MULTICRITERIA DECISION MAKING:

Advances in MCDM Models, Algorithms, Theory, and Applications

INTERNATIONAL SERIES IN OPERATIONS RESEARCH & MANAGEMENT SCIENCE

Frederick S. Hillier, Series Editor
Stanford University

Saigal, R. / *LINEAR PROGRAMMING: A Modern Integrated Analysis*

Nagurney, A. & Zhang, D. / *PROJECTED DYNAMICAL SYSTEMS AND VARIATIONAL INEQUALITIES WITH APPLICATIONS*

Padberg, M. & Rijal, M. / *LOCATION, SCHEDULING, DESIGN AND INTEGER PROGRAMMING*

Vanderbei, R. / *LINEAR PROGRAMMING: Foundations and Extensions*

Jaiswal, N.K. / *MILITARY OPERATIONS RESEARCH: Quantitative Decision Making*

Gal, T. & Greenberg, H. / *ADVANCES IN SENSITIVITY ANALYSIS AND PARAMETRIC PROGRAMMING*

Prabhu, N.U. / *FOUNDATIONS OF QUEUEING THEORY*

Fang, S.-C., Rajasekera, J.R. & Tsao, H.-S.J. / *ENTROPY OPTIMIZATION AND MATHEMATICAL PROGRAMMING*

Yu, G. / *OPERATIONS RESEARCH IN THE AIRLINE INDUSTRY*

Ho, T.-H. & Tang, C. S. / *PRODUCT VARIETY MANAGEMENT*

El-Taha, M. & Stidham , S. / *SAMPLE-PATH ANALYSIS OF QUEUEING SYSTEMS*

Miettinen, K. M. / *NONLINEAR MULTIOBJECTIVE OPTIMIZATION*

Chao, H. & Huntington, H. G. / *DESIGNING COMPETITIVE ELECTRICITY MARKETS*

Weglarz, J. / *PROJECT SCHEDULING: Recent Models, Algorithms & Applications*

Sahin, I. & Polatoglu, H. / *QUALITY, WARRANTY AND PREVENTIVE MAINTENANCE*

Tavares, L. V. / *ADVANCED MODELS FOR PROJECT MANAGEMENT*

Tayur, S., Ganeshan, R. & Magazine, M. / *QUANTITATIVE MODELING FOR SUPPLY CHAIN MANAGEMENT*

Weyant, J./ *ENERGY AND ENVIRONMENTAL POLICY MODELING*

Shanthikumar, J.G. & Sumita, U./*APPLIED PROBABILITY AND STOCHASTIC PROCESSES*

Liu, B. & Esogbue, A.O. / *DECISION CRITERIA AND OPTIMAL INVENTORY PROCESSES*

MULTICRITERIA DECISION MAKING:

Advances in MCDM Models, Algorithms, Theory, and Applications

edited by

Tomas Gal
FernUniversität, Hagen, Germany

Theodor J. Stewart
University of Cape Town, Rep. of South Africa

Thomas Hanne
FernUniversität Hagen, Germany

Springer Science+Business Media, LLC

Electronic Services <http://www.wkap.nl>

Library of Congress Cataloging-in-Publication Data

Multicriteria decision making : advances in MCDM models, algorithms, theory, and applications / edited by Tomas Gal, Theodor J. Stewart, Thomas Hanne.

p. cm. -- (International series in operations research & management science ; 21)

Includes bibliographical references and index.

ISBN 978-1-4613-7283-7 ISBN 978-1-4615-5025-9 (eBook)
DOI 10.1007/978-1-4615-5025-9

1. Multiple criteria decision making. I. Gál, Tomáš. II. Stewart, Theodor J. III. Hanne, Thomas. IV. Series.
T57.95.M843 1999
658.4'03--dc21 99-30336
CIP

CONTENTS

Foreword

by Stanley Zionts

Advances in Multiple Criteria Decision Making - What does this mean? We all know that multiple criteria decision-making (MCDM) has to do with making decisions involving more than one criterion. Virtually all decisions involve more than one criterion. Advances are improvements. So the title of this volume has to do with improving decisions involving more than one criteria. This is an appropriate title for this volume.

Multiple criteria decision making has been a rapidly growing field of management, engineering, and other areas. It has generated many articles, books, and conferences – a huge literature is associated with it. This volume includes contributions by many of the most well known people in the field. I first want to overview the volume, and then speculate on the directions the field should take. The authors of the articles in the volume are all very senior academics in their field -- the book should be required reading for anyone who plans to work in MCDM.

The volume begins with an overview by Bernard Roy, one of the founding fathers of MCDM. His objective is to explore what decision aiding is and should be, from the perspective of a practitioner - someone who uses our methods to make decisions. Johannes Jahn continues with a contribution that develops the theory of vector maximization and explores concepts of efficient solutions. Hirotaka Nakayama explores the role of duality in multi-objective optimization. Carlos Bana e Costa and Jean-Claude Vansnick explore various aspects of preference relations, and actively consider the idea of incomparability. Oleg Larichev considers normative and descriptive aspects of decision making. His is another overview. Thomas Hanne studies meta decision problems in MCDM. Sensitivity Analysis in MCDM is considered by Tetsuzo Tanino. Sang Lee and David Olson present a paper on goal programming, that includes an overview as well as applications of goal programming. Andrzej Wierzbicki explores reference point approaches, an important tool of MCDM that helps transform the desires of a decision-maker into an attainable decision. Theodor Stewart explores the concepts of interactive programming, and categorizes methods that use tradeoffs in one form or another. He also reports on simulation studies that evaluate the convergence of such methods.

Philippe Vincke presents an overview of outranking approaches, which are designed to generate partial orderings of alternatives for decision-makers. He considers roughly ten different methods. Valerie Belton limits her presentation to the certainty case, and the use of elicitation procedures to help people make decisions. Ami Arbel overviews the interior point methods of mathematical programming to help solve multiple criteria problems. Salvatore Greco, Benedetto Matarazzo, and Roman Slowinski explore the use of rough and fuzzy sets in multiple criteria decision making. The subject of artificial intelligence in MCDM is surveyed by Patrice and Jean-Charles Pomerol. Andrew Chipperfield, James Whidborne, and Peter Fleming explore evolutionary approaches and simulated annealing, and their applications to MCDM.

So there you have an overview of this impressive volume. The volume is a "must read" for anyone in the MCDM field, and will serve as a useful reference for students, faculty, and other researchers.

So where is our field and where is it going? What is now needed is a synthesis of the many approaches into easy-to-use methods that are used. We are generating more and more applications of MCDM models. MCDM has made inroads into the education of engineers and managers. But more must be done. We must promote the use of our wares by managers, and respond to the needs of managers with methods that do what they want. I still strongly believe that what is needed is a spreadsheet type of method that will allow ordinary people to use MCDM methods for ordinary decisions.

Enjoy the volume! Go forth and produce!

Stanley Zionts

PREFACE

From the time of our earliest human and pre-human ancestors many millions of years ago, decisions have had to be made and have involved multiple conflicting objectives. Do we commit our resources to hunting the high-risk high-return option of the mammoth, or do we rather hunt a few small antelope? At what point do we begin gathering food for winter, even when that means less for today? Evolution adapted our forbears to be able to take such decisions effectively by intuition. But as the world has become increasingly complex and rapid-changing, the wider the range of objectives to be optimized, and the more difficult the finding of "optimal" solutions has become. Even the word "optimal" must be placed in quotes, as the concept of optimality becomes increasingly ill-defined.

At a practical level, mathematical programming under multiple objectives has emerged as a powerful tool to assist in the process of searching for decisions which best satisfy a multitude of conflicting objectives. In order to justify use of this tool, however, it needs to be supported by a rigorous theoretical foundation regarding the notion of "optimality" when there are several conflicting objectives, and the properties of the methods applied to the finding of the solution. On the other hand, the resulting (generally computer based) multicriteria decision support systems or aids, must effectively be usable by decision makers, enabling them to gain understanding of the choices available, and to express value judgements, without their having to be skilled in these underlying theoretical foundations.

At this time, many distinct methodologies for multicriteria decision making problems exist, which can be categorized in a variety of ways, such as form of model (e.g. linear, non-linear, stochastic), characteristics of the decision space (e.g. finite or infinite), or solution process (e.g. prior specification of preferences or interactive). Scientists from a wide variety of disciplines (such as mathematics, economics and psychology) have contributed to the development of the field of Multicriteria Decision Making (MCDM) (or Multicriteria Decision Analysis (MCDA), Multiattribute Decision Making (MADM), Multiobjective Decision Making (MODM), etc.) over the past 30 years, helping to establish MCDM as an important part of management science, with its own specialist journals, special issues of journals, special interest groups and societies, and conferences. In this way, MCDM is becoming an obligatory component of studies in management science, economics and industrial engineering at many universities around the world.

With this background, the current volume has been conceived, aiming to bring together "state of the art" reviews and the most recent advances by leading experts on the fundamental theories, methodologies and applications of MCDM. This is aimed at graduate students and researchers in mathematics, economics, management and engineering, as well as at practicing management scientists who wish better to understand the principles of this new and fast developing field.

Stanley Zionts, in his Foreword, has already listed the chapters and authors in this volume. The editors have aimed at compiling a volume of long-lasting value, and wish to express their appreciation to the authors for their outstanding presentations and their cooperation in a project that has lasted about two years. The editors would also like to express their gratitude to Frederick S. Hillier for including this book in Kluwer's International Series in OR and MS, for which he is the editor. Last but not least the editors wish to thank Gary Folven, the editor at Kluwer Academic Publishers, and his staff, for coworking on this project.

Tomas Gal, Theodor Stewart, Thomas Hanne

ABOUT THE AUTHORS

AMI ARBEL, Dept. of Industrial Engineering, School of Engineering, Tel-Aviv University, 69978 Tel-Aviv, Israel. ami@eng.tau.ac.il.

Ami Arbel received his B.Sc. in Mechanical Engineering from the Technion in Israel, an M.Sc. degree in Mechanical Engineering from UC Berkeley, and an M.Sc. and Ph.D. degrees from the department of Engineering-Economic Systems at Stanford University. Upon completion of his studies, he worked for consulting companies in the San Francisco bay area performing basic and applied research in systems and decision analysis. In 1981 he moved to Israel were he joined the department of Industrial Engineering at the Tel-Aviv University school of Engineering where he is currently a professor of Industrial Engineering. Over the years he has held a number of visiting appointments which include the University of Arizona, the American University of Armenia, the Helsinki School of Economics and the AT&T Bell Laboratories. His research interests are in the areas of decision analysis, multicriteria decision making, linear programming, dynamic system analysis, and general system modeling and analysis.

PROFESSOR CARLOS ANTÓNIO BANA E COSTA, Economics and Management Department, "Instituto Superior Técnico" (IST), Technical University of Lisbon – IST, Av. Rovisco Pais, 1000, Lisbon, Portugal. cbana@alfa.ist.utl.pt

Carlos A. Bana e Costa received his Ph.D. in Systems Engineering from IST in 1993. His research areas of interest are Operational Research and Management Sciences, in particular Multicriteria Decision Analysis, Problem Structuring Methods and Strategic Management, and he is the co-author of the MACBETH Approach. Dr. Bana e Costa is also a decision-aiding consultant in many private and public organizations, and his main application areas are private and public decision-making, urban and regional planning, transportation, environmental management. Professor Bana e Costa is author of more than 80 theoretical and application articles, half of which published in international scientific journals and books, and he is the editor of the book *Readings in Multiple Criteria Decision Analysis*, Springer, Berlin, 1990. He is member of the International Executive Committee of the *International Society on Multiple Criteria Decision Making*, founding member of ESIGMA - *EURO Special Interest Group on Multicriteria Analysis* and its chairman from 1990 to 1996, member of the Editorial Board of the *Journal of Multi-Criteria Decision Aid*, member of the Permanent Committee of the *International Summer Schools on Multicriteria Decision Aid*, and member of the *Institute for Operations Research and the Management Science* and of the *EURO Working Group "Multicriteria Aid for Decisions"*.

DR VALERIE BELTON, Department of Management Science, University of Strathclyde - 40 George Street, Glasgow, G1 1QE, UK. Val@mansci.strath.ac.uk

Valerie Belton completed her PhD, "A comparative study of methods for multicriteria decision aid" from Cambridge University in 1986. Following on from this work her research has focused on encouraging the theoretically and practically well founded use of multiple criteria decision aiding (MCDA) methodologies to support decision making in practice. Two important themes of this work are the use of visual interactive modelling and integration, of MCDA with other management science methods and methodologies and across MCDA methods. Early research into the communication of MCDA analyses led to the development of V•I•S•A, a multiple criteria decision support system which was at the forefront of the use of Visual Interactive Modelling in MCDA. Dr Belton is very active in the UK and International OR communities; she is currently a Vice President of EURO, the European Federation of OR Societies, and President Elect of the International Society for MCDM.

