approach (Fig. 6.1) for classifying continuously arriving data streams, even for infinite data streams. The whole test-then-train process is synchronized such that when the data stream arrives, one segment at a time, the decision tree is being tested first for prediction output and training (which is also known as updating) of the decision tree then occurs incrementally. The description of testing process will be explained in section "OVFDT Testing Approach" in detail, and the training process will be explained in section "OVFDT Training Approach". Ideally, the node-splitting test updates the tree model in order to improve the accuracy, while a bigger tree model takes longer computation time. The situation to do the node-splitting check

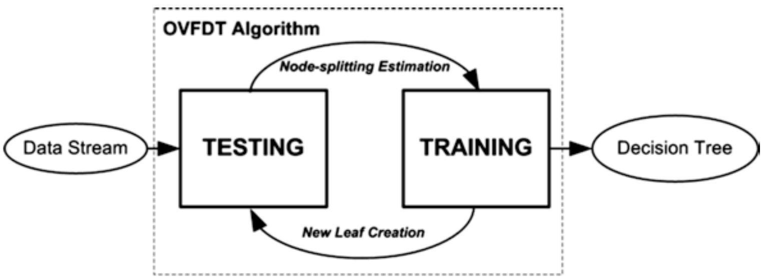

Fig. 6.1 A test-then-train OVFDT workflow

is when the number of instances in a leaf l is greater than the predefined value $n_{\min}$.

Imperfect data streams, including noisy data and bias class distribution, decline the performance of VFDT. Figure 6.2 shows the results of accuracy, tree size, and computation time using VFDT, the same dataset structure added with imperfect values. The ideal stream is free from noise and has a uniform proportion of class samples, which is rare in real world. From the experiment result comparing ideal data streams to imperfect data streams, we conclude lemma 1:

Lemma 1 Imperfections in data streams worsen the performance of VFDT. The tree size and the computation time are increased, but the accuracy is declined. In other words, the optimization goal is to increase the accuracy but not enlarge the tree size, within an acceptable computation time. Naturally a bigger tree size takes longer computation time. For this reason, the computation time is dependent on the tree size.

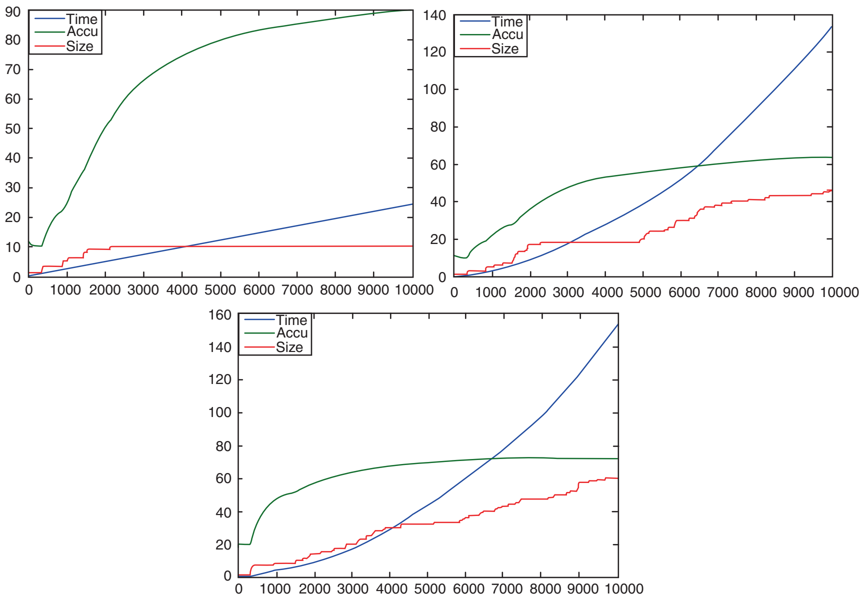

Fig. 6.2 VFDT performance for: (**a**) ideal data, (**b**) data with noise, (**c**) data with noise and bias. X-axis presents the accuracy and y-axis the number of samples

In the decision tree model, each path from the root to a leaf is considered as a way to present a rule. To ensure a high accuracy, there must be sufficient number of rules, which is the number of leaves in the tree model. Suppose Hoeffding Tree (HT) is the decision tree algorithm using Hoeffding bound (HB) as the node-splitting test. Let $\mathrm{Accu}\left(\mathrm{HT}_{m^{th}}\right)$ be the accuracy function for the decision tree structure HT at the *mth* node-splitting estimation, and let $\mathrm{Size}\left(\mathrm{HT}_{m^{th}}\right)$ be the tree size, then:

$$\mathrm{Accu}\left(\mathrm{HT}_{m^{th}}\right) = R\left(\mathrm{Size}\left(\mathrm{HT}_{m^{th}}\right)\right)$$

where $R(\cdot)$ is a mapping function of tree size to accuracy. Most incremental optimization functions can be expressed as the sum of several sub-objective functions:

$$\Phi(x) = \sum_{m=1}^{M} \Phi_m(x)$$

where $\Phi_m : \chi \subset \mathbb{R}^p \to \mathbb{R}$ is a continuously differentiable function whose domain χ is a nonempty, convex and closed set. We consider the following optimization problems:

$$\text{maximize}\ \Phi(x)\ \text{subject}\ to\ x \in \chi$$

Based on Lemma 1, we propose a solution to optimize the decision tree structure by improving the original VFDT that:

$$\Phi_m(x) = \frac{\mathrm{Accu}\left(\mathrm{HT}_{m^{th}+1}\right) - \mathrm{Accu}\left(\mathrm{HT}_{m^{th}}\right)}{\mathrm{Size}\left(\mathrm{HT}_{m^{th}+1}\right) - \mathrm{Size}\left(\mathrm{HT}_{m^{th}}\right)}$$

The tree model is updated when a node-splitting appears. Original VFDT considers the HB as the only index to split node. However, it is not enough. In terms of the above optimization goal, OVFDT proposes an optimized node-splitting control during the tree-building process.

OVFDT Test-Then-Train Process

Data streams are open-ended problem that traditional sampling strategies are not viable in the non-stopping streams scenario. OVFDT is an improved version of the original VFDT and its extensions using *HB* to decide the node-splitting. The most significant contribution is OVFDT that can obtain an optimal tree structure by balancing the accuracy and tree size. It is useful for data mining especially in the events of the tree size explosion, when the decision tree is subject to imperfect streams including noisy data and imbalanced class distribution.

HT algorithms run a test-then-train approach to build a decision tree mode. When new stream arrives, it will be sorted from the root to a predicted leaf. Comparing the predictive class to the true class of this data stream, we can maintain an error matrix for every tree leaf in the testing process. In terms of the stored statistics matrix, the decision tree model is being updated in the training process. Table 6.2 presents the differences between OVFDT and HT algorithms (including the original VFDT and its extensions). Figure 6.3 shows the input parameters and the output of OVFDT and the approach presented as pseudo code.

Table 6.2 The comparison between VFDT and OVFDT

Approach	Hoeffding Tree algorithms	OVFDT
Testing	Sort the new stream by current HT Update the sufficient statistics Construct FTL, by MC, NB, or ADP classifier Assign a predicted class by FTL	Sort the new stream by current HT Update the sufficient statistics Construct FTL, by MC, NB, <u>WNB,</u> or <u>ADP</u> classifier Assign a predicted class by FTL <u>Update incremental sequential-error</u>
Training	Check node-splitting by HB Check node-splitting by fixed τ HT update	Check node-splitting by HB Check node-splitting by <u>adaptive</u> τ Check node-splitting by <u>incremental sequential-error</u> HT update

INPUT:
S : A stream of sample
X : A set of symbolic attributes
G(.) : Heuristic function using for node-splitting estimation
δ : One minus the desired probability of choosing a correct attribute at any given node
n_{min} : The minimum number of samples between check node-splitting estimation
$\mathcal{F}$: A functional tree leaf strategy
OUTPUT:
HT : A decision tree

PROCEDURE: *OVFDT(S, X, G(.), δ, n_{min}, $\mathcal{F}$)*
1. A data stream *S* arrives
2. IF *HT* is null, THEN ***initializeHT(S, X, G(.), δ, n_{min}, $\mathcal{F}$)***
   ELSE ***traverseHT(S, HT, $\mathcal{F}$)*** and update *ΔC*
3. Label *l* as the predicted class among the samples seen so far at the leaf *l*.
4. Let n_l be the number of samples seen at the leaf *l*.
5. IF the samples seen so far at leaf *l* do not all belong to the same class, and (n_l mod n_{min}) is zero, THEN ***doNodeSplittingEstimation(ΔC, S, X, G(.), δ, n_{min})***
6. Return *HT*

Fig. 6.3 Pseudo code of input and the test-then-train approach

OVFDT Testing Approach

Suppose X is a vector of d attributes, and y is the class with k different values included in the data streams. For decision tree prediction learning tasks, the learning goal is to induce a function of $\hat{y}_k = HT_{\mathrm{F}}(X)$, where $\hat{y}_k$ is the predicted class by Hoeffding Tree (HT) according to a functional tree leaf strategy $\mathcal{F}$. When a new data stream (X, y_k) arrives, it traverses from the root of the decision tree to an existing leaf by the current decision tree structure, provided that the root has existed initially. Otherwise, the heuristic function is used to constructs a tree model with a single root node.

When new instance comes, it will be sorted from the root to a leaf by the current tree model. The classifier on the leaf can further enhance the prediction accuracy via the embedded Naïve Bayes classifier. OVFDT embed four different classifiers $\mathcal{F}$ to improve the performance of prediction. They are *Majority Class* ($\mathcal{F}^{\mathrm{MC}}$), *Naïve Bayes* ($\mathcal{F}^{\mathrm{NB}}$), *Weighted Naïve Bayes* ($\mathcal{F}^{\mathrm{WNB}}$) and *Adaptive* ($\mathcal{F}^{\mathrm{Adaptive}}$).

Suppose $\hat{y}_k$ the predicted class value and y_k is actual class in data streams with a vector of attribute X. A sufficient statistics matrix stores the number of passed-by samples, which contain attribute X_i with a value x_{ij} belonging to a certain y_k so far. We call this statistics table Observed Class Distribution (OCD) matrix. The size of OCD is $J \times K$, where J is the total number of distinct values for attribute X_i and K is the number of distinct class values. Suppose n_{ijk} is the sufficient statistic that reflects the number of attribute X_i with a value x_{ij} belonging to class y_k. Therefore, OCD on node X_i is defined as:

$$\mathrm{OCD}_{X_i} = \begin{bmatrix} n_{i11} & \cdots & n_{iJ1} \\ \vdots & \ddots & \vdots \\ n_{i1K} & \cdots & n_{iJK} \end{bmatrix}$$

For a certain leaf that attribute X_i with a value of x_{ij}:

$$\mathrm{OCD}_{x_{ij}} = \left\{ n_{ij1} \ \ldots \ n_{ijK} \right\}$$

Majority Class classifier chooses the class with the maximum value as the predicted class in a leaf. Thus, F^{MC} predicts the class with a value that:

$$\arg\max k = \left\{ n_{ij1} \quad \ldots \quad n_{ijk} \quad \ldots \quad n_{ijK} \right\}$$

Naïve Bayes classifier chooses the class with the maximum possibility computed by *Naïve Bayes*, as the predictive class in a leaf. The formula of *Naïve Bayes* is:

$$p_{ijk} = \frac{\mathrm{P}\left(x_{ij} | y_k\right) \cdot \mathrm{P}\left(y_k\right)}{\mathrm{P}\left(x_{ij}\right)}$$

OCD of leaf with value x_{ij} is updated incrementally. Thus, F^{NB} predicts the class with a value that:

$$\arg\max k = \left\{ p_{ij1} \quad \ldots \quad p_{ijk} \quad \ldots \quad p_{ijK} \right\}$$

Weighted Naïve Bayes classifier proposes to reduce the effect of imbalanced class distribution. It chooses the class with the maximum possibility computed by weighted *Naïve Bayes*, as the predictive class in a leaf:

$$p_{ijk} = \omega_{ijk} \frac{\mathrm{P}\left(x_{ij} | y_k\right) \cdot \mathrm{P}\left(y_k\right)}{\mathrm{P}\left(x_{ij}\right)} \text{ where } \omega_{ijk} = \frac{n_{ijk}}{\sum_{k=1}^{K} n_{ijk}}$$

OCD of leaf with value x_{ij} is updated. Thus, $\mathcal{F}^{\mathrm{WNB}}$ predicts the class with a value that:

$$\arg\max k = \left\{ p_{ij1} \quad \ldots \quad p_{ijk} \quad \ldots \quad p_{ijK} \right\}$$

Adaptive classifier chooses the classifier with the least error from the alternative $\mathcal{F}^{\mathrm{MC}}$, $\mathcal{F}^{\mathrm{NB}}$ and $\mathcal{F}^{\mathrm{WNB}}$. For each time classifier is implemented on the leaf, suppose γ is the index of classifier implementation on leaf assigned to x_{ij}, and suppose Γ is the total number of implementation, where $\Gamma = \sum_{k=1}^{K} n_{ijk}$. The error of a classifier $\mathcal{F}$ to class y_k is calculated by:

$$\mathrm{Err}\left(\mathcal{F}, y_k\right) = \sum_{\gamma=1}^{\Gamma} \mathrm{Error}_k^{\gamma}, \text{where } \mathrm{Error}_k^{\gamma} = \begin{cases} 1, \text{if } \hat{y}_k \neq y_k \\ 0, \text{otherwise} \end{cases}$$

Therefore, $\mathcal{F}^{\mathrm{Adaptive}}$ predicts the class with a value that is chosen by the classifier $\mathcal{F}$ with minimum error:

$$\arg\min \mathcal{F} = \left\{ \mathrm{Err}\left(\mathcal{F}^{\mathrm{MC}}, y_k\right), \quad \mathrm{Err}\left(\mathcal{F}^{\mathrm{NB}}, y_k\right), \quad \mathrm{Err}\left(\mathcal{F}^{\mathrm{WNB}}, y_k\right) \right\}$$

After the stream traverses the whole HT, it is assigned to a predicted class $\hat{y}_k$, which $\hat{y}_k \leftarrow Classifier(HT, \mathcal{F}, X)$ according to the functional tree leaf $\mathcal{F}$. Comparing the predicted class $\hat{y}_k$ to the actual class y_k, the statistics of correctly C_T and incorrectly C_F prediction are updated immediately. Meanwhile, the sufficient statistics n_{ijk}, which is a count of attribute x_i with value j belongs to class y_k, are updated in each node. This series of actions is so called a testing approach in this chapter. Figure 6.4 gives the pseudo code of this approach. According to the functional tree leaf strategy, the current HT sorts a newly arrived sample (X, y_k) from the

PROCEDURE: *traverseHT(S, HT, F)*

1. Sort S from the root to a leaf by HT. Update OCD in each node: $n_{ijk}(l)$ ++
2. Switch ($\mathcal{F}$)
3. Case $\mathcal{F}^{MC}$: predict the class y'_k with max $n_{ijk}(l)$
4. Case $\mathcal{F}^{NB}$: predict the class y'_k with max NB prob.
5. Case $\mathcal{F}^{WNB}$: predict the class y'_k with max WNB prob.
6. Case $\mathcal{F}^{Adaptive}$: predict the class y'_k using $\mathcal{F}$ with $\text{Error}_{\min}$
7. IF y'_k equals to the actual class label in S, THEN C_T ++
8. ELSE C_F ++
9. $\Delta C = C_T - C_F$
10. Return ΔC

Fig. 6.4 Pseudo code of testing approach

root to a predicted leaf $\hat{y}_k$. Comparing the predicted class $\hat{y}_k$ to the actual class y_k, the sequential-error statistics of C_T and C_F prediction are updated immediately.

To store OCD for OVFDT, $\mathcal{F}^{\mathrm{MC}}$, $\mathcal{F}^{\mathrm{NB}}$, and $\mathcal{F}^{\mathrm{WNB}}$ require memory proportional to $O\ (N \cdot I \cdot J \cdot K)$, where N is the number of nodes in tree model; I is the number of attributes; J is the maximum number of values per attribute; K is the number of classes. OCD of $\mathcal{F}^{\mathrm{NB}}$ and $\mathcal{F}^{\mathrm{WNB}}$ are converted from that of $\mathcal{F}^{\mathrm{MC}}$. In other words, we don't require extra memory to store three different OCD for $\mathcal{F}^{\mathrm{Adaptive}}$ respectively. When required, it can be converted from $\mathcal{F}^{\mathrm{MC}}$.

OVFDT Training Approach

Right after the testing approach, the training follows. Node-splitting estimation is used to initially decide if HT should be updated or not; that depends on the amount of samples received so far that can potentially be represented by additional underlying rules in the decision tree. In principle, the optimized node-splitting estimation should apply on every single new sample that arrives. Of course this will be too exhaustive, and it will slow down the tree-building process. Instead, a parameter $n_{\min}$ is proposed in VFDT that only do the node-splitting estimation when $n_{\min}$ examples have been observed on a leaf. In the node-splitting estimation, the tree model should be updated when a heuristic function $H(\cdot)$ chooses the most appropriate attribute with highest heuristic function value $H(x_a)$ as a node-splitting according to HB and tie-breaking threshold. The heuristic function is implemented as an information gain here. This situ of node-splitting estimation constitutes to the so-called training phase.

The node-splitting test is modified to use a dynamic tie-breaking threshold τ, which restricts the attribute splitting as a decision node. The τ parameter traditionally is pre-configured with a default value defined by the user. The optimal value is usually not known until all of the possibilities in an experiment have been tried. An example has been presented in section "Relationship Among Accuracy, Tree Size, and Time". Longitudinal testing of different values in advance is certainly not favorable in real-time applications. Instead, we assign a dynamic tie threshold, equal to the dynamic mean of *HB* at each pass of stream data, as the splitting threshold, which controls the node-splitting during the tree-building process. Tie-breaking that occurs close to the *HB* mean can effectively narrow the variance distribution. *HB* mean is calculated dynamically whenever new data arrives.

The estimation of splits and ties is only executed once for every $n_{\min}$ (a user-supplied value) samples that arrive at a leaf. Instead of a pre-configured tie, OVFDT uses an adaptive tie that is calculated by incremental computing. At the i_{th} node-splitting estimation, the HB estimates the sufficient statistics for a large enough sample size to split a new node, which corresponds to the leaf l. Let T_l be an adaptive tie corresponding to leaf l, within k estimations seen so far. Suppose μ_l is a binary variable that takes the value of 1 if *HB* relates to leaf l, and 0 otherwise. T_l is computed by:

$$T_l = \frac{1}{k}\sum_{i=1}^{k} \mu_l \times HB_i$$

To constrain HB fluctuation, an upper bound TlUPPER and a lower bound TlLOWER are proposed in the adaptive tie mechanism. The formulas are:

$$T_l^{\text{UPPER}} = \arg\max T_l$$

$$T_l^{\text{LOWER}} = \arg\min T_l$$

For lightweight operations, we propose an error-based pre-pruning mechanism for OVFDT, which stops non-informative split node before it splits into a new node. The pre-pruning takes into account the node-splitting error both globally and locally.

According to the optimization goal mentioned in section "Motivation and Overview", besides the HB, we also consider the global and local accuracy in terms of the sequential-error statistics of C_T and C_F prediction computed by functional tree leaf. Let ΔC_m be the difference between C_T and C_F, and m is the index of testing approach. Then ΔC_m is computed by (6.4), which reflects the global accuracy of the current HT prediction on the newly arrived data streams. If $\Delta C_m \geq 0$, the number of correct predictions is no less than the number of incorrect predictions in the current tree structure; otherwise, the current tree graph needs to be updated by node-splitting. In this approach, the statistics of correctly C_T and incorrectly C_F prediction are updated. Suppose $\Delta C_m = C_{T\,-}\ C_F$, which reflects the accuracy of HT. If ΔC declines, it means the global accuracy of current HT model worsens. Likewise, compare ΔC_m and ΔC_{m+1}, the local accuracy is monitored during the node-splitting. If ΔC_m is greater than ΔC_{m+1}, it means the current accuracy is declining locally. In this case, the HT should be updated to suit the newly arrival data streams.

Lemma 2 Monitor Global Accuracy *The model's accuracy varies whenever a node splits and the tree structure is updated. Overall accuracy of current tree model is monitored during node-splitting by comparing the number of correctly and incorrectly predicted samples. The number of correctly predicted instances and otherwise is recorded as global performance indicators so far. This monitoring allows the global accuracy to be determined.*

Lemma 3 Monitor Local Accuracy *The global accuracy can be tracked by comparing the number of correctly predicted samples with the number of wrongly predicted ones. Likewise, comparing the global accuracy measured at the current node-splitting estimation with the previous splitting, the increment* *in accuracy is being tracked dynamically. This monitoring allows us to check whether the current node-splitting is advantageous at each step by comparing with the previous step.*

Figure 6.5 gives an example why our proposed pre-pruning takes into account both the local and the global accuracy in the incremental pruning. At the i_{th} node-splitting estimation, the difference between correctly and incorrectly predicted classes was ΔC_i, and ΔC_{i+1} was at $i+1_{\text{th}}$ estimation. $(\Delta C_i - \Delta C_{i+1})$ was negative that the local accuracy of $i+1_{\text{th}}$ estimation was worse than its previous one, while both were on a global increasing trend. Hence, if accuracy is getting worse, it is necessary to update the HT structure.

Combining the prediction statistics gathered in the testing phase, Fig. 6.6 presents the pseudo code of the training phase in OVFDT in building an upright tree. The optimized node-splitting control is presented in Fig. 6.6 Line 7. In each node-splitting estimation process, *HB* value that relates to a leaf *l* is recorded. The recorded *HB* values are used

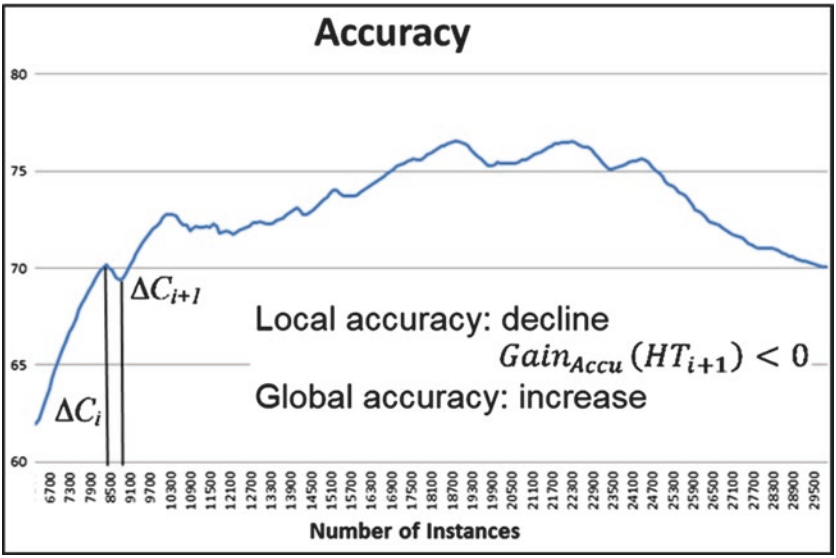

Fig. 6.5 Example of incremental pruning

PROCEDURE *doNodeSplittingEstimation*(ΔC, S, X, $G(.)$, δ)

1. FOR each attribute $X_i \in X_l$ - $\{X_\emptyset\}$ at the leaf l
2. Compute $G_l(X_i)$
3. Let X_a be the attribute with highest $G_l(.)$ and X_b with the 2nd highest $G_l(.)$
4. Compute HB with δ
5. Let $\Delta G_l = G_l(X_a) - G_l(X_b)$
6. END-FOR
7. IF ($\Delta G_l > HB$) or ($\Delta G \leq T_l^{LOWER}$ and $\Delta C_i < \Delta C_{i-1}$) *or* ($\Delta G \leq T_l^{LOWER}$ and $\Delta C_i < 0$) or ($T_l^{LOWER} < \Delta G \leq T_l^{UPPER}$ and $\Delta C_i < \Delta C_{i-1}$)
8. Replace l by an internal node splits on X_a
9. Update Adaptive tie T_l^{LOWER} and T_l^{UPPER}
10. FOR each branch of splitting
11. Add a new leaf l_m and let $X_m = X - \{X_a\}$
12. Let $G(X_\emptyset)$ be $G(.)$ obtained by predicting the class in S, according to F at l_m
13. FOR each class y_k and each value x_{ij} of each attribute
14. $X_i \in X_m - \{X_\emptyset\}$ and reset OCD: $n_{ijk}(l) = 0$
15. END-FOR
16. END-FOR
17. END-IF
18. Return updated HT

Fig. 6.6 Pseudo code of training approach

to compute the adaptive tie, which uses the mean of *HB* to each leaf *l*, instead of a fixed user-defined value in VFDT.

Evaluation

Evaluation Platform and Datasets

A Java package with OVFDT and an embedded MOA toolkit was constructed as a simulation platform for experiments. The running environment was a Windows 7 PC with Intel Quad 2.8GHz CPU and 8G RAM. In all of the experiments, the parameters of the algorithms were $\delta = 10^{-6}$ and $n_{\min} = 200$, which are default values suggested by MOA. *δ is the allowable error in split decision and values closer to zero will take longer to decide; $n_{\min}$ is the number of instances a leaf should observe between split attempts*. The main goal of this section is to provide evidence of the improvement of OVFDT compared to the original VFDT.

The experimental datasets, including pure nominal datasets, pure numeric datasets, and mixed datasets, were either synthetics generated by the MOA Generator or extracted from real-world applications that are publicly available for download from the UCI repository. The descriptions of each experimental dataset are listed in Table 6.3. The generated datasets were also used in previous VFDT-related studies.

The testing ran using a test-then-train approach that is common in stream mining. When a new instance arrived that represented a segment of the incoming data stream, it was sorted by the tree model into a predicted class. This was the testing approach for deriving the predicted class via the latest form of the decision tree. Compared to the actual class label it belonged to, the tree model was updated in the training approach because the prediction accuracy was known. The decision tree can typically take either form of incoming data instances; if the instances are labeled they will be used for training or learning for the decision tree to update itself. If unlabeled, the instances are taken as unseen samples and a prediction is made in the testing phase. In our experiments, all instances were labeled because our objective was to measure the performance of model learning and prediction accuracy.

Synthetic Data *LED24* was generated by MOA. In the experiment, we added 10% noisy data to simulate imperfect data streams. The LED24 problem used 24 binary attributes to classify 10 different classes. *Waveform* was generated by the MOA Generator. The dataset was donated by David

Table 6.3 Description of experimental datasets

Name	*Nom#*	*Num#*	*Cls#*	*Type*	*Instance#*
LED24 10% Noise	24	0	10	Synthetic	10^6
Waveform 21	0	21	3	Synthetic	10^6
Waveform 40	0	40	3	Synthetic	10^6
Random Tree Simple (RTS)	10	10	2	Synthetic	10^6
Random Tree Complex (RTC)	50	50	2	Synthetic	10^6
RBF Simple (RBFS)	0	10	2	Synthetic	10^6
RBF Complex (RBFC)	0	50	2	Synthetic	10^6
Connect-4	42	0	7	UCI	67,557
Person Activity Data (PAD)	2	3	11	UCI	164,860
Cover Type (COVTYPE	42	12	7	UCI	581,012
Nursery	8	0	5	UCI	12,960

Aha to the UCI repository. The goal of the task was to differentiate between three different classes of Waveform. There were two types of waveform: Wave21 had 21 numeric attributes and Wave40 had 40 numeric attributes, all of which contained noise. *Random Tree (RTS and RTC)* was also generated by the MOA Generator. It built a decision tree by choosing attributes to split randomly and assigning a random class label to each leaf. As long as a tree was constructed, new samples were generated by assigning uniformly distributed random values to attributes. Those attributes determined the class label through the tree. *Radial basis Function (RBFS and RBFC)* is a fixed number of random centroids generator. A random position, a single standard deviation, a class label and weight are generated by a centroid.

UCI Data *Connect-4* contained all of the legal 8-ply positions in a two-player game of Connect-4. In the game, the player's next move was not forced and the game was won once four chessmen were connected. *Personal activity data (PAD)* recorded the data streams collected from 4 sensors on the players' bodies. Each sensor collected 3 numeric data. *Cover Type* was used to predict forest cover type from cartographic variables. *Nursery* was from a hierarchical decision model originally developed to rank applications for nursery schools. Because of its known underlying concept structure, this dataset can be useful for testing constructive learning induction and structure discovery algorithms.

Accuracy Comparison

Table 6.4 shows the comparison results of the accuracy tests. On average, OVFDT obtained a higher accuracy for the pure nominal datasets than the mixed datasets. For the numeric Waveform datasets, OVFDT also displayed better accuracy than the other VFDTs. This phenomenon was particularly obvious in OVFDT with the adaptive Functional Tree Leaf. For each dataset, a detailed comparison of its accuracy with the new arrival data streams is illustrated in the Appendix.

Table 6.4 Accuracy (%) comparison

	VFDT tie 0.05				VFDT tie 0.5				OVFDT			
Methods	MC	NB	WNB	ADP	MC	NB	WNB	ADP	MC	NB	WNB	ADP
Datasets												
LED_NP10	63.781	73.73	73.55	73.88	73.68	72.53	73.71	73.26	73.636	73.44	73.82	**73.89**
CONNECT-4	67.3	72.28	73.48	73.48	69.01	72.87	74.26	74.09	69.01	73.55	74.58	**74.98**
NURSERY	83.78	88.5	88.21	89.07	82.09	89.16	86.81	89.82	82.75	89.24	87.56	**90.29**
Nominal AVG	71.62	78.17	78.41	78.81	74.93	78.19	78.26	79.06	75.13	78.74	78.65	**79.72**
WAVE21	76.36	83.15	84.55	83.91	80.9	82.02	82.59	82.47	78.6	83.24	84.47	**84.57**
WAVE40	76.42	83.14	84.37	83.75	80.89	81.31	81.79	81.90	79.1	83.16	84.31	**84.77**
RBFS	82.62	85.09	85.58	86.47	87.71	89.18	89.43	**89.47**	84.87	85.63	86.68	87.73
RBFC	80.75	90.01	90.71	90.74	88.38	92.72	**92.88**	92.82	78.67	89.09	89.45	89.32
Numeric AVG	79.04	85.35	86.30	86.22	84.47	86.31	**86.67**	**86.67**	80.31	85.28	86.23	86.60
RTS	91.78	94.86	94.24	94.77	93.16	95.8	95.52	95.72	92.41	95.56	95.1	**95.84**
RTC	95.14	95.62	95.59	95.63	95.55	95.72	95.71	95.74	95.4	95.67	95.63	**95.76**
COVTYPE	67.45	77.16	78.6	77.77	74.19	90.76	95.52	95.71	92.41	95.55	95.1	**95.84**
PAD	43.75	61.04	59.69	61.01	55.9	72.65	71.56	71.29	51.22	71.07	70.05	**72.7**
Mixed AVG	74.53	82.17	82.03	82.30	79.70	88.73	89.58	89.62	82.86	89.46	88.97	**90.04**
AVG	75.06	81.90	82.25	82.44	79.70	84.41	84.84	85.11	79.43	84.50	84.62	**85.45**

Note: The figures in bold are the greatest numbers obtained per row of algorithms experimented with a specific dataset

Our comparison of the four functional tree leaf strategies revealed that $OVFDT_{ADP}$ generally had the highest accuracy in the most experimental datasets. An improvement comparison of functional tree leaf strategies is given in Table 6.5. The majority class functional tree leaf strategy was chosen as a benchmark. As a result, the adaptive functional tree leaf strategy obtained the best accuracy, with $F^{\text{Adaptive}} > F^{\text{WNB}} > F^{\text{NB}} > F^{\text{MC}}$. This result appeared in both VFDT and OVFDT methods.

Tree Size Comparison

For all of the datasets, a comparison of tree size is shown in Table 6.6. For the pure nominal and mixed datasets, VFDT with a smaller τ generally had smaller tree sizes, but OVFDT obtained the smallest tree size with the pure numeric datasets. For each dataset, a detailed comparison of accuracy with the new arrival data streams is illustrated in the Appendix. The charts in the Appendix essentially show the performance on the y-axis and the dataset samples on the x-axis.

Tree Learning Time Comparison

A comparison of tree learning time is shown in Table 6.7. For all of the datasets, the majority class functional tree leaf consumed the least time in this experiment due to its simplicity. The computation times of the other three Functional Tree Leaves, using the *Naïve Bayes* classifier, were close.

Stability of Functional Tree Leaf in OVFDT

Stability is related to the degree of variance in the prediction results. A stable model is translated into a useful model and its prediction accuracy over the same datasets does not vary significantly, regardless of how many times it is tested. To show the stability of different functional tree leaf mechanisms in OVFDT, we ran the evaluation based on those synthetic datasets ten times. In this experiment, the synthetic datasets were generated using different random seeds. Hence, the generated data streams had

Table 6.5 Accuracy improvement by Functional Tree Leaf

		VFDT tie 0.05				VFDT tie 0.5				OVFDT			
	Methods	MC	NB	WNB	ADP	MC	NB	WNB	ADP	MC	NB	WNB	ADP
Datasets													
Nominal	Accuracy %	71.62	78.17	78.41	78.81	74.93	78.19	78.26	79.06	75.13	78.74	78.65	79.72
	FL Imp. %	0.00	9.15	9.49	10.04	0.00	4.35	4.44	5.51	0.00	9.95	9.82	11.31
Numeric	Accuracy %	79.04	85.35	86.30	86.22	84.47	86.31	86.67	86.67	80.31	85.28	86.23	86.60
	FL Imp. %	0.00	7.98	9.19	9.08	0.00	2.18%	2.61%	2.60	0.00	6.19	7.37	7.83
Mixed	Accuracy %	74.53	82.17	82.03	82.30	79.70	88.73	89.58	89.62	82.86	89.46	88.97	90.04
	FL Imp. %	0.00	10.25	10.06	10.42	0.00	11.33	12.39	12.44	0.00	7.97	7.37	8.66

Table 6.6 Tree size comparison

	VFDT tie 0.05				VFDT tie 0.5				OVFDT			
Methods	MC	NB	WNB	ADP	MC	NB	WNB	ADP	MC	NB	WNB	ADP
Datasets												
LED_NP10	46	46	46	46	2440	2440	2440	2440	370	227	219	172
CONNECT-4	23	23	23	23	437	437	437	437	141	96	92	94
NURSERY	72	72	72	72	13	13	13	13	21	13	13	13
Nominal AVG	47	47	47	47	963	963	963	963	177	112	108	93
WAVE21	160	160	160	160	3557	3557	3557	3557	263	197	181	178
WAVE40	150	150	150	150	3607	2340	3607	3607	413	194	154	152
RBFS	420	420	420	420	2723	2723	2723	2723	398	372	389	391
RBFC	443	443	443	443	2052	3145	3145	3145	604	381	372	396
Numeric AVG	293	293	293	293	2985	2941	3258	3258	420	286	274	279
RTS	1680	1680	1680	1680	2683	2683	2683	2683	1940	1691	1812	1739
RTC	620	620	620	620	1492	1492	1492	1492	670	682	670	678
COVTYPE	127	127	127	127	1882	1882	2683	2683	1940	1691	1812	1739
PAD	167	167	167	167	1829	1614	1614	1614	905	950	855	953
Mixed AVG	649	649	649	649	1972	1918	2118	2118	1364	1254	1287	1277
AVG	330	330	330	330	1973	1941	2113	2113	654	551	556	550

Table 6.7 Tree learning time comparison

	VFDT tie 0.05				VFDT tie 0.5				OVFDT			
Methods	MC	NB	WNB	ADP	MC	NB	WNB	ADP	MC	NB	WNB	ADP
Datasets												
LED_NP10	8.25	13.15	13.23	14.76	14.13	18.21	18.08	19.11	10.58	23.15	22.82	20.63
CONNECT-4	1.28	1.53	1.48	1.53	1.67	1.86	1.83	1.93	1.50	1.95	1.97	2.17
NURSERY	0.34	0.36	0.36	0.37	0.33	0.34	0.34	0.37	0.37	0.39	0.39	0.39
Nominal AVG	3.29	5.01	5.02	5.55	5.38	6.80	6.75	7.14	4.15	8.50	8.39	7.73
WAVE21	17.13	23.18	23.06	26.89	21.54	24.84	24.96	26.94	18.08	34.27	34.38	42.15
WAVE40	30.87	41.90	42.07	49.34	38.19	45.05	45.57	48.61	33.06	64.23	63.37	76.70
RBFS	9.39	11.29	11.37	12.32	11.01	12.43	12.45	13.74	10.06	15.79	15.80	17.34
RBFC	38.20	47.94	47.99	55.43	44.40	52.20	51.71	57.28	42.51	69.40	69.64	86.32
Numeric AVG	23.90	31.08	31.12	36.00	28.79	33.63	33.67	36.64	25.93	45.92	45.80	55.63
RTS	15.43	17.58	17.83	18.78	16.80	18.39	18.30	19.39	16.19	21.93	21.84	22.11
RTC	64.66	74.60	75.19	78.74	67.60	76.47	77.06	80.61	67.03	96.58	96.21	92.84
COVTYPE	11.04	15.24	15.05	18.36	12.68	15.23	18.30	19.39	16.19	21.93	21.84	22.11
PAD	1.31	2.04	1.97	2.48	1.53	1.92	1.90	2.17	1.56	2.68	2.78	3.49
Mixed AVG	23.11	27.37	27.51	29.59	24.65	28.00	28.89	30.39	25.24	35.78	35.67	35.14
AVG	16.77	21.15	21.22	23.71	19.60	22.81	23.10	24.72	18.44	30.07	29.95	32.83

exactly the same data formats, but different random values. The average and its variance of accuracy in the testing are shown in Table 6.8. Generally, datasets that only contained the homogenous attribute types (numeric attributes only or nominal attributes only) had smaller variances. The proposed adaptive functional tree leaf obtained the least variances because it had more stable and comparable accuracy than the other functional tree leaves.