DR ANDREW J. CHIPPERFIELD, Department of Automatic Control and Systems Engineering, University of Sheffield, Mappin Street, Sheffield S1 3JD, UK. A.Chipperfield@Sheffield.ac.uk, http://www.shef.ac.uk/~acse/

Andrew Chipperfield received a B.Sc degree in Computer Systems Engineering from the University of Wales (Bangor) in 1989 and a Ph.D in parallel processing in computer aided control system design in 1995 from the University of Sheffield. He is currently a Lecturer in the Department of Automatic Control and Systems Engineering at The University of Sheffield with research interests in evolutionary computation for control systems engineering, optimization in control system design, adaptive control and systems integration for aerospace applications. He has over 50 Journal and Conference publications in these fields. He also has a number of years experience of computer control in industrial organizations.

PROFESSOR PETER J. FLEMING, Department of Automatic Control and Systems Engineering, University of Sheffield, Mappin Street, Sheffield S1 3JD, UK. P.Fleming@sheffield.ac.uk, http:// www.shef.ac.uk/~acse/

Peter Fleming is Professor of Industrial Systems and Control in the Department of Automatic Control and Systems Engineering at the University of Sheffield. He is currently Head of Department and Director of the Rolls-Royce University Technology Centre for Control and Systems Engineering. He has previously held the positions of Professor of Computer Systems Engineering at the University of Wales (Bangor), Associate Professor at Syracuse University, New York and Research Scientist at NASA Langley, Virginia. His research interests include industrial applications of control, computer aided control system design and control applications of optimization and parallel processing. He has published extensively in these fields and is Chair of the IFAC Technical Committee on Algorithms and Architectures for Real-Time Control.

DR. DR. TOMAS GAL, Professor of Operations Research, Faculty of Economics, FernUniversitaet Hagen, PO Box 940 (Profil 6), D-58084 Hagen, Germany. tomas.gal@fernuni-hagen.de

Born 1926 in the former Czechoslovakia, Tomas Gal received his first Ph.D. in physical chemistry from the Charles University in Prague, 1953, his second Ph.D. in operations research from the Univerity of Agriculture in Prague, 1967, and his „Dr-Habil" from the Academy of Sciences in Prague, 1969. Until 1969 he worked at the Agricultural University, Department for OR, then one year at CORE, Louvain, Belgium, and during 1970 – 1977 at the Department of OR at the Technical University in Aachen. Since 1977, Tomas Gal has been Chair Professor of OR at the Economic Faculty of the FernUniversitaet in Hagen. He has been Professor-Emeritus since 1991, and he continues his research into degeneracy and its impact on various mathematical programming problems. He has published more than 100 papers in the area of sensitivity analysis, parametric programming, multicriteria decision making, redundancy, degeneracy, and related topics. He has been author, co-author or editor of 14 books, and he has organized streams or sessions in various international conferences. Tomas Gal is on the editorial board of of several international journals, and he was the guest editor for a special issue of Annals of OR on the special topic of degeneracy in optimization problems. For many years, Professor Tomas Gal has been invited as visiting professor at several universities, including, among others, Erasmus University (Rotterdam), University of Santa Catarina (Florianopolis), Technion (Haifa), and UNISA (Pretoria). He has been awarded by the Cantor Medal by the Society of MCDM.

PROFESSOR SALVATORE GRECO, Faculty of Economics, University of Catania, Corso Italia 55, 95129 Catania, Italy. salgreco@vm.unict.it

Salvatore Greco received his degree in economics from the Faculty of Economics of Catania University in 1988. He holds a Ph.D. in computer science from the Poznan University of Technology. He had been assistant researcher at Faculty of Economics of the University of Catania since 1994 and he has been associate professor since 1998. His main research interests are in the field of multicriteria decision aid and in the application of the rough set approach to decision analysis. He is a member of the A.M.A.S.E.S. (Italian Association for Mathematics Applied to Economic and Social Sciences), A.I.R.O. (Italian Association of Operational Research) and U.M.I. (Italian Mathematical Society).

THOMAS HANNE, Department of Economics, FernUniversität Hagen, Profilstr. 8, 58084 Hagen, Germany. thomas.hanne@fernuni-hagen.de

Thomas Hanne, born 1965 in Gelsenkirchen/Germany, received a masters degree in Economics from Ruhr-Universität Bochum in 1989 and a masters degree in Computer Science from the University Dortmund in 1991. From 1990 to 1991 he was working in the Deparment of Economics at the Ruhr-Universität Bochum. Since then he is occupied at the FernUniversity Hagen. His current research interests include

multicriteria decision analysis, neural networks, and evolutionary algorithms. He is author of about 15 journal and conference articles.

PROF. DR. JOHANNES JAHN, Angewandte Mathematik II, Universität Erlangen-Nürnberg, Martensstr. 3, D-91058 Erlangen. jahn@am.uni-erlangen.de http://www.am.uni-erlangen.de

Johannes Jahn studied Mathematics at the University of Saarland (Germany) and received the doctorate from the Technical University of Darmstadt in 1978. At the same university he habilitated in 1983 in Mathematics. In 1983 he was a visiting assistant professor at the North Carolina State University in Raleigh (U.S.A.) and since 1986 he is Universitätsprofessor of Applied Mathematics at the University of Erlangen-Nürnberg.

PROFESSOR OLEG I. LARICHEV, Institute for Systems Analysis, Russian Academy of Sciences, 9,pr. 60 let Octjabrja, Moscow, 117312, Russia, larichev@glas.apc.org

Oleg Larichev received his Candidat Degree in Optimal Control Theory in 1965 and his Doctor Degree in Decision Making in 1975 from Institute for Control and Management Problem (Moscow). His fields of interest are: Decision Making, Multicriteria Mathematical Programming, Nonlinear Programming, Artificial Intelligence. Dr. Larichev published 8 books (two books -in English) and 170 papers. In the field of Decision Making Dr. Oleg Larichev developed new methodology for the construction of decision methods and Decision Support Systems having psychological and mathematical foundation - Verbal Decision Analysis. In the field of Artificial Intelligence he proposed a new approach to the study and imitation of experts' knowledge. In 1991 Dr.Larichev was elected as Corresponding Member of Russian Academy of Sciences and in 1997- as the full Member of the Academy. Awarded a Gold Medal by the International Society on Multiple Criteria Decision Making.

PROFESSOR SANG M. LEE, Department of Management, College of Business Administration, University of Nebraska, Lincoln, NE, 68588-0400, USA. slee@unlinfo.unl.edu

Professor Sang M. Lee is currently the University Eminent Scholar, Regents Distinguished Professor, Chair of the Management Department, and Executive Director of the Nebraska Productivity and Entrepreneurship Center. He served as President of the Decision Sciences Institute and is currently serving as President of the Pan-Pacific Business Association. He is a Fellow of the Academy of Management, Decision Sciences Institute, and Pan-Pacific Business Association. He has organized 20 international conferences as the Program Chair. Dr. Lee has authored or co-authored 50 books and more than 170 journal articles. He has received over $4 million funding for his research and outreach activities. He is well known for his pioneering work in Goal Programming and Multiple Objective Decision Making. His current research interests, in addition to MODM, include development of world-class organizations, East European economic transition,

information technology for organizational transformation, and global entrepreneurship. He is on the editorial board of 24 journals and has been listed in more than 50 Who's Who publications including the Who's Who in America.

PROFESSOR BENEDETTO MATARAZZO, Faculty of Economics, University of Catania, Corso Italia 55, 95129 Catania, Italy. matarazz@vm.unict.it

Benedetto Matarazzo received his degree in economics from the Faculty of Economics of Catania University in 1969. He achieved several post-degree qualifications in different universities. He had been assistant, charged professor of Financial Mathematics and associate professor of General Mathematics at Faculty of Economics of Catania University. Currently he is full professor of Financial Mathematics and charged professor of General Mathematics at Faculty of Economics of Catania University, where since 1986 he has been heading of the Mathematical and Statistical Section. In 1990-1996 he has been member of the scientific committee of A.M.A.S.E.S. (Italian Association for Mathematics Applied to Economic and Social Sciences) and in 1990-1992 he has been member of the scientific committee of A.I.R.O. (Italian Association of Operational Research). He is a member of the A.M.A.S.E.S. (Italian Association for Mathematics Applied to Economic and Social Sciences), A.I.R.O. (Italian Association of Operational Research), U.M.I. (Italian Mathematical Society) and I.I.A. (Italian Association of Actuaries). He is organiser and member of the Programme Committee of a lot of domestic and international scientific conferences and member of the Organising Committee of the International Summer School on MCDA. He has been invited as a speaker in a lot of international scientific conferences and he is member of the editorial boards of the *Foundations of Computing and Decision Sciences, J. of Multi-Criteria Decision Analysis, Ricerca Operativa.* He has been member of the Program Committee of EURO XIV (European Conference of Operations Research) (Jerusalem, 1995), EURO XV (Barcelona, 1997) and chairman of the Programme Committee of EURO XVI (Brussels, 1998). His main research interests are in the fields of multicriteria decision aid, quantitative finance, rough set analysis, meta-analysis. He has been invited professor at, and has different scientific co-operation with, several European universities.

PROFESSOR HIROTAKA NAKAYAMA, Department of Applied Mathematics --Konan University, 8-9-1 Okamoto, Higashinada, Kobe 658-8501, Japan. nakayama@edu2.math.konan-u.ac.jp

Hirotaka Nakayama received his Ph.D. in Applied Mathematics and Physics from Kyoto University in 1975. Since then, he has extended his areas of interest to multi-objective decision support and machine learning, in particular, their applications to real problems. He won the Prize of Japanese Society of Operations Research (Application) in 1991 and 1993, and received the Edgeworth-Pareto Award from the society on Multiple Criteria Decision Making in 1995. With about 100 research publications and 6 books (co-authored) he continues his research to develop decision support systems in engineering design problems, resource management in agriculture and financial investment problems.

PROFESSOR DAVID L. OLSON, Department of Business Analysis & Research, Texas A&M University, College Station, TX 77843-4217, USA. dolson@tamvm1.tamu.edu

David L. Olson is professor of management science in the Department of Business Analysis & Research at Texas A&M University. He received his Ph.D. in Business from the University of Nebraska in 1981. He has been at Texas A&M University since then. He has published research in over 45 refereed journals, primarily on the topic of multiple objective decision making. He teaches in the management science, management information systems, and operations management areas. He has authored the book *Decision Aids for Selection Problems*, and coauthored the books *Decision Support Models and Expert Systems, Introduction to Management Science, Introduction to Simulation*, and *Business Statistics*. He has made over 80 presentations at international and national conferences on research topics. He is a member of the Decision Sciences Institute, the Institute for Operations Research and Management Sciences, and the Multiple Criteria Decision Making Society. He has coordinated the Decision Sciences Institute Dissertation Competition, Innovative Education Competition, chaired the Doctoral Affairs Committee, served twice as nationally elected vice president, and as National Program Chair. He has received a Research Fellow Award from the College of Business and Graduate School of Business at Texas A&M University, and held the Business Analysis Faculty Excellence Fellowship for two years.

PATRICE PERNY, LIP6, Université Pierre et Marie Curie - Case 169, 4 Place Jussieu, 75252 Paris Cedex 05, France. Patrice.Perny@lip6.fr

Patrice Perny is 35 years old. He is associate professor at the Paris 6 University and member of the Computer Sciences Laboratory of papers in international OR/AI conferences and journals (e.g. Fuzzy Sets and Systems, EJOR, Annals of OR, JMCDA the University. He works on Decision Theory and Preference modeling, including both theoretical researches and practical applications. His favorite domains of interest are Multicriteria Decision Making, Group Decision Making (Social Choice Theory), Decision under uncertainty, and Fuzzy sets Theory. He also investigates the interest of AI techniques for preference modeling and decision support. Patrice Perny is the author of many) He also contributed to several chapters in collective books on Fuzzy Preference Modeling and Multicriteria Decision Making.

JEAN CHARLES POMEROL, LIP6, Université Pierre et Marie Curie - Case 169, 4 Place Jussieu, 75252 Paris Cedex 05. France. Jean-Charles.Pomerol@lip6.fr

J.-Ch. Pomerol is 55 years old. After a "doctorat d'état" at the University of Paris in convex analysis, he turns to decision analysis, AI and "Intelligent" Decision Support Systems (IDSS). J.-Ch. Pomerol headed for six years the AI Lab of the University of Paris 6 (Pierre and Marie Curie) and he is now with the National Center for Scientific Research (CNRS). Now, his main fields of interest are human decision making,

organization modeling and computer aided decision. He is the author of many papers in various scientific journals and the co-author of three books about Experts Systems, IDSS and multicriteria decision making (this latter being translated in Spanish, Roman, and English (in press)). J.-Ch. Pomerol is also the editor in chief of the journal of Decision Systems.

PROFESSOR BERNARD ROY, LAMSADE, Université Paris-Dauphine, Place du Maréchal De Lattre de Tassigny, 75775 Paris Cedex 16, France. roy@lamsade.dauphine.fr

Bernard Roy is Docteur ès-Sciences Mathematics (Université de Paris). He is Professor of applied mathematics and management sciences at the Université Paris-Dauphine, Director of LAMSADE (associated team to CNRS) and Conseiller Scientifique at RATP. He is Chairman of the European Working Group "Aid for Multicriteria Decisions". He has been President of EURO (1985-1986). Among others, he has published "Algèbre moderne et théorie des graphes orientées vers les sciences économiques et sociales" (Dunod, Paris, Volume 1: 1969, Volume 2: 1970), "Méthodologie multicritère d'aide à la décision" (Economica, Paris, 1985), Aide multicritère à la décision : Méthodes et cas (Economica, Paris, 1993) (in collaboration with Denis Bouyssou), "Multicriteria methodology for decision aiding" (Nonconvex Optimization and its Applications, Volume 12, Kluwer Academic Publishers, Dordrecht, 1996). He is Doctor Honoris Cause of Brussels, Liège, Freiburg, Poznan and Laval Universities, Laureate of the EURO Gold Medal (1992) and of the "MCDM Gold Medal" (1995).

PROFESSOR ROMAN SLOWINSKI, Institute of Computing Science, Poznan University of Technology, ul. Piotrowo 3a, 60-965 Poznan, Poland. slowinsk@sol.put.poznan.pl

Roman Slowinski received the M.S. degree in control engineering, Ph.D. degree in operations research and Habilitation degree in computer science from the Poznan University of Technology (PUT), in 1974, 1977 and 1981, respectively. Currently, he is professor of operations research and decision science in the Institute of Computing Science of the PUT. Within this institute, he is heading Laboratory of Intelligent Decision Support Systems. He has been professor on European Chair at the University of Paris Dauphine and invited professor at the Swiss Federal Institute of Technology in Lausanne and at the University of Catania. He is Editor of *European J. of Operational Research, Foundations of Computing and Decision Sciences* and area editor of *Fuzzy Sets and Systems* for decision analysis. He is also member of editorial boards of the *J. of Multi-Criteria Decision Analysis*, *Belgian J. of Operations Research, Statistics and Computer Sci.*, *J. of Decision Systems*, *J. of Applied Mathematics and Computer Sci.*, and *J. of Computing and Information*. Roman Slowinski has done extensive research on the methodology and techniques of decision support, including multicriteria decision analysis, preference modeling, managing uncertainty and ambiguity in decision support systems, fuzzy and stochastic optimization, project scheduling, water supply system programming and medical decision support. He has published 10 monographs, 3 textbooks and more than 200 papers. In recent years, his research interest has been focused on knowledge and data

engineering using rough sets theory. Roman Slowinski is member of the American Mathematical Society, INFORMS, Polish Information Processing Society, and vice-president of the Association of Polish Operational Research Societies. He has been Program Committee chair of the 1995 European Conference of Operations Research in Jerusalem. He is laureate of the State Award for scientific achievements (Warsaw, 1988), Sixth EURO Gold Medal, by European Association of OR Societies (Aachen, 1991) and of the Edgeworth-Pareto Award, by International Society on Multiple Criteria Decision Making (Cape Town, 1997).

PROFESSOR THEODOR J STEWART, Department of Statistical Sciences, University of Cape Town, Rondebosch 7701, South Africa. tjstew@maths.uct.ac.za

Theo Stewart did his undergraduate training in chemical engineering at the University of Cape Town, and completed his MSc and PhD degrees in operations research and statistics at the University of South Africa, while working at the South African Council for Scientific and Industrial Research (CSIR). He was head of the division of operations research and statistics at the CSIR in Pretoria for 8 years before taking up his current post at the University of Cape Town in 1984. He has been president of both the Operations Research Society of South Africa and the South African Statistical Association. He is on the editorial boards of four international journals, and is chair of the European Special Interest Group on Multicriteria Analysis (ESIGMA). He is the author of about 50 refereed research publications, and currently devotes much of his efforts towards the application of MCDM in natural resource management problems.

PROFESSOR TETSUZO TANINO, Dept. of Electronics & Information Systems, Graduate School of Engineering, Osaka University, Suita, Osaka 565, Japan. tanino@pwr.eng.osaka-u.ac,jp

Tetsuzo Tanino received his B.E., M.E. and Ph. D. degrees in applied mathematics and physics from Kyoto University in 1973, 1975 and 1978, respectively. From 1979 to 1988 he was an associate professor of Tohoku
University. He was a professor of Okayama University from 1989 to 1996. Since July, 1996 he has been a professor at the Department of Electronics and Information Systems in Osaka Unviersity. His principal research interests are in optimization theory, game theory, decision making and soft computing.

PROFESSOR JEAN-CLAUDE VANSNICK, "Quantitative Methods" Department, Warocqué Faculty of Economics, University of Mons-Hainaut, Place du Parc, 20, 7000 Mons, Belgium. Jean-Claude.Vansnick@umh.ac.be

Jean-Claude Vansnick received, in 1973, his Ph.D. in Mathematics from the Free University of Brussels and, in 1974, the Royal Academy of Belgium award (Section "Mathematical Sciences"). Since then, he has extended his areas of interest to Measurement Theory and Decision Aid. His paper "Strength of Preference. Theoretical and Practical Aspects" was selected as National Contribution of Belgium

for the Tenth Triennal Conference of the International Federation of Operational Research Societies (IFORS'84). He is member of several groups on Multicriteria Analysis and was guest professor at the International Summer Schools on Multicriteria Decision Aid in Portugal (1988), Canada (1991) and Greece (1994). He is co-author of the MACBETH Approach which is now currently used in many real world applications of decision aid.

PROFESSOR PHILIPPE VINCKE, Institute of Statistics and Operations Research and Faculty of Applied Sciences - Université Libre de Bruxelles, Service de Mathématiques de la Gestion, CP 210/01, Boulevard du Triomphe, 1050 BRUSSELS, BELGIUM. pvincke@smg.ulb.ac.be - http://smg.ulb.ac.be/personal/Pvincke.html

Philippe Vincke received his Ph.D. in Mathematics from the Free University of Brussels in 1976. He became professor in this University and is today Head of the Institute of Statistics and Operations Research and Vice-Rector for Research. He published 3 books and about 80 papers in the field of preference modelling and multicriteria decision-aid. He is director of the Service de Mathématiques de la Gestion, where about 30 researchers work on theoretical foundations of Operations Research and Decision-Aid and collaborate with companies on practical applications of these fields.

DR JAMES F WHIDBORNE, Department of Mechanical Engineering, King's College London, Strand, London WC2R 2LS UK. james.whidborne@kcl.ac.uk, http://www.eee.kcl.ac.uk/ mecheng/jfw/jfw.html

James Whidborne was born in Zimbabwe in 1960. He received a B.A. degree in Engineering from Cambridge University, England in 1982, and an M.Sc. and a Ph.D. in Systems and Control from UMIST, Manchester, England in 1987 and 1992 respectively. From 1982 to 1985 he worked in industry, and from 1986 to 1991, he was with the Control Systems Centre, UMIST. From 1991 to 1994, he held a position of Research Associate with the Department of Engineering, University of Leicester, England. He joined the Department of Mechanical Engineering, King's College London, as a Lecturer in 1994. He is the co-author of a book, ``Critical Control Systems: Theory, Design and Applications", and is the author or co-author of some 40 journal and conference publications. His current research interests include multi-objective and computer aided control system design, robust control systems, L1/l1 optimisation, automotive control applications and finite word length controller implementation.

PROFESSOR ANDRZEJ P. WIERZBICKI, Institute of Telecommunications, Szachowa 1, 04-894 Warsaw, Poland. a.wierzbicki@itl.waw.pl

Andrzej Piotr Wierzbicki received his Ph.D. degree in automatic control (1964) and D.Sc. degree for his work in computational techniques of optimization (1968) from the Warsaw University of Technology. Specialized in nonlinear dynamic systems,

sensitivity of dynamic and control systems, dynamic and nonlinear optimization, vector optimization, decision support theory and applications. Author of over 150 publications, including 10 books and several industrially implemented patents. Originated reference point approaches to vector optimization and multicriteria model analysis, for which he obtained the Georg Cantor Award of the International Society of Multiple Criteria Decision Making (1992). Served as the dean of the Faculty of Electronics at the Warsaw University of Technology (1975-78), as the chairman of Systems and Decision Sciences Program of the International Institute of Applied Systems Analysis in Laxenburg n. Vienna (1979-84), as the scientific secretary of Future Studies Committee ``Poland 2000" of the Polish Academy of Sciences (1986-91), as the chairman of the Commission of Applied Sciences of the State Committee for Scientific Research of Poland (1991-94). Visited and lectured at many leading universities in the world. Currently serves as the director of national Institute of Telecommunications in Poland, with interests in the developments and forecasts of information society.

1 DECISION-AIDING TODAY: WHAT SHOULD WE EXPECT?

Bernard Roy

CONTENTS

Abstract: The object of this chapter is to present an overall view of what Decision-Aiding (DA) is today, or what it seeks to be. The standpoint adopted here is more that of the practitioner than that of the theoretician. Above all, we shall attempt to emphasize what a scientific approach might claim to contribute to illuminating decision-making and how this approach might facilitate the proper functioning of the decision-making process as a whole. It is possible, in one way, to consider that decision-aiding was the natural outgrowth of operational research. This discipline, as it was conceived of and implemen-

T. Gal et al. (eds.), *Multicriteria Decision Making*

ted particularly during the 1960s, proved to have a relatively limited field of application. The current conception of DA aims to free itself from these limitations, as will be explained in the first two paragraphs of Section 1.1. This allows us to understand why DA is most often Multicriteria Decision Aiding (MCDA). In the final paragraph of this Section we shall highlight, on these bases, what it is reasonable to expect from DA in actual practice. Section 1.2 presents the conceptual architecture which constitutes the foundation of DA (and, more precisely, MCDA). A glossary (see Appendix) provides definitions for the principle terms used, and furnishes some additional methodological information. The final section is devoted to what the practitioner should expect from using computerized procedures and tools to develop recommendations and/or point the way towards a decision. In order to allow the reader to appreciate the scope and variety of existing applications, we have included in the bibliography a very limited selection of applications which in no way claims to be scientifically representative. This sample of related work points out the diverse number of both countries and sectors of activity which have shown an interest in the methods explored in this book.

1.1 FROM OPERATIONAL RESEARCH TO DECISION-AIDING

... And if optimization consists of integrating disorders, uncertainties, unknown factors, challenges and antagonisms, then such optimization comprises unoptimizable; from that moment on, should we not revise, reformulate and open up our notion of optimization? Should we not understand that true optimization is always complex, risky, composed of disorders and conflicts and that its enemy is pseudo-rationalism which claims to drive away conflict, disorder and risk? Any ideal conception of an organization which would be only order, harmony and coherence is the dream of a demented ideologue and/or technocrat (Morin, [46] translated from the original French text).

In a way which some might deem a bit provocative, the above text underlines some of the limitations of any approach – presented as scientific – to questions relative to decision-aiding. Works undertaken over the past fifty years have profoundly changed our conception of what we can expect from such an approach and have, in particular, sensitized us to the necessity of getting out of two ruts, preconceptions inherited from the so-called hard sciences. It is only after clarifying this double necessity that we shall seek, in Subsections 1.1.1 and 1.1.2, to delimit what the user should be able to expect from decision-aiding today.

1.1.1 Getting away from an overly descriptive vision

Discovering or, failing that, getting close to the optimal decision through a scientific approach supported by models describing an objective reality, this was the ambition of Operational Research (OR) during the two decades immediately following World War II. Yet decision-making situations verifying the necessary conditions for this ambition to make sense proved to be far less frequent than pioneers of OR had believed them to be. There were a number of reasons for this, but primarily the two given below:

1°) It is rare for the decision to be the instantaneous act of a single readily identified, rational and omnipotent decision-maker. To the contrary, it is much more likely that arriving at a decision forms a part of an ongoing process, for which a given decision provides a stopping point [1]; often, this process is not very rational. It draws in multiple actors who play out their parts within a complex context, with more or less conflicting stakes.

2°) It is very unusual for preference, which must be described in order to give meaning and legitimacy to the optimum, to pre-exist in an objective way with the requisite properties; for this situation to obtain, we must, in particular, be able to describe *a priori* all possibilities of action and to get a single criterion accepted as an objective criterion, with which all those possibilities of actions can be evaluated and ranked from best to worst.

Observing these aspects of decision-making reality, decision-aiding (DA) seeks to put science in the service of shedding light on managerial decisions and/or guiding complex decision-making processes within organized systems, but DA places limits on its ambitions. It aims, above all, through a constructive, and not simply a descriptive, approach:

– to provide elements of a response to certain questions which an actor involved in a decision-making process will consider, particularly within the framework of working hypotheses which allow actors to take account of an imperfect description of past, present and future reality; when they are mobilized in this way, the models and algorithms of operational research are extremely useful;

– to provide means for giving greater coherence between the final decision reached and the objectives and/or systems of values which are those of actors involved in a decision-making process.

This is to say that even if the DA approach remains essentially scientific (with formalized models, hypothetico-deductive reasonings or inference, optimization computations, ...) the point is no longer – or only very infrequently – to discover or even to approach an ideal decision whose optimality should be obvious to any intelligent actor participating in good faith. As we shall point out, all reference to this kind of ideal may disappear. As a corollary, because DA helps to construct, and not simply to describe, it should give pride of place to a dynamic approach facilitating easy insertion of DA practitioners into the decision-making process. In some cases, DA can thus contribute to legitimating the final decision.

---

[1] This stopping point may be a temporary one and may be concerned with only a fragment of the overall decision; thus it is generally a point of departure for another process linked either to implementing the decision arrived at, or, due to the partial and/or temporary nature of this decision, linked to another phase of the same decision-making process.