Optimal Tree Model

Figure 6.7 presents a comparison of the optima ratio, which was calculated by optimization function accuracy/tree size in (6.5). The higher this ratio is, the better the optimal result. Comparing VFDT to OVFDT, the ratios of OVFDT were clearly higher than those of VFDT. In other words, the optimal tree structures were achieved in the pure numeric and mixed datasets.

Conclusion

Imperfect data stream leads to tree size explosion and detrimental accuracy problems. In original VFDT, a tie-breaking threshold that takes a user-defined value is proposed to alleviate this problem by controlling the node-splitting process that is a way of tree growth. But there is no single default value that always works well and that user-defined value is static throughout the stream mining operation. In this chapter, we propose an extended version of VFDT which we called it Optimized-VFDT (OVFDT) algorithm that uses an adaptive tie mechanism to automatically search for an optimized amount of tree node-splitting, balancing the accuracy, the tree size and the time, during the tree-building process. The optimized node-splitting mechanism controls the attribute-splitting estimation incrementally. Balancing between the accuracy and tree size is important, as stream mining is supposed to operate in limited memory computing environment and a reasonable accuracy is needed. It is a known contradiction that high accuracy requires a large tree with many

Table 6.8 The average and variance of accuracy in four types of Functional Tree Leaves

	Average				Variance			
Method	MC	NB	WNB	ADP	MC	NB	WNB	ADP
Data								
Led_np10	73.6454	73.4987	73.8416	73.9080	0.0106	0.0067	0.0019	0.0027
Led_np15	61.9010	60.7587	61.8477	61.9785	0.0025	0.0255	0.0013	0.0014
Led_np20	51.1407	47.5278	50.7018	51.0752	0.0026	0.1113	0.0029	0.0011
Led_np25	41.4070	35.7713	39.9215	41.1294	0.0027	0.1118	0.0255	0.0029
AVG.Nom.	57.0235	54.3891	56.5782	57.0228	0.0046	0.0638	0.0079	0.0020
Waveform21	79.0604	83.2396	84.5739	84.5634	0.0785	0.0151	0.0025	0.0025
Waveform40	79.0791	83.1607	84.3698	84.3527	0.0697	0.0067	0.0051	0.0058
RBFS	89.8142	90.8911	91.6129	92.4029	2.3677	2.0308	2.3580	1.7588
RBFC	98.1270	98.0033	98.3046	98.9693	0.0710	0.1518	0.1020	0.0329
AVG.Num.	86.5202	88.8237	89.7153	90.0721	0.6467	0.5511	0.6169	0.4500
RTS	89.9290	93.0060	92.4382	93.0394	10.0790	9.6065	9.7757	9.3052
RTC	89.5048	84.9352	87.1039	87.4002	44.1081	84.7361	80.1807	80.7949
AVG.Mix.	89.7169	88.9706	89.7711	90.2198	27.0936	47.1713	44.9782	45.0501

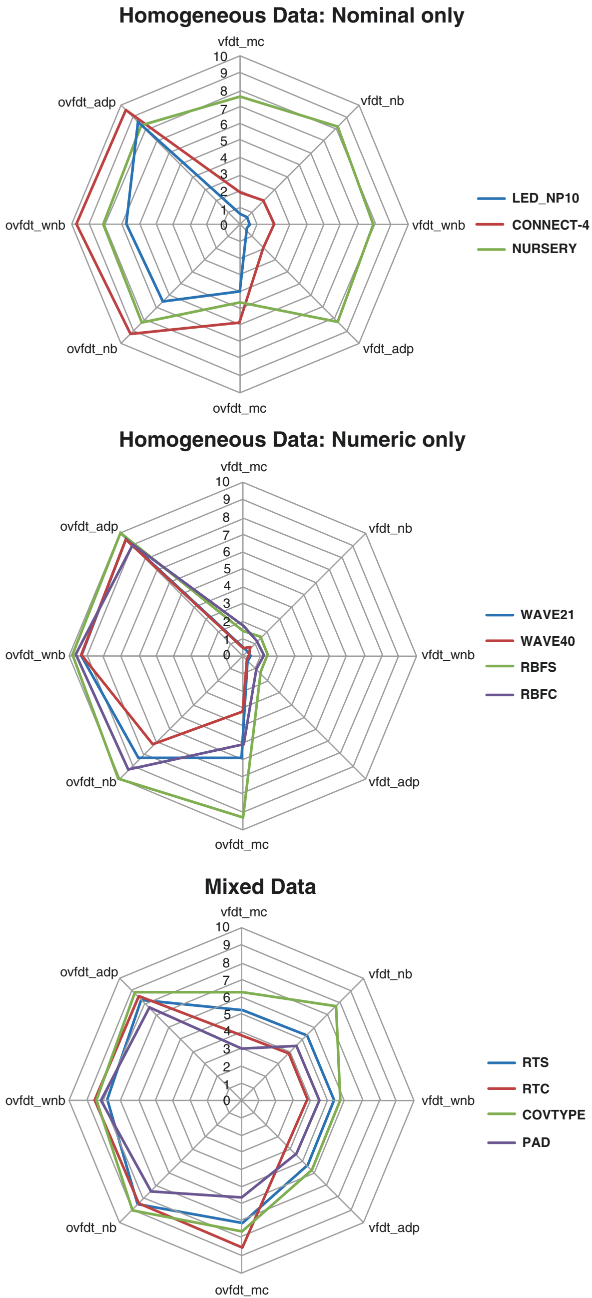

Fig. 6.7 Comparison of optimal tree structures between VFDT and OVFDT

decision paths, and too sparse the decision tree results in poor accuracy. The experiment results show that OVFDT meet the optimization goal and achieve a better performance gain ratio in terms of high prediction accuracy and compact tree size than the other VFDTs. That is, with the minimum tree size, OVFDT can achieve the highest possible accuracy. This advantage can be technically accomplished by means of simple incremental optimization mechanisms as described in this chapter. They are light-weighted and suitable for incremental learning. The contribution is significant because OVFDT can potentially be further modified into other variants of VFDT models in various applications, while the best possible (optimal) accuracy and minimum tree size can always be guaranteed.

Acknowledgment The authors are thankful for the financial support from the research grants "Temporal Data Stream Mining by Using Incrementally Optimized Very Fast Decision Forest (iOVFDF)", Grant no. MYRG2015-00128-FST offered by the University of Macau, FST, and RDAO, and "A scalable data stream mining methodology: stream-based holistic analytics and reasoning in parallel", Grant no. FDCT-126/2014/A3, offered by FDCT Macau.

Reference

Bifet, A., & Gavalda, R. (2007). Learning from Time-Changing Data with Adaptive Windowing. In *Proceedings of SIAM International Conference on Data Mining* (pp. 443–448).

Bifet A., Geoff, H., Bernhard, P., Jesse, R., Philipp, K., Hardy, K., Timm, J., & Thomas, S. (2001). MOA: A Real-Time Analytics Open Source Framework. In *Machine Learning and Knowledge Discovery in Databases* (pp. 617–620). Lecture Notes in Computer Science, Volume 6913/2011.

Bifet, A., Holmes, G., Pfahringer, B., Kirkby, R., & Gavalda, R. (2009). New Ensemble Methods for Evolving Data Streams. In *Proceedings 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 139–147). New York: ACM.

Elomaa, T. (1999). *The Biases of Decision Tree Pruning Strategies, Advances in Intelligent Data Analysis* (pp. 63–74). Lecture Notes in Computer Science, Volume 1642/1999. Berlin/Heidelberg: Springer.

Gama, J., & Kosina, P. (2011). Learning Decision Rules from Data Streams. In T. Walsh (Ed.), *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence – Volume Two* (Vol. 2, pp. 1255–1260). Menlo Park: AAAI Press.

Gama J, Rocha R., & Medas P. (2003). Accurate Decision Trees for Mining High-Speed Data Streams. In *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 523–528). ACM, New York.

Geoffrey H., Richard K., & Bernhard P. (2005). Tie Breaking in Hoeffding Trees. In *Proceedings Workshop W6: Second International Workshop on Knowledge Discovery in Data Streams* (pp. 107–116).

Hartline J. R. K. (2008). *Incremental Optimization* (PhD Thesis). Faculty of the Graduate School, Cornell University.

Hashemi, S., & Yang, Y. (2009). Flexible Decision Tree for Data Stream Classification in the Presence of Concept Change, Noise and Missing Values. *Data Mining and Knowledge Discovery, 19*(1), 95–131.

Hulten G., & Domingos P. (2003). *VFML – A Toolkit for Mining High-Speed Time-Changing Data Streams*. http://www.cs.washington.edu/dm/vfml/

Hulten G., Spencer L., & Domingos P. (2001). Mining Time-Changing Data Streams. In *Proceedings* of the *7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 97–106).

Mladenic D., & Grobelnik M. (1999). Feature Selection for Unbalanced Class Distribution and Naive Bayes, In *Proceeding ICML '99 Proceedings of the Sixteenth International Conference on Machine Learning* (pp. 258–267). ISBN 1-55860-612-2, Morgan Kaufmann.

Nitesh, C., Nathalie, J., & Alek, K. (2004). Special Issue on Learning from Imbalanced Data Sets. *ACM SIGKDD Explorations, 6*(1), 1–6.

Oza N., & Russell S. (2001). Online Bagging and Boosting. In *Artificial Intelligence and Statistics* (pp. 105–112). San Mateo: Morgan Kaufmann.

Pedro D., & Geoff H. (2000). Mining High-Speed Data Streams. In *Proceeding of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 71–80).

Pfahringer B., Holmes G., & Kirkby R. (2007). New Options for Hoeffding Trees. In *Proceedings in Australian Conference on Artificial Intelligence* (pp. 90–99).

Stefan H., Russel P., & Yun S. K. (2009). *CBDT: A Concept Based Approach to Data Stream Mining* (pp. 1006–1012). Lecture Notes in Computer Science, Volume 5476/2009.

Yang H., & Fong S. (2011). Moderated VFDT in Stream Mining Using Adaptive Tie Threshold and Incremental Pruning. In *Proceedings of the 13th International Conference on Data Warehousing And Knowledge Discovery* (pp. 471–483). Berlin/Heidelberg: Springer-Verlag.

7

Bass Model with Explanatory Parameters

Mladen Sokele and Luiz Moutinho

Introduction

The diffusion of innovation and new technology, market adoption of consumer durables, products/services that do not include repeat sales, subscription services (e.g. telecommunications services) and allocations of restricted resources are examples of the S-shaped bounded growth. In the rest of the text the focus will be on these products/services. In general, during its life cycle, after the design phase, every product/service passes through the following phases: introduction, growth, maturity and decline, resembling the profile of the technology life cycle and its associated market-growth profile. The understanding of each segment of the product/service life cycle (P/SLC) for the business planning purposes is especially important in highly competitive market environment and for products/services resulting from emerging technologies (e.g. ICT services).

M. Sokele (✉)
Faculty of Electrical Engineering and Computing, University of Zagreb, Zagreb, Croatia

L. Moutinho
University of Suffolk, Suffolk, England, UK

The University of the South Pacific, Suva, Fiji

L. Moutinho, M. Sokele (eds.), *Innovative Research Methodologies in Management*,
https://doi.org/10.1007/978-3-319-64394-6_7

Only at the beginning of the P/SLC, there is no interaction with other products/services regarding market adoption; therefore, its growth may be described with simple S-shaped growth models such as Logistic model, Bass model and Richards model. These models are widely used in quantitative research for time series data analysis and enable better understanding of forces that influence growth in a sense of its dynamics, market capacities as well as forecasting of a future growth (Meade and Islam 2006).

In general, **quantitative** growth forecasting relies on the principle that a growth model will be valid in the perceivable future, and forecasting result could be obtainable by extrapolation of the observed time series data sequentially through time and supplementary information. In general, this principle is valid only for stable markets (e.g. same market segment boundaries, customer base, competition, cause-and-effect among similar products/services, etc.) without changes of external influences (e.g. technology, macroeconomics, purchasing power, regulation, etc.) and without changes of internal business operations (strategy, business plan, resources, ability of vendors and partners, etc.) (Sokele 2015).

To encompass market changes that can be perceived only by the **qualitative** forecasting methods, the growth models should be able to accept environmental variables and information from business operations as explanatory model parameters (Meade and Islam 2006). The ability of growth model to accept such qualitative estimated inputs is especially important in cases of forecasting prior to product/service launch where little or no time series data is available (Sönke 2004). The result could be growth model as an optimal combination of qualitative and quantitative methods for the forecasting purposes, which flowchart is presented in Fig. 7.1.

Introduction of explanatory parameters $\{\beta_i\}$, obtained by qualitative forecasting, reduces the number of $\{\alpha_i\}$ parameters in the growth model that needed to be determined from time series data fitting. Moreover, the growth models should accept auxiliary parameters $\{\gamma_i\}$ that do not make the model more complex but enable customisation of the model to the specific practical requirements. The above described growth model has the following general form:

$$N(t) = f(t;\ \{\alpha_i\},\ \{\beta_i\},\ \{\gamma_i\})$$

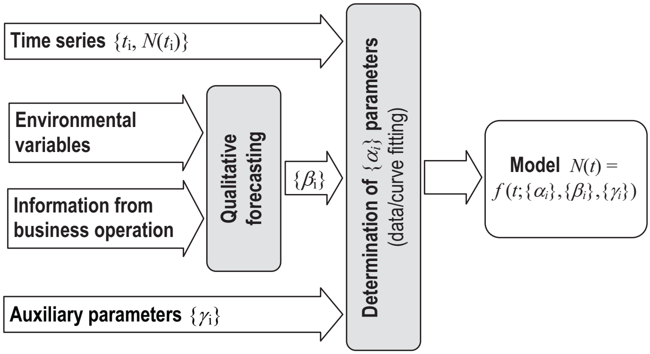

Fig. 7.1 Flowchart of growth model for the growth forecasting purposes

where

- $N(t)$ Cumulative volume (sales) of market adopted products/services (or similar) at time t;
- $\{\alpha_i\}$ Set of model parameters—resulting primarily from fit of time series data $\{t_i, N(t_i)\}$, $i = 1,\ldots,n$;
- $\{\beta_i\}$ Set of explanatory parameters—resulting from qualitative forecasting, for example: t_s time of launch; (t_e, N (t_e)) target point in the future; M (current) market capacity of product/service; t_m time of sales maximum and so on;
- $\{\gamma_i\}$ Set of auxiliary parameters in model which allows forecasting practitioner to customise model to her/his specific needs.

An example of simple growth model that accepts explanatory parameters is Logistic model through two fixed points (Meyer and Ausubel 1999). All parameters of the model are explanatory: market capacity, time when product/service has starting penetration level, time period needed for penetration grows from the starting penetration level to the saturation level and auxiliary parameters: chosen level of starting and saturation penetration. The model is useful for forecasting of new products/service adoption prior to launch assuming M, and two points but due to the limitations of logistic model. However, it is unsuitable for modelling of the beginning phase of P/SLC as well as the point of sales maximum (inflexion point) is fixed at a half of the market capacity.

Bass Model

In distinction from the Logistic growth model $L(t)$, the Bass model $B(t)$ introduces the effect of innovators via coefficient of innovation p, in differential equation of growth (7.1) which corrected deficiency of simple logistic growth ("hardly starts to grow up" problem and that t for which $L(t) = 0$ does not exist). The model considers a population of M adopters who are both innovators (with a constant propensity to purchase) and imitators (whose propensity to purchase is influenced by the amount of previous purchasing) (Sokele 2008).

$$\frac{dB(t)}{dt} = \underbrace{qB(t)\left(1 - \frac{B(t)}{M}\right)}_{\substack{\text{Effect of imitators} \\ \text{(Logistic growth)}}} + \underbrace{p\left(M - B(t)\right)}_{\substack{\text{Effect of} \\ \text{innovators}}} \tag{7.1}$$

Solution of differential Eq. (7.1) gives Bass diffusion model (7.2) defined by four parameters: M, market capacity; p, coefficient of innovation, $p > 0$; q, coefficient of imitation, $q \geq 0$; and t_s, time when product/service is introduced, $B(t_s) = 0$. To emphasise model dependence of its parameters, it is convenient to indicate the model as $B(t;\, M,\, p,\, q,\, t_s)$, $t \geq t_s$:

$$B\left(t; M, p, q, t_\mathrm{s}\right) = M \frac{1 - e^{-(p+q)(t-t_\mathrm{s})}}{1 + \frac{q}{p} e^{-(p+q)(t-t_\mathrm{s})}} \tag{7.2}$$

The Bass model has a shape of S-curve, identical to the Logistic growth model, but shifted down on y-axis. Figure 7.2 shows the effects of different values of parameters p and q on form of S-curve, with fixed values for M and t_s:

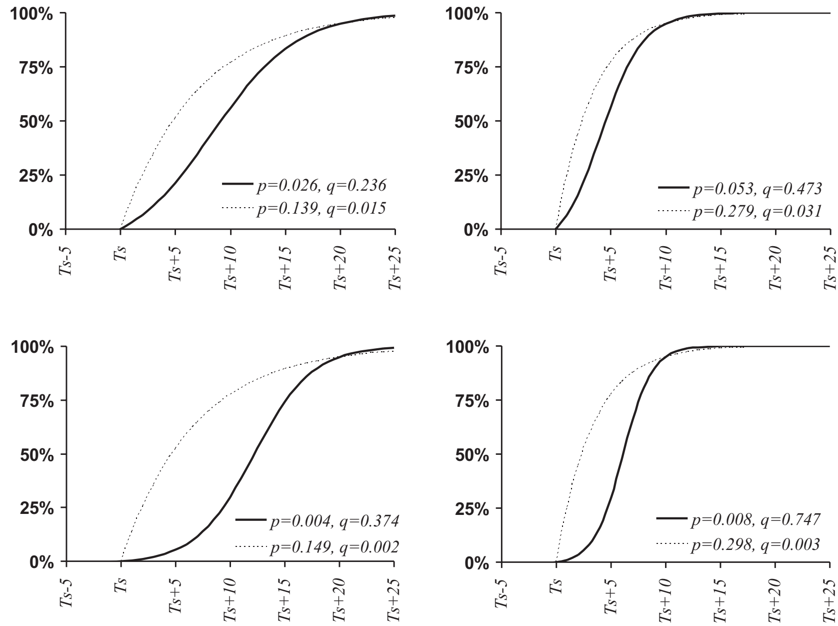

Fig. 7.2 Effects of different values of parameters *p* and *q* (*Chosen values are explained under section* "Bass Model with Explanatory Parameters")

Analysis of Bass Model

Discrete recursive form of the Bass model follows from (7.1), which is useful approximation of (7.2) for small time intervals Δt:

$$B(t) = B(t-\Delta t) + \Delta t \cdot \left(p + q\frac{B(t-\Delta t)}{M} \right) \cdot \left(M - B(t-\Delta t) \right) \Leftrightarrow \Delta t \to 0 \tag{7.3}$$

First derivative of $B(t)$ is given in (7.4):

$$B'(t) = \frac{dB(t)}{dt} = M\frac{(p+q)^2}{p} \frac{e^{-(p+q)(t-t_s)}}{\left[1 + \frac{q}{p} e^{-(p+q)(t-t_s)} \right]^2} \tag{7.4}$$

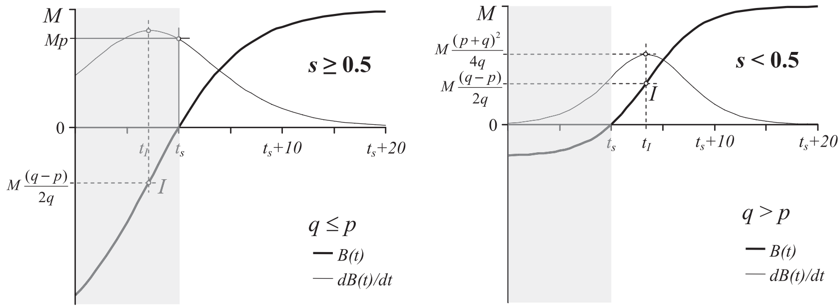

Fig. 7.3 Characteristic values and points of the Bass model of growth

Contrary to the S-shaped cumulative adoption $B(t)$, an adoption per period (sales) is bell-shaped curve (see Fig. 7.3), and it is proportional to the first derivative $B'(t)$ of cumulative adoption:

$$\text{Sales}(t_1,t_2) = B(t_2) - B(t_1) \approx (t_2 - t_1)\cdot B'\left(\frac{t_2 + t_1}{2}\right) \tag{7.5}$$

Maximum of $B'(t)$, as well as the time point when $B(t)$ has inflexion, is obtained from the solution of equation $B''(t) = 0$, where $B''(t)$ is the second derivative of $B(t)$:

$$B''(t) = \frac{d^2B(t)}{dt^2} = M\frac{(p+q)^3}{p}\frac{\left(\frac{q}{p}e^{-(p+q)(t-t_s)} - 1\right)\cdot e^{-(p+q)(t-t_s)}}{\left[1+\frac{q}{p}e^{-(p+q)(t-t_s)}\right]^3} \tag{7.6}$$

From (7.6) follows that $B(t)$ has inflexion for $t = t_I$:

$$t_I = t_s + \frac{1}{p+q}\ln\left(\frac{q}{p}\right) \Leftrightarrow B''(t_I) = 0 \tag{7.7}$$

and maximum of $B'(\mathrm{t})$ also occurs for $t = t_I$, when it has value of:

$$\max B'(t) = M\frac{(p+q)^2}{4q} \Leftrightarrow t = t_I \tag{7.8}$$

Value of the Bass model at point of inflexion is (see Fig. 7.3):

$$B(t_I) = M\frac{(q-p)}{2q} \tag{7.9}$$

In cases when $q < p$, inflexion point and maximum of $B'(\mathrm{t})$ occur before the product/service starts ($t_I < t_s$), and value of the Bass model at that point is negative according to (7.9); therefore interior maximum of $B'(t)$, which represents the real point of maximum sales t_{m}, occurs at $t_{\mathrm{m}} = t_s > t_I$. Similarly, in cases when $q = p$, inflexion point and maximum of $B'(\mathrm{t})$ occur when the product/service starts ($t_{\mathrm{m}} = t_I = t_s$). For $q > p$, sales peak occurs in conventional sense of a P/SLC ($t_{\mathrm{m}} = t_I > t_s$). The abovementioned is summarised in (7.10):

$$\max B'(t) = B'(t_m) = \begin{cases} M\frac{(p+q)^2}{4q} \Leftrightarrow q > p, \quad t_m = t_I \\ \\ Mp \Leftrightarrow q \le p, \; t_m = t_s \end{cases} \tag{7.10}$$

Accordingly, maximum of sales t_{m} occurs when penetration is $B(t_m)/M = (q - p)/2q$ in cases when $q > p$ (at $t_m = t_I$), and in cases when $q \le p$, maximum of sales occurs at $t_m = t_s$ when penetration is $B(t_m)/M = 0$.

In cases when $q > p$, for t_1 and t_2 near point of maximum sales $t_m = t_I$, from (7.5) follows that sales in time interval $[t_1, t_2]$ can be approximated by:

$$\mathrm{Sales}(t_1,t_2) \approx (t_2 - t_1)\cdot M\frac{(p+q)^2}{4q} \Leftrightarrow q > p \tag{7.11}$$

And in cases when $q \le p$, for t_1 and t_2 near point of maximum sales $t_m = t_s$, sales in time interval $[t_1, t_2]$ can be approximated by:

$$\text{Sales}(t_1, t_2) \approx (t_2 - t_1) \cdot Mp \Leftrightarrow q \le p \tag{7.12}$$

The Bass model is centrosymmetric regarding inflexion point I (t_I, $M(q - p)/2q$), see Fig. 7.3:

$$M\frac{(q-p)}{2q} - B(t_I - \Delta t; M, p, q, t_s) = B(t_I + \Delta t; M, p, q, t_s) - M\frac{(q-p)}{2q} \tag{7.13}$$

Growth rate GR for time interval Δt is always positive:

$$GR_{\Delta t} = \frac{B(t) - B(t - \Delta t)}{B(t - \Delta t)}$$

Due to the fact that the Bass model starts from t_s, $B(t_s) = 0$, the growth rate for $t \to t_s$ goes to infinity.

The above described characteristics of the Bass model with its explanatory attributes can be used as helpful input for estimation or assessment of model parameters for the forecasting purposes.

Bass Model with Explanatory Parameters

Although values of parameters p and q describe the process of how new product/service gets adopted as an interaction between users and potential users, their explanatory meaning is deficient in comparison with perceptible M and t_s parameters. Besides that, p and q are mutually dependent while they shape the Bass model S-curve (see Fig. 7.2). Namely, value of characteristic duration of product/service Δt (time interval to reach certain saturation level $v \cdot M$ measured from t_s) is provided only indirectly through values of p and q parameters:

$$\Delta t = \frac{1}{p+q} \ln\left(\frac{1+vq/p}{1-v}\right) \tag{7.14}$$

The idea is to replace p and q with two explanatory parameters: parameter that describes slope/shape of S-curve and characteristic duration Δt. As shown in the previous section, a candidate for the description S-curve slope/shape can be time of sales maximum t_m because of its direct relationship with the point of S-curve inflexion t_I. Unfortunately, simple algebraic expression for determination of $\{p, q\}$ values from explanatory pair $\{\Delta t, t_m\}$ does not exist. Namely, nonlinear system of Eqs. (7.7 and 7.14) needs iterative numerical methods to be performed for its solution (Sokele 2011). Therefore, required modifications of the Bass model will be carried out in two steps:

1. Reparametrisation of the Bass model to be able to accept characteristic duration of product/service Δt
2. Mapping of time of sales maximum t_m to reparametrised model from step 1

Reparametrisation of the Bass Model

Reparametrisation of the Bass model is carried out by replacing coefficient of innovation (p) and coefficient of imitation (q) with shape parameter (s) that describes slope/shape of S-curve and characteristic duration of product/service Δt (Sokele 2008). Shape parameter s is chosen in order to encompass relation between amplitude of positive S-curve part and amplitude of negative S-curve part. Asymptotes of the Bass model are:

$$\text{Negative asymptote}: \lim_{t \to -\infty} B(t) = -\frac{p}{q} M$$
$$\text{Positive asymptote}: \lim_{t \to +\infty} B(t) = M$$

Ratio between negative asymptote and distance of these asymptotes lies in range $\langle 0,1]$ which is convenient to choose as the shape parameter *s*, and which can be measured in percentages. In fact, according to the value of *s*, S-curve is stretched in vertical direction (on *y*-axis) preserving the total market capacity *M*. Distance between these asymptotes is $M \cdot (1 + p/q)$, so shape parameter *s* is:

$$s = \frac{pM/q}{M + pM/q} = \frac{p}{q+p}, \quad p > 0, q \geq 0 \tag{7.15}$$

Characteristic values of *s* are:

$s \to 0$	negative asymptote → 0, imitation prevails, curve is similar to a simple logistic growth model, ($q >> p > 0$)
$s \in \langle 0, 0.5 \rangle$	$q > p$, see the right side of Fig. 7.3
$s = 0.5$	sales peak occurs at time when product/service starts ($q = p > 0$)
$s \in \langle 0.5, 1 \rangle$	$p > q$, see the left side of Fig. 7.3
$s = 1$	negative asymptote → $-\infty$, innovation prevails; curve is similar to an exponential saturation growth model, ($q = 0$, $p > 0$).

From (7.18) follows:

$$p = (p+q) \cdot s; \quad q = (p+q) \cdot (1-s) \tag{7.16}$$

Information about saturation point level $B(t_s + \Delta t) = v \cdot M$ and (7.2) give (7.17) and (7.18):

$$p + q = \frac{1}{\Delta t} \ln\left(1 + \frac{v}{s(1-v)}\right), \quad \Delta t = \frac{1}{p+q} \ln\left(1 + \frac{v}{s(1-v)}\right) \tag{7.17}$$

$$B(t; M, s, v, \Delta t, t_s) = M \frac{1 - \left(1 + \frac{v}{s(1-v)}\right)^{-\frac{t-t_s}{\Delta t}}}{1 + (1/s - 1) \cdot \left(1 + \frac{v}{s(1-v)}\right)^{-\frac{t-t_s}{\Delta t}}} \tag{7.18}$$

Expression (7.18) is the reparametrised Bass model with the following parameters:

M—market capacity
t_s—time when product/service is introduced, $B(t_s) = 0$, $t_s \leq t$
Δt—characteristic duration of product/service, $\Delta t > 0$
s—shape parameter, $0 < s \leq 1$
v—penetration at time point $t_s + \Delta t$, $0 \leq v < 1$

Reparametrised Bass model (7.17), $B(t;\ M;\ s,\ v,\ \Delta t,\ t_s)$, needs four parameters: M, t_s, Δt and s to be determined. Value of auxiliary parameter v does not need to be determined, it just allows forecasting practitioner to choose which level of penetration he/she wants to deal with (i.e. 90%, 95%, etc.).

Characteristics and Special Cases of (7.18):

- For $t = t + t_s$, value of model $B(t)$ is $v{\cdot}M$ regardless of the value of the parameter s:

$$B\left(t_s + \Delta t;\ M,\ s,\ v,\ \Delta t,\ t_s\right) = vM$$

- For $v = 0$, value of model $B(t)$ is zero:

$$B\left(t;\ M,\ s,\ v = 0,\ \Delta t,\ t_s\right) = 0$$

- For $s \rightarrow 0$ Bass model degrades into simple logistic model:

$$B\left(t;\ M,\ s \rightarrow 0,\ v,\ \Delta t,\ t_s\right) \rightarrow \frac{M}{1 + \frac{1}{s} \cdot \left(\frac{v}{s\left(1 - v\right)}\right)^{-\frac{t - t_s}{\Delta t}}} - sM \approx L\left(t;\ M,\ a,\ b\right)$$

where parameters of logistic growth model a and b are:

$$a = \frac{1}{\Delta t}\ln\left(\frac{v}{s(1-v)}\right), \quad b = t_s - \frac{\ln s}{a}$$

- For s = 0.5 Bass model gets a form of a halved logistic model:

$$B(t;M,s=0.5,v,\Delta t,t_s) = M\frac{1-\left(\frac{1+v}{1-v}\right)^{-\frac{t-t_s}{\Delta t}}}{1+\left(\frac{1+v}{1-v}\right)^{-\frac{t-t_s}{\Delta t}}}$$

$$= \frac{2M}{1+\left(\frac{1+v}{1-v}\right)^{-\frac{t-t_s}{\Delta t}}} - M = L(t;2M,a,b) - M$$

This curve has a shape of the logistic model with double market capacity M but vertically shifted down by M. Parameters a and b of this "halved" logistic model are:

$$a = \frac{1}{\Delta t}\ln\left(\frac{1+v}{1-v}\right), \quad b = t_s$$

- For s = 1 Bass model degrades into an exponential saturation growth model:

$$B(t;M,s=1,v,\Delta t,t_s) = M\left(1-(1-v)^{\frac{t-t_s}{\Delta t}}\right)$$

Table 7.1 gives the explanation of chosen values for parameters p and q presented in Fig. 7.2, which are selected according to shape parameter and characteristic duration:

Table 7.1 Parameter values used for curves in Fig. 7.2

Graph in Fig. 7.2	*Shape parameter* s_1 (%)	*Shape parameter* s_2 (%)	*Characteristic duration to 95% penetration* Δt
Top-left	10	90	20 years
Top-right	10	90	10 years
Bottom-left	1	99	20 years
Bottom-right	1	99	10 years

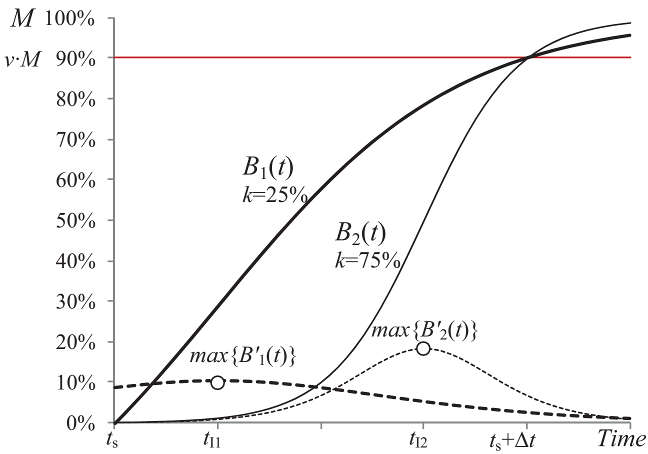

Fig. 7.4 Illustration of the impact of the coefficient k on the shape of the S-curve (Thick line, k = 25%; thin line, k = 75%; dashed lines—corresponding $B'(t)$; v = 90%; Δt is constant)

Mapping of Time of Sales Maximum

Mapping of time of sales maximum t_m to reparametrised Bass model is carried out via dimensionless coefficient k—relative position of sales maximum point within characteristic duration:

$$k = \frac{t_m - t_s}{\Delta t} \tag{7.19}$$

Values for coefficient k are illustrated in Fig 7.4.

To enable higher diversity in cases when $p > q$, it is convenient to allow that t_m exists even before t_s (when significant demand for certain product/services exists before its introduction on a market), or in other words t_m should fully correspond to the inflexion point of S-curve t_I. From (7.7), (7.15), (7.17) and (7.19) follows:

$$k = \frac{t_I - t_s}{\Delta t} = \ln\left(\frac{1-s}{s}\right) / \ln\left(1 + \frac{v}{s(1-v)}\right) = f(s; v) \tag{7.20}$$

It is important to note that the relative position of sales maximum k (7.19) depends only on the value of shape parameter (s) and the value of auxiliary parameter v. Figure 7.5 shows the relationship between relative position of sales maximum (k) and shape parameter (s) for different values of penetration level v.

The main difficulty rises from the fact that nonlinear relationship (7.20) is without possibility of analytical solution in a form of $s = f^{-1}(k; v)$.