1.1.2 Getting away from arithmo-morphism

With this term, borrowed from Alain Schärlig, we designate the turn of mind which consists of using arithmetic to keep account of heterogeneous factors and phenomena on a single scale on a common unit (Dollar for instance); arithmo-morphism underlies the belief which holds that in any decision-making context, there should be an "optimal" decision, i.e., a decision which is better or at least as good as any other.

In many cases arithmo-morphism is extremely reductive:

– it can lead to neglecting, wrongly, certain aspects of reality;

– it facilitates setting up equivalencies, the fictitious nature of which remains invisible;

– it tends to present features of one particular system of values as objective.

These are the unwanted effects against which we must arm ourselves. This is especially true since their deceitful nature allows them to pass unobserved, as the examples given below will demonstrate.

Arithmo-morphism encourages actors in a decision-making process to attempt to place a monetary value on effects and attributes of a purely qualitative nature. In order to do so, we may resort to various procedures which are often very artificial. For example, we construct a behavioral model incorporating one or more interpretable parameters such as a price, for which we seek values such as to reproduce as accurately as possible certain types of behavior observed in the field: these values are referred to as *revealed prices*. We may also seek to provide an explanation for responses given in surveys in which participants are asked to compare fictitious situations, in which case we speak of *state preferences*. We can, likewise, place individuals in the context of an artificial market; here we refer to *contingent evaluations*. Within a more theoretical context, the monetarisation of external effects has long been a subject of concern to economists dealing with Neo-classical theory. As interesting as all these attempts may be (see especially Grégory *et al.* [27]; Johansson [31]; Le Pen [38]; Mishan [43]; Perez [51]; Ray [57]; Sugden and Williams [76]), it would appear that in many cases they try to give explanations of very diverse value systems through a single, or even several, unitary prices interpretable as an average value within the scope of a rather unrealistic model, because founded on more or less falsified hypotheses. The idea of approximation is often given prominence, at the risk of losing sight of the fact that what is approximated remains subjective and extremely ill defined.

In France, decisions concerning the choice of infrastructures to be developed for public transportation systems gives high priority to the internal rate of return. Although the definition of this rate is very finely honed to take into account very complex real-like situations, it can be criticized from several standpoints; it fails to deal with certain consequences of the final decision which deserve to be taken into account (see especially STP [74] [2]). It was for this reason that a working group composed of government representatives, transportation firms and some research

[2] Official report published by Syndicat des Transports Parisiens, 11 avenue de Villars, 75007 Paris, France.

institutes [3] acted to substitute this single criterion in favour of a family of multiple criteria reflecting eight different points of view (STP [75]). Henceforth, no investment project should be characterised by a single figure (rate of return) but by a series of evaluations designed to bring out separately:

– the way in which the project contributes to reaching certain overall objectives;

– the impact it is thought to have on travel, the environment, employment, public finance and the financial equilibrium of public transportation firms;

– its economic and social profitability;

– factors of resistance likely to interrupt completion of the project.

The preceding considerations show (as do many of the examples of applications cited in the bibliography) that it is often appropriate to avoid the dangers of arithmo-morphism by delimiting a broad spectrum of viewpoints likely to structure the decision-making process with regard to the actors involved. A criterion family can then be constructed which is capable of dealing with these structure-providing points of view and of allowing decision-makers to debate the respective roles (weight, veto, ...) that each criterion might be called upon to play during the decision-making process.

Thus, in order to escape arithmo-morphism, decision-aiding (DA) becomes in many cases, multicriteria decision-aiding (MCDA). This means that it is supported by multiple scales, which, in general, cannot, in any objective way, be reduced to, or converted in a single one. In such conditions, rather than dismissing or concealing subjectivity, it is important to make an objective place for it which will be compatible with a plurality of expression.

1.1.3 Decision-aiding: to what end?

It evolves from what we have said thus far that the goal of DA is not to set forth objective truths. Rather more modestly, DA aims at establishing, on recognized scientific bases, with reference to working hypotheses, formulations of propositions (elements of responses to questions, a presentation of satisfying solutions or possible compromises, ...) which are then submitted to the judgment of a decision-maker and/or the various actors involved in the decision-making process. In order to accomplish this goal, DA draws its support from models (see Appendix). These models are not necessarily (more of less simplified) descriptions of hard-liner reality. We could imagine, for example, the preferences that one of the actors might have in mind relative to numerous potential actions with very complex ensuing consequences. These preferences might evolve under the influence of decision-aiding or under the influence of other actors. Which is to say that the role of decision-aiding is not to discover hidden truths, but rather to contribute to constructing [4] individual

---

[3] The author of the present chapter served as a member of this commission, representing the LAMSADE Research Center.

[4] For more details on this constructivist aspect of DA, see Roy [59, 61].

convictions, collective decisions and compromises between multiple, and often conflicting, rationalities, stakes and values. To situate DA's contribution correctly, it is important to remember (as underlined by Thépot [78]) that "The modeling procedure presupposes nothing concerning the rationality of the individual, whose capacity of investigation and observation are accepted for what they are, including their limitations and imperfections. On the other hand, it postulates the full and entire rationality of the solicited third party" in decision-aiding activity. Thus we see the nature and scope of the role played in this activity by hypothetico-deductive reasoning and, in particular, optimization computations.

Due to its rigourous conceptual architecture, decision-aiding (see Glossary in Appendix), decision-aiding can contribute to structuring and organizing the decision-making process. It may follow that this aiding will both respond to questions and help to choose or construct solutions.

In this respect, it is important to underline that DA can, in certain cases, result in reorienting the analysis of solutions envisaged early in the decision-making process towards a deeper understanding of the problem, and in provoking a debate focusing on other types of questions, namely: who are the actors involved and what are the stakes, the points of view and the axes of meaning for the criteria used?; where are the constraints, the possibilities?; how can we anticipate the effects of a given action under consideration?; what causal links allow us to evaluate an action's consequences? how do we weight up the pros and cons of their advantages and disadvantages?, ... The original formulation of a problem can thus be significantly modified as a result of the framework for "concertation" [5] provided by DA.

DA may facilitate concertation between stakeholders and contribute to legitimating the decision. Although this legitimization could, of course, be obtained in other ways, decision-aiding presents the advantage of being open-ended. The sources of legitimization can, in a certain way, be considered to come from different rationalities, each one characterized by its own system of values and internal logic. DA cannot claim to unify or synthesize these systems of values, logical approaches to dealing with information, rationalities or the foundations of legitimacy when two or more clash within the same decision-making process. Nevertheless, in a certain number of cases, DA should allow participants to structure debate and facilitate concertation, especially by helping to establish a climate of confidence and by providing a common understanding of the problem. Discussion could, for example, be organized around the aspiration and rejection levels on each of the scales, the relative weights attributed to each criterion, the possibilities of compensation between poor and excellent performance levels, veto thresholds, etc. The comparison of results stemming from these different options may, of course, bring to light deep, unresolvable disagreements. It may also facilitate partial agreements, for example

[5] "Concertation" is a French word which has no equivalent in English. In a decision aiding situation, this word covers a conception of dialogue between actors and/or group reflection so as to progress towards a good and common understanding of positions of each actor and/or of the questions under examination. The objective of these dialogue and/or reflection is to make a decision or/and to choose a way to formulate and solve a specific problem, as far as possible in a consensus way.

concerning either the elimination of certain potential actions (see Bana e Costa [5]) or the relative position of some of them (see Bollinger *et al.* [11]). In such cases, decision-aiding may constitute a source of legitimation for recommendations and/or conclusions arrived at during the decision-making process. Indeed, whatever the degree of scientific character the different parties may attribute to a given source of legitimation, its pertinence to the problem may be recognized by many of the stakeholders. Thus between the two extreme positions, legitimacy based on relative positions of power (in which context decision-aiding plays no role), and legitimacy sought above all on bases we call scientific, grounded in a single system of values (the special sphere of activity for traditional DA), there is room for more complex situations in which different rationalities are taken into account and in which DA should be able to play a significant role.

Research undertaken over the past thirty years has produced a number of procedures well suited to a variety of decision-making contexts. Easy-to-use computerized tools [6] facilitate using these procedures. Making use of one or several working hypotheses, these procedures allow us to reply to questions and elaborate recommendations. In real time, they may also serve to guide the unfolding of a decision-making process. As correctly emphasized by David [18], decision-aiding should not be seen as aid intervening only at the very moment a choice is made. He concludes the first part of his article with the following words: "In what follows, we must consider decision-aiding from a functional perspective in order to understand what kind of aid actors in a decision-making process need, as well as from a critical standpoint to take into account the fact that any DA procedure involves taking steps which entail the irruption of a new actor into the decision-making process" (Translated from the original French text). Although the angle from which David looks at decision-aiding differs significantly from the one adopted in this chapter, the division of what follows here into two sections is not unrelated to David's dual perspective.

1.2 RIGOROUS CONCEPTS FOR ANALYSING AND COMMUNICATING

As far as science is concerned, inaptitude in the selection of the appropriate words chosen to restore a line of argumentation will most often lead to a complete misunderstanding of the latter (d'Espagnat [23], translated from the French text).

Someone who wants to be convincing should put his faith, not in the right argument, but in the right word, for words will always have more power than meaning (Conrad, translated from the French, following Flaubert and Voltaire).

The work involved in analysis, as well as the requirements of communication among actors of different backgrounds, is often hindered by confusions engendered by an insufficiently rigorous vocabulary in which the terms used do not necessarily have the same meaning for all participants. Enlarging the field of operational research, decision-aiding today offers a conceptual architecture (see the glossary),

[6] To avoid any confusion, the word tool will be used here only for computer-based methods.

which has developed gradually through contact with concrete problems and aims at responding, in a more or less limited way, to at least six large categories of concerns; the following paragraphs provide a rapid survey of this.

1.2.1 Identifying potential actions (or alternatives) which deserve consideration

Depending on the decision-making contexts, the object under examination can be of a widely varying nature. The concept of potential action and the somewhat more restricted one of alternative should help to make this more precise. In practice, however, these terms are replaced by others which are more concrete or more suitable, such as: scenario, plan, programme, project, proposal, variant, dossier, operation, investment, solution, ...

In decision-aiding, we endeavor to specify contents which are as clear and precise as possible for each of the potential actions (or alternatives) which deserve to be considered within a given phase of the decision-making process. Let us consider the case of a company S concerned with developing its research and development (R & D) program for the following year. It is possible to ask each of S's divisions to establish R & D proposals following a given model. Each proposal can be regarded, at this stage, as a potential action. It may be accepted as written in the next R & D program, or it may be rejected. It might be sent back for modifications.

The potential action should enable us to formalize the object of the decision; in other words, what deserves to be evaluated in order to shed light on the decision. There are two ways in which the potential action can be envisaged:

– a fragmented conception: the action is related to a fragment of a more comprehensive decision; this is the case in the preceding example in which each proposal received is a possible fragment of the research program (comprehensive decision); in this conception, two distinct potential actions can be conjointly implemented;

– a comprehensive conception: the action is concerned with the whole of the decision; in the preceding example, this conception would result in conceiving of each potential action as a complete, feasible program; in this conception, two distinct potential actions are mutually exclusive; under these conditions, the potential actions become alternatives.

At a given stage in the decision-making process (in other words, in a defined phase of study), it is fitting to reach an agreement on exactly what the concept of action covers, as defined within DA. We must then delimit the set of all actions worthy of interest for this phase of study. This is the set A of potential actions. It may be represented by a list, a series of maps, a stack of dossiers, ... If the potential action lends itself to a coding process, using a certain number of variables, the set A can then be represented (this is the case in mathematical programming) as the set of solutions which verify a family of constraints relating to these variables. Whatever the definition, we should emphasize that A will frequently be endowed with an evolutionary character as it will espouse the dynamics of the process in which decision-aiding is brought to bear.

1.2.2 Itemizing and structuring consequences relevant to decision evaluation

Any factor, aspect or attribute attached to potential actions (price, quality, time limits, safety, image, market share, ...) and likely to influence the final decision in one direction or another is designated by the general term consequence. Before each individual action can be specifically assessed and evaluated with respect to these consequences, it is essential to delimit and structure all of them.

To accomplish this, one of two following approaches (which are not mutually exclusive) is customarily used.

The top-down approach consists of deriving the consequences from one or several general objectives according to increasingly specific and refined points of view, dimensions and terms. A tree diagram of that type presented in Fig. 1.1 results from this process.

The bottom-up approach begins by gathering information from the diverse stakeholders about the concrete, precise elements which, in their eyes, should inform and orient the decision-making process. The material thus gathered will enable us to establish a first list of "elementary consequences". We then seek to cluster progressively the elementary consequences selected according to increasingly more comprehensive families of points of view deemed to be significant by the stakeholders. This approach, which finds support in certain concepts which we cannot present here [7], may also produce a tree diagram.

Whatever the approach used, the evaluation of actions on the basis of the consequences it has brought to light relies on, according to more or less explicit schemas, the supposed existence of causal links, usually complex and difficult to determine. Toulemonde [79] asks the following question: "Should we free evaluation from its causal links ?". Which, we observe, leads him (see [79], page 87) to underline that: evaluation "reveals factors which influence efficiency" but that evaluation "does not demonstrate its efficiency"; evaluation brings together different points of view on the theory of action and helps to sketch out an acceptable schema of causality (but only occasionally does it compare this schema with the facts). From this he concludes that "in most cases, what evaluation can contribute should be the construction of a schema of causality (a theory of action) accepted by all parties involved. Chains of causality whose scientific robustness is deficient will, nonetheless, be able to resist political and social pressure inasmuch as they have been forged by common consent. When causal links cannot be proven by either observation or analysis, an attempt should be made to represent them as accurately as possible and to encourage a collective understanding of them... The community of evaluators can usefully integrate this "evaluation-theorising" into its professional standards by giving it a place at least as eminent as the one it occupies, for example, in cost-efficiency analysis. In fact, the idea seems to be well on its way to acceptance and has encountered favourable reaction within professional community... Schemas of

[7] For more details, consult Roy [62] (chapters 8 and 9).

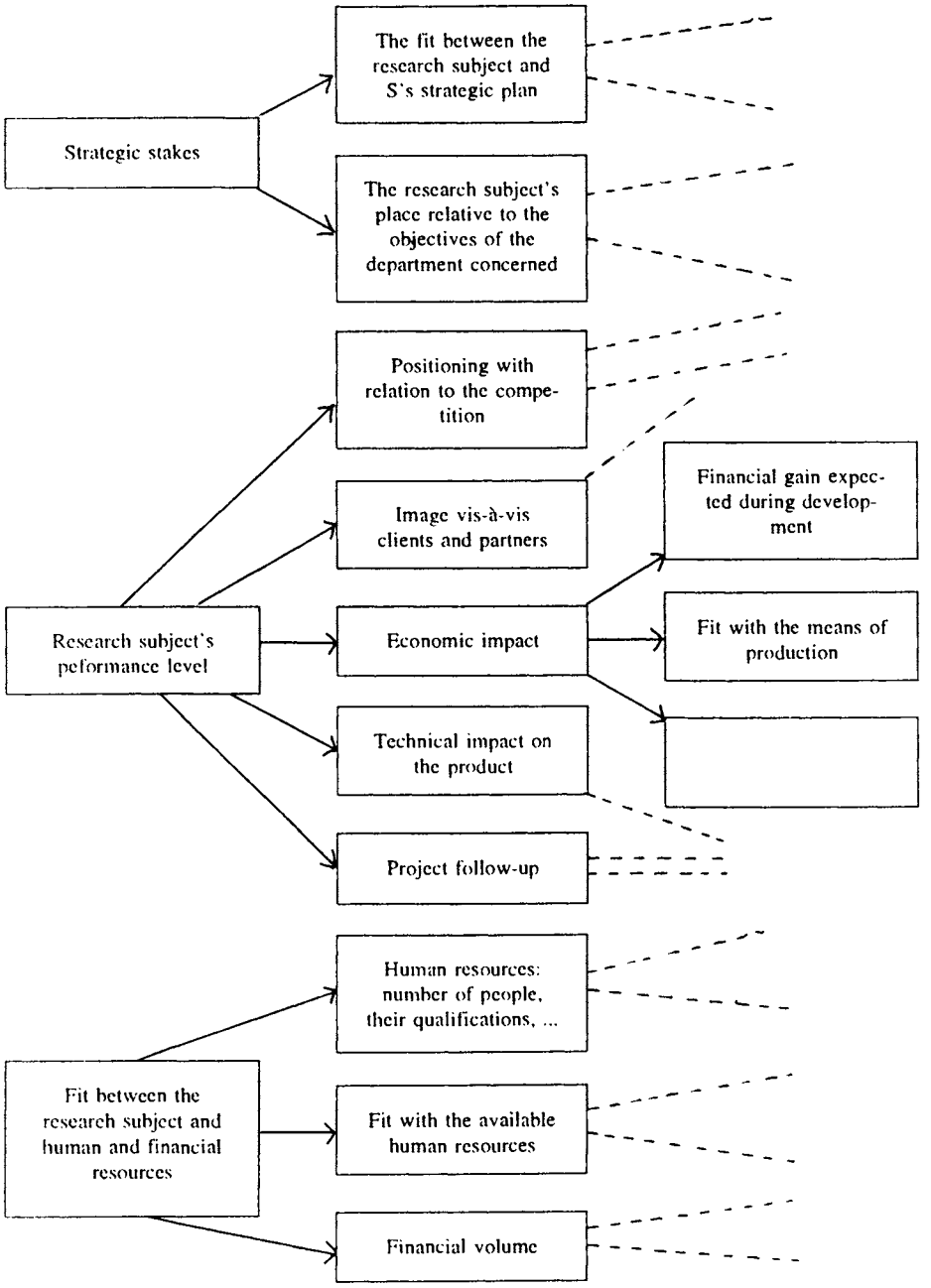


Fig. 1.1 Examples of tree diagrams: the case of R and D proposals

causality, however, are still too often developed by technicians "in chamber". It is preferable to construct them by gathering together the implicit hypotheses of all the stakeholders (political decision-makers, operators in the field, clients)". (Translated from the French text).

1.2.3 Conceiving of scales for evaluating without necessarily quantifying

We should be able to evaluate each potential action relative to each elementary consequence and to each aggregate of these consequences present in the structure which has been elaborated (see Section 1.2.2) and, to whatever extent possible, according to each point of view. If the nature of what must be tackled allows it, the evaluation may be carried out in quantitative terms, using a monetary, duration or length scale.

The mere fact of using a numerical indicator should not automatically lead us to look upon evaluation as quantitative. The very idea of quantity presupposes, first of all, that the total absence of quantity is meaningful, hence the importance of the role played by 0 in the scale. When the thermometer reads 20°C it is not twice as hot as when the mercury stops at 10°C. In addition, when there is "quantity" equal differences in terms of evaluation reflect equal differences relative to what we are seeking to tackle. If we are concerned with a human being's thermal comfort, the variation between 5° and 8°C can be perceived as non equal to the variation between 15° and 18°C, even if there is equality with reference to certain underlying physical quantities. Likewise, with a percentage indicator the variation of satisfaction produced by an increase of 10% can be different according to whether the increase is from 20% to 30%, 50% to 60% or 70% to 80%. If this is true, the scale is not quantitative. This implies that certain basic arithmetic calculations (notably calculating averages) have no real meaning.

In many cases, evaluation cannot be carried out in quantitative terms. We then use a qualitative ordinal scale such as the one shown in Fig. 1.2. It is often useful to assign numbers to the degrees: for example, 0 for the worst, 1, 2, etc. for the following degrees. Nonetheless, a numerical scale thus defined can lead us to think, wrongly, that the difference separating two consecutive degrees, such as the passage from 1 to 2 and from 4 to 5, reveal preference variations of the same magnitude. Whenever this is not true, these numbers have only ordinal significance. It is, therefore, senseless to add them up, subtract them or multiply them by a coefficient. With certain precautions, this may not apply notably to scales known as interval or ratio scales (see Appendix).

0	1	2	3	4	5
Completely unsuited	Ill-suited	Fairly ill-suited	Fairly well-suited	Well-suited	Perfectly suited

Fig. 1.2: Example of a qualitative ordinal scale

1.2.4 Validating a criterion family [8] to base and argue comparisons

In decision-aiding, the term *criterion* [9] designates a way of evaluating which serves to position a potential action (or an alternative) on a preference scale corresponding to a well-identified point of view. This mode of evaluation may rest on a more or less complex *indicator* or even an aggregation of diverse indicators used to tackle *consequences* pertinent to the *point of view* considered.

Validating a *criterion family* is above all making the *significance axis*, for each of the criteria intelligible for each of the *stakeholders*. It must be made clear what it allows us to apprehend, but also on what terms: on what scale? We must make sure that each criterion is perceived to be a relevant instrument for comparing potential actions; in other words, that it is capable of playing a role (the relative importance of which could vary considerably from one stakeholder to another) in laying the foundations for convictions, communicating concerning these, in debating and orienting the process towards a decision and, if the occasion arises, in contributing to its legitimation.

Validating a criterion family also means verifying that this family satisfies a certain number of logical demands which ensure coherence (see Appendix: Coherent criterion family).

1.2.5 Discussing the role allotted to each criterion in order to "weigh the pros and cons"

A given actor will consider that the criterion g is "more important" than the criterion h. Another actor may have the opposite opinion. The first actor may go so far as to translate his system of values into numerical terms saying that the criterion g should be weighted at 3 if the criterion h is weighted at 1. The weight metaphor is often deceptive. This does not stem from the eminently subjective character of the value assigned to the weight; it comes above all from the fact that this figure has meaning only relative to its formal use in weighing the pros and cons in a comparison of different actions (see Mousseau [47], Roy and Mousseau [65]).

Frequently, weights are perceived as coefficients which we should be able to use to multiply the numerical performance level of an action according to the criterion considered; the products thus obtained being then added on the set of criteria used in order to assign a score, utility or comprehensive value (called an average or weighted sum procedure) to the action. In this conception, assigning values to weights cannot be carried out without taking into close consideration the manner in which the scales associated with each criterion were defined. This is because weights operate like rates of exchange (usually called substitution rates or tradeoffs) which

---

[8] This family may be reduced to a single criterion.

[9] For more information on the terms in italics, see the Appendix.

determine the magnitude of the gain we should obtain on a criterion in order to compensate with precision a unit loss on another criterion.

This is entirely different when weights are interpreted as a number of votes given to each criterion in a voting procedure. This type of procedure leads to comparing actions exclusive of any notion of compensating possible losses with possible gains. Moreover, it allows us to integrate a veto mechanism into the comparison procedure and thus to apprehend another aspect of the importance allotted to each criterion.

1.2.6 Apprehending the sources of incomplete knowledge in order to avoid imputing more to data than they really signify

The sources of incomplete knowledge are legion. They stem primarily from (see Roy [58]):

– the imprecision of instruments used to obtain measurements when we must give an account of present or past facts and/or events;

– the uncertainty inherent in any assessment of future facts or situations;

– the inevitable presence of some degree of ambiguity and/or arbitrariness in the way in which we take into account complex phenomena, whether they be in the past, the present or the future.

Certain characteristics of the organizational context may contribute to reinforcing the impact of each of these three sources, whether they be related to phenomena of (conscious or unconscious) self-censuring or various actors' strategic behavior, and more generally all the obstacles which might hinder the free circulation of information.

The main instruments available to us (see especially Bouyssou [12], Perny and Roy [52]) to take this incomplete knowledge into account in our reasoning, as well as in more formal calculations are: probability distributions, dispersion and discrimination thresholds (see Appendix), fuzzy numbers, rough sets, ...

1.3 PROCEDURES AND TOOLS FOR ELABORATING RECOMMENDATIONS AND/OR FOSTERING COOPERATION WITH A VIEW TO REACHING A DECISION

Cooperation is the "raison d'être" of organizations, but observation has shown that maintaining cooperative behavior remains a recurrent problem, an objective which is constantly threatened... As soon as a minimum of complexity is reached, cooperation necessitates instruments of coordination acceptable to all actors. "Common languages" and communication techniques (data exchanges, working groups, ...) set up and naturally extend the scope of cooperation. This is also true of information technologies or any approach to formalized modeling of activity which constitutes supports for collective action and allows actors to coordinate better with one another. Even with such instruments, actors cannot avoid the task of coming to an agreement concening the information they possess and the choice or impact of their actions. A grammatically "comprehensible" statement, accurate knowledge or

a coherence rule will not be legitimate, valid or acceptable per se to those who receive it (Hatchuel [28], translated from the original French text).

The conceptual framework presented in Section 1.1.2 constitutes an environment within which numerous procedures and computerized tools have been conceived of in the context of decision-aiding, as introduced in Section 1.1.3. By referring to this conceptual framework and, if necessary, to certain appropriate procedures and tools, it is in many cases possible to contribute to creating a climate of confidence amongst actors and to facilitate concertation.

This applies to preparing a decision-making process of an exceptional nature, as well as to the conception of a tool designed to provide time and again elements of a response to decision-making processes of a repetitive nature. It also applies to intermediate situations somewhere between these two types.

The references provided below will enable the reader to form an idea of both the geographic and sectorial diversity of decision-aiding applications as practiced today: Anandalingam and Olsson [1], d'Avignon and Sauvageau [2], Balestra and Norese [3], Bana e Costa and Dinis das Neves [6], Bell *et al.* [9], Benson *et al.* [10], Bollinger *et al.* [11], Cerny [15], Dimitras *et al.* [20], Duckstein *et al.* [21], Eder *et al.* [22], Gomes [26], Hokkanen and Salminen [29], Jacquet-Lagrèze [30], Krasojevic and Haller [34], Larichev *et al.* [36], Le Pen [38], Mareschal and Brans [39], McCord *et al.* [41, 42], Mladineo *et al.* [44], Ostanello [48], Özelkan and Duckstein [49], Perny and Vanderpooten [53], Pictet [54], Pomerol *et al.* [55], Roy *et al.* [66], Siskos *et al.* [71], Stathopoulos [72], Stewart and Scott [73], Teich *et al.* [77], Urli and Beaudry [81]. Inevitably selected somewhat arbitrarily, these works constitute only a small sample of the innumerable decision-aiding applications in progress today in a great variety of countries.