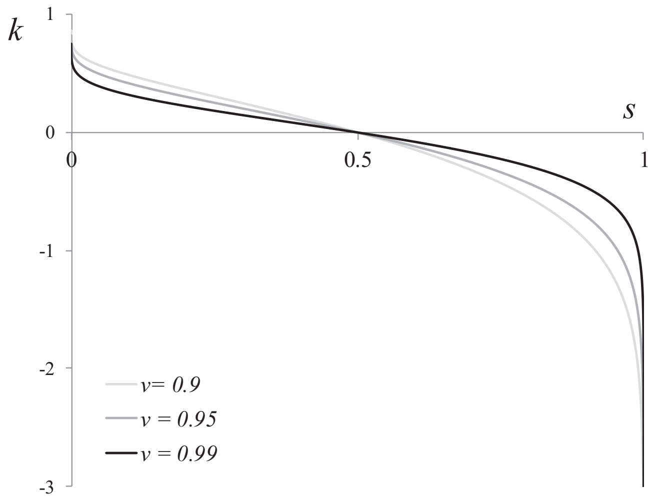

Fig. 7.5 Relationship between relative position of sales maximum k and shape parameter s

Solution is to find an analytical function $g(k; v)$ that approximate $= f^{-1}(k; v)$ with the minimal error ε_m on the whole interval of possible values for k and s:

$$s = f^{-1}(k; v) = g(k; v) + \varepsilon_m \tag{7.21}$$

Resulting from (7.20), function $g(k, v)$ should satisfy the following conditions:

$$\begin{aligned} s &= g(k = 0; v) = ½ \\ s &= g(k \to 1; v) \to 0 \\ s &= g(k \to -\infty; v) \to 1 \end{aligned} \tag{7.22}$$

An analytical function (7.23) that satisfies conditions (7.22) and has two parameters a and b that depend on a chosen value for v:

$$s = 2^{-\exp\left(ak/(1-k)^b\right)} \tag{7.23}$$

In the rest of the text, function (7.23) is chosen for mapping function. Optimal values for parameters a and b are determined by the ordinary least squares method for interval of values for k, $-1 < k < 0.75$ and presented in Table 7.1. In the next section it will be shown that used interval for k is in accordance with examples for ICT service growth.

Based on all abovementioned, the procedure for usage the Bass model with explanatory parameters is as follows:

1. Obtain estimations for M, market capacity; Δt, characteristic duration of product/service; t_s, time when product/service will be/is introduced; and t_m, time of sales maximum;
2. Choose auxiliary parameter v—penetration at time point $t_s + \Delta t$ which is in the relation with estimated characteristic duration of the product/service;

3. Calculate coefficient k by (7.19) and put its value into mapping function (7.23) to obtain corresponding shape parameter s;
4. Use reparametrised Bass model (7.18) for modelling of desired time interval.

Although it is actually a numerical solution of nonlinear system of equations, this procedure and developed mapping function do not require iterative procedures. Mapping function introduces error, but features of reparametrised Bass model help in reducing them at the beginning and at the end of the time interval, that is, near t_s, and $t_s + \Delta t$ where errors are 0%. In addition, the inherent structure of chosen mapping function (7.23) reduces errors near $k = 0$. Figure 7.6 shows percentage errors for $v = 90\%$. Percentage errors for k in interval $-1 < k < 0.75$ are less than 0.5% on a whole time interval $[t_s, t_s + \Delta t]$ which makes the whole concept suitable framework for the forecasting of new products/service adoption prior to its launch.

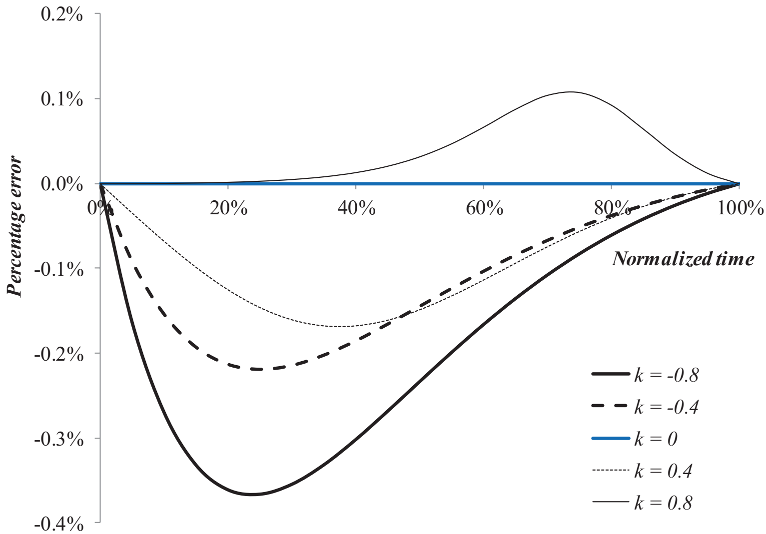

Fig. 7.6 Percentage errors for $v = 90\%$ (x-axis normalised time 0% = t_s, 100% = $ts + \Delta t$; y-axis percentage error of Bass model with explanatory parameters obtained via mapping function vs. ordinary Bass model)

Table 7.2 Optimal values for parameters *a* and *b* for model (7.23)

	$v = 90\%$	$v = 95\%$	$v = 99\%$
$a =$	1.001649	1.699404	3.246991
$b =$	1.082064	0.915601	0.399523

Example: Modelling of Fixed (Wired)-Broadband Subscriptions

The concept of the Bass model with explanatory parameters will be examined with data for fixed (wired)-broadband subscriptions. This service is prerequisite and/or infrastructure for all ICT services. Therefore, its market diffusion dynamics is important to be identified through explanatory variables/parameters (Meade Islam 2015).

Table 7.2 presents results of modelling where all parameters of the Bass model (M, t_s, p and q) are determined by means of Ordinary Least Squares method. From all available data, selected are those countries whose data sets modelling have correlation coefficient r greater than 0.9970.

The remaining table columns contain the calculated explanatory parameters that can be easily interpreted by a forecasting practitioner: Δt, characteristic duration of product/service, time needed to reach v = 90% of market capacity M and k, $k = (t_I - ts)/\Delta t$—relative position of sales maximum point within characteristic duration Δt. Fixed (wired)-broadband subscriptions analysed in a sense of P/SLC reaches mature phase for most countries, without any presence of decline phase so results of its modelling can be the basis for demand forecasting of new ICT products/services. Results show that for all encompassed countries characteristic duration Δt lies in the interval from 8 to 18 years for saturation level of 90%. The relative position of sales maximum point k is in the interval from −40% to 67% which confirms the possibility of applying mapping function (7.23) with parameters given in Tables 7.1 and 7.3.

Table 7.3 Modelling of fixed (wired)-broadband subscriptions

Country	M	t_s	p	q	r	Δt	t_I	k	s
Austria	2 446 973	1998.7	0.0407	0.2871	0.9976	13.1	2004.6	0.455	0.12
Belgium	4 423 284	1999.7	0.0828	0.1431	0.9993	14.3	2002.2	0.169	0.37
Bosnia and Herzegovina	544 422	2003.1	0.0097	0.7352	0.9985	8.8	2008.9	0.662	0.01
Brazil	29 051 365	2000.4	0.0126	0.3356	0.9988	15.9	2009.8	0.595	0.04
Canada	13 604 299	1998.3	0.0595	0.1896	0.9984	14.7	2003.0	0.317	0.24
Chile	3 233 257	2000.4	0.0385	0.1872	0.9981	17.7	2007.5	0.397	0.17
Colombia	6 584 087	2001.8	0.0124	0.3684	0.9987	14.8	2010.7	0.603	0.03
Denmark	2 293 262	1999.6	0.0526	0.4434	0.9970	9.0	2003.9	0.479	0.11
France	27 076 716	2000.5	0.0445	0.3327	0.9988	11.5	2005.8	0.463	0.12
Germany	29 685 391	1998.4	0.0093	0.5293	0.9987	11.6	2005.9	0.646	0.02
Hong Kong	2 416 477	1998.8	0.1631	0.0589	0.9979	11.6	1994.3	−0.394	0.73
Hungary	2 675 724	2000.7	0.0284	0.4463	0.9980	10.6	2006.5	0.549	0.06
Iceland	118 354	1999.9	0.0736	0.4182	0.9976	8.4	2003.4	0.422	0.15
Italy	14 354 830	2000.5	0.0481	0.4508	0.9986	9.1	2005.0	0.492	0.10
Japan	39 507 075	2000.0	0.1276	0.1427	0.9989	11.1	2000.4	0.037	0.47
Netherlands	6 795 317	1999.8	0.0552	0.4636	0.9980	8.6	2003.9	0.478	0.11
Norway	1 947 801	2000.2	0.0513	0.4290	0.9980	9.3	2004.7	0.478	0.11
Singapore	1 570 360	1998.4	0.0267	0.3356	0.9986	13.3	2005.4	0.526	0.07
Spain	13 476 625	2000.3	0.0509	0.2961	0.9980	11.9	2005.4	0.426	0.15
Switzerland	3 751 600	2000.3	0.0733	0.2155	0.9976	12.5	2004.0	0.300	0.25
USA	101 998 962	1998.9	0.0467	0.2930	0.9981	12.4	2004.3	0.438	0.14
Venezuela	2 749 057	2000.8	0.0160	0.3744	0.9988	13.8	2008.9	0.584	0.04
Minimum	–	**1998.3**	**0.0093**	**0.0589**	**0.9970**	**8.4**	**1994.3**	**−0.394**	**0.01**
Maximum	–	**2003.1**	**0.1631**	**0.7352**	**0.9993**	**17.7**	**2010.7**	**0.662**	**0.73**
Average	–	**2000.0**	**0.0511**	**0.3398**	**0.9983**	**12.0**	**2004.9**	**0.415**	**0.16**
St. Deviation	–	**1.1**	**0.0376**	**0.1516**	**0.0006**	**2.6**	**3.5**	**0.233**	**0.17**

Data source: ITU development statistics (http://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx, December 2016); data span: 2000–2015

M, market capacity [subscribers]; t_s, time when product/service is introduced [*calendar year*]; p and q, Bass model coefficients of innovation and imitation [$year^{-1}$]; r, correlation coefficient; Δt, characteristic duration of product/service for saturation level $v = 90\%$ [*years*]; t_I, inflexion point (time of sales maximum) [*calendar year*]; k, relative position of sales maximum point [*dimensionless*]; s, shape parameter [*dimensionless*]

Conclusions and Further Research

The Bass model is the most convenient model for market adoption forecasting of a new product/service in the sense of flexibility versus number of free parameters needed to be determined. Estimation of the parameters when limited time series data are available can be improved by introducing the Bass model with explanatory parameters. Direct inclusion of explanatory parameters in the Bass model is not possible, but solution with an approximate mapping function in combination with reparametrised Bass model makes it achievable. Bass model with explanatory parameters is suitable for the forecasting of new product/service adoption where all model parameters can be assumed by means of analogy with the existing products/services.

Future research could continue in two directions: enhancements of mapping function and making database of explanatory diffusion parameters for the existing ICT products/services. Database would be used by forecasting practitioners as a guide for the forecasting by means of analogy for the new ICT products/services.

References

Meade, N., & Islam, T. (2006). Modelling and Forecasting the Diffusion of Innovation—A 25-Year Review. *International Journal of Forecasting, 22*(3), 519–545.

Meade, N., & Islam, T. (2015). Forecasting in Telecommunications and ICT—A Review. *International Journal of Forecasting, 31*(4), 1105–1126.

Meyer, S. P., & Ausubel, J. H. (1999). Carrying Capacity: A Model with Logistically Varying Limits. *Technological Forecasting and Social Change, 61*(3), 209–214.

Sokele, M. (2008). Growth Models for the Forecasting of New Product Market Adoption. *Telektronikk, 3/4*, 144–154.

Sokele, M. (2011). Bass Model. In L. Moutinho & D. G. Hutcheson (Eds.), *The SAGE Dictionary of Quantitative Management Research* (pp. 18–23). London: SAGE Publications Ltd..

Sokele, M. (2015). *Growth Models, Wiley Encyclopaedia of Management Volume 9* (N. Lee & A. M. Farrell, Eds.). New York: John Wiley and Sons (Online ISBN: 9781118785317).

Sönke, A. (2004). Forecasting the Diffusion of an Innovation Prior to Launch. In *Cross-Functional Innovation Management Perspectives from Different Disciplines* (pp. 243–258). Gabler: Wiesbaden.

8

A Brief Introduction to Evolutionary Algorithms from the Perspective of Management Science

Volker Nissen

Introduction

Evolutionary Algorithms (EA) are a family of powerful search and optimization methods inspired by the mechanisms of natural evolution. They imitate, on an abstract level, evolutionary principles such as a population-based approach, the inheritance of information, the variation of information via crossover/mutation, and the selection of individuals based on fitness. We will look at EA from the perspective of management science, indicating that EA are of interest as methods to facilitate management decisions and support management processes in organizations.

EA evolve a set of solutions to solve a given constrained or unconstrained problem. Important mainstream classes of EA include:

- Genetic Algorithms (GA), the most popular approach among EA,
- Genetic Programming (GP), a variant of GA that evolves computer programs,

V. Nissen (✉)
Chair of Information Systems Engineering in Services,
University of Technology Ilmenau, Ilmenau, Germany

L. Moutinho, M. Sokele (eds.), *Innovative Research Methodologies in Management*,
https://doi.org/10.1007/978-3-319-64394-6_8

- Evolution Strategies (ES), an approach that focuses more on mutation than classical GA.

Other popular variants exist, such as Estimation of Distribution Algorithms (EDA), Learning Classifier Systems (LCS), and Differential Evolution (DE). For a recent overview see Eiben and Smith (2015). Also Evolutionary Programming (Fogel et al. 1966; Fogel 2006) as one of the founding branches of EA must be mentioned here. Moreover, there are other nature-inspired approaches, many of which are somehow close in spirit to EA. Examples include Particle Swarm Optimization (PSO) and Ant Colony Optimization (ACO). A classification of nature-inspired heuristics for optimization can be found in Fister et al. (2013). More details are provided in Yang (2014). However, one should be aware of the critical remarks on the "metaphor-centric period of heuristics" in Sörensen et al. (2016, p. 12).

Besides biologically inspired terminology (explained in Table 8.1), EA differ from more classical search methods in a number of ways (Biethahn and Nissen 1995, pp. 4–5):

- EA operate on a representation (e.g. in the form of binary strings or vectors of real numbers) of the decision variables. Results are later mapped to the original solution space.

Table 8.1 Basic EA terminology

Individual	Contains the adequately represented elements of a solution to the given problem
Population	Set of individuals
Parents	Individuals selected for reproduction
Children, offspring	New solutions generated from the parents
Fitness	Solution quality w.r.t. the given objectives
Crossover	Search operator that mixes elements from different individuals
Mutation	Search operator that modifies a given single individual
Generation	EA-iteration
Chromosome	Basically identical with individual; occasionally an individual is composed of several chromosomes; common form: character string
Genotype	Coded solution
Phenotype	Decoded solution

- EA generally process a set of solutions, exploring the search space from many different points simultaneously (population-based approach).
- The basic EA-operators imitate processes of replication, variation, and selection as the driving forces of natural evolution. Favoring the fittest individuals in selection, combined with a variation of inherited material in the offspring will iteratively lead to good solutions for the given problem.
- For goal-directed search, only information on the quality (fitness) of solutions is required. This fitness information is frequently calculated from an objective function, but may also be obtained by other measures, for example, simulation. However, incorporating available domain knowledge in the problem representation, initialization, fitness function, search operators, or decoding scheme may substantially increase the competitiveness of an EA at the cost of a reduced scope of application.
- Stochastic elements are deliberately employed. This means no simple random search, though, but an intelligent exploration of the search space. Information from already discovered good solutions is *exploited* while promising new areas of the search space continue to be *explored.*

A nice feature of EA is their ability to continue the optimization as long as resources are available. However, the convergence toward good solutions is generally asymptotic. Moreover, it must be stressed that EA aim at good but not necessarily optimal solutions for a given application problem. This is in line with other heuristic approaches. In management practice it is often sufficient to find good solutions within reasonable time.

Evolutionary Algorithms are also termed "Evolutionary Computation", and rank among computational intelligence (or soft computing) methods. They also belong to the more general class of metaheuristics, which Sörensen and Glover (2013) define in the following way: "A metaheuristic is a high-level problem-independent algorithmic framework that provides a set of guidelines or strategies to develop heuristic optimization algorithms. The term is also used to refer to a problem-specific implementation of a heuristic optimization algorithm according to the guidelines expressed in such a framework."[1]

In their "history of metaheuristics", Sörensen et al. (2016, p. 3) identify five distinct periods, starting with the pre-theoretical period (until c. 1940), during which heuristics and even metaheuristics are used but not formally studied, followed by the early period (c. 1940–c. 1980), during which the first formal studies on heuristics appear, and the method-centric period (c. 1980–c. 2000), during which the field of metaheuristics truly takes off and many different methods are proposed. They argue that we currently witness the framework-centric period (c. 2000–now), during which the insight grows that metaheuristics are more usefully described as frameworks, and not as methods. The final scientific period (the future) then is characterized by the design of metaheuristics becoming a science instead of an art, with the focus shifting from performance to understanding.

Particularly important for our understanding here is the framework-centric view (Sörensen et al. 2016, p. 10). It rests upon the idea that opening up the individual algorithmic frameworks allows researchers to combine a metaheuristic with any auxiliary method available. Metaheuristics, such as EA, in this sense can be seen as a more or less coherent set of ideas, which could be freely mixed with other ideas. By combining the most efficient operators and carefully tuning the resulting heuristic, solution methods can be created that solve any real-life application problem efficiently. We shall come back to this perspective in section "How to Apply". However, before discussing EA from a framework-centric view it is useful, in this introductory contribution, to understand what "pure" (canonical) EA are. Therefore, the next section first introduces a general flow scheme of EA, before highlighting Genetic Algorithms and Evolution Strategies in their archetypical forms.

Description

Generalized Scheme

The flow scheme of a generalized EA, which contains the different EA-variants (including those with population size one) as special cases, is shown in Fig. 8.1 and will be explained below.

```
Choose strategy parameter values (such as population size, crossover rate)
Initialize population P(0)
t ← 0
Evaluate individuals of P(0)
Repeat  (*Generation Cycle*)
        Selection for reproduction (*Parent Selection*)
        Replication                (*Generate ...)
        Variation          (...Offspring*)
        Evaluate offspring
        Select new population P(t+1)
        t ← t+1
Until t_max or some other termination criterion holds
Output result
```

Fig. 8.1 Generalized EA-scheme

One starts with an initial set (population) of alternative solutions (individuals) for the given problem. Frequently, these initial solutions are randomly generated or evenly scattered over the region of interest in search space. However, a biased initialization is possible when useful prior knowledge about the application problem is available. For the initial solutions their quality (fitness values) are determined.

An iterative cycle follows, during which new, modified solution proposals are repeatedly produced from the prior solutions of the last cycle. Here, the evolutionary operators of replication, variation, and selection are applied. In this generation cycle, at first individuals are determined by selection for reproduction. Within the scope of creating offspring the solution information is copied from these parents (replication) and changed by the use of one or more variation operators, such as mutation or crossover. The resulting offspring are evaluated, and afterward individuals are selected and transferred to the new population (selection for survival). Depending on the EA-variant the parents as well as their descendants may compete for survival. The interaction between the stochastic changes introduced through the variation operators and the preference of the best solutions in the selection process leads to successively better solution proposals in the course of many generation cycles. This process will proceed until a termination criterion holds. Finally, the user is given the result, usually the best found solution of a run, and the procedure is terminated.

Genetic Algorithms and Evolution Strategies are particularly prominent variants of EA, which originated independent of one another in the

USA (GA; Holland 1975; Goldberg 1989) and in Germany (ES; Rechenberg 1973; Schwefel 1975). In short, the main differences between ES and GA are (Nissen and Gold 2008):

- A GA focuses on recombination (crossover) as the main search operator, while an ES puts the focus more on mutation.
- In an ES, λ offspring solutions are created from μ parents by copying and modifying the parents through mutation and, if desired, recombination. Common ratios of μ and λ are between 1:5 and 1:7. This differs from the generation cycle in a GA that generally keeps the population size constant. Moreover, selection is deterministic in ES while it is frequently stochastic in GA.
- Mutation in the classical ES is performed using normally distributed random variables so that small changes in a solution are more frequent than large changes. For GA, many different mutation operators and problem representations have been developed that address continuous parameter optimization as well as combinatorial problems.
- The concept of self-adapting strategy parameters was introduced early in ES-research. In particular, the mutation step size in an ES can be made self-adaptive by incorporating it in the representation and applying a log-normally distributed random variable to mutate the step size before using this step size to mutate decision parameters. Later concepts have de-randomized ES, though (Hansen and Ostermeier 2001).

Outline of a Basic Genetic Algorithm

GA are the most popular EA-variant. They imitate evolutionary processes with the particular emphasis on genetic mechanisms. From the manifold of published procedure variants only a basic binary GA-concept (Goldberg 1989; Nissen 1997) will be presented here.

Every individual $\vec{a}$ in the GA-population is a character string which contains L bits, where L is an application dependent value: $\vec{a} = (a_1, a_2, \ldots, a_{L-1}, a_L) \in \{0,1\}^L$. Every string is broken down into n segments ($n \leq L$). Every segment corresponds to a variable of the considered optimization problem. Segment j ($j = 1,2,\ldots,n$) contains, in binary coded form, a value for the decision variable j of the optimization problem. The segments can be equally long or can contain bit strings of different lengths.

In the following it is supposed that the objective function $F(\vec{x})$ of n continuous decision variables $\vec{x} = x_1, x_2, \ldots, x_n$ should be maximized. The binary coding on a string with finite length requires a lower and upper boundary value $[u_j, o_j] \in \mathbf{R}$, $u_j < o_j$ for every variable x_j ($j = 1, 2,\ldots,n$) of the optimization problem. Thus the search space is limited. The function $F(\vec{x})$ therefore delivers the following mapping:

$$F : \prod_{j=1}^{n} \left[u_j, o_j\right] \to \mathbf{R}$$

By means of a segmentally proceeding decoding function

$$\Gamma : \{0,1\}^L \to \prod_{j=1}^{n} \left[u_j, o_j\right]$$

the decoded variable value can be gained from the binary strings: $\vec{x} = \Gamma(\vec{a})$.

If we now focus on string segment j, it has the length L_j bits and codes for the value of the variable x_j. Let a_{jz} denote the bit with the number z (starting on the left side, $z = 1, 2, \ldots L_j$) of the string segment j.

The decoding of this segment runs as follows:

$$\Gamma^j\left(a_{j1}, \ldots, a_{jL_j}\right) = u_j + \frac{o_j - u_j}{2^{L_j} - 1} \cdot \left(\sum_{z=1}^{L_j} a_{j(L_j - z + 1)} \cdot 2^{z-1}\right) = x_j$$

Real numbers can only be represented in a binary code with limited precision. The approach during coding and decoding is shown by the following example.

A continuous variable x with the definition region of $-1 \leq x \leq 2$ ($x \in \mathbf{R}$) is binary coded. At first the required precision needs to be set. Here the first decimal place should be adequate, so that the definition region x is divided into 30 intervals with a width of 0.1. This degree of precision can only be coded when L is at least equal to five bits, because $16 = 2^4 < 30 < 2^5 = 32$. A higher degree of precision would need a longer string.

The lower interval boundary, −1, is presented by the string 00000 and the upper interval boundary +2 by the string 11111. All remaining strings are linearly depicted in the definition region between these boundaries. Within the scope of decoding, every binary string $a_1\ a_2\ a_3\ a_4\ a_5$ is first converted to the basis 10. The decoded x-value is given by:

$$x = -1 + \frac{2-(-1)}{2^5 - 1} \cdot \sum_{z=1}^{5} \left(a_{5-z+1}\right) \cdot 2^{z-1}$$

The string 11001 would thus be decoded as:

$$x = -1 + \frac{3}{31} \cdot 25 \approx 1{,}4$$

This approach is not without problems, because it often does not provide a clear mapping between binary strings and decoded values. Moreover the true optimum of a function of continuous variables can generally only be approximated by binary coding, that is, this basic GA realizes a form of grid search. In the following, the common steps during the execution of a GA are highlighted:

Step 1: Initialization

During the initialization phase a starting population $P(t = 0)$ of μ individuals $\vec{a}_i$ $(i = 1,2,\ldots,\mu)$ is generated.[2] Common values for μ are between 30 and 500. Generally μ is an even number. The initialization of the starting population is often carried out at random so that the bits of all individuals of the population are stochastically independent of each other set to either 1 or 0. If prior knowledge about the structure of the problem is available, the initialization could be done in a biased way, but care should be taken to achieve a sufficient diversity in the population.

Step 2: Evaluation of the Initial Solutions
This step is necessary for the following selection so that the solution alternatives in the population can qualitatively be differentiated. The individuals are decoded and assessed by means of a fitness function Φ, derived from the objectives. This fitness function is made up of the objective function F and the decoding function Γ so that $\Phi = F \circ \Gamma$. Then the following applies ($i = 1,2,\ldots,\mu$): $\Phi\left(\vec{a}_i\right) = F\left(\Gamma\left(\vec{a}_i\right)\right)$.

Step 3: Stochastic Parent Selection and Replication
In this sub-step μ individuals are stochastically selected from the actual population in order to become parents for reproduction. Depending on its fitness value the same individual may be selected multiple times. Different selection forms are possible. It is, however, always important to favor the good solutions during the selection process so that their solution elements can preferentially be passed onto the next generation (*exploitation*). In the case of fitness-proportional selection, this results in the selection probability p_s of an individual $\vec{a}_i \in P$ as follows ($i = 1,2,\ldots,\mu$):

$$p_s\left(\vec{a}_i\right) = \frac{\Phi\left(\vec{a}_i\right)}{\sum_{j=1}^{\mu}\Phi\left(\vec{a}_j\right)}$$

A selective algorithm is now necessary, which selects the parents for the next procedure step on the basis of given selection probabilities. Stochastic Universal Sampling (Baker 1987) in particular is suitable for this procedure. This algorithm can be pictured as a wheel-of-fortune with μ sections where each corresponds with a population member. The width of each section on the wheel correlates with the selection probability of the relevant individual. There are μ pointers at equal intervals around the wheel. By one turn of the wheel-of-fortune all parents can be selected at the same time. The pointers determine how many copies of each individual go into the mating pool.

Fitness-proportional selection is widely used, but has important disadvantages. Special measures are required when negative fitness values occur, or a minimizing task is processed. It is in general more advantageous to use tournament selection or rank-based selection (ranking). Tournament selection goes back to Wetzel's unpublished considerations. The common procedure is to extract ξ individuals ($2 \leq \xi < \mu$) with equal selection probability $p_s = 1/\mu$ from the population and to copy the best individual among them in the mating pool afterward. This procedure is repeated μ times until the mating pool is complete. The rigorousness of the selection and hence the extent of exploitation can be controlled through the value of ξ. If the selection is too rigorous, premature convergence of the GA on suboptimal solutions is possible. On the other hand a soft selection causes long computing times and the risk of a low target orientation during the search. A common choice for the tournament parameter is $\xi = 2$. Tournament selection can be used unmodified also for negative fitness values or for the case of minimizing tasks. It is, however, a disadvantage that an individual can theoretically be represented in the mating pool by up to μ copies.

In the case of ranking (Baker 1985), the selection probability of each individual is dependent on its position in a fitness-based ranking order of all members in the actual population. The absolute extent of fitness differences among individuals are thus subsidiary. Initially, the individuals of the population are selected according to declining fitness. The first individual $\vec{a}_1$, therefore, has the highest fitness value during maximizing. An expectation value of $E_{\max}$ copies in the mating pool is assigned to this individual. The last, and therefore worst, individual in the ranking receives the expectation value of $E_{\min}$ copies. For each individual $\vec{a}_i$ ($i = 1,2,\ldots,\mu$) in the population, the ranking is given by $r(\vec{a}_i)$. In that case the following selection probabilities are given:

$$p_s(\vec{a}_i) = \frac{1}{\mu} \cdot \left(E_{\max} - (E_{\max} - E_{\min}) \cdot \frac{r(\vec{a}_i) - 1}{\mu - 1} \right)$$

$$p_s(\vec{a}_i) \geq 0 \;\; \forall i \in \{1,2,\ldots,\mu\}, \sum_i p_s(\vec{a}_i) = 1$$

The selection algorithm Stochastic Universal Sampling should be used again. As the selection probability is still only depending on the ranking, the rigorousness of the selection can be directly controlled with the value of E_{max}. Ranking can be used with negative fitness values without modification taking place. In the case of minimizing, the individuals should be sorted differently.

Step 4: Generating Offspring

The following sub-steps have to be run through $(\mu/2)$ times.

Sub-step 4-1: Stochastic Partner Choice

Two parents are now extracted from the mating pool with the equal probability of $1/\mu$ and without laying back individuals. Two descendants are produced from these parents by the variation operators, crossover and mutation, as discussed in the following sub-steps.

Sub-step 4-2: Crossover

Crossover is the main operator in GA when searching for new, improved solutions (*exploration* of the search space). Solution elements of the parents are hereby mixed and transferred to the offspring. The N-point crossover is popular (Fig. 8.2). Initially, based on the predetermined crossover probability p_c (recommended $p_c \geq 0.6$), it is assessed whether

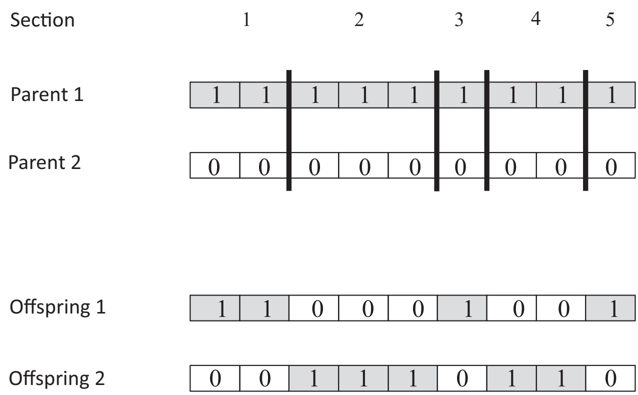

Fig. 8.2 N-point crossover

a crossover should take place. Thus the manifestation of a uniformly spread random variable **U** is determined in the interval [0,1]. Crossover will only take place in the case of U $\leq p_c$. If no crossover takes place, the two unmodified strings are given over to the mutation operator (sub-step 4.3).

During N-point crossover $N > 1$, crossover points are stochastically determined on the base of a uniform distribution between 1 and $L-1$. They are identical for both parent strings and have to be newly determined for each parental couple. In general, N is an even number for symmetric reasons. All sections with even numbers are exchanged between the involved strings. The case $N = 4$ is illustrated in Fig. 8.2.

In case of the likewise widely distributed Uniform Crossover every bit position on the strings needs to be tested individually to find out whether an exchange between both parents should take place. The exchange takes place with a bit-oriented probability of p_{ux}. Common values are $0.5 \leq p_{ux} \leq 0.8$. Figure 8.3 illustrates this principle.

Sub-step 4-3: Mutation

Mutation as a search operator is rated as secondary important in GA. It is though used to prevent that individual string positions of all individuals in a population have the same value. In this case the optimization process could only take place in a subspace of the original search space.

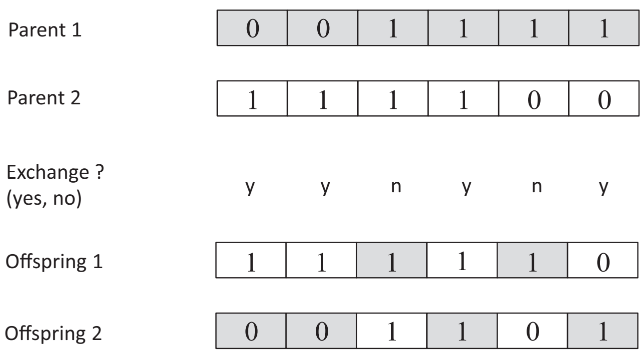

Fig. 8.3 Uniform crossover

In the scope of mutation, every bit of an individual is inverted with a probability of $p_m \approx 1/L$, where L is the length of the string.

Sub-step 4-4: Evaluation of the Offspring and Supplementing the New Population

Both produced descendants are now evaluated:

$$\Phi\left(\vec{a}_1'\right) = F\left(\Gamma\left(\vec{a}_1'\right)\right)$$
$$\Phi\left(\vec{a}_2'\right) = F\left(\Gamma\left(\vec{a}_2'\right)\right)$$

They are taken over into the new population which is empty at the beginning. When this population is complete (contains μ individuals) it replaces the previous population (generational replacement).

Step 5: Continue with Step 3 Until a Termination Criterion Holds

Finally, the relevant results are given to the user. In general this will be the best solution found during the entire run.

In many of practical applications binary solution coding is dispensed with in favor of representations of higher cardinality, which should possibly derive "naturally" from the given problem. This includes, for example, representations in the form of vectors of real numbers, permutations, matrices, or tree structures. Non-binary solution representation in principle needs other forms of crossover and mutation. More information can be found in introductory text books, such as Eiben and Smith (2015) or Bäck (1996).

Outline of a Basic Evolution Strategy

The subsequent description of an Evolution Strategy (ES) (Bäck 1996; Nissen 1997; Beyer and Schwefel 2002) relates to optimizing of a function F with n continuous decision variables: $F : \mathbf{R}^\mathbf{n} \rightarrow \mathbf{R}$

Without loss of generality a minimizing process is presumed. Every ES-individual corresponds to a vector and contains values for all decision variables $x_j \in \mathbf{R}$ $(j = 1,2,\ldots,n)$ of the given optimization problem.

Furthermore every individual contains $n_\sigma (1 \leq n_\sigma \leq n)$ standard deviations $\sigma_k \in \mathbf{R}_+$ ($k = 1,2,\ldots,n_\sigma$) as well. These are important concerning the mutation operator and are described as mutation step sizes. They form strategy parameters, which are optimized by a self-adaptive procedure. Provided $1 < n_\sigma < n$ applies, then the standard deviations $\sigma_1, \sigma_2,\ldots, \sigma_{n\sigma-1}$ are linked to the decision variables $x_1, x_2,\ldots,x_{n\sigma-1}$, while the standard deviation $\sigma_{n\sigma}$ applies to the remaining variables $x_{n\sigma}, x_{n\sigma+1},\ldots, x_n$. It is common to use either $n_\sigma = 1$ or $n_\sigma = n$ standard deviations. The running of an ES can be divided into the following steps:

Step 1: Initialization
In the initialization phase a starting population $P(t = 0)$ of μ individuals $\vec{a}_i = (\vec{x}_i, \vec{\sigma}_i)$ $(i = 1,2,\ldots,\mu)$ is generated.[3] As long as there is no prior knowledge about the location of the global optimum, the individuals of the initial population should be uniformly distributed over the search space. It is recommended (Bäck 1996, p. 83) to choose a standard deviation of $n_\sigma = n$ with a consistent starting value of $\sigma_k^{(0)} = 3.0$.

Step 2: Evaluation of the Initial Solutions
A fitness value $\Phi(\vec{a}_i)$ is assigned to each individual $\vec{a}_i$. Commonly the objective function value and the fitness value are identical: $\Phi(\vec{a}_i) = F(\vec{x}_i)$ $(i = 1,2,\ldots,\mu)$.

Step 3: Generating Offspring
In the scope of an ES-generation, λ descendants are produced from μ parents. In this process parents are repeatedly selected, their components are recombined, and the resulting offspring are mutated afterward. It is traditionally recommended that a ratio μ/λ of about 1/7 is chosen where μ is distinctly greater than 1. Standard values would be $\mu = 15$ and $\lambda = 100$ (Beyer and Schwefel 2002). The ES population with μ individuals contains *selected parents* other than in the case of the population of GA. The following sub-steps have to be run through λ- times.

Sub-step 3-1: Stochastic Parent Selection
In the basic ES a descendant is produced by two parents respectively. Initially two individuals, $\vec{a}_{P1}$ and $\vec{a}_{P2}$, are stochastically determined as

parents of the offspring. All individuals have the same selection probability $1/\mu$.