Many models, procedures and tools mobilized by these applications are explained in the following chapters of this book. We therefore refer the reader to these chapters for information concerning formalism and the techniques of description and of producing results. In the four sub-sections which follow, we shall emphasize those fundamental aspects linked either to the object of these procedures and tools, or to the conditions in which they are used.

1.3.1 The problem of aggregation

Whenever a given potential action is rated higher than another one according to certain criteria, it will frequently be rated lower according to other criteria. As a result, in order to compare potential actions to one another, decision-aiding must allow us to reason by using, in a comprehensive way, all performance levels of each

potential action according to the n [10] criteria from the coherent criterion family used.

Reasoning which thus takes into account all criteria comprehensively necessarily requires the support from a logic of aggregation (see especially Roy and Bouyssou [64], Tsoukiàs [80]). This logic assigns to each criterion a role characterized by the importance attributed to it within the working hypothesis under consideration (weights, aspiration levels, veto, ...). This mechanism taken as a whole constitutes what is called a *multicriteria aggregation procedure* (MCAP).

The average or the weighted sum constitutes the simplest example of a first type of such procedures. It results in aggregating all the performance levels of a potential action into a single number (see Belton, chapter 12 in this volume). Many other procedures, generally more sophisticated, translate the aggregation result into similar terms (score, utility, value, ...). Among the procedures of this type we could cite MAUT (see Keeney and Raiffa [33]), AHP (see Saaty [68, 69]) and MACBETH (see Bana e Costa and Vansnick [7]).

We should observe that the procedures of this first type lead us to define a single criterion synthesizing the n criteria of the coherent family. By the very fact that such a criterion family has been conceived of and explained, these procedures cannot be confused with monocriterion procedures (such as, for example, cost-benefit; see Johansson [31], Mishan [43], Ray [57], Sugden and Williams [76]) which skip over this multicriteria phase of analysis. It is not only the form of the single criterion which may be different, but also the way the team of analysts is integrated into the decision-making process, as well as the opportunities it offers for interpreting and discussing results.

The methods known as the ELECTRE type (see Brans *et al.* [14], Diez de Castro *et al.* [19], Goicoechea *et al.* [25], Maystre *et al.* [40], Roy [60], Schärlig [70], Vanderpooten [82], Vincke [83] and in this volume chapter 11) find their support in a second type of MCAP, which proceeds by pairwise comparisons. The result of this comparison is formulated in terms such as indifference, weak preference, strict preference, incomparability, ... It is obtained by applying rules which are generally not very or not at all compensatory and which remind us of a voting mechanism with a possible veto.

The procedures most commonly employed today to structure the decision-aiding approach come from one of the two preceding types. Others, however, have been suggested (see especially Bana e Costa [4], Gardiner and Vanderpooten [24], Pardalos *et al.* [50] and in this volume Wierzbicki, chapter 9; Stewart, chapter 10; Slowinski *et al.*, chapter 14; Pomerol and Perny, chapter 15; Chipperfield, chapter 16).

---

[10] It is possible that a single criterion (n = 1) might have been constructed straightaway on the bases of considerations presented in 1.2.2 and 1.2.3 above, in which case, we could speak of monocriterion approaches and procedures; under these conditions, the problem of aggregating, in the terms in which it is envisaged in this paragraph, becomes pointless.

1.3.2 From results to recommendations [11]

Decision-aiding activity frequently comes to an end without any recommendation being formulated. The results arrived at by agreeing on a coherent criterion family, obtaining one or more performance tableaux, bringing to light levels of aspiration, of rejection, of discrimination thresholds, of different sets of weights, ... all form a positive contribution that the actors (other than the team of analysts) involved in the decision-making process often judge to be satisfying. It is nevertheless desirable, in many cases, to go further, which presupposes using more or less formalized procedures (especially aggregation), in order to arrive at less rudimentary results. The procedures and tools conceived of in this perspective lead to results which usually take one of the following forms:

– discovering one or even all of the optimal potential actions in a working hypothesis defined by a fixed set A and an optimization criterion;

– selecting as restricted a number as possible of potential actions which justify eliminating all the others;

– assigning each potential action to one category among those of a family previously defined, such as those shown in Fig. 1.3 below (see also 1.3.3.1 below);

Actions for which implementation is fully justified	Actions for which implementation could be advised after only slight changes	Actions for which implementation can be advised only after significant changes	Actions for which implementation is inadvisable

Fig. 1.3: Example of families of categories suitable for sorting procedures

– ranking actions from the best to the worst, incomparabilities and ex aequo remaining possible (see also 1.3.3.2 below).

The way in which DA is envisaged (the problematic [12]) conditions the form of results which it is suitable to look for. The four problematics mentioned above are not the only possible ones (see Bana e Costa [5]).

Except under very unusual conditions, the result arrived at by treating a set of data through any individual procedure should not be confused with a well-founded scientific recommendation (see Roy [61]). Repeated calculations, using different, but

---

[11] This term is increasingly used in DA to replace "prescription". The latter is, in many cases, inappropriate (see Hatchuel [28], Roy [61]) for designating what a team of analysts accompanying a decision-making process might achieve.

[12] As in Roy [62], we will use this term for translating the French word "problématique". We considered translating this term as problem statement, problem type, or problem formulation, but felt that these could give the wrong impression.

equally realistic sets of data (given the imprecise, uncertain and indeed ill-determined nature of certain factors or parameters; see end of section 1.2.2), are generally necessary to elaborate a recommendation on the basis of robust conclusions stemming from the multiple results thus obtained (see Roy [63] and *infra* Tanino, chapter 7). The statements of the proposals which make up the recommendations should be submitted to the assessment and discernment of the decision-maker [13] and/or the actors involved.

1.3.3 Decision-aiding and working in a committee

DA procedures and tools can be mobilized to facilitate working in groups. As in all group work, this presupposes that a minimum of "meta-rules" be accepted. We shall limit the discussion of DA's contribution here to two specific cases.

1.3.3.1 Case number 1 [14]

For strategic reasons, a company we shall refer to as *R* carries out a considerable part of its research and innovation work in the area of markets it can obtain following a tender for bids. *R* receives numerous tenders. Replying to them necessitates the equivalent of several months of work and sometimes calls for starting up research which is costly in terms of material. For this reason every week a committee *C*, presided over by the sales manager in charge of the tender budget examines each week the "new business" files received. Each of these files comes from a service within *R* which has suggested replying to a tender it has received. Each of these is treated as a potential action. *C* must decide, for each of them, whether to accept or refuse and, in the case of acceptance, how much money to allocate to the service in question for developing a response on behalf of *R*.

Using information contained in the file, the reply proposal is evaluated according to nine criteria, covered by three structure-providing points of view:

- chances of being awarded the contract;
- strategic interest for *R*;
- economic interest.

On this basis, the file is assigned by the ELECTRE TRI method to one of the following categories:

- no restriction concerning acceptance;
- some hesitation concerning acceptance (a doubtful "yes");
- hesitation concerning refusal (a doubtful "no");
- unhesitating refusal.

---

[13] In DA, this term usually designates the entity (individual or group of individuals) on whose account and in whose name DA is used.

[14] The rule of client confidentiality prevents us from revealing our sources here.

For each item of business the committee has the file, the nine evaluations and the category assigned to it. Any member of the committee may query any of the information leading to the evaluation according to one or more criteria (different assessment concerning certain risks, allocation of resources other than those requested, ...). When the bases are thus modified, new evaluations are immediately recalculated and the resulting reassignment is immediately made known. The committee remains sovereign: it makes its decision straightaway or defers it until the following week.

1.3.3.2 Case number 2 [15]

In 1993, the Swiss Confederation's Federal Office of the Environment (OFEFP) [16] began coordinating an inter-cantonal project for the incineration of urban wastes. The project primarily involved three cantons. One of these, Geneva, already possessed its own waste incinerator (WI), "Les Cheneviers"; according to the agreement concluded in 1993, each of the other two cantons, Vaud and Fribourg, would ultimately be equipped with its own waste incinerator.

In February of 1996, after an official appraisal of the Cheneviers facility, the canton of Geneva offered to incinerate, on an annual basis, an additional 130,000 tonnes of waste from the cantons of Vaud and Fribourg. This offer brought into question the necessity of constructing simultaneously the waste incinerators planned for Vaud ("Tridel") and Fribourg ("Posieux").

Confronted with the difficulty of either justifying the construction of both these projects or of choosing which of the two should be delayed, or even completely abandoned, the political authorities for the three cantons, together with the Director of the OFEFP appointed a technical commission to examine and compare the Tridel and Posieux projects as objectively as possible.

Composition of the commission

(...) The commission was composed of 9 people representing 4 actors:

– the canton of Geneva, represented by two officials from the Cheneviers WI;

– the canton of Vaud, represented by the person in charge of the Tridel project and the head of the canton's Office of the Environment;

– the canton of Fribourg, represented by two officials from the national Office of the Environment in Fribourg and an attorney assigned to the Posieux project;

– two representatives from the Federal Office of the Environment, appointed as technical advisors, who assumed the task of presiding over the commission.

In addition to these representatives, an analyst was appointed from the Institut de Génie de l'Environnement (IGE) associated with the Ecole Polytechnique Fédérale de Lausanne to act as a methodology advisor. His task was that of overseeing all

---

[15] Translated from excerpts of the French text by D. Bollinger et al. [11].

[16] Office fédéral de l'environnement de la Confédération suisse.

phases of the decision-making process, including the results and how to interpret them.

It was indispensable for this specialist in decision-aiding methods to be impartial. It would obviously have been undesirable if he could have been accused of manipulating the data. Moreover, each of the other actors in the process had to have total confidence in the analyst. His participation had to be limited to methodological aspects of the problem and their implementation. Elements influencing the results could come only from the actors themselves.

Certain technical experts were contacted during the course of the commission's work in order to calculate costs or provide figures for other data (cardinal data).

Working methodology

In order to win over all the actors' confidence, the analyst must act in a wholly transparent manner, presenting the methodology of multicriteria methods in language that is clear and accessible. Thus a number of 15-minute talks were presented over time to the commission members at appropriate stages of the decision-making process. Certain basic principles of the multicriteria methods, however, were put forth from the outset: the interest of non-compensatory procedures and of using various weightings to take into account the importance of different points of view.

All the actors must also agree on using the same criterion family. Theoretically, structuring scenarios [17] and constructing criteria constitute two separate, successive phases of the decision-making process. In practice, however, it is difficult to disassociate these two phases in real time. The first stage of the commission's work, therefore, focuses on these two phases simultaneously as "parallel" phases. Work sessions allowed commission members to discuss the evolution of the scenarios as well as that of the criteria. In the minds of the actors, who were not specialists in decision-aiding, these two aspects of the commission's work seemed very closely linked.

Due to time constraints, it was necessary to opt for a simple, efficient starting point which would allow the actors to delineate the problematic quickly and to formulate it in terms of the multicriteria approach used. The technical representative from the federal ministry of the environment thus proposed a first rough draft structuring criteria and scenarios so that they could be discussed within an existing framework. The principle of "constructive demolition" was adopted, which is to say that this first draft was designed *a priori* to be thoroughly reworked, but it gave opportunities for structuring ideas on the basis of arguments often born of conflicts and agreements.

Conclusions

The general conclusion adopted unanimously by all members of the commission was the following:

It appears that the three types of scenarios 1, 2 and 3 can be recommended. They are nonetheless not equivalent. Two principal paths are likely to result in an appropriate solution for eliminating urban wastes in southern Switzerland: scenario

[17] i.e., potential actions.

1.1, the construction of both Tridel and Posieux (the original project); and one or the other of scenarios 3.1 or 3.2, in which only the Tridel WI would be constructed. Despite the widely divergent weightings of criteria proposed by the four main actors, these scenarios were systematically well-ranked.

(...) Multicriteria analysis was able to show that scenarios of type 0 and 4 were not suitable for the problem being evaluated. The commission was able to justify the results of the ELECTRE III method by analysing the decisive criteria which ranked these scenarios last.

(...) The federal office of the environment had called for the formation of a commission in order to be able to choose which waste incinerator would be built and which project would be abandoned. Yet the work of the commission resulted in recommendations that were clear and unanimous for all the actors. The scenario concerning the construction of both facilities according to the original project is certainly not perfect, but it is just as envisageable as the scenarios involving only the construction of Tridel. The OFEFP, as the ultimate decision-maker, is faced with the following possible courses of action:

– maintaining its initial position and rejecting outright the construction of two new WIs; in this case, a type 3 scenario would be chosen, probably without the consent of the Fribourg cantonal government;

– relaxing its original position and leaving the choice of scenarios recommended by the commission to the political decision-makers;

– repudiating the commission's report and calling for a fresh evaluation.

(...)

1.3.4 Judicious use of software

There is a wide variety of software commercially available. In order to make good use of it, it is important to perceive clearly:

– The nature of the aid each specialized computer program can provide. This may be a matter of: structuring the consequences with a view towards elaborating a coherent criterion family, assigning a weight to each of the criteria from a family, determining or approaching an optimum, selecting among potential actions, assigning each potential action to a category, ranking potential actions, orienting exploration so as to find a compromise, ...

– The nature and meaning of the data to be elaborated, in order to use a specialized computer program correctly.