Sub-step 3-2: Recombination (Including Replication)

Subsequently the components of both parent vectors recombine and the result is considered as the offspring. It has been shown empirically that it is advantageous to use different recombination schemes for the decision variables and the strategy parameters. To determine the values of the decision variables of the offspring (Index O), for each decision variable the value of the one or the other parent (indices P_1 and P_2) is stochastically selected with equal probability. This approach is referred to as *discrete recombination*. It strongly resembles the Uniform Crossover in GA. On the other hand the parental manifestation of the mutation step widths is averaged in the offspring. This is called *intermediary recombination*:

$$x_{O,j} = x_{P_1,j} \text{ or } x_{P_2,j} \qquad \left(j = 1,2,\ldots,n;\, P_1, P_2 \in \left(1,2,\ldots,\mu\right)\right)$$

$$\sigma_{O,k} = 0{,}5 \cdot \left(\sigma_{P_1,k} + \sigma_{P_2,k}\right) \; \left(k = 1,2,\ldots,n_\sigma\right)$$

Figures 8.4 and 8.5 illustrate the working of the recombination forms. Further recombination variants exist.

Sub-step 3-3: Mutation of the Offspring

In this step the offspring is mutated. Initially the mutation step sizes are changed by multiplication with a logarithmic normally distributed random variable. Subsequently the manifestation of each decision variable is

Fig. 8.4 Discrete recombination

Parent 1	4.0	6.2	1.8	0.3
Parent 2	8.0	3.6	0.8	0.7
Offspring	6.0	4.9	1.3	0.5

Fig. 8.5 Intermediary recombination

mutated by adding a $(0, \sigma_j{}')$-normally distributed random variable to its value[4]:

$$\sigma_k' = \sigma_k \cdot \exp\left(\tau_1 \cdot N(0,1) + \tau_2 \cdot N_k(0,1)\right)$$
$$x_j' = x_j + \sigma_j' \cdot N_j(0,1) \qquad \forall j \geq n_\sigma : \sigma_j' = \sigma'_{n_\sigma}$$

$N(0,1)$ refers to the unique realization of a standard normally distributed random variable while $N_j(0,1)$ expresses that the manifestation of a standard normally distributed random variable has to be newly determined for each value of the counter j. The values τ_1 and τ_2 are exogenous constants. The global factor $\exp(\tau_1 \cdot N(0,1))$ has a uniform influence on the changes of all mutation step sizes, whereas the factor $\exp(\tau_2 \cdot N_k(0,1))$ makes an individual adaptation for single step size possible. Favorable values for both strategy parameters τ_1 and τ_2 are in the range 0.1 up to 0.2 (Kursawe 1996).

The mutation is simplified if only one standard deviation is used.

$$\sigma' = \sigma \cdot \exp\left(\tau_0 \cdot N(0,1)\right)$$
$$x_j' = x_j + \sigma' \times N_j(0,1)$$

The use of normally distributed mutations in the case of decision variables are commonly justified by the observation that children resemble their parents and that small changes occur more often than large ones in natural inheritance. On the other hand the logarithmic normal distribution for the multiplicative mutation of the standard deviations has the following advantages (Schwefel 1995, p. 143):

- The standard deviation stays automatically positive.
- Small changes occur more often than large ones.
- The median of the modification is 1, so that in the absence of selection, it does not come to a drift. In this case the mutation process will be neutral. Nevertheless it can happen that the standard deviation during multiplicative mutation is reduced to a value of close to zero (lower than a minimum value $\varepsilon > 0$). In this case, the respective standard deviation is set on ε. This happens because the optimization process of the original problem would otherwise be reduced to a search process in a subspace.

The ES can, therefore, find favorable values for the mutation step sizes in a self-adapting manner.

Sub-step 3-4: Evaluation of the Offspring
The resulting offspring $\vec{a}' = \left(\vec{x}', \vec{\sigma}'\right)$ is now evaluated:

$\Phi(\vec{a}') = F(\vec{x}')$

Subsequently this descendant is added to the initially empty pool of the new generation.

Step 4: Deterministic Selection for Survival
In the next selection step, the μ best individuals in regard to their fitness from λ offspring provide the new population. This procedure is described as (μ,λ)- selection or "comma selection". The lifespan of each individual is, thus, limited to one generation. Alternatively, the parents and their offspring can compete for survival. This variant is called $(\mu+\lambda)$-selection ("plus selection"). The ratio μ/λ controls the rigorousness of the selection (selection pressure). The smaller the value of the fraction the more rigorous is the selection. In the case of a multimodal function, it is recommended to use a softer selection than with unimodal functions. In this way, the danger of premature convergence to local sub-optima is reduced.

Step 5: Continue with Step 3 Until a Termination Criterion Holds
Further procedure variants of the Evolution Strategy can be found in the literature (Bäck 1996; Hansen and Ostermeier 2001; Beyer and Schwefel 2002; Hansen 2006; Bäck et al. 2013).

Extensions and Modern EA-Variants

Numerous extensions to the basic variants of EA have been proposed over the years. Due to space limitations we can only briefly point to some of them and refer the reader to suggested literature for a more detailed understanding.

An important modern stream of EA is Estimation of Distribution Algorithms (EDA), originally proposed by Mühlenbein and Paaß (1996). EDA replace the creation of offspring via recombination and mutation in classical EA by the sampling of a distribution previously learnt from the selected individuals. Thus, in EDA offspring are created based on an explicit probability distribution. More specifically, first the dependencies between variables describing a candidate solution are modeled. Then the parameters of this model are estimated from the current population to create a conditional probability distribution over the variables. In the last step, offspring are created by sampling this distribution. More details can be found in the books of Lozano et al. (2006) as well as Eiben and Smith (2015). For a comparison with contemporary ES see in particular Hansen (2006).

Differential Evolution (DE) is another acknowledged modern form of EA, going back to the Technical Report of Storn and Price (1995). Basically, in DE the current population members are perturbed with the scaled differences of randomly selected and distinct population members. Therefore, no probability distribution has to be used for generating the offspring. A survey of the state-of-the-art in DE can be found in Das and Suganthan (2011). Sarker et al. (2014) propose a new mechanism in DE to dynamically select the best performing combinations of parameters (amplification factor, crossover rate, and the population size) for an application during the course of a single run, thus addressing the important problem of optimally selecting EA control parameters.

Genetic Programming (GP), originally created by Koza (1992, 1994) has grown to become one of the major EA-variants today. GP follows the general procedure of GA, but with a specific representation of solutions in the population. Koza started from the observation that representation is a key issue in GA, because it is actually the coding of the underlying problem that a GA can manipulate. He concluded that for many problems in machine learning and artificial intelligence, the most natural known representation for a solution is a hierarchical computer program of indeterminate size and shape. Thus, GP aims to evolve computer programs that solve complex application problems. GP frequently starts the evolutionary process with an initial population of randomly generated programs of functions and terminals appropriate to the problem domain. Each program is evaluated in terms of fitness by running it on a number of representative test problems and averaging the results. Many extensions and variants to this basic form of GP have been created in recent years. For a helpful overview of the field see Poli et al. (2008).

Many practical problems in management are characterized by multiple objectives. For instance, economic goals might have to be combined with objectives derived from aspects of customer satisfaction. The simplest approach then is to have a fitness function that additively includes all objectives, possibly prioritized via individual weights. However, this approach is often unsatisfactory and a true Pareto-optimization is sought after. Different variants of EA have been developed to treat such multi-objective problems, termed multi-objective EA (MOEA). A brief introduction to MOEA can be found in Chap. 12 of the textbook by Eiben and Smith (2015). Other, more extensive treatments of MOEA include Deb (2001), Abraham et al. (2005), Coello Coello et al. (2007), and Gaspar-Cunha et al. (2015). Ishibuchi et al. (2008) discuss issues related to scaling-up MOEA to many-objective optimization problems.

Most application problems we encounter in management practice are somehow constrained. Constraint handling in EA can be approached in different ways. One option is to repair invalid solutions so that the result is valid, that is, respects the constraints of the problem. The repair heuristic is generally specific for the application problem investigated. Another approach is penalizing invalid solutions. Usually the penalty is related to the degree of constraint violation in a given solution, thus leading the search back to a valid area of the solution space. A third option would be to convert constraints to objectives and use a MOEA to find good solutions for the transformed problem. Other approaches to cover constraints are conceivable. An introductory text on constraint handling is again contained in the textbook of Eiben and Smith (2015, Chap. 13). An overview of constraint-handling techniques in nature-inspired optimization can be found in Mezura-Montes and Coello Coello (2011). A very recent treatment of the subject is the book of Datta and Deb (2015). Finally, the bibliography compiled by Alander (2015) contains many references to publications on MOEA and constrained optimization with EA.

The last aspect to be mentioned here are stochastic problems. As uncertainty is all around us, it is often encountered in management practice that some stochastic influence (noise) is present. Consequently, deterministic models may become oversimplified versions of real-life systems. With a stochastic fitness function, the calculated fitness value of a particular solution varies, depending on the current stochastic influence, that is, different evaluations of the same solution generate different results. On a general level, this makes optimization more difficult. However, it has been shown (Nissen and Propach 1998) (Beyer 2000) that population-based approaches, such as EA, are more robust to noisy fitness functions than point-based optimization methods, such as Simulated Annealing and Threshold Accepting. A useful approach in noisy optimization is taking the mean of several fitness evaluations to be the "true" fitness of a given solution during optimization. In a related line of thought, Juan et al. (2015) describe a general methodology that allows for extending metaheuristics through simulation to solve stochastic combinatorial optimization problems.

Benefits and Limitations

Evolutionary Algorithms have a number of *advantages* in practical management applications.

Wide Applicability (Domain Independence) As Eiben and Smith (2015, p. 21) point out, we are faced with the challenge of deploying automated solution methods for more and more problems, which are ever more complex, in less and less time. In this situation, there is an urgent need for robust algorithms that are applicable to a wide range of problems, do not need much tailoring for specific problems, and deliver good (not necessarily optimal) solutions within acceptable time.

The basic forms of EA assume almost no application-specific prior knowledge. Thus, even when no foreknowledge is available, EA can be used to attain improvements. Moreover, such canonical EA do not make restrictive assumptions concerning the problem structure. They are therefore widely applicable optimization methods. There is, however, a trade-off between the application width and the solution quality.

Suitability for Highly Complex Search Spaces Due to the population-based search approach and stochastic procedure elements, EA are suitable for complex nonlinear problems that would pose a problem for classical optimization approaches. Michalewicz (1996, p. 16) highlighted that, unlike other heuristics, EA may process simultaneously several points of the search space, thus providing information formation and exchange between search directions. Their tolerance toward a temporary decline of fitness values makes it possible for EA to leave local sub-optima again. Moreover, the EA-results are barely dependent on the initial population, provided it is sufficiently heterogeneous.

No Restrictive Requirements Concerning the Objective Function EA do not require the objective function to be continuous and differentiable as classical derivative-based approaches do. Rothlauf (2011, p. 93) points out that usually only two basic requirements are relevant:

- One must be able to represent complete solutions to a problem such that variation operators can be applied to them.
- One must be able to compare the quality of two solutions and indicate which of the two solutions has a better objective function value.

Furthermore EA are resilient toward stochastic influences in the fitness calculations (Nissen and Propach 1998). Since many management problems are characterized by noise and changing conditions, this is an advantage that should not be underestimated. Also, highly nonlinear problems may successfully be solved with EA.

Basic Principles Are Easily Understood Often in practice, advanced optimization methods are in principle available (for instance, as part of IT-based planning systems). However, since users do not understand how the results were produced these methods are often not applied, leading to a waste of resources. The basic idea and approach of EA are easily understood by laymen, which improves the practical acceptance.

Many Modification and Hybridization Possibilities EA can be adapted to the given application in many ways as will be outlined in section "How to Apply", raising the attainable solution quality. Moreover, there are manifold possibilities to combine EA with local hill-climbing, heuristic initialization, intelligent decoding procedures, neural networks or fuzzy approaches, and so on by which the solution quality of the total system can be further improved. It is today common standard to recombine and tune the most efficient operators of existing metaheuristic frameworks to achieve the best possible solution method for a given application.

Good Parallelizability Based on their inherent parallel processing structure, EA can efficiently be implemented on parallel computer architecture, thus shortening the elapsed time to produce good results. For this purpose there are various parallelization models.

As opposed to the benefits of EA, the following *disadvantages* need also to be mentioned:

Heuristic Character For practical purposes, it cannot be ensured that EA will find the global optimal solution for a given problem in limited time. In this sense EA should be seen as heuristics. However, it can be argued that in a practical setting it is usually not important to prove global optimality, but a good solution in reasonable time is what constitutes a satisfying solution for many management problems.

Computational Intensity As EA are population-based methods with fairly complex search operators, often requiring the generation of many random numbers, they are generally CPU-intensive. They share this characteristic with competing metaheuristic approaches, such as Simulated Annealing or Tabu Search. With ever-increasing hardware performance, this disadvantage is declining in importance, though. It remains an issue particularly in the context of real-time optimization.

Low Efficiency in the Final Search Phase The search operators, in particular those of GA and GP, are not designed for fast local optimization, as should be efficient for the final phase of a run. Therefore, combinations of EA and local hill-climbing procedures are often used to compensate for this disadvantage.

Difficult Adaptation to the Given Application Problem As will be detailed in section "How to Apply", the application of "out-of-the-box" optimization methods will generally produce at best mediocre results. More specifically, an informed and integrated choice of heuristic design elements and the tuning of strategy parameters (control parameters) are necessary to arrive at really good results. The many degrees of freedom in the case of the concrete arrangement of EA can become a problem, in particular for inexperienced developers and users.

I like to compare the situation of the EA-designer with a cook who wants to prepare an excellent meal. The cook must choose proper ingredients and combine them well in order to achieve the goal. Some ingredients will go together well while others will spoil the result when used to a great extent or in the wrong context. In this situation, current research can offer some helpful advice, but still, designing high-quality heuristics is a difficult task that Rothlauf calls "an art" (2011, p. 12). In the end, it is necessary to gain practical experience in designing and implementing EA. One must develop a feeling for algorithmic behavior, just as a cook may learn much from reading recipes in books, but in the end it is own experience, creative ideas, and the informed combination of ingredients that will make the difference.

How to Apply

General Approach

On a general level, the application of EA follows the procedure of applying heuristics, as, for instance, outlined by Rothlauf (2011)[5]:

1. Recognizing the problem
2. Defining the problem
3. Constructing a model for the problem
4. Solving the model
5. Validating the obtained solutions
6. Implementing a solution

In the following, we shall briefly comment on these steps. For further details on applying modern heuristics the reader is referred to Rothlauf (2011) and Michalewicz and Fogel (2004). A more general introduction to the heuristic problem-solving approach can be found, for instance, in Foulds (1983).

Recognizing and Defining the Problem

Initially, the focus is on realizing that there are alternatives to the way of doing business which may provide a better quality of solutions. It is surprising (and frustrating) to encounter how much potential for optimization is actually *not* realized in companies today. Often this is due to people getting used to and content with certain ways of working, which may have been adequate at one time, but leave room for improvement as time goes by. Also, there is often a tendency to plan manually or with easily understandable low-end tools like MS Excel™ instead of employing powerful but more complex optimization tools. Moreover, focusing too much on small fractions of a greater business process may also hinder to become aware of potential for improvement on a larger scale.

Defining the problem at hand includes describing the different decision alternatives, identifying possible constraints, defining objectives of planning, and deciding on reasonable evaluation criteria to differentiate between alternative solutions. At this stage it should be clear whether one is faced with continuous or discrete decision alternatives. This in turn, influences how the problem should be modeled in the next step and which operators might be relevant for solving the model.

Constructing a Model for the Problem

Once the problem is defined, a model is to be constructed that keeps the characteristics of the real-world problem one is interested in but abstracts from unnecessary details and irrelevant aspects of reality in order to keep the model solvable in reasonable time. This step is critical, as a model that is too simplified will yield results that might not lead to a beneficial solution being implemented in the end.

At the core of this step is defining the decision variables that will allow the modeling of alternative solutions, but also defines the search space for the later solution method. Restrictions that hold for certain decision variables are added as constraints. Finally, an objective function that adequately represents the goals of the planning process must be defined. Here, one must also decide if a single objective is optimized or multiple objectives are relevant. Multiple objectives could be integrated as weighted components in an additive objective function or a true multi-objective optimization (searching for the Pareto front) might be required.

Solving the Model

General Upfront Considerations

Modeling the real-world problem in a valid but also clever way is crucial for solving this model. In particular, the modeler should be aware of efficient algorithms for solving certain classes of models to optimality in reasonable time (such as the Simplex Method for Linear Programming), as this can greatly speed up the solution process and improve the quality of results. It is definitely not a useful approach to start an optimization project with the idea of applying certain forms of heuristics in mind. In fact, if it is possible to construct a model in such a way, that it is a valid representation of the original problem and can be solved to optimality with an efficient algorithm in a reasonable time frame then no heuristic approach should be applied.

Moreover, as many authors have pointed out (e.g. Newell (1969), Wolpert and Macready (1997), Rothlauf (2011)), there is a trade-off between the generality (application range) and the power (solution quality) of an optimization method. Therefore, it is necessary to devise a problem representation and an adapted solution method that can exploit properties of the model in order to arrive at good results that really generate a competitive benefit in business practice. Put in different words, it is not useful to simply take some general purpose (black-box) evolutionary (or other) metaheuristic implementation and run it on the problem at hand. The results will in general be disappointing as no problem-specific tuning has been done.

Design of Heuristics

Consequently, the problem-specific design of the heuristic is at the core of succeeding in evolutionary optimization. The main design elements of EA are (Rothlauf 2011, p. 94):

1. Problem representation
2. Fitness function
3. Search strategy
4. Variation (search) operators
5. Initialization

Representing the value of decision variables at the genotype-level and devising a proper mapping from genotype to phenotype, is often a non-trivial activity. It is closely tied to the design of adapted variation operators that must be able to work on this representation and find (near-) optimal solutions. Good representations can make it easier to find good solutions while bad representations make it more difficult or even impossible. Therefore, deciding on the representation is an important step in designing and applying EA. Rothlauf (2011, p. 4) recommends a high locality of representations, meaning similarities between phenotypes must correspond to similarities between genotypes.

While classical GA often used binary representations, many management problems are of a combinatorial nature, that is, the decision variables can take values from bounded, discrete sets and various constraints may apply. Moreover, while the number of solution alternatives is finite and can in principle be enumerated, there is usually a very large number of alternative solutions of different quality. In this case, frequently "direct representations" are used, that is, the problem is represented in some form of "natural representation" (often permutations) and problem-specific search operators are directly applied at this phenotype level. For a more detailed discussion of representation-related issues see Rothlauf (2006).

The search space of alternative solutions may be viewed as a topological space with hills and valleys in terms of associated *fitness* values. In metric search spaces the similarity of solutions can be measured by some distance metric, such as the Hamming distance that measures on how many bits two given binary vectors differ. Intuitively, for a successful EA-application it is important to have a "smooth" search space where similar solutions to the underlying problem are associated with similar fitness. Note that the fitness of some solution is not necessarily identical to the objective function value, but could be some transformation of it, or even be determined by completely different approaches, such as simulation. Any mapping introduced here can have an influence on the smoothness of the search space and, thus, make it easier or harder to find good solutions.

Of major importance for the success of an EA is the correct choice of the dominant *search strategy*. Rothlauf (2011, p. 131) distinguishes between two fundamental concepts that are found in different modern heuristics: local search methods and recombination-based search methods. Local search methods exploit the locality of a problem, that is, the fact that distances between solutions correspond to their fitness difference. Mutation is such a variation operator that is based on local search. Recombination-based search operators, like crossover in GA, assume that the problem at hand can be decomposed in smaller sub-problems that can be solved more or less independently and put together to form a solution for the overall problem.

Search strategies are implemented through the choice of certain *variation operators* that are iteratively employed at the genotype-level to identify better solutions. The chosen variation operators implicitly define a neighborhood structure on the search space that characterizes solutions as "similar" to each other.[6] Operators that are based on local search, such as mutation, and operators that are based on decomposing the overall problem in sub-problems, such as crossover, create very different neighborhood structures in the search space. Thus, if the problem demonstrates high locality, this fact is not well exploited through the neighborhood structure created by crossover and vice versa. It must be stressed that not all practical problems (and associated models) are decomposable and, thus, the use of recombination as a variation operator is not advisable on these types of problems. More generally, the decision on the dominant search strategy and associated variation operators is a problem-specific choice, and often real-world problems will show locality as well as a certain degree of decomposability.

Equally important is the choice of selection pressure, as too much pressure can lead to premature convergence in local optima, while a selection that is not rigorous enough may hinder a focused search, particularly in the final stage of an optimization run. Selection consists of two steps, the selection of individuals for reproduction (parent selection), and the complementing replacement strategy that, after offspring have been created, decides which members of the population are replaced by the new individuals. As Eiben and Smith (2015, p. 33) point out, parent selection is often stochastic while replacement is frequently deterministic. A sizable number of selection operators are differentiated in the literature that are associated with different selection pressure, such as rank-based selection, tournament selection and fitness-proportional selection (see Eiben and Smith 2015 for more details). It could actually be a good idea to vary the selection pressure throughout an optimization run, starting with a low selective pressure to allow for an initial full exploration of the search space.

EA require the creation of an initial set of solutions from which the search starts. If no prior knowledge about the application problem is available, the *initialization* is frequently done at random so that initial solutions are uniformly distributed in the search space. If knowledge about the problem at hand is available, this should be used during initialization (and possibly also for other design elements of EA) as to bias the search to promising regions of the solution space. However, care should be taken not to bias the initialization too strongly in the direction of high-quality solutions as this could lead to premature convergence on local sub-optima. This is particularly relevant when the population size of the EA is low and the selection pressure is high. Controlling diversity of the EA population is a delicate issue that is associated with creating a balance between broadly *exploring* the search space to find attractive regions and *exploiting* good solutions that have already been found. A variety of measures have been devised in the literature to keep diversity up in order to make recombination operators work properly. For a discussion of exploration and exploitation as well as a brief overview of diversity-related measures see, for instance, Eiben and Smith (2015).

Next to the main heuristic design elements discussed briefly above, the EA-designer must also decide on further *strategy parameters (control parameters)*, such as the population size, crossover and mutation rate (or some adaptive procedure to change these rates in the course of an optimization run). The book of Lobo et al. (2007) contains a helpful collection of papers on issues associated with parameter setting in EA.

Moreover, it might be necessary to design a repair heuristic that converts infeasible into similar feasible solutions or, alternatively, create penalty functions that will lead the optimization to feasible regions of the search space. Finally, hybridizing EA with other optimization methods (such as local hill-climbing) is a design option that should be taken into account. Summarizing, there is a wealth of interdependent aspects the designer of a powerful evolutionary heuristic must be aware of and handle adequately.

Validation and Implementation

Validating what is a (near-)optimal solution to the model can be done through sensitivity analysis or retrospective tests (Rothlauf 2011, p. 12).

While sensitivity analysis aims to reveal how the optimal solution depends on variations of the model, a retrospective test uses historical data to measure how well the model and its solution would have performed if they had been used in the past. One should keep in mind here that in business practice often the context of an implementation is subject to gradual (or even abrupt) change over the years, and, thus, a certain robustness of a solution is required under these conditions. This is of particular importance, when a solution is only implemented once (e.g. in factory layout planning) as opposed to a model that is used and solved repeatedly (e.g. in production planning).

Self-Adaptation of Strategy Parameters and Hyper-Heuristics

To make heuristic design easier and create better optimization results, next to some rules of thumb, self-adaptive mechanisms and hyper-heuristics are suggested in the literature. We can only briefly mention a few examples here and refer for further reading to the cited references, the book on parameter optimization in EA by Lobo et al. (2007), as well as general textbooks on EA, such as Eiben and Smith (2015), De Jong (2006), and Fogel (2006).

Self-adaptation of the mutation rate has long been used in Evolution Strategies (see e.g. Schwefel 1981). Hansen and Ostermeier (2001) developed a de-randomized form of self-adaptation in ES. In other forms of EA, self-adaptation came later. An early account of approaches for parameter control by self-adaptation can be found in Bäck (1998). Bäck (1992) also studied self-adaptation in GA. More recent overviews on self-adaptation in EA are Meyer-Nieberg and Beyer (2007), and Smith (2008) for combinatorial optimization (as is most relevant for management applications). Smith not only discusses mechanisms to control the parameters defining crossover and mutation, but also the very definition of local search operators used within hybrid EA. Kühn et al. (2013) propose a mechanism for balancing control parameters to the fitness of individual chromosomes in a multi-chromosome representation of a complex scheduling problem.

A parameter-free GA was proposed by Sawai and Kizu (1998). This approach is applied in Matsui et al. (2002) to solve a Job-Shop Scheduling Problem without the need to set control parameters for the GA in advance. Dahal et al. (2008) in the context of a real-world workforce scheduling problem propose a variable fitness function that describes how the weights of a weighted sum fitness function change over the iterations of a search process. The authors then use an evolutionary approach to evolve weights for each of the (multiple) objectives. The variable fitness function can potentially enhance any search-based optimization heuristic where multiple objectives can be defined.

A different approach to heuristic design is taken by hyper-heuristics. According to Burke et al. (2010) hyper-heuristics share the common goal of automating the design and adaptation of heuristic methods to solve hard computational search problems. The authors differentiate two main classes of hyper-heuristic approaches: "heuristic selection", which are methodologies for choosing or selecting existing heuristics, and "heuristic generation", namely methodologies for generating new heuristics from components of existing heuristics. The higher-level strategy for selecting or generating heuristics need not be a heuristic, but could be any kind of knowledge-based technique (such as case-based reasoning). Moreover, online learning mechanisms may be employed to improve the higher-level strategy over time. In Kubalik (2012) the idea of a hyper-heuristic is applied to a set of standard combinatorial optimization problems with relevance for management applications, such as the bin packing problem, personnel scheduling, flowshop scheduling, traveling salesman, and the vehicle routing problem.

Bezerra et al. (2014) suggest an automatic algorithmic configuration of EA for multi-objective combinatorial optimization. More specifically, they use an offline algorithm configuration tool, to automatically select components from a framework of different multi-objective EA (MOEA) that are more effective for the particular problem variant at hand, thus producing novel MOEA. They demonstrate the usefulness of their approach by direct comparison to traditional MOEA on four variants of the multi-objective permutation flowshop problem.

Hybridization and Memetic Algorithms

Hybridization of EA with other solution approaches, such as gradient-based optimization, other heuristics or machine learning approaches, is getting popular due to their capabilities in handling real-world problems involving complexity, noisy environment, imprecision, uncertainty, and vagueness (Grosan and Abraham 2007). Modern heuristic frameworks are actually built on the idea that design elements from different methods of resolution may have to be combined to produce a truly powerful solution approach.

The idea of hybridizing EA with other solution methods is actually not new. In Biethahn and Nissen (1994) various options to combine EA with simulation are discussed in the context of economics and management applications. There have also been many suggestions how to combine EA with Artificial Neural Networks (ANN). Basically EA can be used to evolve ANN connection weights, architectures, learning rules, and input features. An early overview of these topics is given by Yao (1999). A more recent account of research can be found in Floreano et al. (2008). As a practical example from the management domain, in Harrald and Kamstra (1997) Evolutionary Programming is used to evolve ANN for combining financial forecasts. Desell et al. (2014) evolve ANN weights for time-series prediction of general aviation flight data, using large, noisy, realistic data sets.

In an interesting co-evolutionary approach for discovering fuzzy classification rules Mendes et al. (2001) combine a GP algorithm evolving a population of fuzzy rule sets and an (1+5)-ES evolving a population of membership function definitions. Both populations co-evolve, so that the final result of the co-evolutionary process is a fuzzy rule set and a set of membership function definitions that are well adapted to each other. This approach produces competitive results and has several advantages, for instance, discovering fuzzy rules, which tend to be more intuitive for the user than the rules discovered by benchmark methods.

Grosan, Abraham, and Ishibuchi in their 2007 book give a nice account of recent results concerning hybrid EA. A helpful overview of hybrid metaheuristics with EA specializing in intensification and diversification is also given by Lozano and Garcia-Martinez (2010).

Memetic Algorithms, initially proposed by Moscato (1989), are based on the assumption that optimization problems can be tackled more efficiently by hybridizing and combining existing algorithmic structures rather than using existing paradigms in isolation. So, instead of deciding between alternative modern heuristics, a solution method can be generated by combining the strong points of various paradigms and obtaining a solver which is capable to outperform each paradigm individually (Neri et al. 2012, p. IX). The term "memetic" is derived from the notion of a "meme" as introduced by R. Dawkins (1976) as the concept of self-reproducing ideas, an analogy to the gene but at a higher level in the context of cultural evolution. In Xhafa (2007) a memetic algorithm is successfully applied to the problem of efficiently assigning application jobs to grid resources, a hard multi-objective optimization problem.

Using Computational Frameworks

Basically, one could start developing an Evolutionary Algorithm from scratch and implement it with some standard programming language, say Java or C#. The author did this quite a few times in the 1990s, and at that time it was a reasonable approach to develop a tailored solution. However, nowadays many good technical frameworks are available that deliver a variety of ready-to-use EA-components, such as selection and variation operators, also addressing issues such as multiple objectives, hard and soft constraints as well as genotype-phenotype distinction (e.g. Lukasiewycz et al. 2011). Developing a tailored heuristic to cover a given application problem then mainly consists of choosing the components one wants to use, customizing some parameters and building a structure similar to playing with Lego blocks. If a standard operations research problem (such as the traveling salesman problem) is to be solved, even complete implementations or code may be found online. Relevant sources are listed, for instance, in the publication "An Indexed Bibliography of Genetic Algorithm Implementations" (available at http://lipas.uwasa.fi/~TAU/reports/report94-1/IMPLEbib.pdf). In the following, we briefly mention some frameworks for developing EA in different programming languages. Many more can be found, but here our intention is only to demonstrate the available variety.

1. **Evolving Objects** (EO, accessible at http://eodev.sourceforge.net/) is a template-based, C++ evolutionary computation framework that aims at speeding up the development process of EA. For classical problems, such as a black-box problem with real-valued variables, code already exists. Thus, the user only needs to choose components to form the heuristic and connect it to the relevant fitness function. In more complex applications, the user has to code a class that describes how the individuals are represented, and perhaps some variation operators, but most of the other operators (selection, replacement, stopping criteria, etc.) are available in EO. More information can be found in Keijzer et al. (2002).
2. **Opt4J** is a modular framework for metaheuristic optimization, implemented in Java and available at http://www.opt4j.org. It provides an efficient design and development approach for complex optimization tasks by decomposing these into correlated subtasks that are optimized concurrently. For this purpose, a strict distinction between the genotype and phenotype is imposed that separates genetic representation and solution representation of an optimization task. The optimization tasks are decomposed into subtasks that might be designed and developed separately. In order to enable this modular design of optimization tasks, compositional genotypes and appropriate operators are proposed. A reference implementation of the proposed concept is presented in Lukasiewycz et al. (2011).
3. **Distributed Evolutionary Algorithms in Python** (DEAP, accessible at: http://deap.readthedocs.io/en/master/) is a Python-based EA-framework for rapid prototyping that seeks to make algorithms explicit and data structures transparent. It works with parallelization mechanisms such as multiprocessing and SCOOP. More information on DEAP is available in Fortin et al. (2012).
4. **Multiobjective Evolutionary Algorithms** (MOEA, accessible at: http://moeaframework.org/) is an open source Java library for developing and experimenting with MOEA and other general purpose multi-objective optimization algorithms. A number of methods are

provided out-of-the-box. In addition, the MOEA framework provides the tools necessary to rapidly design, develop, execute, and statistically test optimization algorithms. Also included are major test problems from the literature. Additionally, new problems written in Java or other languages can be incorporated. Recently, the framework was used in Cocaña-Fernández et al. (2016) for improving the eco-efficiency of high-performance computing clusters.

A detailed account on benefits and limitations of using an EA-framework is given in Fink and Voß (2003). These authors also claim, supported by references to applications of the framework, that results are frequently as good as fully tailored approaches. A survey and benchmarking of various EA-frameworks can be found in Parejo et al. (2012).

Using EA in Commercial Application Software

Ready-to-use EA are today part of optimizers in various types of business software. In the following, we will focus on the GA available in MathWorks Global Optimization Toolbox™ as well as in SAP's Advanced Planner and Optimizer™. These heuristics can be used with some customizing only and no individual code need to be developed (or put together as with an EA-framework).

MathWorks Global Optimization Toolbox

MathWorks (2016) in its MATLAB-related product portfolio offers the so-called Global Optimization Toolbox, an interactive tool to define and solve optimization problems by specifying the problem and picking from a pre-defined set of available solvers. Problems can be imported from MATLAB. Intermediate or final results can be stored, visualized in different ways, or exported to MATLAB for further processing.

The toolbox consists of methods for global optimization problems with continuous, discrete, or stochastic objective functions. The solvers include GA, Simulated Annealing, Pattern Search, and a multi-start heuristic. The GA is available in a basic and a multi-objective version; the latter can create a Pareto front of solutions. The two GA can be adapted to the user's requirements up to a certain degree. In particular, the user can influence the population size, number of offspring, creation of the initial population as well as the fitness-scaling. Moreover, selection, crossover and mutation operators can be chosen by the user, and the migration between (optional) sub-populations can be customized. It is even possible to create customer-specific algorithmic functions and data formats. Constraints can be customized and various termination criteria applied. The Global Optimization Toolbox by MathWorks offers not quite the degree of flexibility as the previously mentioned EA-frameworks, but can be used with less expert knowledge. However, the toolbox is still more versatile than the GA in our next example, a complex Advanced Planning and Scheduling System (APS).

SAP Advanced Planner and Optimizer (APO)

In 1998, SAP AG introduced its APS, the SAP Advanced Planner and Optimizer (APO), which today is part of the solution SAP Supply Chain Management (SCM). SAP APO contains advanced features for supply chain planning that go beyond the available features in its Enterprise Resource Planning System SAP ERP. Basically, SAP APO receives master data and transaction data from SAP ERP which is then used for the planning. Finally, results, such as planned orders and purchase requisitions, are transferred from SAP APO to SAP ERP and executed there. The planning in SAP APO is done at different levels, usually starting with long-term demand planning, then medium-term supply network planning, short-term production planning and detailed scheduling, and optionally distribution and transportation planning (vehicle routing and scheduling).

As logistics and supply chain planning are complex optimization domains, SAP decided to include a sizable number of classical and modern algorithms and heuristics in SAP APO (SAP 2016). In the module PP/DS (production planning/detailed scheduling) a GA is available that allows for some customization to better fit the customer's needs. A maximum runtime for the GA can be entered and weights can be assigned to different objectives in the cumulative objective function. Also, setting certain flags indicates whether particular conditions (such as production line priorities) apply or not. During optimization, the system carries out finite scheduling to achieve a feasible production plan. This means the system optimizes the production dates and the resource allocation for operations, based on criteria like make-span, setup costs, and delay costs. During optimization, the system considers the various constraints in the schedule, where hard and soft constraints can be differentiated. When a hard constraint is violated and the GA cannot resolve this issue, the GA run is automatically terminated. The sequence of orders then remains as before the optimization started. Moreover, a simulative optimization can be executed and the optimized schedule can be adopted into the operative planning version

While this allows SAP customers to employ a potentially powerful optimization tool without further programming, some caution appears necessary. First of all, the possibilities of tailoring the GA are lower than with standard EA-frameworks. However, the GA in SAP APO is only used for limited pre-defined purposes (such as detailed scheduling for single or multiple production lines), so a lower flexibility in algorithmic design is acceptable. Second, even the customization options that *are* available will generally extend beyond the knowledge of SAP users and even most SAP consultants. Thus, it is recommended to employ a person with special knowledge in optimization and metaheuristics to customize the GA and other algorithms in SAP APO to generate maximum benefit from these advanced planning features.