– A minimum of understanding concerning the essential hypotheses giving meaning to the results obtained, in order to avoid unfortunate errors of interpretation.

Anyone using software must remember that it is no more than a support, useful in implementing a method or manipulating a model. Its function is not to resolve the problem. It is only designed to facilitate the work of whoever is using the method or model.

The conveniences provided by a specialized computer program (user-friendliness, ergonomy) for recording and transforming data, manipulating various phases of

calculations and presenting results often constitute advantages which can influence the user in his or her choice of method or model. But do the latter correspond to the problem under consideration? Are they well-suited to the decision-making context? These aspects must not be given a back seat. Of course, the nature of the interaction which the computer program helps to create between user and method or model constitutes an important means of appropriating the latter. Yet, unfortunately, this has perverse effects which Thépot [78] rightly underlines: "With the development of computer tools, computation and representation are now possible in ways we could have scarcely imagined twenty years ago. Yet, at the same time, this has brought about a popularization of methods which, when ill-advisedly applied, discredit modeling by confining it to an instrumental role." (translated from the original French text).

1.4 CONCLUSIONS

First of all, it would seem necessary to emphasize that decision-aiding, as presented here, cannot claim to play a positive role in all decision contexts. Characterizing those decision-making contexts in which DA should be pertinent will necessitate further research and experimentation.

Secondly, we should remember that even in the so-called hard sciences, there is no known objective criterion (see Chalmers [16], Popper [56]) which allows us to affirm that a theory is true. Similarly, we cannot hope to prove scientifically, in a given decision-making context, that a given decision is the best. DA can, however, make clear that a solution thought to be good is, in fact, bad.

The concepts, procedures and tools which we have spoken of here should be seen as keys capable of opening doors which give access to pathways allowing us to bring the implicit to the surface, to bring unavoidable subjective elements out into the open and to deduce consequences from those that are objective, within the framework of working hypotheses. We are within our rights to expect of DA that it will illuminate choices, not only by the results thus brought to light, but also in organizing debate on objectives and ends. More generally, the roles of investigating organizational operations, of conformation, of accompanying change and of exploring what is new that management instruments are usually called upon to play can in certain circumstances, according to Moisdon [45], be given over to decision-aiding.

In conclusion, it would seem reasonable to expect from decision-aiding that, depending on the case, it would contribute to:

– analysing the decision-making context by identifying the actors, the various possibilities of action, their consequences, the stakes, ...

– organising and/or structuring how the decision-making process unfolds in order to increase coherence among, on the one hand, the values underlying the objectives and goals, and, on the other hand, the final decision arrived at;

– getting the actors to cooperate by proposing keys to a better mutual understanding and a framework favorable to debate;

– elaborating recommendations using results taken from models and computational procedures conceived of within the framework of a working hypothesis;
– participating in the final decision legitimization.

REFERENCES

[1] Anandalingam, G., C.E. Olsson: A multi-stage multi-attribute decision model for project selection. European Journal of Operational Research 43(1989)271-283

[2] d'Avignon, G., M. Sauvageau: L'aide multicritère à la décision : un cas d'intégration de critères techniques, économiques et environnementaux à Hydro-Québec. RAIRO Recherche Opérationnelle 30,3(1996)317-332

[3] Balestra, G., M.F. Norese: "Model structuring by means of ELECTRE III: a clinical research and application context." Proc. of the International Conference on MCDM, Mons, Belgium, 1997, pp. 239-242

[4] Bana e Costa, C.A. (ed.): *Readings in Multiple Criteria Decision Aid.* Springer-Verlag 1990

[5] Bana e Costa, C.A.: *Structuration, construction et exploitation d'un modèle multicritère d'aide à la décision*, Thèse de doctorat pour l'obtention du titre de Docteur en Ingénierie de Systèmes, Universidad Técnica de Lisboa, Instituto Superior Técnico, Portugal, 1992

[6] Bana e Costa, C.A., C.F. Dinis das Neves: "Describing and formalizing the evaluation process of Portuguese navy officers". In: A.G. Lockett, G. Islei (eds.): *Improving Decision Making in Organisations.* Proceedings, Manchester, UK, August 1988, Springer-Verlag, 1989, pp. 355-369

[7] Bana e Costa, C.A., J.C. Vansnick: MACBETH – An interactive path towards the construction of cardinal value functions. International Transactions in Operations Research 1, 4(1994)489-500

[8] Banville, C., M. Landry, J.M. Martel, C. Boulaire: A stakeholder approach to MCDA. Syst. Res. 15(1998)15-32

[9] Bell, D.E., R.L. Keeney, H. Raiffa (eds.): *Conflicting objectives in decision.* IIASA, Laxenburg and Wiley, New York 1977

[10] Benson, H.P., D. Lee, J.P. McClure: A multiple-objective linear programming model for the citrus rootstock selection problem in Florida. Journal of Multi-Criteria Decision Analysis 6(1997)289-295

[11] Bollinger, D., L.Y. Maystre, R. Slowinski: Un cas concret de la contribution de l'aide multicritère à la décision à la coordination inter-cantonale pour l'incinération des déchets urbains. Paper presented at the *45e Journées du Groupe de Travail Européen "Aide Multicritère à la Décision"*, Čelàkovice, République Tchèque, 20-21 mars 1997

[12] Bouyssou, D.: "Modelling inaccurate determination, uncertainty, imprecision using multiple criteria". In: A.G. Lockett, G. Islei (eds): *Improving Decision Making in Organisations.* Lecture Notes in Economics and Mathematical Systems 335, Springer-Verlag, 1989, pp. 78-87

[13] Bouyssou, D., M. Pirlot, A. Tsoukiàs, Ph. Vincke: A manifesto for the new MCDA era. Journal of Multi-Criteria Decision Analysis 2(1993)125-127

[14] Brans, J.P., B. Mareschal, Ph. Vincke: "PROMÉTHÉE: A new family of outranking methods in multicriteria analysis". In: J.P. Brans (ed.): *Operational Research '84*, Elsevier Science Publishers B.V. (North-Holland), 1984, pp. 408-421

[15] Cerny, M.: Decision making in human resources management. CEJORE 2(1993)161-175

[16] Chalmers, A.F.: *What is this thing called science? An assessment of the nature and status of science and its methods.* Second edition, Open University Press 1982

[17] Chin, C., L. Duckstein, M.L. Wymore: Factory automation project selection using multicriterion Q-analysis. Applied Mathematics and Computation 46(1991)107-126

[18] David, A.: L'aide à la décision entre outils et organisation. Entreprises et Histoire 13(1996)9-26

[19] Diez de Castro, J.A., C. Redondo Lopez, B. Barreiro Fernandez: *Ayuda a la decision: Un nuevo instrumento de gestion.* Universidade de Santiago de Compostela, Servicio de Publicacions e Intercambio Cientifico, 1997

[20] Dimitras, A.I., C. Zopounidis, C. Hurson: A multicriteria decision aid method for the assessment of business failure risk. Foundations of Computing and Decision Sciences 20, 2(1995)99-112

[21] Duckstein, L., W. Treichel, S. El Magnouni: Ranking ground-water management alternatives by multicriterion analysis. Journal of Water Resources Planning and Management 120, 4(1994)546-565

[22] Eder, G., L. Duckstein, H.P. Nachtnebel: Ranking water resources projects and evaluating criteria by multicriterion Q-analysis: An Austrian case study. Journal of Multi-Criteria Decision Analysis 6(1997)259-271

[23] d'Espagnat, B. (1990), *Penser la science.* Dunod, Paris, 1990

[24] Gardiner, L.R., D. Vanderpooten: "Interactive multiple criteria procedures: Some reflections". In: J. Climaco (ed): *Multicriteria Analysis*, Springer Verlag, Berlin 1997 pp. 290-301

[25] Goicoechea, A., D.R. Hansen, L. Duckstein: *Multiobjective Decision Analysis with Engineering and Business Applications.* John Wiley and Sons, New York, 1982

[26] Gomes, L.F.A.M.: Comparing two methods for multicriteria ranking of urban transportation system alternatives. Journal of Advanced Transportation 23, 2 & 3(1989)217-219

[27] Grégory, R., S. Lichtenstein, P. Slovic: Valuing environmental resources: A constructive approach. Journal of Risk and Uncertainty 7(1993)177-197

[28] Hatchuel, A.: Coopération et conception collective – Variété et crises des rapports de prescription. In: G. De Terssac, E. Friedberg (eds): *Coopération et conception*, Octares Editions, 1996, pp. 101-121

[29] Hokkanen, J., P. Salminen: Choosing a solid waste management system using multicriteria decision analysis. European Journal of Operational Research 98, 1(1997)19-36

[30] Jacquet-Lagrèze E.: "An application of the UTA discriminant model for the evaluation of R & D projects". In: P.M. Pardalos, Y. Siskos, C. Zopounidis (eds): *Advances in Multicriteria Analysis*. Kluwer Academic Publishers, Dordrecht 1995, pp. 203-211

[31] Johansson P.: *The economic theory and measurement of environmental benefits*. Cambridge University Press, Cambridge, UK, 1987

[32] Kahneman D., A. Tversky: The psychology of preferences. Scientific American 1(1982)160-173

[33] Keeney, R.L., H. Raiffa: *Decisions with multiple objectives – Preferences and value tradeoffs*. John Wiley and Sons, New York, 1976

[34] Krasojevic, M., P. Haller: Catalogue paramétré des assemblages bois. In: *EuropIA '95 "Approches Critiques des Applications des Technologies Avancées en Architecture, Génie Civil et Génie Urbain"*, Europia Production, 1995, pp. 123-132

[35] Landry, M., C. Banville, M. Oral: Model legitimation in operations research. European Journal of Operational Research 92(1996)443-457

[36] Larichev, O., R. Brown, E. Andreyeva, N. Flanders: "Categorical decision analysis for environmental management: A Siberian gas distribution case". In: J.P. Caverny, M. Bar-Hillel, F.H. Barron, H. Jungermann (eds): Contribution to Decision Making – I, Elsevier Science B.V., 1995, pp. 255-279.

[37] Larichev, O.I., H.M. Moshkovich: *Verbal analysis decision for unstructured problems*. Kluwer Academic Publishers, Dordrecht, 1997

[38] Le Pen, C: Théorie de l'utilité et mesure des états de santé : le débat QALYs-HYEs. Economie et Prévision (1997)129-130

[39] Mareschal, B., J.P. Brans: BANKADVISER: an industrial evaluation system. European Journal of Operational Research 54, 3(1991)318-324

[40] Maystre, L.Y., J. Pictet, J. Simos: *Les méthodes multicritères ELECTRE*. Presses Polytechniques et Universitaires Romandes, Lausanne, 1994

[41] McCord, M.R., O. Franzese, X.D. Sun: Attribute importance in the supply of aeromedical service. Transportation Research Record 1359(1992)95-102

[42] McCord, M.R., O. Franzese, X.D. Sun: Multicriteria analysis of aeromedical fleet expansion. Journal of Applied Mathematics and Computing 54, 2 & 3(1993)101-129

[43] Mishan, E.: *Cost benefit analysis*. Fourth edition, Routledge, 1988

[44] Mladineo, N., S. Stosic, T. Tomic, A. Baric: The applications of multicriteria analysis for estimating pollution in urban and coastal zones. Revue des Systèmes de Décision 1(1992)401-419

[45] Moisdon, J.C.: *Du mode d'existence des outils de gestion – Les instruments de gestion à l'épreuve de l'organisation*. Editions Seli Arslan, Paris, 1997

[46] Morin, E.: *La méthode 2. La vie*. Editions du Seuil, Paris, 1980

[47] Mousseau, V.: "Eliciting information concerning the relative importance of criteria". In: P.M. Pardalos, Y. Siskos, C. Zopounidis (eds): *Advances in Multicriteria Analysis*. Kluwer Academic Publishers, Dordrecht 1995, pp. 17-43.

[48] Ostanello, A.: "Action evaluation and action structuring: Different decision aid situations reviewed through two actual cases". In: C.A. Bana e Costa (ed): *Readings in Multiple Criteria Decision Aid*, Springer-Verlag, 1990, pp. 36-57

[49] Özelkan, E.C., L. Duckstein: Analysing water resources alternatives and handling criteria by multi criterion decision techniques. Journal of Environmental Management 48(1996)69-96

[50] Pardalos, P.M., Y. Siskos, C. Zopounidis (eds): *Advances in Multicriteria Analysis*. Kluwer Academic Publishers, Dordrecht, 1995

[51] Perez, M.: *Coûts externes et expression des préférences collectives dans le marché des transports*. Thèse pour le doctorat en Sciences Economiques, mention Economie des Transports, Université Lumière Lyon 2, 1996

[52] Perny, P., B. Roy: The use of fuzzy outranking relations in preference modelling. Fuzzy Sets and Systems 49(1992)33-53

[53] Perny, P., D. Vanderpooten: An interactive multiobjective procedure for selections medium-term counter measures. Journal of Multi-Criteria Decision Analysis 7(1998)48-60

[54] J. Pictet: *Dépasser l'évaluation environnementale – Procédure d'étude et insertion dans la décision globale*. Presses Polytechniques et Universitaires Romandes, Lausanne, 1996

[55] Pomerol, J.Ch., B. Roy, C. Rosenthal-Sabroux: Developing an "intelligent" DSS for the multicriteria evaluation of railway timetables: problems and issues. Revue des Systèmes de Décision 5, 3-4(1996)249-267

[56] Popper, K.: *The logic of scientific discovery*. Hutchinson, London, 1968

[57] Ray, M.: *Cost-benefit analysis: Issues and methodologies*. Hopkins, Baltimore, 1984

[58] Roy, B.: Main sources of inaccurate determination, uncertainty and imprecision in decision models. Mathematical and Computer Modelling 12, 10/11(1989)1245-1254

[59] Roy, B.: Decision-aid and decision-making. European Journal of Operational Research 45(1990)324-331.