However, the user is not *required* to use complex heuristics like GA during optimization in SAP APO. For instance, in model-mix-planning a number of alternative, partly simpler algorithms are available. Some of them (e.g. Linear Programming) can also be used in (sequential) combination with the GA.

Conclusions

Summarizing, it is useful to differentiate between several perspectives. From a methodological point of view the myriad of nature-inspired heuristics is rather confusing for casual users and definitely not helpful in creating more acceptance for metaheuristics in practical applications. Moreover, there is evidence (e.g. Weyland 2015; Sörensen 2015) that at least some of the nature-inspired concepts published recently (such as Harmony Search) are rather old wine in new skins. One should better look for over-arching and bearable concepts within EA and related metaheuristics, putting together a common framework (or toolbox) that integrates different options of solution representation, search operators, selection mechanisms, constraint-handling techniques, termination criteria, and so on. Fortunately, such frameworks are available today. Then, properly choosing the components for a hybrid heuristic from such a framework requires a deep understanding of which components actually fit together and work well for certain classes of problems or search spaces.

Moreover, following the No-Free-Lunch Theorem (Wolpert and Macready 1997), problem-specific fine-tuning of heuristics remains important to achieve truly good results. Today, much of this is still more an art than a science, despite helpful textbooks like Rothlauf (2011). As a consequence, there is lots of interesting research issues to be solved along these lines. This process is indeed ongoing for several years now, but as Sörensen et al. (2016) point out also bypassed by useless initiatives to invent ever "new" nature-inspired heuristics.

For more than 20 years now, EA have become integrated in business software products (e.g. for production planning), so that, as a result, the end user is often unaware that an evolutionary approach to problem solving is employed (Nissen 1995). Today, large software companies like SAP use EA in their enterprise software. However, since customizing options are limited in these systems, it appears fair to say that the full power of EA is frequently not unleashed by such standardized approaches. This confronts us with a dilemma. Simple forms of EA that can be fairly easily understood and applied widely are of only limited power. If we want to use metaheuristics like EA to full extent then this requires knowledge and experience in their design and application. Most users have neither the qualification nor the time to dive that deep into the matter. According to my own observations as an IT-consultant, this unfortunately also holds for most consultants that could potentially help customers in applying modern heuristics. Thus, creating really powerful applications of EA today is frequently an issue for only a small number of highly specialized IT-companies.

This situation is unsatisfactory and could only be changed if the design and application of modern heuristics becomes an important topic in general management studies at universities and related higher learning institutions. I would argue that this should indeed be the case, because in today's digital era there is strong evidence that we are entering an age of knowledge-based competition where a qualified workforce that is able to creatively use modern tools for data mining, ad hoc reporting, heuristic optimization, artificial intelligence, and so on will make the difference in many branches of industry.

Further Information

The current submission (mail) address of the Evolutionary Computation Digest, which is one of the central means of communication among researchers in EA, is EC-Digest-L@listserv.gmu.edu. The archive of this list can be found at http://listserv.gmu.edu/cgi-bin/wa?A0=EC-DIGEST-L. A list of journals that publish articles in evolutionary computation and related areas, compiled by Michael Lones, can be found at http://www.macs.hw.ac.uk/~ml355/journals.htm.

Notes

1. For a stricter understanding of the term "metaheuristic" see Lones (2014), who points out that a metaheuristic is like a design pattern, in that it encapsulates knowledge that may be applied to the design of a range of specific optimization algorithms. Likewise, there is no reason why a particular algorithm cannot implement multiple metaheuristics, making use of complementary ideas of how to search for optima within solution landscapes. This is very much in line with the framework-centric view on metaheuristics of Sörensen et al. (2016).
2. In the following the generation index is dropped.
3. In the following the generation index is dropped.
4. The index O for the offspring is neglected in the following.
5. For an edited book on designing metaheuristics see Borenstein and Moraglio (2014).

6. It is also possible to explicitly define the relevant neighborhood of a solution—for example, see Günther and Nissen 2009.

References

Abraham, A., Jain, L. C., & Goldberg, R. (Eds.). (2005). *Evolutionary Multiobjective Optimization: Theoretical Advances and Applications*. Berlin: Springer.

Alander, J. (2015). *An Indexed Bibliography of Genetic Algorithms & Pareto and Constrained Optimization. University of Vaasa*. Technical Report, (Available at http://lipas.uwasa.fi/~TAU/reports/report94-1/gaPARETObib.pdf).

Bäck, T. (1992). Self-Adaptation in Genetic Algorithms. In *Proceedings of the 1st European Conference on Artificial Life* (pp. 263–271). Cambridge, MA: MIT Press.

Bäck, T. (1996). *Evolutionary Algorithms in Theory and Practice*. New York: Oxford University Press.

Bäck, T. (1998). An Overview of Parameter Control Methods by Self-Adaptation in Evolutionary Algorithms. *Fundamenta Informaticae, 35*, 51–66.

Bäck, T., Foussette, C., & Krause, P. (2013). *Contemporary Evolution Strategies*. Berlin: Springer.

Baker, J. E. (1985). Adaptive Selection Methods for Genetic Algorithms. In J. J. Grefenstette (Ed.), *Proceedings of an International Conference on Genetic Algorithms and Their Applications* (pp. 101–111). Hillsdale: Lawrence Erlbaum.

Baker, J. E. (1987). Reducing Bias and Inefficiency in the Selection Algorithm. In J. J. Grefenstette (Ed.), *Proceedings of the Second International Conference on Genetic Algorithms* (pp. 14–21). Hillsdale: Lawrence Erlbaum.

Beyer, H.-G. (2000). Evolutionary Algorithms in Noisy Environments: Theoretical Issues and Guidelines for Practice. *Computer Methods in Mechanics and Applied Engineering, 186*, 239–267.

Beyer, H.-G., & Schwefel, H.-P. (2002). Evolution Strategies—A Comprehensive Introduction. *Natural Computing, 1*, 3–52.

Bezerra, L. C. T., Lopez-Ibanez, M., & Stützle, T. (2014). Automatic Design of Evolutionary Algorithms for Multi-Objective Combinatorial Optimization. In *Proceedings of PPSN XIII, LNCS 8672* (pp. 508–517). Berlin: Springer.

Biethahn, J., & Nissen, V. (1994). Combinations of Simulation and Evolutionary Algorithms in Management Science and Economics. *Annals of Operations Research, 52*, 183–208.

Biethahn, J., & Nissen, V. (Eds.). (1995). *Evolutionary Algorithms in Management Applications*. Berlin: Springer.

Borenstein, Y., & Moraglio, A. (2014). *Theory and Principled Methods for Designing Metaheuristics*. Berlin: Springer.

Burke, E. K., Hyde, M., Kendall, G., Ochoa, G., Özcan, E., & Woodward, J. R. (2010). A Classification of Hyper-Heuristic Approaches. In M. Gendreau & J. Y. Potvin (Eds.), *Handbook of Metaheuristics* (pp. 449–468). Heidelberg: Springer.

Cocaña-Fernández, A., Sánchez, L., & Ranilla, J. (2016). Improving the Eco-efficiency of High Performance Computing Clusters Using EECluster. *Energies, 9*, 3. https://doi.org/10.3390/en9030197.

Coello Coello, C., Lamont, G. B., & van Veldhuizen, D. A. (2007). *Evolutionary Algorithms for Solving Multi-Objective Problems*. Berlin: Springer.

Dahal, K., Remde, S., Cowling, P., & Colledge, N. (2008). Improving Metaheuristic Performance by Evolving a Variable Fitness Function. In J. Hemert & C. Cotta (Eds.), *Proc. EvoCOP 2008. LNCS 4972* (pp. 170–181). Berlin: Springer.

Das, S., & Suganthan, P. N. (2011). Differential Evolution: A Survey of the State-of-the-Art. *IEEE Transactions on Evolutionary Computation, 15*, 4–31.

Datta, R., & Deb, K. (2015). *Evolutionary Constrained Optimization*. Berlin: Springer.

Dawkins, R. (1976). *The Selfish Gene*. London: Oxford University Press.

De Jong, K. (2006). *Evolutionary Computation: A Unified Approach*. Cambridge, MA: MIT Press.

Deb, K. (2001). *Multi-Objective Optimization Using Evolutionary Algorithms*. Hoboken: Wiley.

Desell, T., Clachar, S., Higgins, J., & Wild, B. (2014). Evolving Neural Network Weights for Time-Series Prediction of General Aviation Flight Data. In *Proceedings of PPSN XIII* (pp. 771–781). Berlin: Springer.

Eiben, A. E., & Smith, J. E. (2015). *Introduction to Evolutionary Computing* (2nd ed.). Berlin: Springer.

Fink, A., & Voß, S. (2003). Anwendung von Metaheuristiken zur Lösung betrieblicher Planungsprobleme. Potenziale und Grenzen einer softwaretechnischen Unterstützung. *Wirtschaftsinformatik, 45*(4), 395–407.

Fister, I., Jr., Yang, X. S., Fister, I., Brest, J., & Fister, D. (2013). A Brief Review of Nature-Inspired Algorithms for Optimization. *Elektrotehniˇski Vestnik* (English Edition), *80*(3), 1–7.

Floreano, D., Dürr, P., & Mattiussi, C. (2008). Neuroevolution: From Architectures to Learning. *Evolutionary Intelligence, 1*, 47–62.

Fogel, D. B. (2006). *Evolutionary Computation. Toward a New Philosophy of Machine Intelligence* (3rd ed.). Piscataway: IEEE Press.

Fogel, L., Owens, A. J., & Walsh, M. J. (1966). *Artificial Intelligence Through Simulated Evolution*. New York: John Wiley & Sons.

Fortin, F. A., De Rainville, F. M., Gardner, M. A., Parizeau, M., & Gagné, C. (2012). DEAP: Evolutionary Algorithms Made Easy. *Journal of Machine Learning Research, 13*, 2171–2175.

Foulds, L. R. (1983). The Heuristic Problem-Solving Approach. *Journal of the Operational Research Society, 34*(10), 927–934.

Gaspar-Cunha, A., Henggeler, A., & Coello Coello, C. (Eds.). (2015). *Evolutionary Multi-Criterion Optimization. LNCS 9019*. Berlin: Springer.

Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*. Reading: Addison-Wesley.

Grosan, C., & Abraham, A. (2007). Hybrid Evolutionary Algorithms: Methodologies, Architectures, and Reviews. In C. Grosan, A. Abraham, & H. Ishibuchi (Eds.), *Hybrid Evolutionary Algorithms* (pp. 1–17). Berlin: Springer.

Grosan, C., Abraham, A., & Ishibuchi, H. (Eds.). (2007). *Hybrid Evolutionary Algorithms*. Berlin: Springer.

Günther, M., & Nissen, V. (2009). A Comparison of Neighbourhood Topologies for Staff Scheduling with Particle Swarm Optimisation. In B. Mertsching, M. Hund, & A. Zaheer (Hrsg.), *KI 2009: Advances in Artificial Intelligence, LNCS 5803* (pp. 185–192). Berlin: Springer.

Hansen, N. (2006). The CMA Evolution Strategy: A Comparing Review. In J. A. Lozano, P. Larranaga, I. Inza, & E. Bengoetxea (Eds.), *Towards a New Evolutionary Computation: Advances in Estimation of Distribution Algorithms* (pp. 75–102). Berlin: Springer.

Hansen, N., & Ostermeier, A. (2001). Completely Derandomized Self-Adaptation in Evolution Strategies. *Evolutionary Computation, 9*(2), 159–195.

Harrald, P. G., & Kamstra, M. (1997). Evolving Artificial Neural Networks to Combine Financial Forecasts. *IEEE Transactions on Evolutionary Computation, 1*(1), 40–52.

Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems*. Ann Arbor: University of Michigan Press.

Ishibuchi, H., Tsukamoto, N., & Nojima, Y. (2008). Evolutionary Many-Objective Optimization. A Short Review. In *Proceedings of 2008 IEEE Congress on Evolutionary Computation* (pp. 2424–2431). Piscataway: IEEE.

Juan, A. A., Faulin, J., Grasman, S. E., Rabe, M., & Figueira, G. (2015). A Review of Simheuristics: Extending Metaheuristics to Deal with Stochastic Combinatorial Optimization Problems. *Operations Research Perspectives, 2*, 62–72.

Keijzer, M.; Merelo, J.J.; Romero, G.; Schoenauer, M., Evolving Objects: A General Purpose Evolutionary Computation Library. In: Collet, P.; Fonlupt, C.; Hao, J.-K.; Lutton, E.; Schoenauer, M.: Artificial Evolution. LNCS 2310. Springer, Berlin, 2002, 231 – 242.

Koza, J. R. (1992). *Genetic Programming*. Cambridge, MA: MIT Press.

Koza, J. R. (1994). *Genetic Programming II*. Cambridge, MA: MIT Press.

Kubalik, J. (2012). Hyper-Heuristic Based on Iterated Local Search Driven by Evolutionary Algorithm. In J. K. Hao & M. Middendorf (Eds.), *Proceedings EvoCOP 2012* (pp. 148–159). Berlin: Springer.

Kühn, M., Severin, T., & Salzwedel, H. (2013). Variable Mutation Rate at Genetic Algorithms: Introduction of Chromosome Fitness in Connection with Multi-Chromosome Representation. *International Journal of Computer Applications, 72*(17), 31–38.

Kursawe, F. (1996, March). Unpublished Presentation. Dagstuhl Seminar on EA and Their Applications.

Lobo, F. G., Lima, C. F., & Michalewicz, Z. (Eds.). (2007). *Parameter Setting in Evolutionary Algorithms*. Berlin: Springer.

Lones, M. A. (2014). Metaheuristics in Nature-Inspired Algorithms. In *Proceedings of GECCO 2014* (pp. 1419–1422). New York: ACM.

Lozano, M., & Garcia-Martinez, C. (2010). Hybrid Metaheuristics with Evolutionary Algorithms Specializing in Intensification and Diversification: Overview and Progress Report. *Computers & Operations Research, 37*, 481–497.

Lozano, J. A., Larranaga, P., Inza, I., & Bengoetxea, E. (Eds.). (2006). *Towards a New Evolutionary Computation: Advances in Estimation of Distribution Algorithms*. Berlin: Springer.

Lukasiewycz, M., Glaß, M., Reimann, F., & Teich, J. (2011, July 12–16). Opt4J—A Modular Framework for Meta-Heuristic Optimization. In *Proceedings of GECCO'11* (pp. 1723–1730). Dublin, Ireland.

MathWorks. (2016). Product Information for Global Optimization Toolbox on MathWorks Website. Accessed 27 Sept 2016.

Matsui, S., Watanabe, I., & Tokoro, K. I. (2002). Real-Coded Parameter-Free Genetic Algorithm for Job-Shop Scheduling Problems. In *Proceedings of PPSN VII, LNCS 2439* (pp. 800–810). Berlin: Springer.

Mendes, R. R. F., Voznika, F., Freitas, A. A., & Nievola, J. C. (2001). Discovering Fuzzy Classification Rules with Genetic Programming and Co-evolution. In L. De Raedt & A. Siebes (Eds.), *Proceedings of 5th European Conference on Principles and Practice of Knowledge Discovery in Databases* (pp. 314–325). Berlin: Springer.

Meyer-Nieberg, S., & Beyer, H. G. (2007). Self-Adaptation in Evolutionary Algorithms. In F. G. Lobo, C. F. Lima, & Z. Michalewicz (Eds.), *Parameter Setting in Evolutionary Algorithms* (pp. 47–75). Berlin: Springer.

Mezura-Montes, E., & Coello Coello, C. A. (2011). Constraint-Handling in Nature-Inspired Numerical Optimization: Past, Present, and Future. *Swarm and Evolutionary Computation, 1*(4), 173–194.

Michalewicz, Z. (1996). *Genetic Algorithms + Data Structures = Evolution Programs* (3rd ed.). Berlin: Springer.

Michalewicz, Z., & Fogel, D. B. (2004). *How to Solve It: Modern Heuristics* (2nd ed.). Berlin: Springer.

Moscato, P. (1989). On Evolution, Search, Optimization, Genetic Algorithms and Martial Arts: Towards Memetic Algorithms. TR 826, California Institute of Technology, USA.

Mühlenbein, H., & Paaß, G. (1996). From Recombination of Genes to the Estimation of Distributions I. Binary Parameters. In *Proceedings of PPSN IV. LNCS 1141* (pp. 178–187). Berlin: Springer.

Neri, F., Cotta, C., & Moscato, P. (2012). *Handbook of Memetic Algorithms*. Berlin: Springer.

Newell, A. (1969). Heuristic Programming: Ill Structured Problems. In J. Aronofsky (Ed.), *Progress in Operations Research* (Vol. 3, pp. 361–414). New York: Wiley.

Nissen, V. (1995). An Overview of Evolutionary Algorithms in Management Applications. In J. Biethahn & V. Nissen (Eds.), *Evolutionary Algorithms in Management Applications* (pp. 44–97). Berlin: Springer.

Nissen, V. (1997). *Einführung in Evolutionäre Algorithmen*. Wiesbaden: Vieweg.

Nissen, V., & Gold, S. (2008). Survivable Network Design with an Evolution Strategy. In A. Yang, Y. Shan, & L. T. Bui (Eds.), *Success in Evolutionary Computation* (pp. 263–283). Berlin: Springer.

Nissen, V., & Propach, J. (1998). On the Robustness of Population-Based Versus Point-Based Optimization in the Presence of Noise. *IEEE Transactions on Evolutionary Computation, 2*(3), 107–119.

Parejo, J. A., Ruiz-Cortés, A., Lozano, S., & Fernandez, P. (2012). Metaheuristic Optimization Frameworks: A Survey and Benchmarking. *Soft Computing, 16*(3), 527–561.

Poli, R., Langdon, W. B., McPhee, N. F., & Koza, J. R. (2008). A Field Guide to Genetic Programming. Freely available at http://www.gp-field-guide.org.uk.

Rechenberg, I. (1973). *Evolutionsstrategie. Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*. Stuttgart: Frommann-Holzboog.

Rothlauf, F. (2006). *Representations for Genetic and Evolutionary Algorithms* (2nd ed.). Heidelberg: Springer.

Rothlauf, F. (2011). *Design of Modern Heuristics. Principles and Application*. Berlin: Springer.

SAP SE. (2016). SAP Help SCM 7.0 Optimization. Accessed 27 Sept 2016.

Sarker, R. A., Elsayed, S. M., & Ray, T. (2014). Differential Evolution with Dynamic Parameters Selection for Optimization Problems. *IEEE Transactions on Evolutionary Computation, 18*(5), 689–707.

Sawai, H., & Kizu, S. (1998). Parameter-Free Genetic Algorithm Inspired by "Disparity Theory of Evolution". In *Proceedings of PPSN V* (pp. 702–711). Berlin/Heidelberg: Springer.

Schwefel, H.-P. (1975). Evolutionsstrategie und numerische Optimierung, Dissertation, TU Berlin.

Schwefel, H.-P. (1981). *Numerical Optimization of Computer Models*. Chichester: Wiley.

Schwefel, H.-P. (1995). *Evolution and Optimum Seeking*. New York: Wiley.

Smith, J. S. (2008). Self-Adaptation in Evolutionary Algorithms for Combinatorial Optimisation. In C. Cotta, M. Sevaux, & K. Sörensen (Eds.), *Adaptive and Multilevel Metaheuristics* (pp. 31–57). Berlin: Springer.

Sörensen, K. (2015). Metaheuristics—The Metaphor Exposed. *International Transaction in Operational Research, 22*, 3–18.

Sörensen, K., & Glover, F. (2013). Metaheuristics. In S. Gass & M. Fu (Eds.), *Encyclopedia of OR/MS* (3rd ed., pp. 960–970). Hoboken: Wiley.

Sörensen, K., Sevaux, M., & Glover, F. (2016). A History of Metaheuristics. In R. Marti, P. Pardalos, & M. Resende (Eds.), *Handbook of Heuristics*. Berlin: Springer. (to appear).

Storn, R., & Price, K. (1995). Differential Evolution: A Simple and Efficient Adaptive Scheme for Global Optimization over Continuous Spaces. Int. Comp. Sci. Inst., Berkeley, CA, Tech. Rep. TR-95-012.

Weyland, D. (2015). A Critical Analysis of the Harmony Search Algorithm—How Not to Solve Sudoku. *Operations Research Perspectives, 2*, 97–105.

Wolpert, D. H., & Macready, W. G. (1997). No Free Lunch Theorems for Optimisation. *IEEE Transactions on Evolutionary Computation, 1*, 67–82.

Xhafa, F. (2007). A Hybrid Evolutionary Heuristic for Job Scheduling in Computational Grids. In J. Kacprzyk (Ed.), *Studies in Computational Intelligence* (Vol. 75). Berlin: Springer, (Chapter 10).

Yang, X. S. (2014). *Nature-Inspired Optimization Algorithms*. Amsterdam: Elsevier.

Yao, X. (1999). Evolving Artificial Neural Networks. *Proceedings of the IEEE, 87*(9), 1423–1447.

9

Applications of Evolutionary Algorithms to Management Problems

Volker Nissen

Introduction

Evolutionary Algorithms (or Evolutionary Computation) are nature-inspired metaheuristics. They represent a well-established field of research at the intersection of operations research and computational intelligence. Their working is based on a rough abstraction of the mechanisms of natural evolution. Evolutionary Algorithms (EA) imitate biological principles, such as a population-based approach, the inheritance of information, the variation of solutions via crossover and mutation, and the selection of individual solutions for reproduction based on fitness (quality). Different variants of EA exist, such as Genetic Algorithms (GA), Evolution Strategies (ES), Evolutionary Programming, and Genetic Programming (De Jong 2006; Fogel 2006; Poli et al. 2008; Eiben and Smith 2015).

This contribution reviews EA from the perspective of management applications where "management" indicates that predominantly economic targets are pursued. The general objective of the paper is to give an

V. Nissen (✉)
Chair of Information Systems Engineering in Services,
University of Technology Ilmenau, Ilmenau, Germany

L. Moutinho, M. Sokele (eds.), *Innovative Research Methodologies in Management*,
https://doi.org/10.1007/978-3-319-64394-6_9

introductory overview of applications of EA to management problems for researchers and practitioners alike who are not familiar to this class of metaheuristics. An introduction to the algorithmic working of EA, as well as their advantages and disadvantages, is given in a different chapter of this volume (Nissen 2017), where also some computational frameworks that support the EA implementation are mentioned as well as examples of the integration of EA in business software (e.g. SAP APO™).

In general terms, the preferred area of application for EA and other metaheuristics as well are optimization problems that cannot be solved analytically or with efficient algorithms, such as the simplex method for linear programming, in reasonable time or without making strong simplifying assumptions on the problem. Many of these applications are of a combinatorial nature, such as job shop scheduling, timetabling, nurse rostering, and vehicle routing, to name just a few. In practical settings, often the issue of "robustness" of a solution is equally important as "optimality", because the optimization context is characterized by uncertainty and changing conditions.

In the following section the power and versatility of EA in management applications is demonstrated through brief practical examples. The first subsection gives an overview of sample applications differentiated by the abstract problem class. The second subsection highlights EA applications differentiated by the branch of industry. Thereafter, a full example from staff scheduling is presented, using an Evolution Strategy with an integrated repair heuristic that vastly outperforms established commercial software on this type of application. Finally, an assessment of the current state of EA applications to management problems is attempted. We also compare the result with our own assessment of the field in the year 1995 (Nissen 1995) and point to some future developments.

Examples

Management Applications Ordered by Problem Class

Forecasting

The accurate ex ante estimation of software cost is an important task in software production. Gharehchopogh et al. (2015) combine Tabu Search and a Genetic Algorithm to form a hybrid heuristic for this purpose that outperforms a classical COCOMO model. Petrlik et al. (2014) employ a multi-objective GA in order to select a subset of input sensors for a support vector regression model which is used for traffic prediction. The multi-objective approach allows finding a good trade-off between the predictive error and the number of sensors in real-world situations when many traffic data measurements are unavailable.

Classification

Chen and Huang (2003) address the task of credit admission evaluation, which significantly affects risk and profitability of banks. This application uses an artificial neural net-based credit scoring model, which classifies applicants as accepted (good) or rejected (bad) credits. In an attempt to better understand the rejected credits and trying to reassign them to the preferable accepted class, a GA-based inverse classification technique is devised. The GA-based reassignment technique balances between adjustment cost and customer preference. On this basis, creditors can suggest the conditional acceptance and further explain the conditions to rejected applicants.

In Cao et al. (2016) a typical problem concerning the issue of classifying credit card datasets is addressed. In particular, there is often an imbalance between the problem (fraud or non-payment) class (minority class) and the non-problem (non-fraud or full payment) class (majority class). Then, classifiers tend to yield poor accuracy on the minority class despite of high overall accuracy. The authors propose two fitness functions to evolve Genetic Programming classifiers with significantly increased classification accuracy on the minority class while retaining high overall classification accuracy.

Modeling

Ognjanovic et al. (2012) address a configuration problem in the domain of business process families. Such process families provide an overarching representation of the possible business processes of a target domain by capturing the similarities and differences among the possible processes.

Through bounding the variability points of the process family, a precise business process is realized. The authors employ a GA to optimally configure a business process family. The GA optimizes the degree of satisfaction of nonfunctional stakeholder requirements and preserves the behavioral correctness of each business process which can be derived from the business process model family. The proposed approach not only achieves the selection of process activities performed, but also the selection of services with the most suitable quality attributes for each activity.

Alves et al. (2006) suggest a genetic planner for the automatic generation of process models when a large number of activities must be considered as well as interdependencies between activities. An additional simulation module is responsible for identifying the most suitable model among those generated before. This analysis takes into account the resources capable of performing each activity defined in the respective process models and applies scheduling techniques based on a GA.

Optimization

Kellner et al. (2012) successfully solve the logistics problem to decide on vertical and horizontal parking positions of trains in a rail–rail transshipment yard, so that the workload of loaded container moves is evenly shared among gantry cranes. Among other solution approaches, they employ a GA with a population size of 200 that runs for 700 generations until the best solution found during the solution process is returned. Based on simulations of real-world yard operations they conclude that, depending on the specific yard setting and workload, between 44% and 65% of processing time per bundle of trains can be saved whenever parking positions are determined by the GA, resulting in a tremendous economic potential for this application.

Nissen and Gold (2008) develop a combinatorial (μ, λ)-ES within the framework of survivable telecommunication network design when considering both cost efficiency and reliability. The approach keeps the overall working of a standard-ES, while simultaneously adapting mutation as the main search operator to the combinatorial problem domain. A repair heuristic is applied within the ES generation cycle when the reliability constraint is violated by a solution. The proposed ES clearly outperforms the best-so-far results by GA and finds new best solutions for every problem instance. Moreover, the ES scales better for large problems instances than the benchmark GA, underlining the application potential of ES to combinatorial problems.

Scheduling

Scheduling (including Timetabling and Vehicle Routing) is clearly a primary application area of EA in management. A list of related publications was compiled by Alander (2014). Conferences like PATAT (International Conference on the Practice and Theory of Automated Timetabling) exclusively deal with this type of application, where many contributions apply metaheuristics like EA. As an example, Derigs and Jenal (2005) propose a GA-based decision support system for professional course scheduling. It allows the planner to generate, evaluate, and compare different schedules obtained by different runs based on variations of the objective function and different strategies of blocking, that is, preassigning certain subsets of courses. The approach was developed for the specific planning situation at Ford Service Organization in Germany and has shown to significantly improve the planning process with respect to quality of schedules, time-to-plan, and flexibility.

Urquhart (2015) optimizes the delivery of dairy products to households in three urban areas. Using data from an existing customer base, he utilizes an EA for ordering deliveries and a multi-agent approach for reassigning deliveries between rounds. The case study suggests that distance traveled and the deviation between round lengths may be considerably reduced, with only 10% of customers being moved between rounds.

Machine Learning

In Ghandar et al. (2012) the authors apply a multi-objective evolutionary fuzzy system to produce and evaluate a Pareto set of rule bases that balance the conflicting criteria of complexity and utility in an IT system for algorithmic financial trading. They conclude there is evidence that stochastic systems based on natural computing techniques can deliver results that outperform the market. Moreover, computational intelligence can lead to novel ways of controlling performance in algorithmic trading. For instance, as solution complexity is found to be a strong driver of risk and return, performance can be reliably shaped through identifying the locus where additional return starts to generate higher risk.

As Lipinski (2015) points out, recent trends in computational intelligence often try to minimize the expert knowledge put in the learning system and replace it with artificial knowledge discovered, using very large models from very large training datasets. Lipinski proposes an improvement of the training process of rule-based intelligent systems. He starts from the observation that the number of decision rules included in the automatically acquired knowledge from very large training datasets is usually large. Thus, the search space of integration parameters has a very high dimension. This in turn leads to long chromosomes and consequently large populations as well as many iterations until convergence. As a remedy for this problem, the author enhances a Differential Evolution Algorithm by embedding dimensionality reduction based on Principal Component Analysis. He applies this procedure to combine 5000 decision rules in a financial decision support system. The results confirm that the method may significantly improve the search process.

Management Applications in Different Branches of Industry

Primary Production Sector

Compared to the other branches of industry, the number of publications applying EA in primary production is relatively low. This may change in the future. For instance, with low prices and intensive competition optimization issues are increasingly prevalent in agriculture. An example is given in the dissertation of Matthews (2001), who starts from the observation that rural land managers are faced with an increasingly complex decision-making situation where a varied mix of goals have to be achieved.

Thus, the author develops single- and multi-objective GA to assist land managers in providing plans which define the mix and spatial pattern of land uses to meet one or more objectives. He tests his heuristic approach on problems typical in scale of real-world applications. The solutions achieved by the multi-objective GA were considered useful for land-use planning by the practitioners. Matthews (2001, p. 128) concludes that his GA are sufficiently robust and efficient that they may be employed for practical planning tasks. Moreover, they are flexible enough to support an iterative investigation of a land-use planning problem by the land manager rather than operating as a black box just providing a single answer.

In a different application, Falcao and Borges (2001) propose a GA-approach to solve very large integer forest management scheduling models. A chromosome in the GA consists of a sequence of genes each representing a management unit. The size of a chromosome is thus defined by the number of forest stands. The value of each gene identifies the management alternative assigned to the respective management unit. This heuristic was applied to a forest management scheduling problem in Portugal. It encompassed an area of about 8700 ha classified into 696 stands, a planning horizon extending over 70 one-year periods and target annual volume flows for two products. An average of 175 management alternatives was available for each stand. Model building thus involved the definition of over 121,000 binary variables and about 1000 constraints. Tests confirmed that the proposed heuristic is both efficient and robust.

Industrial Sector

Industrial applications of EA were among the first suggested in the early EA literature and still play an important role—resource management and scheduling of production being good examples. Part 2 of the book by Biethahn and Nissen (1995) presents some early examples. A recent overview of EA in manufacturing applications can be found in Alander (2015), a similar overview of EA in Economics is Alander (2009).

Schneider et al. (2013) address the problem to match uncertain demand with production capacities. Operations managers can use mix-flexible resources to shift excess demands to unused capacities. To find the optimal configuration of a mix-flexible production network, a flexibility design problem (FDP) is solved, using a tailored GA. FDP is a difficult stochastic optimization problem, for which traditional exact approaches are not able to solve but the smallest instances in reasonable time. The proposed GA outperforms Simulated Annealing as well as a commercial solver on this problem with respect to both computing time and solution quality.

There is a substantial literature on the management of electric power generation, using metaheuristics like EA. As an example, in Soares et al. (2013) the authors use a GA to optimize the scheduling of domestic electric loads, according to technical and user-defined constraints and input signals. The objective is to minimize the end-user's electricity bill according to his/her preferences. The load scheduling is done for the next 36 hours assuming that a dynamic pricing structure is known in advance. In a case study with real data, the suggested GA resulted in a noticeable bill reduction for the customers as compared to a traditional approach without automated scheduling.

Service Sector

If one examines the literature on management applications of EA, a growing number of publications describe applications in different branches of the service sector. This corresponds to the ever-growing relevance of this sector in terms of gross national product share in developed countries. Early examples (with a focus on financial services and traffic management) can be found in Parts 4–6 of the book by Biethahn and Nissen (1995).

In a more recent contribution, Duran et al. (2012) apply three different multi-objective EA to the problem of selecting investment components according to two partially opposed measures: the portfolio performance and its risk. They consider the case of mutual funds market in Europe until July 2010, with data comprising the stock value of the funds, sampled on a monthly basis for five years. The authors conclude that the problem of portfolio optimization is a natural scenario for the use of multi-objective EA, in which their power and flexibility can be readily exploited.

Quite a few publications address the issue of admission scheduling in hospitals. This is due to the fact that admission planning involves a large number of decisions that have complex and uncertain consequences for hospital resource utilization and patient flow. Thus, it has significant impact on the hospital's profitability, access, and quality of care (Helm et al. 2010). Using real-world datasets and a simulation model of the hospital, Kühn et al. (2012) employ a GA for admission scheduling where each individual codes patient IDs and associated appointment times. The authors find evidence for a particular treatment center that waiting times for patients could be reduced up to 40% by using GA-based admission planning. A similar work dealing with the optimization of the admission scheduling strategy in an ophthalmology department is presented in Chen et al. (2010).

Lesel et al. (2016) in an approach to design energy optimal railway timetables, implement heuristics based on particle swarm optimization and a GA to define timetables that will maximize braking energy recovery. The investigation is based on an electrical modeling of a metro line and a Newton-Raphson method to calculate power flows inside the network so as to assess the amount of dissipated braking energy. Both metaheuristics lead to substantial energy gains during a normal operational day.

Trade

Early application examples of EA in trade are described in Part 3 of the book by Biethahn and Nissen (1995) with a certain focus on inventory management issues. Modern contributions often deal with aspects of financial markets. For instance, a multi-objective evolutionary fuzzy system for algorithmic trading is proposed in Ghandar et al. (2012).

The paper by Gypteau et al. (2015) points in a similar direction. A Genetic Programming approach is proposed that aims to find an optimal trading strategy to forecast the future price moves of a financial market. Traditionally, financial forecasting tools are based on physical time to forecast the future price movement and probe the market activity at a fixed interval of time. Gypteau et al. do not assume such discrete time steps but use the notion of directional changes in the context of the so-called intrinsic time. A directional change (DC) event is characterized by a fixed threshold of different sizes and time in price time series, eliminating any irrelevant details of price evolution. In this way, the price fluctuations are described by the frequency of directional change events over a sampling period, which provides an alternative measure of the risk. Gypteau et al. use Genetic Programming to automatically generate trading strategies that make use of DC thresholds. Strategies are created by combining the output of multiple thresholds in a single expression. The authors demonstrate that the new paradigm of DC leads to effective and profitable trading strategies, which can potentially outperform the traditional physical-time strategies that are currently widely used.

Alves et al. (2016) propose a bi-level programming model for modeling the interaction between electricity retailers and consumers endowed with energy management systems, capable of providing demand response to variable prices. A bi-level problem is a programming problem where an (lower level) optimization problem is embedded as a constraint in another (upper level) optimization problem. In the present model, the electricity retailer is the upper-level decision-maker, which buys energy in the spot market and sells it to consumers. The retailer determines prices to be charged to the consumer, subject to the regulation framework, with the aim of maximizing its profit. The consumer (lower-level decision-maker) reacts to the prices by scheduling loads to minimize his electricity bill. The model intends to determine the optimal pricing scheme to be established by the retailer and the optimal load schedule adopted by the consumer under this price setting. A hybrid approach is suggested that uses a GA where individuals of the population represent the retailer's choices (electricity prices). For each price setting, the exact optimal solution to the consumer's problem is obtained in a very efficient way using a mixed-integer linear programming solver. An illustrative case is studied considering real data for the loads that demonstrates the potential of this hybrid approach.

Public Administration

Flores-Revuelta et al. (2008) develop a Memetic Algorithm for the delineation of local labor markets (LLMA). To test their proposal, the authors use a case study: the delineation of a set of LLMA in the Region of Valencia, Spain. Given the complexity of typical real-world situations, the authors ascertain that conventional evolutionary operators hardly ever lead to valid solutions. To avoid this problem specialized operators are designed. These allow obtaining good final delineations. However, the percentage of generated invalid individuals continues to be very high. A memetic extension of the EA then includes a repair procedure that turns these solutions into valid ones that can contribute to the evolutionary process. Moreover, a local search procedure is added. The evolutionary heuristic substantially improves the results obtained by traditional methods. Although some improvement in solution quality is achieved by the memetic extension as compared to the pure tailored EA approach, the time consumed by the memetic approach is comparatively high.

Chang and Chang (2009) apply a popular multi-objective evolutionary algorithm, the non-dominated sorting genetic algorithm (NSGA-II), to examine the operations of a multi-reservoir water system in Taiwan. The main goal is to increase the efficiency of use of water resources from two reservoirs. To this end, the authors develop an integrated simulation model of the two parallel reservoir systems, coupled with the NSGA-II to assess optimal joint operating strategies for these reservoirs. A daily operational simulation was applied and shortage indices calculated for a 49-year period. Chang and Chang conclude that there is a great potential to alleviate the water shortage problem of this parallel reservoir system through optimizing their operation strategies using this approach. Ultimately, the decision-maker is provided with a wide range of practical alternatives for system management.

Full Example (Workforce Management Problem)

Introduction

The following complete example, first published in Nissen and Günther (2010), deals with the problem of an automatic generation of optimized working time models in personnel planning. The ability to adapt personnel assignment to changing requirements is of critical importance in workforce management (WFM). An interesting option is automatically generated flexible working time models that are based on demand while respecting certain constraints. The targeted benefits are cost reduction through improved utilization of employee time, reduction of overtime and idle time, a rise in employee motivation and, thus, an increase of turnover through a higher level of service.

The traditional approach to WFM in separate planning steps can be very inefficient. Therefore, an integrated design of working time models and staff schedules is suggested in our solution approach. More specifically, we adapt an Evolution Strategy and compare the results to a commercial constructive approach for this integrated planning task. The research goals we pursue are twofold. First, we aim for good solutions to a meaningful practical application that is relevant in industries such as logistics, retail, and call centers. Second, we want to contribute to the comparison of modern metaheuristics on problems of realistic complexity.

Description of the Real-World Problem from a Retailer

This practical case concerns personnel planning in the department for ladies' wear at a department store. We assume a set of employees $E = \{1, \ldots, E\}$, a set of workstations $W = \{1, \ldots, W\}$ and a discrete timeframe T with index $t = 0, \ldots, T-1$. Each period t of the range has a length l_t greater than zero.

$$l_t > 0 \qquad \forall t \in T$$

The assignment of an employee to a workstation is controlled using the binary variable x_{ewt}.

$$x_{ewt} = \begin{cases} 1 & \textit{if employee } e \textit{ is assigned to workstation } w \textit{ at period } t \\ 0 & \textit{otherwise} \end{cases}$$

The store is open Monday to Saturday from 10:00 to 20:00 and closed on Sunday and holidays. The availability of the employees is determined using the binary variable a_{et}.

$$a_{et} = \begin{cases} 1 & \textit{if employee e is available at period t} \\ 0 & \textit{otherwise} \end{cases}$$

There are 15 employees, assigned to two workstations (till and sales), with all employees trained for both stations. The personnel demand d_{wt} is given in one-hour intervals and is determined based on past data. A minimal and maximal number of employees per workstation and period are set. The demand d_{wt} of employees per workstation and period cannot be negative.

$$d_{wt} \geq 0 \qquad \forall w \in W \; and \forall t \in T$$

There are many factors influencing planning, such as regulations, employee availability, and time sheets. Because of fluctuations in demand, sub-daily workstation changes are allowed. As a hard constraint, working time models must begin and end on the hour. Also, sub-daily workstation changes are only possible on the hour. Moreover, an employee *e* can only be associated with a workstation *w* in the period *t* if he or she is actually present. Additionally, an employee can only be designated to one workstation at a time.

$$\sum_{w=1}^{W} x_{ewt} = a_{et} \qquad \forall e \in E \text{ and} \forall t \in T$$

There are also soft constraints. Their violation is penalized with error points that reflect the company's requirements as inquired through interviews. If a discrepancy arises from the workstation staffing target d_{wt}, error points P_d are generated for the duration and size of the erroneous assignment.

$$P_d = \sum_{t=0}^{T-1} \sum_{w=1}^{W} \left(c_{dn} + c_{do} + c_{du}\right) l_t \left| \left(\sum_{e=1}^{\mathrm{E}} x_{ewt} \right) - d_{wt} \right|$$

with:

$$c_{do} > 0 \, if \; w \, is \; overstaffed \; at \; t \; and \; d_{wt} > 0, \; else \; c_{do} = 0$$
$$c_{dn} > 0 \, if \; w \, is \; overstaffed \; at \; t \; and \; d_{wt} = 0, \; else \; c_{dn} = 0$$
$$c_{du} > 0 \, if \; w \, is \; understaffed \; at \; t \; and \; d_{wt} > 0, \; else \; c_{du} = 0$$

Six employment contracts exist with a planned weekly working time between 10 and 40 hours. During weeks with bank holidays the planned working time s_e is reduced by a proportional factor h. The effective weekly working time i_e for an employee should not exceed the contractually agreed number of hours. Each minute in excess is punished with error points.

$$P_w = c_w \sum_{week=1}^{52} \sum_{e=1}^{E} \left(i_e - s_e^{\,*} h \right)$$

with $c_w = 0 \; if \, s_e^* \, h - i_e \geq 0$, $c_w = 1 \; else$.

The automatically generated working time models should not be shorter than 3 hours or longer than 9 hours. Any violation leads to error points C_t per employee and day. The sum of these error points for the planning horizon is P_t. Working time models must not be split up during a working day, with violations leading to error points C_c per employee and day. The sum of these error points for the planning horizon is P_c. For an optimal coverage of personnel demand sub-daily workstation changes are required. However, to avoid an excessive number r_e of rotations for any employee C_r error points arise for such workstation changes.

$$P_r = c_r \sum_{e=1}^{E} r_e$$

Therefore, the objective function to be minimized becomes:

$$minP = P_d + P_w + P_t + P_c + P_r$$

Historical data is available for a complete calendar year, so that an entire year (8.760 one-hour timeslots) can be planned ahead, resulting in a very complex search space of 131,400 decision variables. In practice, also shorter planning horizons (month) are employed. However, the full year plan helps the company to better understand on a more strategic level how well it can cope with demand using current staff.

To apply the ES, the problem needs to be conveniently represented. A two-dimensional matrix is applied (see Fig. 9.1). The rows represent the employees and the columns the timeslots. The meaning of the matrix elements is as follows:

- 0: Store is closed or employee is absent (holiday, training, illness).
- 1: Employee is assigned to workstation 1.
- 2: Employee is assigned to workstation 2.
- 3: Employee is generally available but not dispatched in staffing.

Benchmark: Commercial Constructive Method

As a benchmark, we use commercial workforce management software. (We cannot mention the name for legal reasons.) This tool delivers an adequate constructive method, capable of solving the problem. The software is in regular use at around 300 companies. All restrictions are supported and the error points associated with soft constraints can be entered into the software. Unfortunately, no code was available and no information was provided as to how the working time models are generated.

employee	**period**										
	1	2	3	4	5	6	7	8	9	10	...
1	0	3	3	1	1	1	2	2	2	0	
2	0	1	2	2	2	2	3	3	3	0	
3	0	3	3	3	3	3	3	3	3	0	
4	0	2	2	2	3	3	3	3	3	0	
5	0	3	3	3	3	1	1	1	3	0	
6	0	3	1	1	1	1	1	3	3	0	
...											

Fig. 9.1 Assignment of workstations in a two-dimensional matrix

Thus, it must be considered a black box. However, in applying the software we were supported by the software manufacturer, so errors in handling can be excluded.

Evolution Strategy

The ES was traditionally applied to continuous parameter optimization problems. Our application is of a combinatorial nature, though, which requires some adaptation of the ES. The pseudocode in Fig. 9.2 presents an overview of the implemented ES.

The ES population is initialized with valid solutions with regard to the hard problem constraints. Ten alternative recombination variants were evaluated in a pretest. The best performance was achieved with a rather simple form that is based on the classical one-point crossover. The same crossover point is determined at random for all employees (row) of a solution, and the associated parts of the parents are exchanged to create an offspring (see Fig. 9.3).

An offspring is mutated by picking an employee at random and changing the workstation assignment for a time interval chosen at random. It must be ensured, though, that valid assignments are made with regard to the hard problem constraints. The number of employees selected for mutation follows a $(0;\sigma)$-normal distribution. Results are rounded and converted to positive integer numbers. The mutation stepsize sigma is controlled self-adaptively using a log-normal distribution and intermediary recombination, following the standard scheme of ES (Beyer and Schwefel 2002).

```
Initialize the Population with μ Individuals
Repair the Population
Evaluate the μ Individuals
Loop
      Recombination to generate λ Offspring
      Mutate the λ Offspring
      Repair the λ Offspring
      Evaluate all repaired Individuals
      Selection ((μ + λ) or (μ,λ))
Until Termination Criterion
```

Fig. 9.2 ES implementation in pseudocode

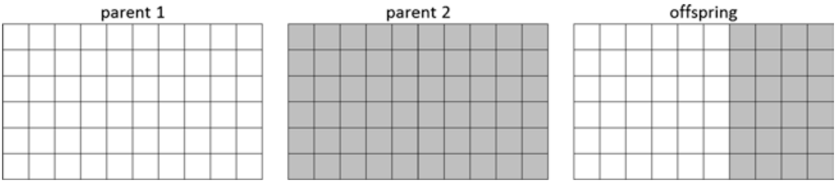

Fig. 9.3 Recombination in the proposed ES

After mutation, a repair heuristic is applied to individuals to remove constraint violations. This heuristic is outlined in the following subsection.

(μ, λ)-selection (comma selection) as well as $(\mu + \lambda)$-selection (plus selection) are used and different population sizes. The best solution found during an experimental run is always stored and updated. It represents the final solution of the run. Following suggestions in the literature (Beyer and Schwefel 2002), the ratio μ/λ is set to 1/5–1/7 during the experiments.

Repair Heuristic

Preliminary tests showed that for the highly complex retailing problem repairing the solutions created by the ES could be beneficial. Therefore, a repair heuristic was devised to reduce the total error points of a solution. This repair heuristic corrects violations of soft constraints in the following order, based on the observed frequency of error occurrences:

- *Overstaffing*: If too many employees are assigned to a workstation in a timeslot, some are reassigned to the dummy workstation. Overstaffing cannot always be completely eliminated due to legal regulations, though.
- *Understaffing*: If too few employees are assigned in a timeslot, an attempt is made to reassign employees from the dummy workstation. Understaffing cannot always be fully eliminated, because there are still phases where personnel demand exceeds personnel availability, which inevitably leads to coverage errors.
- *More than one working time model per day*: If an employee is assigned two models for one day (e.g. first start at 10:00, first end 14:00, second start 16:00, second end 20:00), an attempt is made to transfer one of

the time periods to another employee who does not yet have an assignment for those timeslots.

- *Minimum length of a working time model*: If a model falls below the minimum length, the algorithm attempts to shorten the model of another employee and add those timeslots to the short model. If that does not succeed, the algorithm attempts to add the timeslots to another employee's model.
- *Maximum length of a working time model*: If a model exceeds the allowable length, it is first split into two models of permissible length. Subsequently, an attempt is made to transfer one of the models to another employee or to add it to the model of another employee.
- *Workstation rotations*: Workstation rotations are an integral part of demand-oriented personnel assignment planning. However, unnecessary switching is obviously something to be avoided. For each employee and timeslot a check is performed of whether rotating workstations with another employee leads to a reduction of total rotation instances. Only in this case the switch is actually carried out.
- *Final correction of the working time models*: After all of the corrections have been completed, it is still possible that the minimum and maximum model lengths have been breached or gaps can still be found in the models. Adherence to the basic assignment rules, given by the planner, is of the highest priority, though. Consequently, this is assigned the highest error points for violations. In view of that, the working time models containing errors are corrected again without regard to over- and understaffing as well as weekly working hours. Models which are too short are either extended or discarded, depending on which alternative causes fewer error points. Long models are simply shortened. Gaps in the models are either filled or the ends are dropped, whatever causes fewer error points.

Results and Discussion

All heuristics were tested on the retailer problem with the objective of minimizing error points under the given constraints. The implementation was done in C# on a 2.66 GHz quad core PC with 4 GB

RAM. Table 9.1 presents the results. The runs using the ES were repeated 30 times for each parameter set. One calculation (10 minutes) is sufficient for the deterministic constructive method, because the same result would be achieved each repetition. The ES requires roughly 6 hours for a single run with 400,000 fitness calculations.

The constructive heuristic can be regarded as a benchmark, since it is actually used in 300 companies for staff planning. Even though it is capable of handling sub-daily workstation rotations (which was explicitly activated), no switching takes place. Additionally, the effective weekly working time of employees is frequently in excess of their contracts. A high effort of replanning would be required to remove these errors.

The ES performs significantly better on this minimization problem. The best mean result is achieved with (1,5)-ES. The assignment plans generated with the (1,5)-ES can hardly be improved upon, even with highly complex manual changes. For this reason, and because of the vast improvement over the commercial software, these plans can be regarded as very usable.

Generally, the comma selection performs better than plus selection (for the same μ and λ). The comma strategy "forgets" the parent values after each generation, which allows for a temporary deterioration of objective function values. This is helpful in escaping from a local

Table 9.1 Results of the heuristics with various parameter settings (mean of 30 runs for the ES)

Heuristic	Mean error	Minimal error	Std. dev.	Job changes	Understaff. in minutes	Overstaff. in minutes	Too much weekly working time in min.
Constructive method	84690.0	84690	–	0.0	1500.0	0.0	83190.0
ES(1,5)	**8267.1**	**5924**	**1265.8**	**214.3**	**834.0**	**8.0**	**7210.8**
ES(1+5)	47198.1	13720	20486.4	444.5	1870.0	28.0	44855.6
ES(10,50)	17528.5	8459	4094.2	247.3	1170.0	12.0	16099.2
ES(10+50)	49794.5	24609	26010.7	451.7	2022.0	24.0	47296.8
ES(30,200)	22222.7	13579	5780.6	283.9	1130.0	10.0	20798.8
ES(30+200)	39491.5	25706	14801.5	460.7	1720.0	12.0	37298.8

optimum. With regard to improving solutions, a tendency can be seen in the comma selection toward smaller populations. Because of the uniform termination criterion of 400,000 fitness calculations, a smaller population means more iteration cycles. Many steps are required to arrive at a good plan. Thus, it seems preferable to track changes for more iterations as compared to richer diversity (through larger populations) of the solution space in each iteration. The effect is not so clearly visible for the ES with plus selection. A plus strategy is more easily trapped in a local optimum, which is particularly pronounced when the population is small. So there are two opposite effects, resulting in an overall medium performance of the plus strategy.

The computational requirements of the ES are high. This is acceptable, though, since the task is not time-critical. Moreover, if the planning is performed for a month instead of an entire year the computational effort would be reduced accordingly.

Conclusions

Regarding the current state of practical applications of EA to management problems (Fig. 9.4), the overview given in this paper demonstrates that EA have meanwhile reached the state of maturity.[1] While industrial applications dominated in the early years of EA, nowadays we also find many interesting EA-based approaches in the service sector, nicely reflecting the ever-growing importance of services for the gross national product of developed economies. A sector that is just beginning to identify potentials for optimization in the author's view is agriculture and related primary production. As was shown in this paper, some pioneering applications of EA in agriculture and forest management already exist, but this is to be expanded in the next years. One would also wish to see EA more often applied in public administration and other nonprofit organizations.

It is interesting to compare the situation with my assessment some 20 years ago (Nissen 1995, pp. 62–63). At that time, there were only few real-world applications of EA in management, some of them collected in Biethahn and Nissen (1995). Many researchers, however, were testing the potential of their

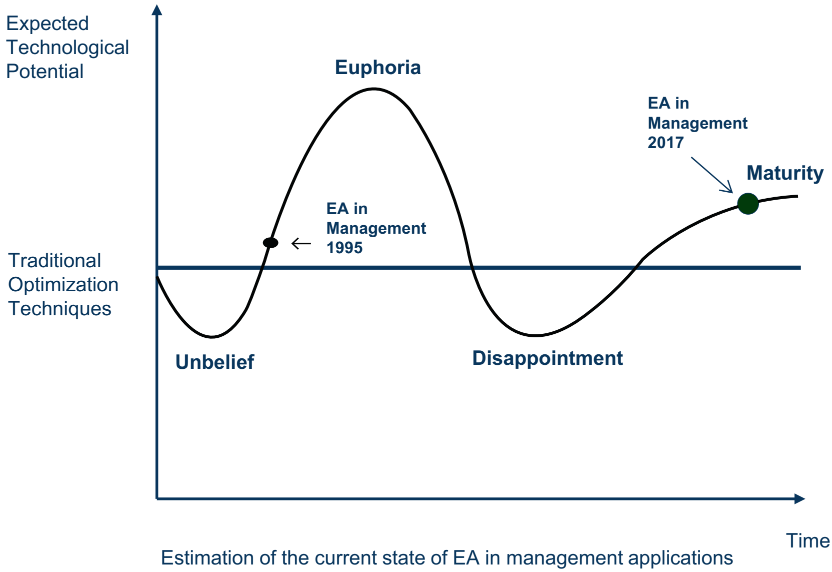

Fig. 9.4 Estimation of current state of EA in management applications in a hype cycle notation as compared to Nissen (1995)

heuristic approaches on standard OR-problems like Travelling Salesman, Bin Packing, and Quadratic Assignment that somehow relate to management applications. One could view this as "warming-up" for more practical things to come later. One other observation in 1995 was that GA were the dominant approach within EA, which still is true to some extent, but nowadays many researchers think more in terms of a common EA framework rather than focusing on the classical (canonical) EA in isolation.

Notes

1. In EC Digest V29 (2015) W.B. Langdon announced that the bibliography had reached 10,000 genetic programming entries. Thus, it can be estimated that the total number of publications with regard to the entire field of EA is well beyond 50,000 as of 2016. A substantial part of these deals with problems and applications related to Management Science.

References

Alander, J. (2009). *An Indexed Bibliography of Genetic Algorithms in Economics* (Technical Report). University of Vaasa. Available via Anonymous ftp: site https://www.researchgate.net/publication/2647009_An_Indexed_Bibliography_of_Genetic_Algorithms_in_Economics or from ResearchGate.

Alander, J. (2014). *An Indexed Bibliography of Genetic Algorithms in Scheduling* (Technical Report). University of Vaasa. Available at: http://www.uva.fi/~TAU/reports/report94-1/gaSCHEDULINGbib.pdf

Alander, J. (2015). *An Indexed Bibliography of Genetic Algorithms in Manufacturing* (Technical Report). University of Vaasa. Available at http://www.uva.fi/~TAU/reports/report94-1/gaMANUbib.pdf

Alves, F. S. R., Guimaraes, K. F., & Fernandes, M. A. (2006). Modeling Workflow Systems with Genetic Planner and Scheduler. In *Proceedings of the 18th International Conference on Tools with Artificial Intelligence* (pp. 381–388). Piscataway: IEEE.

Alves, M. J., Antunes, C. H., & Carrasqueira, P. (2016). A Hybrid Genetic Algorithm for the Interaction of Electricity Retailers with Demand Response. In *Proceedings of EvoApplications 2016* (LNCS 9597, pp. 459–474). Berlin: Springer.

Beyer, H.-G., & Schwefel, H.-P. (2002). Evolution Strategies – A Comprehensive Introduction. *Natural Computing, 1*, 3–52.

Biethahn, J., & Nissen, V. (Eds.). (1995). *Evolutionary Algorithms in Management Applications*. Berlin: Springer.

Cao, V. L., Le-Khac, N. A., O'Neill, M., Nicolau, M., & McDermott, J. (2016). Improving Fitness Functions in Genetic Programming for Classification on Unbalanced Credit Card Data. In *Proceedings of EvoApplications 2016* (LNCS 9597, pp. 35–45). Berlin: Springer.

Chang, L. C., & Chang, F. J. (2009). Multi-Objective Evolutionary Algorithm for Operating Parallel Reservoir System. *Journal of Hydrology, 377*, 12–20.

Chen, M. C., & Huang, S. H. (2003). Credit Scoring and Rejected Instances Reassigning Through Evolutionary Computation Techniques. *Expert Systems with Applications, 24*, 433–441.

Chen, N., Zhan, Z., Zhang, J., Liu, O., & Liu, H. (2010). A Genetic Algorithm for the Optimization of Admission Scheduling Strategy in Hospitals. In *Proceedings of the IEEE Congress on Evolutionary Computation 2010* (pp. 1–5). Piscataway: IEEE.

De Jong, K. (2006). *Evolutionary Computation: A Unified Approach*. Cambridge, MA: MIT Press.

Derigs, U., & Jenal, O. (2005). A GA-Based Decision Support System for Professional Course Scheduling at Ford Service Organisation. *OR Spectrum, 27*, 147–162.

Duran, F. E. C., Cotta, C., & Fernandez-Leiva, A. J. (2012). A Comparative Study of Multi-objective Evolutionary Algorithms to Optimize the Selection of Investment Portfolios with Cardinality Constraints. In *Proceedings of EvoApplications 2012* (LNCS 7248, pp. 165–173). Berlin: Springer.

Eiben, A. E., & Smith, J. E. (2015). *Introduction to Evolutionary Computing* (2nd ed.). Berlin: Springer.

Falcao, A. O., & Borges, J. G. (2001). Designing an Evolution Program for Solving Integer Forest Management Scheduling Models: An Application in Portugal. *Forest Science, 47*(2), 158–168.

Flores-Revuelta, F., Casado-Diaz, J. M., Martinez-Bernabeu, L., & Gomez-Hernandez, R. (2008). A Memetic Algorithm for the Delineation of Local Labour Markets. In *Proceedings of PPSN X* (LNCS 5199, pp. 1011–1020). Berlin: Springer.

Fogel, D. B. (2006). *Evolutionary Computation. Toward a New Philosophy of Machine Intelligence* (3rd ed.). Piscataway: IEEE Press.

Ghandar, A., Michaelwicz, Z., & Zurbruegg, R. (2012). Enhancing Profitability Through Interpretability in Algorithmic Trading with a Multiobjective Evolutionary Fuzzy System. In *Proceedings of PPSN XII* (pp. 42–51). Berlin: Springer.

Gharehchopogh, F. S., Rezaii, R., & Arasteh, B. (2015). A New Approach by Using Tabu Search and Genetic Algorithms in Software Cost Estimation. In *Proceedings of 9th International Conference on Application of Information and Communication Technologies (AICT)* (pp. 113–117). New York: ACM. Rostov on Don.

Gypteau, J., Otero, F. E. B., & Kampourides, M. (2015). Generating Directional Change Based Trading Strategies with Genetic Programming. In *Proceedings of EvoApplications 2015* (LNCS 9028, pp. 267–278). Berlin: Springer.

Helm, J. E., Lapp, M., & See, B. D. (2010). Characterizing an Effective Hospital Admissions Scheduling and Control Management System: A Genetic Algorithm Approach. In *Proceedings of the 2010 Winter Simulation Conference* (pp. 2387–2398).

Kellner, M., Boysen, N., & Fliedner, M. (2012). How to Park Freight Trains on Rail–Rail Transshipment Yards: The Train Location Problem. *OR Spectrum, 34*, 535–561.

Kühn, M., Baumann, T., & Salzwedel, H. (2012). Genetic Algorithm for Process Optimization in Hospitals. In *Proceedings of the 26th European Conference on Modelling and Simulation* (pp. 103–107).

Lesel, J., Claisse, G., Debay, P., & Robyns, B. (2016). Design of Daily Energy Optimal Timetables for Metro Lines Using Metaheuristics. In *Proceedings of the 18th Mediterranean Electrotechnical Conference*. Piscataway: IEEE. https://doi.org/10.1109/MELCON.2016. 7495456.

Lipinski, P. (2015). Training Financial Decision Support Systems with Thousands of Decision Rules Using Differential Evolution with Embedded Dimensionality Reduction. In *Proceedings of EvoApplications 2015* (LNCS 9028, pp. 289–301). Berlin: Springer.

Matthews, K. B. (2001). *Applying Genetic Algorithms to Multi-objective Land-Use Planning* (PhD Dissertation). Robert Gordon University, Aberdeen.

Nissen, V. (1995). An Overview of Evolutionary Algorithms in Management Applications. In J. Biethahn & V. Nissen (Eds.), *Evolutionary Algorithms in Management Applications* (pp. 44–97). Berlin: Springer.

Nissen, V. (2017., in this volume). Chapter 8: A Brief Introduction to Evolutionary Algorithms from the Perspective of Management Science. In L. Moutinho & M. Sokele (Eds.), *Innovative Research Methodologies in Management*. Cham: Palgrave Macmillan.

Nissen, V., & Gold, S. (2008). Survivable Network Design with an Evolution Strategy. In A. Yang, Y. Shan, & L. T. Bui (Eds.), *Success in Evolutionary Computation* (pp. 263–283). Berlin: Springer.

Nissen, V., & Günther, M. (2010). Automatic Generation of Optimised Working Time Models in Personnel Planning. In *Proceedings of ANTS 2010 – 7th Int. Conf. on Swarm Intelligence* (LNCS 6234, pp. 384–391). Berlin: Springer.

Ognjanovic, I., Mohabbati, B., Gasevic, D., Bagheri, E., & Boskovic, M. (2012). A Metaheuristic Approach for the Configuration of Business Process Families. In *Proceedings of the Ninth International Conference on Services Computing* (pp. 25–32). Piscataway: IEEE.

Petrlik, J., Fucik, O., & Sekanina, L. (2014). Multiobjective Selection of Input Sensors for SVR Applied to Road Traffic Prediction. In *Proc. of PPSN XIII* (LNCS 8672, pp. 802–811). Berlin: Springer.

Poli, R., Langdon, W. B., McPhee, N. F., & Koza, J. R. (2008). *A Field Guide to Genetic Programming*. Freely available at http://www.gp-field-guide.org.uk

Schneider, M., Grahl, J., Francas, D., & Vigo, D. (2013). A Problem-adjusted Genetic Algorithm for Flexibility Design. *International Journal of Production Economics, 141*(1), 56–65.

Soares, A., Gomes, A., Antunes, C. H., & Cardoso, H. (2013). Domestic Load Scheduling Using Genetic Algorithms. In *Proceedings of EvoApplications 2013* (LNCS 7835, pp. 142–151). Berlin: Springer.

Urquhart, N. (2015). Optimising the Scheduling and Planning of Urban Milk Deliveries. In *Proceedings of EvoApplications 2015* (LNCS 9028, pp. 604–615). Berlin: Springer.

10

An Exposition of the Role of Consideration Sets in a DS/AHP Analysis of Consumer Choice

Malcolm J. Beynon, Luiz Moutinho, and Cleopatra Veloutsou

Introduction

For an individual consumer, the ability to undertake choice is an empowerment of his or her own thought process. Any technique utilised should attempt to allow the consumer to control their contribution without being forced to undertake judgements simply to obey the respective technique's remit, an acknowledgement of the well-known bounded rationality problem (Simon 1955; Miller 1956). Both internal and external

The authors would like to thank Dr. Helena Pestana and Ms. Camila Mello for their outstanding assistance during the last stage of preparation and submission of the paper. We are grateful to both.

M.J. Beynon
Cardiff Business School, Cardiff University, Cardiff, Wales, UK

L. Moutinho (✉)
University of Suffolk, Suffolk, England, UK

The University of the South Pacific, Suva, Fiji

C. Veloutsou
Adam Smith Business School, University of Glasgow, Glasgow, Scotland, UK

L. Moutinho, M. Sokele (eds.), *Innovative Research Methodologies in Management*,
https://doi.org/10.1007/978-3-319-64394-6_10

constraints may exist which inhibit a consumer's choice process. Internal constraints relate from the consumer, including ignorance and non-specificity in their knowledge to a problem. External constraints are those placed on the consumer such as time constraints and information overload (Hogarth 1980).

Consumers use various criteria to analyse their options when they are making a purchase decision. Within multi-criteria decision-making (MCDM), a number of methods have been developed to aid a decision-maker (consumer) in their choice process (Manrai 1995), including multi-attribute utility models (Lock and Thomas 1979; Arora and Huber 2001; Analytis et al. 2014) and the Analytic Hierarchy Process—AHP (Saaty 1977; Benitez et al. 2015). In the case of Analytic Hierarchy Process (AHP), it contends to enable a consumer to deconstruct the problem in question, with the judgement making between the considered decision alternatives (DAs), made sequentially over the different criteria. In this paper, a nascent method of multi-criteria decision-making analysis (Park et al. 2015; Dede et al. 2016) is exposited, namely, Dempster-Shafer/Analytic Hierarchy Process (DS/AHP) (Beynon et al. 2000; Beynon 2002; Wang et al. 2016), with a model structure similar to AHP but analytical foundation based on the Dempster-Shafer theory of evidence—DST (Dempster 1968; Shafer 1976; Taroun and Yang 2013).

Within the extant literature, the issue of consideration sets is an active and ongoing area of consumer research (Roberts and Nedungadi 1995; Roberts and Lattin 1997; Beaman 2013; Goodman et al. 2013; Carson and Louviere 2014). Using DS/AHP, two directions of understanding consideration sets are exposited. Firstly, those which are memory-based and brought to the problem by a consumer (Desai and Hoyer 2000). Secondly, the results from a DS/AHP analysis are in the form of levels of preference on different-sized groups of decision alternatives, DAs (future consideration and choice sets). An aim of this paper is to highlight the notion of consideration sets as a fundamental aspect to the DS/AHP methodology. Further, the elucidation of the benefit of utilising DS/AHP with respect to the internal and external constraints associated with the understood choice process. With a number of consumers considered in the problem, the DS/AHP method incorporates a novel approach for the aggregation of evidence from consumers.

The chapter describes related consumer choice theory with respect to consideration sets, followed by discussion on evaluation and choice. The research focus is then briefly described. It then briefly introduces the Dempster-Shafer theory followed by an exposition of the DS/AHP method within a car choice problem. Results from the DS/AHP method are then presented, followed by interpretation of DS/AHP as an analysis tool in consumer choice MCDM. Finally conclusions are presented as well as directions for future research.

Formation of Consideration Sets

In order to evaluate alternatives, a consideration set has to be formed in the mind of each potential consumer. Many studies have tackled the subject of consideration sets (CS). They have different focuses and investigate many factors associated with the formation or the evaluation of consideration sets. Punj and Brookes (2002) proposed that the manner in which a purchase decision is initiated has an important influence on subsequent product evaluation and choice. Specifically, they proposed that the problem recognition "event" and the consequent retrieval of pre-decisional constraints from memory significantly influence the ensuing processes of external information search and consideration sets formation. Their results suggested that the type of pre-decisional constraints that are activated as a consequence of the problem recognition event significantly influences the "route" consumers follow through the remainder of the purchase process.

Costly search can result in consumers restricting their attention to a subset of products—the consideration set—before making a final purchase decision. The search process is usually not observed, which creates econometric challenges. Pires (2016) shows that inventory and the availability of different package sizes create new sources of variation to identify search costs in storable goods markets. To evaluate the importance of costly search in these markets, he estimates a dynamic choice model with search frictions using data on purchases of laundry detergent. Pires's estimates show that consumers incur significant search costs, and ignoring costly search overestimates the own-price elasticity for products more often present in consideration sets and underestimates the elasticity of frequently excluded products. Companies employ marketing devices, such as product displays and advertising, to influence consideration sets. These devices have direct and strategic effects, which this author explores using the estimates of the model. Pires finds that using marketing devices to reduce a product's search cost during a price promotion has modest effects on the overall category revenues, and decreases the revenues of some products.

Gensch and Soofi (1995) proposed an information-theoretic algorithm for estimating the CS. The algorithm estimates the choice probabilities in the awareness sets according to the maximum entropy principle which by use of a multinomial logit model, computes an individual information index for each set, identifies the weak or unacceptable alternatives for each individual and reduces each awareness set to a consideration sets. The average of the selected alternative probabilities was proposed as a statistic by which the predictive quality of various consideration sets can be compared. They found that the predictive power of the multinomial logit is in identifying the weak (non-considered) alternatives rather than predicting the chosen alternative amongst the choices in the consideration sets. The proposed algorithm enables researchers to empirically implement choice set reduction for a given data set using an information criterion.

Several heuristics, including conjunctive, disjunctive, lexicographic, linear additive and geometric compensatory, can be used by individuals in the formation of considerations sets. Most results agree that the conjunctive heuristic is the most often used decision model by consumers (Laroche et al. 2003), suggesting that a brand will be included in a consideration set if it meets the cut-off points in all salient dimensions.

The number of alternatives, which will be evaluated, will vary. Consideration is driven by search costs, opportunity costs and evaluation costs; consideration sets are larger as variance of a brand's utility over purchase occasions increases (Hauser and Wernerfelt 1990). The brand and product category experience of a specific consumer may also affect the size of the CS (Johnson and Lehmann 1997). A consideration set does not necessarily contain several brands, since a consumer will decide to stop searching for more alternatives if a further search for possible solutions is not perceived to be potentially cost-effective (Lapersonne et al. 1995; Eliaz and Spiegler 2011). Certain consumers may therefore make a decision with a consideration set of size one, and it has been suggested that this phenomenon is rather frequent and the factors predicting the consideration of a single brand have been examined (Lapersonne et al. 1995). It has been also suggested that consideration sets with larger number of alternatives do not necessarily lead to better decisions (Diehl 2004).

Allenby and Ginter (1995) investigated consideration sets with respect to in-store displays and feature advertising, within which they highlight on occasions consumers do not expend sufficient mental effort to arrive at a well-defined set of considered decision alternatives. Shapiro et al. (1997) focused on a research study of incidental advertisement exposure by examining whether incidental exposure to an advertisement increases the likelihood that a product/service depicted in the ad will be included in a consideration set.

Consumers often have to create consideration sets when purchasing goals are not well defined. In these situations, the contents of a consideration sets depend on a combination of two motives. First, consumers prefer to create consideration sets of easy-to-compare alternatives. It is easier to compare alternatives that have alignable attributes or alternatives that have overlapping features. Second, consumers prefer to create consideration sets that have a high likelihood of containing their optimal alternative. For example, when the set of available alternatives requires the consumer to make trade-offs between benefits (i.e., to be compensatory), the consumer often delays making a decision about which benefits are preferable, and the consideration sets tends to contain a more diverse set of alternatives (Chakravarti and Janiszewski 2003; Hauser et al. 2010).

Horowitz and Louviere (1995) presented evidence that operational consideration sets are simply indicators of preference. They argue that for the choice settings they have investigated, choice need not be modelled as a two-step process in which a consideration step precedes choice. They contended that modelling choice this way may lead to a misspecified model that makes erroneous forecasts.

Evaluation and Choice

Consumers evaluate a number of alternatives. The academic literature on cognitive psychology has also devoted several studies to the analysis of memory sets and the concept of "chunking". For example, the research showing how retention of multiple features as visual chunks may be achieved with and without long-term memory (Raffone and Wolters 2001); the chunking theory of expert memory (Chase and Simon 1973); experiments on latencies and chess relations, inter-chunk intervals, chunk boundaries and retrieval structures (Gobet and Simon 1998a); and motor chunks (Verwey 2003; Wright et al. 2010; Verwey et al. 2016). Verwey (2003) defended the notion that coding of longer keying sequences involves motor chunks for the individual sequence segments and information on how those motor chunks are to be concatenated.

Punj and Brookes (2001) focused on the effects of the consumer decision process on produce evaluation and choice. They have analysed the influence of product-related decision constraints on external information search. They have also formulated some implications on the way consideration sets are formed through their experimental research study. Jeongwen and Chib (1999) analysed the capabilities of accounting for heterogeneity in consideration sets and in the parameters of the brand choice model. The action related to brand choice heterogeneity in ignoring the consideration sets is surely likely to impact on the marketing mix.

Cherenev and Carpenter (2001) examined consumer inferences about product attributes that are unobservable at the time of the decision. Extant research predicts that in the absence of an explicit correlation between product attributes, consumers will infer that the brand that is superior on the observable attributes is also superior on the unobservable attributes. These authors proposed an alternative inference strategy that makes the counter-intuitive prediction that the apparently superior brand is inferior on the unobservable attributes. These authors refer to these inferences as "compensatory inferences" and asserted that they are associated with consumers' intuitive theories about the competitive nature of a market. Results suggested that consumers' reliance on compensatory inference strategy is likely to depend on the strength of their market efficiency beliefs and that the strength of compensatory inferences depends on the availability of other bases for inference.

Hastak and Mitra (1996) reported on an experiment designed to investigate the effects of brand cues on subsequent retrieval and consideration of other brands in a product category. Prabhaker and Sauer (1994) presented a conceptual framework for analysing the process by which consumers evaluate brand quality based on multiple cues. These authors focused on the use of hierarchical versus non-hierarchical heuristics by consumers in making overall brand evaluations, as well as on the modelling of individual differences amongst consumers.

Previous research on brand name utilisation in consumer judgements has yielded mixed results. Maheswaran et al. (1992) attempted to understand brand name effects within the framework of the heuristic-systematic model. Results showed that low-task importance subjects' evaluations were influenced only by brand name valence. High-task importance subjects' evaluations were affected only by attribute importance in the incongruent conditions, whereas both attribute importance and brand name valence influenced evaluations in the congruent conditions (Guest et al. 2016). Their findings indicated that both consumers' level of motivation and the extent to which brand name based expectations are confirmed by subsequent processing of attribute information, moderate brand name utilisation. Mitra (1995) focused on the dynamics of consumers' consideration sets over a series of purchase occasions and suggested some new measures of composition and stability of the consideration sets. The proposed measures were: the number of brands considered at least once, the standard deviation of consideration sets frequencies and the average discordance in consideration sets composition. This author examined how this stability in consideration sets composition was affected by information in advertising. Higher perceived dispersion of brand-utilities resulting from exposure to differentiating advertising is expected to lead to more stable consideration sets over occasions.

Building on the notion that buyers have a category-specific consideration sets of price-quality tiers, Nowlis and Simonson (2000) proposed that sales promotions and the choice set composition (or the choice context) have compensatory effects on brand switching between price-quality tiers. Mehta et al. (2003) offered an econometric framework that models consumer's consideration sets formation as an outcome of the costly information search behaviour. The proposed structural model dealt with price uncertainty and was estimated using scanner data.

Academic researchers for a number of years have examined the impact of familiarity on consumer decision biases and heuristics. For example, Park and Lessig (1981) studied subjects at three different familiarity levels and results revealed interesting differences in perceptual category breadth, usage of functional and non-functional product dimensions, decision time and confidence. Desai and Hoyer (2000) studied the composition of memory-based consideration sets in terms of their descriptive characteristics, namely stability, variety and preference dispersion. These authors assessed the characteristics of memory-sets relative to occasion and location familiarity. Two experiments investigated by Butler and Berry (2001) demonstrated significant priming for unfamiliar labels and established that priming was unaffected by changing the product type with which the brand name was associated. Also, priming on both auditory and visual versions of the preference judgement task was reduced by changes in modality. Aurier et al. (2000) proposed a theoretical framework and an operationalisation of the concept of consideration sets in relation with usage context. Taking into account the consumer usage situation enables the researcher to analyse the influence of two main components on consideration sets size: context of consumption and familiarity (depth and breadth). These authors showed that CS size varies significantly across consumption contexts and were positively correlated to the breadth of familiarity. Moreover, they found an inverted U relation between the depth of familiarity and consideration sets size.

Chiang et al. (1998) proposed an integrated consideration sets-brand choice model that is capable of accounting for the heterogeneity in consideration sets and in the parameters of the brand choice model. The model was estimated by an approximation Free Markov chain Monte Carlo sampling procedure and was applied to a scanner panel data. They found that ignoring consideration sets heterogeneity understates the impact of marketing mix and overstates the impact of preferences and past purchase feedback even when heterogeneity in parameters is modelled; the estimate of consideration sets heterogeneity was robust to the inclusion of parameter heterogeneity; when consideration sets heterogeneity was included, the parameter heterogeneity took on considerably less importance.

Brown and Carpenter (2000) examined how consumers sometimes treat trivial attributes as though they were critically important in the sense that they have a significant impact on choice. The valence of the effect can depend on whether a positive or negative reason provides a clearer justification for preferring a single brand over its competitors. Thus, the same trivial attribute can generate a positive or negative valuation depending on the choice setting. Such valuation is not always driven by inferences about the attribute itself but can reflect transitory reasoning about the brand as a whole based on the way it is differentiated from its competitors.

Brand choice can be viewed as a two-step process. Households first construct a consideration sets, which does not necessarily include all available brands, and then make a final choice from this set. Vroomen et al. (2003) put forward an econometric model for this two-step process, where they have taken into account that consideration sets usually are not observed. Their model is an artificial neural network, where the CS corresponds with the hidden layer of the network. They discussed representation, parameter estimation and inference. Their results showed that the model improves upon one-step models, in terms of fit and out-of-sample forecasting.

Kivetz and Simonson (2000) examined the mechanisms underlying the impact of incomplete information on consumer choice. The studies included within and between subjects, tests of preference intransitivity, written explanations of choices, think-aloud protocols of choices and choice difficulty. Furthermore, buyers tend to interpret missing attribute values in a way that supports the purchase of the option that is superior on the common attribute. Findings also indicated that choosing from sets with missing information could affect buyer tastes and purchase decisions made subsequently.

Luce et al. (1999) explored whether choice patterns were sensitive to the potential of relevant trade-offs to elicit negative emotion. They have examined on how emotional trade-off difficulty may influence choice, as well as discussing on how the potential of a particular attribute to elicit emotional trade-off difficulty can be measured or manipulated.

Mattila (1998) presented a study which examined the propensity of consumers to rely on heuristic ones when making satisfaction judgements in a repeated-purchase context. They analysed the impact of mood states at the information-encoding for on-line and memory-based judgements, as well as examining whether information-processing efficiency can provide insight into the initial-judgement effect in a consumer behaviour context.

The literature indicating that person, context and task-specific factors cause consumers to utilise different decision strategies has generally failed to affect the specialisation of choice models used by practitioners and academics alike, who still tend to assume a utility maximising, omniscient, indefatigable consumer. Swait and Adamowicz (2001) introduced decision strategy selection, within a maintained compensatory framework, into aggregate choice models via latent classes, which arise because of task complexity; it demonstrates that within an experimental choice task, the model reflects changing aggregate preferences as choice complexity changes and as the task progresses.

Arora and Huber (2001) proposed aggregate customisation as an approach to improve individual estimates using a hierarchical Bayes choice model, with two simulation studies to investigate conditions that are most conducive to aggregate customisation. The simulations were validated through a field study showing that aggregate customisation results in better estimates of individual parameters and more accurate predictions of individuals' choices.

Many individual decisions take place in a group context wherein group members voice their choices segmentally. Ariely and Levav (2000) proposed that choices reflect a balancing of two classes of goals: goals that are strictly individual and goals that are triggered by the existence of the group. They found support for goal balancing, and in one of the three studies undertaken, it was demonstrated that individual choices in a group context are also aimed at satisfying goals of information gathering and self-presentation in the form of uniqueness. However, Hamilton (2003) suggested that people are influenced by others in the selection of certain alternatives, but the degree of this influence will vary, depending on the conditions. For example, people often resist influence when they recognise that somebody is attempting to persuade them.

Prices for grocery items differ across stores and time because of promotion periods. Consumers therefore have an incentive to search for the lowest prices. However, when a product is purchased infrequently, the effort to check the price every shopping trip might outweigh the benefit of spending less. Seiler (2013) proposes a structural model for storable goods that takes into account inventory holdings and search. The model is estimated using data on laundry detergent purchases. He finds search costs play a large role in explaining purchase behaviour, with consumers unaware of the price of detergent on 70% of their shopping trips. Therefore, from the retailer's point of view raising awareness of a promotion through advertising and displays is important. This author also finds that a promotion for a particular product increases the consumer's incentive to search. This change in incentives leads to an increase in category traffic, which from the store manager's perspective is a desirable side effect of the promotion.

Erdem et al. (2003) develop a model of household demand for frequently purchased consumer goods that are branded, storable and subject to stochastic price fluctuations. Their framework accounts for how inventories and expectations of future prices affect current-period purchase decisions. The authors estimate their research model using scanner data for the ketchup category. The results indicate that price expectations and the nature of the price process have important effects on demand elasticities. Long-run cross-price elasticities of demand are more than twice as great as short-run cross-price elasticities. Temporary price cuts (or "deals") primarily generate purchase acceleration and category expansion, rather than brand switching.

Research Focus

Most individuals screen alternatives on more than one attributes, mostly in well-known characteristics of the brands rather than novel characteristics (Gilbride and Allenby 2004). From the discussion above, it is clear that individuals are influenced by the characteristics of the product/brand and the price when making purchasing decisions. Furthermore, the personal characteristics and the experience of a certain consumer in a buying situation could be of influence. These aspects of the customer behaviour are the key focus of this study. They are examined with the use of the DS/AHP.

DS/AHP introduces a number of novel measures describing the judgement-making undertaken in the consumer choice process. These measures include the notion of belief and plausibility in the results, together with the levels of conflict and non-specificity in the judgements made. These measures aid in the identification of the awareness, consideration and choice sets associated with the decision-making group. Within consumer behaviour, the belief and plausibility measures are related to additive and subtractive choice framing (Shafir 1993).

The utilisation of Dempster-Shafer (DST) in DS/AHP brings an allowance of ignorance throughout the judgement-making process, which may encapsulate the notions of incompleteness, imprecision and uncertainty (Smets 1994). An example of the incompleteness is in preference judgements not having to be made on individual DAs; this could be due to forestalling or doubt by the consumer (Lipshitz and Strauss 1997). While the antecedents of the possible presence of ignorance are not the subject of this paper, the technique does acknowledge its presence in the preference judgements made. The implication here is that a consumer may not exactly know what the reasons are for their possible non-specificity in the judgements they make.

Dempster-Shafer Theory

Central to the DS/AHP method of multi-criteria decision-making (MCDM) utilised here is the Dempster-Shafer theory of evidence (DST). The origins of DST came from the seminal work of Dempster (1968) and Shafer (1976) and considered as a generalisation of Bayesian theory that can robustly deal with incomplete and imprecise data (Shafer 1990). DST offers a number of advantages (in MCDM), including the opportunity to assign measures of belief to *focal elements* (groups of DAs), and allow for the attachment of belief to the *frame of discernment* (all DAs). Bloch (1996) presents a description of the basic principles of DST including its main advantages (see also Bryson and Mobolurin 1999).

More formally, let $\Theta = \{h_1, h_2, \ldots, h_n\}$ be a finite set of n hypotheses (frame of discernment). A *basic probability assignment* or *mass value* is a function m: $2^\Theta \rightarrow [0, 1]$ such that $m(\varnothing) = 0$, ($\varnothing$—empty set) and $\sum_{x \in 2^\Theta} m(x) = 1$ (the notation 2^Θ relates to the power set of Θ). Any subset x of the frame of discernment Θ for which the mass value $m(x)$ is non-zero is called a *focal element* and represents the exact belief in the proposition depicted by x. A collection of mass values is denoted a *body of evidence* (BOE), with $m(\Theta)$ considered the amount of ignorance (also called uncertainty) within the BOE $m(\cdot)$, since it represents the level of exact belief that cannot be discerned to any proper subsets of Θ (Bloch 1996).

Further measures of total belief can be found. A *belief* measure is a function Bel: $2^\Theta \rightarrow [0, 1]$, and is drawn from the sum of exact beliefs associated with focal elements that are subsets of the focal element x_1 in question, defined by $Bel(x_1) = \sum_{x_2 \subseteq x_1} m(x_2)$ for $x_1 \subseteq \Theta$. It represents the confidence that a proposition y lies in x_1 or any subset of x_1. Moreover, $m(x_1)$ measures the assignment of belief exactly to x_1, with $Bel(x_1)$ measuring the total assignment of belief to x_1 (Ducey 2001). A *plausibility* measure is a function Pls: $2^\Theta \rightarrow [0, 1]$, defined by $Pls(x_1) = \sum_{x_2 \cap x_1 = \varnothing} m(x_2)$ for $x_1 \subseteq \Theta$. Clearly $Pls(x_1)$ represents the extent to which we fail to disbelieve x_1, the total assignment which does not exclude x_1.

DST provides a method to combine different sources of evidence (BOEs), using Dempster's rule of combination. This rule assumes that these sources are independent. Then the function $[m_1 \oplus m_2]$: $2^\Theta \rightarrow [0, 1]$, defined by:

$$[m_1 \oplus m_2](y) = \begin{cases} 0 & y = \varnothing \\ \dfrac{\sum_{x_1 \cap x_2 = y} m_1(x_1) m_2(x_2)}{1 - \sum_{x_1 \cap x_2 = \varnothing} m_1(x_1) m_2(x_2)} & y \neq \varnothing \end{cases}$$

is a mass value, where x_1 and x_2 are focal elements. An important feature in the denominator part of $[m_1 \oplus m_2]$, is $\sum_{x_1 \cap x_2 = \varnothing} m_1(x_1) m_2(x_2)$, often denoted by k, considered representative of conflict between the independent sources of evidence. The larger the value of k the more conflict in the evidence, and less sense there is in their combination (Murphy 2000). In the limit $k = 1$ (complete conflict), it indicates no focal elements intersect between sources of evidence, and the combination function is undefined (Bloch 1996).

Other functions have been constructed which aim to extract further information from a BOE. For a summary discussion of these functions, see Klir and Wierman (1998). A Non-specificity measure (denoted $N(\cdot)$) within DST was introduced by Dubois and Prade (1985), defined as $N(m(\cdot)) = \sum_{x_1 \in 2^\Theta} m(x_1)\log_2|x_1|$. Hence $N(\cdot)$ is considered the weighted average of the focal elements, with $m(\cdot)$ the degree of evidence focusing on x_1, while $\log_2|x_1|$ indicates the lack of specificity of this evidential claim. The general range of this measure (Klir and Wierman 1998) is $[0, \log_2|\Theta|]$, where $|\Theta|$ is the number of DAs in the frame of discernment. Measurements such as non-specificity are viewed as species of a higher uncertainty type, encapsulated by the term *ambiguity*, Klir and Wierman (1998) state:

> "the latter (ambiguity) is associated with any situation in which it remains unclear which of several alternatives should be accepted as the genuine one."

Exposition of DS/AHP Method Within a Car Choice Problem

The research experiment chosen for this study was based on the conduct of a group discussion with 11 consumers—3 couples and 5 single individuals. The focus of the experiment was on their preferences to a number of different makes of cars. The car choice problem considered here is an often investigated problem and closely inset in the general study of consumer brand choice (Punj and Brookes 2001). This problem brings with it the notion of emotional decision-making (Luce et al. 1999), where familiarity with the problem and the social stereotypes are prevalent. Also the implication of brand cues, with the advertisement on the different makes of cars influential in the judgements made by a consumer (Hastak and Mitra 1996; Shapiro et al. 1997; Wedel and Pieters 2000; Sharpanskykh and Zia 2012).

The ages of the members of the decision-making group ranged from 25 to 60. The majority of the participants had a university degree and the group as a whole had a good level of education. The moderation of the group discussion was performed by both of the researchers in order to ensure a required level of investigator triangulation. A number of projective techniques were used in particular through the utilisation of pictorial information and visual aids pertaining to the subject under investigation: the formation of consumer consideration sets with regard to choice criteria utilised in car purchase decision-making.

The pairing of the stimuli focused on three analytical dyads shown to the participants in three folders (one for each dyad), containing extensive information and pictures about each car model under study, as suggested by Raffone and Wolters (2001). Each car model was labelled with a letter and the comparative dyads were designed in terms of level of consumer familiarity (Aurier et al. 2000), product diversity as well as price-quality tiers. Therefore, the first dyad to be analysed included SMART and IGNIS (a new model just launched), whereas the second grouping dealt with ALFA 156 and VOLVO S60, and, finally, the last pairing contained a "sports car" cluster—TOYOTA MR2 and BMW 3. The setting of these research stimuli was also designed to manipulate brand name valence as well as testing the subjects' processing task. Furthermore, the selected research design rested upon the notion that buyers have category specifics based on "mentally defined" price-quality tiers, following the experiment of Gobet and Simon (1998a) and Verwey (2003). The five criteria selected for analysis of consumers' decision-making with regard to car purchase were comfort, economy, performance, price and safety.

Time was spent introducing DS/AHP to the participants (consumers), including the types of judgement-making required (weight allocation). After having analysed all the provided information for the dyads for a considerable period of time, a very short research instrument was applied in order to gauge and quantify their perceptions related to choice criteria leading to the potential formation of consideration sets. The rest of this section exposits the DS/AHP analysis on the judgements made by the 11 consumers. Firstly, this includes an elucidation of the judgements made by a single consumer and the construction of the subsequent results describing the choice process in identifying the best car (or cars).

Each consumer was allowed to control the level of judgement making, to what they felt confident to undertake (Chase and Simon 1973). The participants were informed that the levels of preference for each criterion analysed should be considered in relation to all the available cars known to the respondents. No cars were allowed to appear in more than one group identified over a single criterion, and not all cars needed to have judgements made on them. Within the DS/AHP analysis of the car choice problem, the six cars SMART, IGNIS, ALFA 156, VOLVO S60, TOYOTA MR2 and BMW 3 considered are labelled A, B, C, D, E and F, respectively, collectively defined the frame of discernment Θ. The judgements made from one individual consumer (labelled DM1) are reported in Fig. 10.1.

In Fig. 10.1, a hierarchical structure (as in AHP) is used to present the judgements made by DM1. Moving down from the focus "best car" to the identified groups of DAs over each criterion, there are two different sets of judgements made by DM1. Firstly, there is the set of criterion priority values (CPVs); these indicate the level of importance or perceived knowledge a consumer has towards the criteria.

In this study the participants were asked to allocate a weight of between 0 and 100 towards each criterion based on their perceived importance, which are then normalised, so they sum to unity (Beynon 2002). If a participant decided to assign a value of 0 to any criterion, then he/she was not required to make judgements for that criterion on the cars considered. Normalising the weights shown in Fig. 10.1, the CPVs (from DM1) for the criteria, comfort ($p_{1,\mathrm{C}}$), economy ($p_{1,\mathrm{E}}$), performance ($p_{1,\mathrm{PE}}$), price

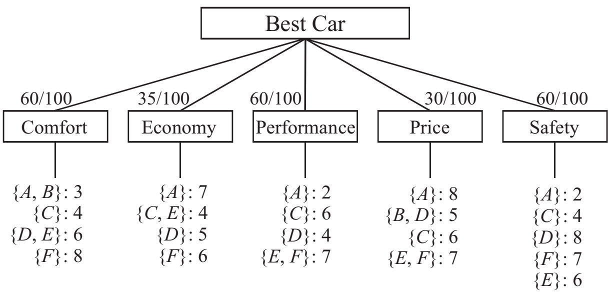

Fig. 10.1 Hierarchy of DS/AHP model of judgements on best car made by DM1

($p_{1,\mathrm{PR}}$) and safety ($p_{1,\mathrm{S}}$) are 0.2449, 0.1429, 0.2449, 0.1224 and 0.2449, respectively. With the CPV assigned for each criterion, it is required for DM1 to make preference judgements towards groups of cars on those criteria with positive CPVs.

For each criterion, a number of groups of cars are identified and assigned a (positive) preference scale value (Fig. 10.1). In this paper a seven-unit scale is used (integer values 2, 3, …, 8), to allow a consumer to discern levels of preference between groups of cars identified (ranging from "moderately preferred" to "extremely preferred"). This positively skewed measurement-scaling procedure was tailored to the prerequisites of DS/AHP and is in line with the well-known work of Miller (1956) and Beynon (2002). Table 10.1 reports a presentation of the relative meaning of the verbal statements to the associated numerical values (with certain verbal statements not given).

In Table 10.1, the numeric values from two to eight indicate from the associated verbal statements an increase in the level of preference on groups of cars. For example, in the case of the price criterion, the group of cars $\{B, D\}$ has been assigned the numerical scale value 5. This indicates the group of cars $\{B, D\}$ has been identified by DM1 as strongly preferred when compared to the whole set of cars considered (frame of discernment Θ) with respect to the price criterion. This approach to preference judgement making, to a frame of reference is not uncommon (Lootsma 1993). Beynon (2002) showed that for a single criterion, if a list of d focal elements (groups of cars) $s_1, s_2, \ldots, s_d$ is identified and assigned with the scale values of $a_1, a_2, \ldots, a_d$, respectively, defining $m(\cdot)$ as the relevant mass values making up the *criterion* BOE for the specific criterion, then

$$m(s_i) = \frac{a_i p}{\sum_{j=1}^{d} a_j p + \sqrt{d}}, \; i = 1,2,\ldots,d \text{ and } m(\Theta) = \frac{\sqrt{d}}{\sum_{j=1}^{d} a_j p + \sqrt{d}},$$

Table 10.1 Connection between numerical values and verbal statements

Numerical value	2	5	8
Verbal statement	Moderately preferred	Strongly preferred	Extremely preferred

where Θ is the frame of discernment and p is the associated CPV. The measure $m(\Theta)$ is defined the level of local ignorance here, since it is the value assigned to Θ, based on the judgements towards a single criterion only. Furthermore, these belief values were found without direct comparison between identified groups of DAs. This relates to the incompleteness in judgements, which is acknowledged and incumbent in the concomitant ignorance.

Results

With respect to the "best car" problem, defining $m_{1,\mathrm{C}}(\cdot)$ as the criterion BOE for the judgements made by DM1 on the comfort criterion, from Fig. 10.1, $s_1 = \{A, B\}$, $s_2 = \{C\}$, $s_3 = \{D, E\}$ and $s_4 = \{F\}$ with $a_1 = 3$, $a_2 = 4$, $a_3 = 6$ and $a_4 = 8$, respectively. For a general CPV $p_{1,\mathrm{C}}$, then

$$m_{1,\mathrm{C}}(\{A,B\}) = \frac{3p_{1,\mathrm{C}}}{21p_{1,\mathrm{C}} + \sqrt{4}},\ m_{1,\mathrm{C}}(\{C\}) = \frac{4p_{1,\mathrm{C}}}{21p_{1,\mathrm{C}} + \sqrt{4}},\ m_{1,\mathrm{C}}(\{D,E\}) = \frac{6p_{1,\mathrm{C}}}{21p_{1,\mathrm{C}} + \sqrt{4}},$$

$$m_{1,\mathrm{C}}(\{F\}) = \frac{8p_{1,\mathrm{C}}}{21p_{1,\mathrm{C}} + \sqrt{4}} \text{ and } m_{1,\mathrm{C}}(\Theta) = \frac{\sqrt{4}}{21p_{1,\mathrm{C}} + \sqrt{4}}.$$

These mass values are dependent only on the value $p_{1,\mathrm{C}}$, for the comfort criterion $p_{1,\mathrm{C}} = 0.2449$, hence $m_{1,\mathrm{C}}(\{A, B\}) = 0.1029$, $m_{1,\mathrm{C}}(\{C\}) = 0.1371$, $m_{1,\mathrm{C}}(\{D, E\}) = 0.2057$, $m_{1,\mathrm{C}}(\{F\}) = 0.2743$ and $m_{1,\mathrm{C}}(\Theta) = 0.2800$. Using the more general values of $m_{1,\mathrm{C}}(\cdot)$, Fig. 10.2a illustrates the effect of the CPV $p_{1,\mathrm{C}}$ on the comfort BOE (also shown in Fig. 10.2b is the respective graphs for the price criterion BOE defined $m_{1,\mathrm{PR}}(\cdot)$ with associated CPV $p_{1,\mathrm{PR}}$).

In Fig. 10.2a, as $p_{1,\mathrm{C}}$ tends to 0 (little importance), more belief value would be assigned to the associated local ignorance $m_{1,\mathrm{C}}(\Theta)$ and less to the identified groups of cars. The reciprocal is true, as $p_{1,\mathrm{C}}$ tends to 1 when there is perceived importance on the comfort criterion so the level of local ignorance decreases. The values of $m_{1,\mathrm{C}}(\cdot)$ for when $p_{1,\mathrm{C}} = 0.2449$ are also confirmed in Fig. 10.2a. In Fig. 10.2b a similar set of graphs are constructed for the mass values making up the BOE of the price criterion (with general CPV $p_{1,\mathrm{PR}}$). The graphs representing the $m_{1,\mathrm{PR}}(\cdot)$ values for

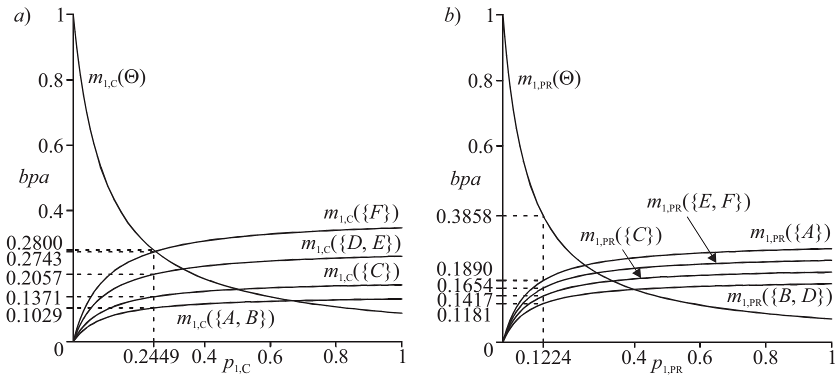

Fig. 10.2 BOE $m_{1,C}(\cdot)$ and $m_{1,PR}(\cdot)$ values as $p_{1,C}$ and $p_{1,PR}$ go from 0 to 1

the identified groups of cars in Fig. 10.2b are closer together than in Fig. 10.2a. Inspection of the judgements made by DM1 in Fig. 10.1 elucidates the range of scale values used on the comfort criterion is larger than those scale values used on the price criterion. For the price criterion with $p_{1,PR} = 0.1224$, then $m_{1,PR}(\{A\}) = 0.1890$, $m_{1,PR}(\{E, F\}) = 0.1654$, $m_{1,PR}(\{C\}) = 0.1417$, $m_{1,PR}(\{B, D\}) = 0.1181$ and $m_{1,PR}(\Theta) = 0.3858$, as shown in Fig. 10.2b.

Criterion BOE can be found for the other three criteria, economy, $m_{1,E}(\cdot)$; performance, $m_{1,PE}(\cdot)$; and safety, $m_{1,S}(\cdot)$, based on the judgements made by DM1 shown in Fig. 10.1 (using their respective CPV: $p_{1,E} = 0.1429$, $p_{1,PE} = 0.2449$ and $p_{1,S} = 0.2449$):

Economy: $m_{1,E}(\{A\}) = 0.1944$, $m_{1,E}(\{C, E\}) = 0.1111$, $m_{1,E}(\{D\}) = 0.1389$, $m_{1,E}(\{F\}) = 0.1667$ and $m_{1,E}(\Theta) = 0.3888$.

Performance: $m_{1,PE}(\{A\}) = 0.0736$, $m_{1,PE}(\{C\}) = 0.2209$, $m_{1,PE}(\{D\}) = 0.1472$, $m_{1,PE}(\{E, F\}) = 0.2577$ and $m_{1,PE}(\Theta) = 0.3000$.

Safety: $m_{1,S}(\{A\}) = 0.0554$, $m_{1,S}(\{C\}) = 0.1107$, $m_{1,S}(\{D\}) = 0.2214$, $m_{1,S}(\{E\}) = 0.1661$, $m_{1,S}(\{F\}) = 0.1937$ and $m_{1,S}(\Theta) = 0.2527$.

The goal for DM1 is to consolidate their evidence on the best car to choose, based on all the criteria considered. Using DS/AHP, this necessitates the combining of the associated criterion BOE using Dempster's combination rule presented in section "Evaluation and Choice". In Table 10.2, the intermediate values from the combination of the two criterion BOEs, comfort $m_{1,\text{C}}(\cdot)$ and price $m_{1,\text{PR}}(\cdot)$, are reported. That is, from Dempster's combination rule the combination is made up of the intersection and multiplication of focal elements and mass values, respectively, from the two different criteria BOE considered.

To illustrate, for the individual mass value $m_{1,\text{C}}(\{A, B\}) = 0.1029$ and $m_{1,\text{PR}}(\{A\}) = 0.1890$ from the comfort and price criterion BOE respectively, their combination results in a focal element $\{A, B\} \cap \{A\} = \{A\}$ with a value $0.1029 \times 0.1890 = 0.0194$. The ∅ present in Table 10.2 is the empty set and the sum of these values (in italics) represents the level of associated conflict (see k definition in section "Dempster-Shafer Theory") in the combination of these two criterion BOE, in this case $k = 0.2875$. The final mass value constructed for a particular focal element is illustrated for the $\{A\}$ focal element, which is given by:

$$\left[m_{1,\text{C}} \oplus m_{1,PR}\right]\left(\{A\}\right) = \frac{m_{1,\text{C}}\left(\{A,B\}\right)m_{1,\text{PR}}\left(\{A\}\right) + m_{1,\text{C}}\left(\Theta\right)m_{1,\text{PR}}\left(\{A\}\right)}{1-0.2875}$$
$$= \frac{0.0194 + 0.0529}{0.7125} = 0.1015.$$

To re-iterate, Dempster's rule of combination is used to aggregate the evidence from a consumer's judgements on the five different criteria

Table 10.2 Intermediate values from combination of comfort and price BOE for DM1

$m_{1,\text{C}}(\cdot)\backslash m_{1,\text{PR}}(\cdot)$	{*A*}, 0.1890	{*E*, *F*}, 0.1654	{*C*}, 0.1417	{*B*, *D*}, 0.1181	Θ, 0.3858
{*A*, *B*}, 0.1029	{*A*}, 0.0194	∅, *0.0170*	∅, *0.0146*	{*B*}, 0.0121	{*A*, *B*}, 0.0397
{*C*}, 0.1371	∅, *0.0259*	∅, *0.0227*	{*C*}, 0.0194	∅, *0.0162*	{*C*}, 0.0529
{*D*, *E*}, 0.2057	∅, *0.0389*	{*E*}, 0.0340	∅, *0.0292*	{*D*}, 0.0243	{*D*, *E*}, 0.0794
{*F*}, 0.2743	∅, *0.0518*	{*F*}, 0.0454	∅, *0.0389*	∅, *0.0324*	{*F*}, 0.1058
Θ, 0.2800	{*A*}, 0.0529	{*E*, *F*}, 0.0463	{*C*}, 0.0397	{*B*, *D*}, 0.0331	Θ, 0.1080

considered. Defining $m_{1,\mathrm{CAR}}(\cdot)$ as the post-combination *consumer* BOE from all the criterion BOEs for DM1, its associated focal elements (groups of cars) and mass values are reported in Table 10.3.

In Table 10.3, 12 groups of cars (focal elements including Θ) and mass value, making up the consumer BOE for DM1. To illustrate, the focal element $m_{1,\mathrm{CAR}}(\{B, D\}) = 0.0084$, implies the exact belief in the group of cars {*B*, *D*} including the best car from the combined evidence is 0.0084. Furthermore, the level of ignorance $m_{1,\mathrm{CAR}}(\Theta) = 0.0276$, from the combination of all the judgements of DM1 towards their choice of best car. To consider total beliefs to groups of cars, the belief (*Bel*) and plausibility (*Pls*) functions utilised on $m_{1,\mathrm{CAR}}(\cdot)$ associated with DM1. Rather than present the belief and plausibility values for each possible subgroup of cars considered (62 in number) a specific reduced number are described. Moreover, Table 10.4 reports those groups of cars that have the largest belief and plausibility values from all those groups of cars of the same size.

To illustrate the results in Table 10.4, considering all groups of cars made up of three cars, those with the largest belief and plausibility values are {*D*, *E*, *F*} in both cases, with $Bel(\{D, E, F\}) = 0.7080$ and with $Pls(\{D, E, F\}) = 0.7519$. These values are calculated from the information reported in Table 10.3 and are constructed as shown below:

$$\begin{aligned}
Bel\left(\{D,E,F\}\right) &= \sum_{x_2 \subseteq \{D,E,F\}} m_{1,\mathrm{CAR}}\left(x_2\right),\\
&= m_{1,\mathrm{CAR}}\left(\{D\}\right) + m_{1,\mathrm{CAR}}\left(\{E\}\right) + m_{1,\mathrm{CAR}}\left(\{F\}\right) + m_{1,\mathrm{CAR}}\left(\{D,E\}\right) + m_{1,\mathrm{CAR}}\left(\{E,F\}\right),\\
&= 0.1807 + 0.1691 + 0.2923 + 0.0203 + 0.0456,\\
&= 0.7080,
\end{aligned}$$

and

Table 10.3 Individual groups of cars and mass values in the $m_{1,\mathrm{CAR}}(\cdot)$ BOE

{*A*}, 0.0902	{*D*}, 0.1807	{*A*, *B*}, 0.0101	{*D*, *E*}, 0.0203
{*B*}, 0.0031	{*E*}, 0.1691	{*B*, *D*}, 0.0084	{*E*, *F*}, 0.0456
{*C*}, 0.1447	{*F*}, 0.2923	{*C*, *E*}, 0.0079	{Θ}, 0.0276

Table 10.4 Subsets of DAs with largest belief and plausibility values from the $m_{1,\mathrm{CAR}}(\cdot)$ BOE

Size of car group	Belief	Plausibility
1	{*F*}, 0.2923	{*F*}, 0.3655
2	{*E*, *F*}, 0.5070	{*D*, *F*}, 0.5749
3	{*D*, *E*, *F*}, 0.7080	{*D*, *E*, *F*}, 0.7519
4	{*C*, *D*, *E*, *F*}, 0.8606	{*C*, *D*, *E*, *F*}, 0.8966
5	{*A*, *C*, *D*, *E*, *F*}, 0.9508	{*A*, *C*, *D*, *E*, *F*}, 0.9969

$$\begin{aligned}
Pls(\{D,E,F\}) &= \sum_{\{D,E,F\}\cap x_2 \neq \varnothing} m_{1,\mathrm{CAR}}(x_2),\\
&= m_{1,\mathrm{CAR}}(\{D\}) + m_{1,\mathrm{CAR}}(\{E\}) + m_{1,\mathrm{CAR}}(\{F\}) + m_{1,\mathrm{CAR}}(\{B,D\})\\
&\quad + m_{1,\mathrm{CAR}}(\{C,E\}) + m_{1,\mathrm{CAR}}(\{D,E\}) + m_{1,\mathrm{CAR}}(\{E,F\}) + m_{1,\mathrm{CAR}}(\Theta),\\
&= 0.1807 + 0.1691 + 0.2923 + 0.0084 + 0.0079 + 0.0203 + 0.0456 + 0.0276,\\
&= 0.7519.
\end{aligned}$$

The results in Table 10.4 highlight the use of DS/AHP to identify a reduced number of cars to possibly further consider. For the "best car" problem here, if considering finding only the single best car, the measures of belief and plausibility both indicate the car *F* (BMW 3) is best, based on all the judgements from DM1. This discussion and results in Table 10.4 illustrate the possible role of DS/AHP as a method to identify choice sets from consideration and/or awareness sets (see section "Dempster-Shafer Theory" for further discussion).

To offer information on the homogeneity and intensity of the consumer's choice process, the conflict levels between the judgements made by DM1 over the different criteria can be calculated (see Table 10.5). With respect to DS/AHP, the level of conflict relates to how different the judgements made are over the different criteria (see section "Evaluation and Choice").

In Table 10.5, the higher the conflict value (within the domain [0, 1]), the more conflict there exists between the criteria. The most conflict evident is between the comfort and safety criteria (with k = 0.3844). From section "Evaluation and Choice", since the conflict levels are relatively low between criteria, it strengthens the validity of the results found from the combination of the five criterion BOEs to produce the consumer BOE.

A further measure defined in section "Evaluation and Choice" is non-specificity; here it relates to the level of grouping apparent in the groups of cars identified for preference by DM1 over the different criteria. From section "Evaluation and Choice", with six cars considered, the domain on the level of non-specificity is [0, 2.5850]. In Table 10.6, the levels of non-specificity on the judgements made by DM1 over the five criterion BOEs and the final consumer BOE are reported.

From Table 10.6, the largest and least levels of non-specificity amongst the criterion BOE are associated with the price (1.2808) and safety (0.6532) criteria respectively. To illustrate the calculation of these non-specificity values ($N(\cdot)$), for the price criterion:

$$\begin{aligned}
N\left(m_{1,\mathrm{PR}}(.)\right) &= \sum_{x_1 \in 2^{\Theta}} m_{1,\mathrm{PR}}(x_1)\log_2|x_1|,\\
&= m_{1,\mathrm{PR}}(\{A\})\log_2|\{A\}| + m_{1,\mathrm{PR}}(\{E,F\})\log_2|\{E,F\}| + m_{1,\mathrm{PR}}(\{C\})\log_2|\{C\}|\\
&\quad + m_{1,\mathrm{PR}}(\{B,D\})\log_2|\{B,D\}| + m_{1,\mathrm{PR}}(\Theta)\log_2|\Theta|,\\
&= 0.1890\log_2 1 + 0.1654\log_2 2 + 0.1417\log_2 1 + 0.1181\log_2 2 + 0.3858\log_2 6,\\
&= 1.2808.
\end{aligned}$$

A comparison between the judgements made on the price and safety (and other) criteria (given in Fig. 10.1) shows the price criterion includes

Table 10.5 Conflict values between criterion BOEs for DM1

Criteria	Economy	Performance	Price	Safety
Comfort	0.3076	0.3117	0.2875	0.3844
Economy	–	0.2965	0.2605	0.3521
Performance	–	–	0.3243	0.3688
Price	–	–	–	0.3471

Table 10.6 Levels of non-specificity on judgements made by DM1

Evidence	Comfort	Economy	Performance	Price	Safety	Consumer
Non-specificity	1.0323	1.1164	1.0347	1.2808	0.6532	0.1596

two groups of cars identified with two cars in each, whereas only singleton groups of cars are identified with the safety criterion. Following the premise that information chunk boundaries have psychological reality (Gobet and Simon 1998a). One further important factor is the value of the associated CPV, since with a low CPV more mass value is assigned to Θ, hence a higher non-specificity value. The non-specificity of the consumer BOE $m_{1,CAR}(\cdot)$, $N(m_{1,CAR}(\cdot))$ = 0.1596, is lower than the non-specificity levels for the individual criterion BOE. This is a direct consequence of the utilisation of Dempster's combination rule, which apportions mass values to smaller groups of cars through the intersection of groups of cars from the different criterion BOE (see Table 10.2).

To set against the analysis on DM1 so far described, a further series of results are briefly reported based on the judgements of a second consumer (labelled DM2), see Fig. 10.3.

DM2s' judgements are considerably less specific than those with DM1 (see Figs. 10.1 and 10.3), with larger-sized groups of cars were identified by DM2 over the five criteria. Incidentally the judgements made by DM2 are consistent with the dyad grouping of the cars presented to the consumers. With the cars grouped by {A, B}, {C, D} and {E, F} suggesting a level of brand name valence by this consumer, their judgements exhibit influence by the price-quality tiers of the three dyad groups of cars, reinforcing the notion of flat chunk organisation and its relation to retrieval structures. As with DM1, the criterion BOE graphs for the comfort and price criteria for DM2 are reported in Fig. 10.4.

Comparing the results reported in Figs. 10.2 and 10.4, the separation between the $m_{2,C}(\{E, F\})$ and $m_{2,C}(\{A, B, C, D\})$ lines in Fig. 10.4a is a

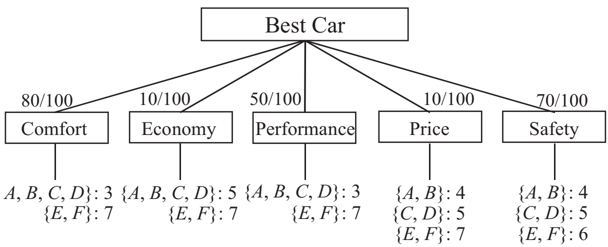

Fig. 10.3 Best car judgements over the five criteria from DM2

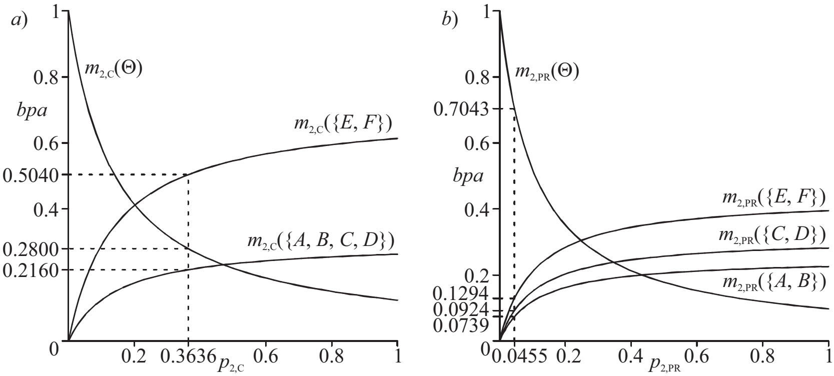

Fig. 10.4 BOE $m_{2,C}(\cdot)$ and $m_{2,PR}(\cdot)$ mass values as $p_{2,C}$ and $p_{2,PR}$ go from 0 to 1

consequence of the large difference between the scale values 3 and 7 assigned to the two groups of cars {*A*, *B*, *C*, *D*} and {*E*, *F*}, respectively. The non-specificity levels on the criterion BOEs for DM2 are reported in Table 10.7 and exhibit consistently higher values than those associated with DM1 (see Table 10.6). This is a consequence of the larger-sized groups of cars identified across all the criteria by DM2. A further consequence of the less specific judgements made is the non-specificity of consumer BOE for DM2 (1.1368) is considerably larger than that for DM1 (0.1596).

While the views of the individual consumers are of interest, the combination of the evidence from the eleven consumers would offer information on the overall levels of belief towards the identification of the best car(s) from the six cars considered over the five criteria. That is, the combination of all the consumer BOEs from the 11 consumers enables a novel approach to the evaluation of results from group decision-making with DS/AHP. This is undertaken by the utilisation of Dempster's combination rule described in section "Evaluation and Choice". For brevity we do not present the final group BOE from all consumers; instead the

Table 10.7 Levels of non-specificity on judgements made by DM2

Evidence	Comfort	Economy	Performance	Price	Safety	Consumer
Non-specificity	1.6598	2.2598	1.7929	2.1163	1.4220	1.1368

best groups of cars of different sizes based on the belief and plausibility measures are reported in Table 10.8.

From Table 10.8, irrespective of whether belief and plausibility measures are considered, the same group of cars is identified for each specific size of group. The best single car is identified as the car *D* (VOLVO S60), if a choice set of say three cars was considered, then the group of cars {*C*, *D*, *F*} should be chosen. The results in Table 10.8 exhibit the possible consideration or choice sets that the consumers could further consider (see section "Dempster-Shafer Theory" for further discussion).

At each stage of the DS/AHP analysis certain BOE are constructed and can be combined in a number of different ways to allow further understanding of the prevalent judgements made. For example, each consumer BOE was found from the combination of criterion BOEs, and the group BOE found from the combination of the consumer BOEs. To gauge a measure on the judgements made specifically over the different criteria, the criterion BOEs associated with a single criterion from the 11 consumers can be combined. The result is five BOEs. Table 10.9 reports their concomitant levels of non-specificity.

An inspection of the results in Table 10.9 shows the criterion with overall least and largest levels of non-specificity in the judgements made are safety (0.0510) and economy (0.3215), respectively. This result is interesting in that overall safety was judged on most discernibly in terms of both

Table 10.8 Subsets of cars with largest belief and plausibility values from final group BOE

Size of car group	Belief	Plausibility
1	{*D*}, 0.6651	{*D*}, 0.6552
2	{*D*, *F*}, 0.9747	{*D*, *F*}, 0.9748
3	{*C*, *D*, *F*}, 0.9911	{*C*, *D*, *F*}, 0.9913
4	{*C*, *D*, *E*, *F*}, 0.9999	{*C*, *D*, *E*, *F*}, 1.0000
5	{*B*, *C*, *D*, *E*, *F*}, 0.9999	{*B*, *C*, *D*, *E*, *F*}, 1.0000

Table 10.9 Levels of non-specificity for the different criteria

Evidence	Comfort	Economy	Performance	Price	Safety
Non-specificity	0.1302	0.3215	0.0881	0.1758	0.0510

the grouping of cars under this criterion and the level of CPV each consumer assigned to it, whereas the economy criterion was most non-specific. This could be a direct consequence of the information made available to the consumers not including all that was necessary for them to make more specific judgements. The combined judgements of the 11 consumers over the different criteria are next exposited in Table 10.10. That is, for each criterion using the defined combined BOE, the different-sized groups of cars with highest belief and plausibility values (not given) are shown.

From Table 10.10, in terms of a single best car to identify, three of the five criteria (comfort, economy and safety) all suggest the car *D* as best choice of car. With cars *F* and *A* identified as best from the criteria performance and price respectively (based on belief or plausibility values). The results from the price criterion are interesting and also in some way different to those from the other criteria. That is (considering only the belief value), the best two cars to consider under the price criterion are *A* and *B*—the cheapest two of the six cars considered. Also (for the price criterion) the best four cars to further consider are *A*, *B*, *C* and *D*, the cheapest four of the six cars. The reader is reminded the six cars considered were presented to the consumers in the dyad groups {*A*, *B*}, {*C*, *D*} and {*E*, *F*} based primarily on their prices.

The results presented here show the individual consumers generally followed this dyadic grouping. This highlights the effect of brand cues, which in this case were in the form of the folders containing extensive information and pictures about each car. Indeed, with the price clearly included in the cue information, the results on the price criterion indicate the consumers have exhibited "mentally defined" price-quality tiers during their judgement making. This finding is supported by the research study conducted by Mehta et al. (2003).

Interpretation of DS/AHP as an Analysis Tool in Consumer Choice MCDM

A DS/AHP analysis was undertaken on a car choice problem. Throughout, the judgements made and results considered have been with respect to groups of cars. This places the notion of consideration

Table 10.10 Subsets of cars with largest belief and plausibility values from different criteria

Belief	Comfort	Economy	Performance	Price	Safety
1	{*D*}	{*D*}	{*F*}	{*A*}	{*D*}
2	{*D*, *F*}	{*C*, *D*}	{*C*, *F*}	{*A*, *B*}	{*D*, *F*}
3	{*D*, *E*, *F*}	{*C*, *D*, *E*}	{*C*, *E*, *F*}	{*A*, *B*, *C*}	{*D*, *E*, *F*}
4	{*C*, *D*, *E*, *F*}	{*B*, *C*, *D*, *E*}	{*C*, *D*, *E*, *F*}	{*A*, *B*, *C*, *D*}	{*C*, *D*, *E*, *F*}
5	{*B*, *C*, *D*, *E*, *F*}	{*B*, *C*, *D*, *E*, *F*}	{*A*, *C*, *D*, *E*, *F*}	{*A*, *B*, *C*, *D*, *E*}	{*A*, *C*, *D*, *E*, *F*}
Plausibility	**Comfort**	**Economy**	**Performance**	**Price**	**Safety**
1	{*D*}	{*D*}	{*F*}	{*A*}	{*D*}
2	{*D*, *F*}	{*B*, *D*}	{*C*, *F*}	{*A*, *C*}	{*D*, *F*}
3	{*D*, *E*, *F*}	{*B*, *D*, *E*}	{*C*, *E*, *F*}	{*A*, *B*, *C*}	{*C*, *D*, *F*}
4	{*C*, *D*, *E*, *F*}	{*B*, *D*, *E*, *F*}	{*C*, *D*, *E*, *F*}	{*A*, *B*, *C*, *D*}	{*C*, *D*, *E*, *F*}
5	{*B*, *C*, *D*, *E*, *F*}	{*B*, *C*, *D*, *E*, *F*}	{*A*, *C*, *D*, *E*, *F*}	{*A*, *B*, *C*, *D*, *E*}	{*A*, *C*, *D*, *E*, *F*}

sets as a fundamental tool in DS/AHP. The allowance of a consumer to discern groups of cars does place the control in the level of judgement making squarely with the consumer. This is strengthened with the allowance to not include in their judgements all cars present over particular criteria. Importantly on a criterion, these cars are not simply ignored, but included with all the other cars and a value assigned to them, represented by local ignorance. The criteria priority values representing the level of knowledge (or importance) of the criteria enable a consumer to indicate no knowledge on a criterion and hence requiring no further judgements on this criterion (Aurier et al. 2000).

The results from the DS/AHP analysis is in the form of a BOE (Racioppi et al. 2015; Beynon 2006; Han et al. 2013) which enables the identification of a single best car or best group of cars of a certain size (number of cars). This identification is based on the two measures, belief and plausibility. While relatively novel to consumer choice theory, they do attempt to elucidate different approaches to the identification of best groups of cars (consideration sets). The belief function represents the confidence that the best car does exist in the group of cars it describes. Whereas the plausibility function represents the extent to which we fail to disbelieve the best car does exist in the group of cars the value describes. These two functions have connection with the consumer choice process as suggested in Park et al. (2000), whose title included the phrase "choosing what I want versus rejecting what I do not want" (Chakravarti and Janiszewski 2003; Sharpanskykh and Zia 2012). Indeed, a decision on the choice of whether to utilise the belief or plausibility values may came from whether the consumer has undertaken their judgements in a subtractive or additive option framing (Shafir 1993).

Within the area of consideration sets, the ability to find the levels of belief or plausibility on groups of cars could aid in the elucidation of groups of cars representing a consumer's awareness, consideration and choice sets. That is, from investigating the changes in the largest levels of belief and plausibility on groups of cars of different sizes, this evidence can suggest the sizes and content of groups of cars associated with the varying consideration sets. To illustrate, for DM1 whose judgements were elucidated in section "Research Focus", Fig. 10.5 reports the results from Table 10.4 in a form to elucidate the level of change in belief and plausibility values between best groups of cars of different sizes.

In Fig. 10.5, the relative change in the levels of (largest) belief or plausibility to groups of cars of different sizes is clearly exposited. A notional attempt in Fig. 10.5 is given to the understanding of the effect of the belief and plausibility values in elucidating the possible awareness, consideration and choice sets. That is, the shaded regions relate to bounds on levels of belief or plausibility, which may discern groups of cars into awareness, consideration and choice sets. In general, as the levels of these measures increase so more cars are included in each of the best groups of cars identified. In this case, in Fig. 10.5, two (notional) bounds are utilised; these are between awareness and consideration sets with a boundary value 0.9 and between consideration and choice sets with a boundary value 0.55.

With these boundary values, the group of objects immediately below the boundary values is of particular importance. Also with this convention defined here, the awareness set is the group of all cars considered in the problem, the frame of discernment Θ. It follows, considering only the belief values, the group of cars $\{C, D, E, F\}$ is the first group of cars with a belief value below the awareness and consideration sets boundary value of 0.9 (*Bel*({*C*, *D*, *E*, *F*}) = 0.8606) and is then defined the consideration set. To explain this result from the view of the judgements made by DM1, the two cars not included in this identified consideration set are *A* and *B*, which from Fig. 10.1 overall were the two cars given the least amount of preference judgement on them. This may be due to non-inclusion in many identified groups (in the case of car *B*) or on the generally low levels of preference assigned to them (in the case of car *A*, with the exception of the price criterion). Using the plausibility value the group of cars {*C*, *D*, *E*, *F*} with *Pls*({*C*, *D*, *E*, *F*}) = 0.8966 would also be considered the concomitant consideration set.

A similar argument follows for the boundary between the regions describing the consideration and choice sets (Gobet and Simon 1998b). From Fig. 10.5, for the belief value, the group of objects {*E*, *F*} would be identified (*Bel*({*E*, *F*}) = 0.5070) as the choice set since its level of belief is the largest below the 0.55 boundary value employed here. In this case, based on the plausibility value the group of cars {*F*} would be identified as the choice set, which is different from (included in) the choice set based on the belief value. As a general rule, if the same boundary values were used for belief and plausibility levels then the number of cars in a plausibility-based best group of cars is less than or equal to the number of cars in a group based on the belief value.

Due to the novelty of the method employed here, the boundary values utilised are notional. In the future with the application of this method increasing, a fuller understanding on the boundary values to consider

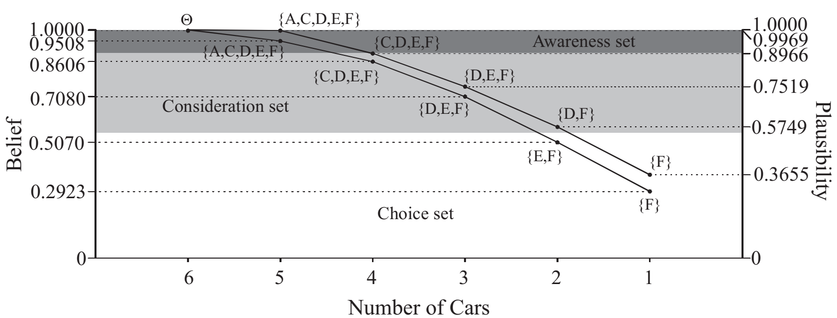

Fig. 10.5 Visual representation of the groups of cars with largest levels of *Bel* and *Pls*

would be apparent. However, this would depend on the number of cars in the awareness set and so on.

Conclusions

Understanding how consumers choose specific brands is critical for companies, especially when the available number of brands competing with one another is large. This paper has utilised a nascent approach to multi-criteria decision-making, namely, DS/AHP in the area of consumer choice. The paper has attempted to convey a more realistic approach for the individual consumer to undertake the required judgement-making process. Importantly, the DS/AHP method allows the consumer to control the intensity of the judgement making they perform. The results (intermediate and final) elucidate a plethora of information for the consumer choice problem to be gauged on.

The central element in the DS/AHP analysis is the body of evidence (BOE), with certain BOE constructed at different stages in the analysis, then a number of different sets of results can be found. The descriptive measures conflict and non-specificity allow a novel insight into the judgement making by the individual members of a decision-making group. Further analysis could include the investigation of the levels of conflict between the individual members of the group and looking into the possible identification of subgroups of a group with the most similar series of judgements.

Allowance exists for each consumer to assign levels of positive preference to groups of cars. The results also included information on groups of cars; hence the notion of consideration sets is firmly implanted in the fundamentals of DS/AHP. Moreover, the idea of consideration sets is exhibited in the judgement-making opportunities of the consumer and in the interpretation of the final results. A notional approach to the identification of awareness, consideration and choice sets is described, based on the levels of belief and plausibility in the best car existing in a group of cars, which could be compared with the algorithm developed by Gensch and Soofi (1995).

It is hoped this first DS/AHP analysis in a consumer choice problem, has shown it to be a novel and lucrative method of analysis. Its ability to allow the decision-maker to make judgements to the level of their ability as well as offer results that can identify a number of different aspects of the whole decision-making process.

References

Allenby, G. M., & Ginter, J. L. (1995). The Effects of In-store Displays and Feature Advertising on Consideration Sets. *International Journal of Research in Marketing, 12*, 67–80.

Analytis, P., Kothiyal, A., & Katsikopoulos, K. (2014). Multi-Attribute Utility Models as Cognitive Search Engines. *Judgment and Decision making, 9*(5), 403–419.

Ariely, D., & Levav, J. (2000). Sequential Choice in Group Settings: Taking the Road Less Travelled and Less Enjoyed. *Journal of Consumer Research, 27*(3), 279–290.

Arora, N., & Huber, J. (2001). Improving Parameter Estimates and Model Prediction by Aggregate Customization in Choice Experiments. *Journal of Consumer Research, 28*(2), 273–283.

Aurier, P., Jean, S., & Zaichkowsky, J. L. (2000). Consideration Set Size and Familiarity with Usage Context. *Advances in Consumer Research, 27*, 307–313.

Beaman, C. P. (2013). Inferring the Biggest and Best: A Measurement Model for Applying Recognition to Evoke Consideration Sets and Judge Between Multiple Alternatives. *Cognitive Systems Research, 24*, 18–25.

Benitez, J., Delgado-G, X., Izquierdo, J., & Pérez-G, R. (2015). Consistent Completion of Incomplete Judgments in Decision Making Using AHP. *Journal of Computational and Applied Mathematics, 290*, 412–422.

Beynon, M. (2002). DS/AHP Method: A Mathematical Analysis, Including an Understanding of Uncertainty. *European Journal of Operational Research, 140*(1), 149–165.

Beynon, M. J. (2006). The Role of the DS/AHP in Identifying Inter-Group Alliances and Majority Rule Within Group Decision Making. *Group Decision and Negotiation, 15*(1), 21–42.

Beynon, M. J., Curry, B., & Morgan, P. H. (2000). The Dempster-Shafer Theory of Evidence: An Alternative Approach to Multicriteria Decision Modelling. *Omega, 28*(1), 37–50.

Bloch, B. (1996). Some Aspects of Dempster-Shafer Evidence Theory for Classification of Multi-Modality Images Taking Partial Volume Effect into Account. *Pattern Recognition Letters, 17*, 905–919.

Brown, C. L., & Carpenter, G. S. (2000). Why Is the Trivial Important? A Reasons-Based Account for the Effects of Trivial Attributes on Choice. *Journal of Consumer Research, 26*(4), 372–385.

Bryson, N., & Mobolurin, A. (1999). A Process for Generating Quantitative Belief Functions. *European Journal of Operational Research, 115*(3), 624–633.

Butler, L. T., & Berry, D. C. (2001). Transfer Effects in Implicit Memory and Consumer Choice. *Applied Cognitive Psychology, 15*(6), 587–601.

Carson, R., & Louviere, J. (2014). Statistical Properties of Consideration Sets. *Journal of Choice Modelling, 13*, 37–48.

Chakravarti, A., & Janiszewski, C. (2003). The Influence of Macro-Level Motives on Consideration Set Composition in Novel Purchase Situations. *Journal of Consumer Research, 30*(September), 244–258.

Chase, W. G., & Simon, H. K. (1973). Perception in Chess. *Cognitive Psychology, 4*, 55–81.

Cherenev, A., & Carpenter, G. S. (2001). The Role of Market Efficiency Intuitions in Consumer Choice: A Case of Compensatory Inferences. *Journal of Marketing Research, 38*(3), 349–361.

Chiang, J., Chib, S., & Narasimhan, C. (1998). Markov Chain Monte Carlo and Models of Consideration Set and Parameter Heterogeneity. *Journal of Econometrics, 89*(1–2), 223–248.

Dede, G., Kamalakis, T., & Sphicopoulos, T. (2016). Theoretical Estimation of the Probability of Weight Rank Reversal in Pairwise Comparisons. *European Journal of Operational Research, 252*(2), 587–600.

Dempster, A. P. (1968). A Generalization of Bayesian Inference (with Discussion). *Journal of the Royal Statistical Society. Series B, 30*(2), 205–247.

Desai, K. K., & Hoyer, W. D. (2000). Descriptive Characteristics of Memory-Based Consideration Sets: Influence of Usage Occasion Frequency and Sage Location Familiarity. *Journal of Consumer Research, 27*(3), 309–323.

Diehl, K. (2004,August). When Two Rights Make a Wrong: Searching Too Much in Ordered Environments. *Journal of Marketing Research, XLII*, 213–322.

Dubois, D., & Prade, H. (1985). A Note on Measures of Specificity for Fuzzy Sets. *International Journal of General Systems, 10*(4), 279–283.

Ducey, M. J. (2001). Representing Uncertainty in Silvicultural Decisions: An Application of the Dempster-Shafer Theory of Evidence. *Forest Ecology and Management, 150*, 199–211.

Eliaz, K., & Spiegler, R. (2011). Consideration Sets and Competitive Marketing. *Review of Economic Studies, 78*(1), 235–262.

Erdem, T., Imai, S., & Keane, M. P. (2003). Brand and Quantity Choice Dynamics Under Price Uncertainty. *Quantitative Marketing and Economics, 1*(1 March), 5–64.
Gensch, D. H., & Soofi, E. S. (1995). Information-Theoretic Estimation of Consideration Sets. *International Journal of Research in Marketing, 12*, 25–38.
Gilbride, T., & Allenby, G. (2004). A Choice Model with Conjunctive, Disjunctive, and Compensatory Screening Rules. *Marketing Science, 23*(3), 391–406.
Gobet, F., & Simon, H. A. (1998a). Expert Chess Memory: Revisiting the Chunking Hypothesis. *Memory, 6*(3), 225–255.
Gobet, F., & Simon, H. A. (1998b). Pattern Recognition Makes Search Possible: Comments on Holding (1992). *Psychological Research, 61*(3), 204–209.
Goodman, J., Broniarczyk, S., Griffin, J., & McAlister, L. (2013). Help or Hinder? When Recommendation Signage Expands Consideration Sets and Heightens Decision Difficulty. *Journal of Consumer Psychology, 23*(2), 165–174.
Guest, D., Estes, Z., Gibbert, M., & Mazunrsky, D. (2016). Brand Suicide? Memory and Linking of Negative Brand Names. *PloS One, 11*(3), e0151628.
Hamilton, R. (2003). Why Do People Suggest What They Don Not Want? Using Context Effects to Influence Others' Choices. *Journal of Consumer Research, 29*, 492–506.
Han, D., Han, C., & Deng, Y. (2013). Novel Approaches for the Transformation of Fuzzy Membership Function into Basic Probability Assignment Based on Uncertainty Optimization. *International Journal of Uncertainty Fuzziness and Knowledge-Based Systems, 21*(*2*), 289–322.
Hastak, M., & Mitra, A. (1996). Facilitating and Inhibiting Effects of Brand Cues on Recall, Consideration Set and Choice. *Journal of Business Research, 37*(2), 121–127.
Hauser, J. R., & Wernerfelt, B. (1990). An Evaluation Cost Model of Evoked Sets. *Journal of Consumer Research, 16*(March), 383–408.
Hauser, J., Toubia, O., Evgeniou, T., Befurt, R., & Dzyaburra, D. (2010). Disjunctions of Conjunctions, Cognitive Simplicity, and Consideration Sets. *Journal of Marketing Research, 47*(3), 485–496.
Hogarth, R. M. (1980). *Judgement and Choice* (2nd ed.). New York: Wiley.
Horowitz, J. L., & Louviere, J. J. (1995). What Is the Role of Consideration Sets in Choice Modeling? *International Journal of Research in Marketing, 12*(1), 39–54.
Jeongwen, C., & Chib, S. (1999). Markov Chain, Monte Carlo and Models of Consideration Set and Parameter Heterogeneity. *Journal of Econometrics, 89*(1/2), 223–249.

Johnson, M. D., & Lehmann, D. R. (1997). Consumer Experiences and Consideration Sets for Brands and Product Categories. *Advances in Consumer Research, 24*, 295–301.

Kivetz, R., & Simonson, I. (2000). The Effects of Incomplete Information on Consumer Choice. *Journal of Marketing Research, 37*(4), 427–448.

Klir, G. J., & Wierman, M. J. (1998). *Uncertainty-Based Information: Elements of Generalized Information Theory*. Heidelberg: Physica-Verlag.

Lapersonne, E., Laurent, G., & Le Goff, J.-J. (1995). Consideration Sets of Size One: An Empirical Investigation of Automobile Purchases. *International Journal of Research in Marketing, 12*(1), 55–66.

Laroche, M., Kim, C., & Marsui, T. (2003). Which Decision Heuristivs Are Used in Consideration Set Formation? *Journal of Consumer Marketing, 20*(3), 192–209.

Lipshitz, R., & Strauss, O. (1997). Coping with Uncertainty: A Naturalistic Decision-Making Analysis. *Organisational Behaviour and Human Decision Processes, 69*(2), 149–163.

Lock, A. R., & Thomas, H. (1979). Appraisal of Multi-Attribute Utility Models in Marketing. *European Journal of Marketing, 13*(5), 294–307.

Lootsma, F. A. (1993). Scale Sensitivity in the Multiplicative AHP and SMART. *Journal of Multi-Criteria Decision Analysis, 2*, 87–110.

Luce, M. F., Payne, J. W., & Bettman, J. R. (1999). Emotional Trade-off Difficulty and Choice. *Journal of Marketing Research, 36*(2), 143–159.

Maheswaran, D., Mackie, D. M., & Chaiken, S. (1992). Brand Name as a Heuristic Cue: The Effects of Task Importance and Expectancy Confirmation on Consumer Judgments. *Journal of Consumer Psychology, 1*(4), 317–336.

Manrai, A. K. (1995). Mathematical Models of Brand Choice Behaviour. *European Journal of Operational Research, 82*, 1–17.

Mattila, A. (1998). An Examination of Consumers' Use of Heuristic Cues in Making Satisfaction Judgments. *Psychology and Marketing, 15*(5), 477–501.

Mehta, N., Rajiv, S., & Srinivasan, K. (2003). Price Uncertainty and Consumer Search: A Structural Model of Consideration Set Formation. *Marketing Science, 22*(1), 58–84.

Miller, G. A. (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *The Psychological Review, 63*, 81–97.

Mitra, A. (1995). Advertising and the Stability of Consideration Sets Over Multiple Purchase Occasions. *International Journal of Research in Marketing, 12*(1), 81–94.

Murphy, C. K. (2000). Combining Belief Functions When Evidence Conflicts. *Decision Support Systems, 29*, 1–9.
Nowlis, S. M., & Simonson, I. (2000). Sales Promotions and the Choice Context as Competing Influences on Consumer Decision Making. *Journal of Consumer Psychology, 9*(1), 1–16.
Park, C., & Lessig, V. P. (1981). Familiarity and Its Impact on Consumer Decision Biases and Heuristics. *Journal of Consumer Research, 8*(12), 223–331.
Park, C. W., Jun, S. Y., & MacInnis, D. J. (2000). Choosing What I Want Versus Rejecting What I Do Not Want: An Application of Decision Framing to Product Option Choice Decisions. *Journal of Marketing Research, XXXVII*, 187–202.
Park, D., Kim, Y., Um, M. J., & Choi, S. U. (2015). Robust Priority for Strategic Environmental Assessment with Incomplete Information Using Multi-Criteria Decision Making Analysis. *Sustainability, 7*(8), 10233–10249.
Pires, T. (2016, September). Costly Search and Consideration Sets in Storable Goods Markets. *Quantitative Marketing and Economics, 14*(3), 157–193.
Prabhaker, P. R., & Sauer, P. (1994). Hierarchical Heuristics in Evaluation of Competitive Brands Based on Multiple Cues. *Psychology and Marketing, 11*(3), 217–235.
Punj, G., & Brookes, R. (2001). Decision Constraints and Consideration-Set Formation in Consumer Durables. *Psychology and Marketing, 18*(8), 843–863.
Punj, G., & Brookes, R. (2002). The Influence of Pre-decisional Constraints on Information Search and Consideration Set Formation in New Automobile Purchases. *International Journal of Research in Marketing, 19*, 383–400.
Racioppi, V., Marcarelli, G., & Squillante, M. (2015). Modelling a Sustainable Requalification Problem by Analytic Hierarchy Process. *Quality and Quantity, 49*(4), 1661–1677.
Raffone, A., & Wolters, G. (2001). A Cortical Mechanism for Binding in Visual Working Memory. *Journal of Cognitive Neuroscience, 13*(6), 766–785.
Roberts, J. H., & Lattin, J. M. (1997). Consideration: Review of Research and Prospects for Future Insights. *Journal of Marketing Research, XXXIV*, 406–410.
Roberts, J., & Nedungadi, P. (1995). Studying Consideration in the Consumer Decision Process: Progress and Challenges. *International Journal of Research in Marketing, 12*, 3–7.
Saaty, T. L. (1977). A Scaling Method for Priorities in Hierarchical Structures. *Journal of Mathematical Psychology, 15*, 59–62.

Seiler, S. (2013, June). The Impact of Search Costs on Consumer Behavior: A Dynamic Approach. *Quantitative Marketing and Economics, 11*(2), 155–203.
Shafer, G. (1976). *A Mathematical Theory of Evidence*. Princeton: Princeton University Press.
Shafer, G. (1990). Perspectives in the Theory of Belief Functions. *International Journal of Approximate Reasoning, 4*, 323–362.
Shafir, E. (1993). Choosing Versus Rejecting: Why Some Options are Better and Worse Than Others. *Memory and Cognition, 21*(4), 546–556.
Shapiro, S., MacInnis, D. J., & Heckler, S. E. (1997). The Effects of Incidental ad Exposure on the Formation of Consideration Sets. *Journal of Consumer Research, 24*(1), 94–104.
Sharpanskykh, A., & Zia, K. (2012). Emotional Decision Making in Large Crowds. In Y. Demazeau, J. Müller, J. Rodríguez, & J. Pérez (Eds.), *Advances on Practical Applications of Agents and Multi-Agent Systems* (pp. 191–200). Heidelberg: Springer.
Simon, H. A. (1955). A Behavioral Model of Rational Choice. *Quarterly Journal of Economics, 69*(1), 99–118.
Smets, P. (1994). What Is Dempster-Shafer's Model? In R. R. Yager, M. Fedrizzi, & J. Kacprzyk (Eds.), *Advances in the Dempster-Shafer Theory of Evidence* (pp. 5–34). New York: Wiley.
Swait, J., & Adamowicz, W. (2001). The Influence of Task Complexity on Consumer Choice: A Latent Class Model of Decision Strategy Switching. *Journal of Consumer Research, 28*(1), 135–148.
Taroun, A., & Yang, J.-B. (2013). A DST-Based Approach for Construction Project Risk Analysis. *Journal of the Operational Research Society, 64*(8), 1221–1230.
Verwey, W. B. (2003). Effect of Sequence Length on the Execution of Familiar Keying Sequences: Lasting Segmentation and Preparation? *Journal of Motor Behavior, 35*(4), 343–354.
Verwey, W., Groen, E., & Wright, D. (2016). The Stuff that Motor Chunks Are Made of: Spatial Instead of Motor Representations? *Experimental Brain Research, 234*(2), 353–366.
Vroomen, B., van Nierop, E., & Franses, P. H. (2003). Modeling Consideration Sets and Brand Choice Using Artificial Neural Networks. *European Journal of Operational Research, 154*, 206–217.
Wang, J., Hu, Y., Xiao, F., Deng, X., & Deng, Y. (2016). A Novel Method to Use Fuzzy Soft Sets in Decision Making Based on Ambiguity Measure and

Dempster-Shafer Theory of Evidence: An Application in Medical Diagnosis. *Artificial Intelligence in Medicine, 69*, 1–11.
Wedel, M., & Pieters, R. (2000). Eye Fixations on Advertisements and Memory for Brands: A Model and Findings. *Marketing Science, 19*(4), 297–312.
Wright, D., Rhee, J., & Vaculin, A. (2010). Offline Improvement During Motor Sequence Learning Is Not Restricted to Developing Motor Chunks. *Journal of Motor Behavior, 42*(5), 317–324.

Index[1]

[1] Note: Page numbers followed by "n" refer to notes.

L. Moutinho, M. Sokele (eds.), *Innovative Research Methodologies in Management*,
https://doi.org/10.1007/978-3-319-64394-6

L

M

N