[60] Roy, B.: The outranking approach and the foundations of ELECTRE methods. Theory and Decision 31(1991)49-73

[61] Roy, B.: Decision science or decision-aid science?. European Journal of Operational Research 66(1993)184-203

[62] Roy, B.: *Multicriteria Methodology for Decision Aiding*. Kluwer Academic Publishers, Dordrecht, 1996

[63] Roy, B.: A missing link in OR-DA: Robustness analysis. Foundations of Computing & Decision Sciences (to appear)

[64] Roy, B., D. Bouyssou: *Aide multicritère à la décision : Méthodes et cas*, Economica, Paris, 1993

[65] Roy, B., V. Mousseau: A theoretical framework for analysing the notion of relative importance of criteria. Journal of Multi-Criteria Decision Analysis 5(1996)145-159.

[66] Roy, B., R. Slowinski, W. Treichel: Multicriteria programming of water supply systems for rural areas. Water Resources Bulletin 28, 1(1992)13-31

[67] Roy, B., D. Vanderpooten: The European School of MCDA: Emergence, basic features and current works. Journal of Multi-Criteria Decision Analysis 5(1996)22-38

[68] Saaty, T.L.: *The analytic hierarchy process.* McGraw Hill, 1980

[69] Saaty, T.L.: *Décider face à la complexité – Une approche analytique multicritère d'aide à la décision*, EME, Paris, 1984

[70] Schärlig, A.: *Pratiquer ELECTRE et PROMÉTHÉE – Un complément à Décider sur plusieurs critères.* Presses Polytechniques et Universitaires Romandes, Lausanne, 1996

[71] Siskos, Y., E. Grigoroudis, N.F. Matsatsinis, G. Gaourakis: "Preference disaggregation analysis in agricultural product consumer behaviour". In: P.M. Pardalos, Y. Siskos, C. Zopounidis (eds): *Advances in Multicriteria Analysis*, Kluwer Academic Publishers, Dordrecht, 1995, pp. 185-202

[72] Stathopoulos, N.: *La performance territoriale des réseaux de transport.* Presses de l'Ecole Nationale des Ponts et Chaussées, 1997

[73] Stewart, T.J., L. Scott: A scenario-based framework for multicriteria decision analysis in water resources planning. Water Resources Research 31, 11(1995)2835-2844

[74] STP (Syndicat des Transports Parisiens), Direction des Investissements: Méthodes d'évaluation socio-économique des projets d'infrastructures de transports en commun en région Ile de France (Rapport d'Etape), 1996

[75] STP (Syndicat des Transports Parisiens), Direction des Investissements: Méthodes d'évaluation des projets d'infrastructures de transports collectifs en région Ile de France (Rapport Final), 1998

[76] Sugden R., A. Williams: *The principles of practical cost-benefit analysis.* Oxford University Press, Oxford, 1983

[77] Teich, J.E., H. Wallenius, M. Kuula, S. Zionts: A decision support approach for negotiation with an application to agricultural income policy negotiations. European Journal of Operational Research 81(1995)76-87

[78] Thépot, J.: La modélisation en sciences de gestion ou l'irruption du tiers. Revue Française de Gestion 102(1995)66-70

[79] Toulemonde, J.: Faut-il libérer l'évaluation de ses liens de causalité ? Annales des Mines, Gérer et Comprendre 47(1997)76-88

[80] Tsoukiàs, A.: *Sur la généralisation des concepts de concordance et de discordance en aide multicritère à la décision.* Université Paris-Dauphine, Mémoire pour l'obtention de l'habilitation à diriger des recherches, 1997

[81] Urli, B., D. Beaudry: Une approche multicritère d'application des ressources financières dans le domaine de la santé. RAIRO Recherche Opérationnelle 29, 4(1995)373-389

[82] Vanderpooten, D.: "The construction of prescriptions in outranking methods". In: C.A. Bana e Costa (ed): *Readings in Multiple Criteria Decision Aid*, Springer-Verlag, 1990, pp. 184-215

[83] Vincke, Ph.: *Multicriteria decision-aid.* Wiley, New York, 1992

APPENDIX: GLOSSARY [18]

The first occurrence of a term used elsewhere as an item in the Glossary is italicized.

ACTION (potential action)

A generic term used especially in theoretical presentations to designate that which constitutes the object of the decision or that which decision-aiding is directed towards. In practice, the term action may be replaced by such terms as scenario, plan, program, project, proposal, variant, dossier, operation, investment or solution, depending on the situation.

The concept of action does not, *a priori*, incorporate any notion of feasibility or implementation. An action is qualified as "potential" when it is deemed possible to implement it or simply when it deserves our interest within the decision-aiding context. A potential action may thus be a fictitious one.

ACTOR

A very general term designating any individual, entity or community likely to play any role whatsoever, directly or indirectly, in the unfolding of the decision-making process.

AGGREGATION

See Multicriteria Aggregation Procedure.

ALTERNATIVE

A term used instead of *action* when modeling is such that two (distinct) potential actions can in no way be conjointly implemented. This mutual exclusion comes from a conception of potential action which tackles the object of the decision or that towards which decision-aiding is directed in a comprehensive way. In some cases, a different conception called fragmented may be adopted so that several potential actions can be implemented conjointly. Alternative is not an appropriate word in such cases.

ASPIRATION LEVEL

A degree on a criterion scale marking a performance level which, if achieved by an action according to this criterion, indicates sufficient satisfaction. This is to say that any improvement on this scale is deemed non-significant of a real increase in satisfaction.

---

[18] For more specifics, the reader may refer to Roy [62], Roy and Bouyssou [64].

COHERENT CRITERION FAMILY

A family of n (> 1) criteria which satisfies the three following requirements:

a) Exhaustiveness: No argument acceptable to all stakeholders can be put forward to justify a preference in favour of action a vis-à-vis action b when a and b have the same performance level on each of the n criteria of the family.

b) Cohesiveness: Stakeholders unanimously recognize that action a must be preferred to action b whenever the performance level of a is significantly better than that of b on one of the criteria of positive *weight*, performance levels of a and b being the same on each of the other $n - 1$ criteria.

c) Nonredundancy: One of the above requirements is violated if one of the n criteria is left out from the family.

Remark: None of the above requirements implies that the criteria of a coherent criterion family are independent. Independence, however, may be desirable and sought after. We must specify what type of independence we hope to attain. The concept of independence is complex and multicriteria analysis has allowed us to give prominence to some important distinctions in this regard (structural independence, preferential independence, utility independence, ...).

In many cases, the coherent criterion family aims first of all at bringing to each stakeholder elements of judgment able to facilitate concertation and debate. So, the criterion family must respond to two additional requirements:

d) Understanding: the meaning of each criterion seems sufficiently intelligible to each stakeholder.

e) Commitment: the set of the n criteria seems appropriate to each stakeholder for tackling the main pertinent consequences.

CONSEQUENCE

This term is used to designate any effect or attribute inherent in or stemming from the implementation of any potential action and which should be taken into consideration to illuminate a decision. An effect or attribute should be taken into consideration whenever it may interfere with an *actor*'s objectives or system of values as a primary element capable of influencing the way in which he/she conceives of, modifies or argues his/her preferences.

CONSTRAINT

A condition imposed on an action in order for it to belong to a set of potential actions which are of interest at a given point in the decision-making process.

In mathematical programming, potential actions correspond traditionally only to feasible actions. It follows that the term takes on a more restrictive meaning in this case.

CRITERION

A tool constructed for evaluating and comparing potential actions according to a well-defined *point of view*.

The evaluation of an action according to a criterion may bring into play more or less complex computation rules, a more or less extensive survey or the opinion of one or more experts. Whatever procedure is used, we must take into account the pertinent effects and attributes (i.e., consequences) according to the point of view considered. To do so, it is often convenient to use one or more *indicators*.

The evaluation of an action according to a criterion is made concrete through its *performance* level which positions it on a preference *scale*. Two actions are compared according to the point of view considered just as their performance levels are compared.

The evaluation instrument made up of a criterion and *a fortiori* a coherent criterion family aims above all to provide some elements of assessment capable of facilitating concertation and debate to the different stakeholders involved. It is thus necessary that the *coherent criterion family* selected be deemed legitimate (commitment requirement) by each of the stakeholders and that the latter understand (understanding requirement) the way in which their concerns are translated in the *performance tableau*. This correct understanding presupposes, in particular, that the terms in which the performance is formulated for each of the criteria (unit for a quantitative scale, description of the degree if the scale is qualitative, ...) should be readily intelligible to the various stakeholders, and not only to the initiates.

All these performance levels cannot necessarily be seen as the approximation of an objective reality that the criterion would aim to apprehend as well as possible. The criteria can, in fact, sometimes be conceived so as to give an account of subjective aspects. When the reality to be tackled is very complex, the search for a good approximation is likely to result in illusory refinements, which, moreover, may cloud understanding and debate. Such refinements could also make it easier to give a "little push" in a certain direction. The absence of objective bases for anchoring the real value of a performance level according to certain points of view should not lead us to eliminate these points of view, in particular if the actors make use of them to conceive, modify and argue for their preferences. In the same way, the subjective character of certain effects or attributes only occasionally constitutes a valid argument for reducing their relevance and scope.

To be accepted by all stakeholders, a criterion should not bring into play in a way which might be determinant, any aspects of the system of values that certain of the stakeholders would find necessary to reject. This implies, in particular, that the direction in which preference is varied along the scale is not open to contest. On the other hand, this does not exclude significant divergences among stakeholders as to the relative importance which can be assigned to each criterion: a given criterion, important for some, could be judged to be of no interest for others (see entry on "Relative importance of criteria").

CRITERION FAMILY

See Coherent criterion family.

EFFICIENT (action)

An action a is efficient in a set A according to a criterion family F if any other action from A which is better than a according to at least one criterion proves to be not as good as a to at least one other criterion. The action a is consequently efficient if and only if within A there is no action which is at least as good as a on all the criteria of F and strictly better than a on at least one of these.

IMPORTANCE (relative importance of criteria)

This is a complex notion concerned with the differentiation of roles that a *stakeholder* would like different criteria to play in the elaboration and argumentation of comprehensive preferences. Consequently, this notion refers to the system of values of the stakeholder in question.

To define the idea of relative importance we frequently use the weight metaphor, according to which the greater the weight of a criterion, the more important role this criterion will play in forming comprehensive preferences. This metaphor is often misleading. The way in which the weights operate depends on the logic behind the *multicriteria aggregation procedure*. Thus in compensatory logics (in elementary weighted sum or in AHP or MAUT), assigning a greater numerical value to the weight of criterion g than to the weight of criterion h does not mean that the importance of g is greater than that of h.

Because the notion of relative importance of criteria has meaning only relative to a stakeholder whose value system it reflects, this notion is necessarily infused with a measure of subjectivity. In most cases, this means that any search for a perfectly objective value or for a procedure allowing us to compare objectively any given action with any other, is illusory. This impossibility does not stem from multicriteria analysis. It is just as present, although often hidden (or even masked under the appearance of objectivity) in any monocriterion analysis. The single criterion's goal is, in fact, to evaluate heterogeneous effects and attributes with a common unit. The different criteria of multicriteria analysis, however, aim to provide structuring elements to apprehend these effects and attributes in relatively homogeneous categories.

Thus cost-benefit analysis, for example, involves more or less implicit weighting, using techniques which translate values into monetary terms with reference to more or less fictitious markets. The weights thus arrived at are defined with reference to a very elegant economico-mathematical theory. The values assigned to these weights reflect the subjectivity of those who believe this theory is pertinent in guiding decisions. Other practitioners rightly refuse its legitimacy. In fact, the

theory rests upon rather unrealistic hypotheses and proves to be falsified in many decision contexts.

Unlike what happens with monocriterion analysis, multicriteria analysis avoids to prejudge any aggregation logic, nor does it favor the value system of any individual stakeholders. The criterion family should form a framework for structuring dialogue and debate. This framework seeks to give the most fundamental subjective aspects their due in order to foster a confrontation among different rationalities.

INDICATOR

An instrument which synthesizes, in qualitative or quantitative terms, certain information which should lay the foundation for a judgment of an action relative to certain of its characteristics, attributes or effects (*consequences*) which might arise from its implementation. An indicator might lead to associate to an action:

– either a simple state (urban motorway, ring-road, by-pass, ...),

– or a number (length, surface, cost, ...).

The set of possible states or values should not necessarily be conceived of to be a preference scale. If this set constitutes such a scale, the indicator can be used as a criterion. Several indicators may also be synthesized to define a criterion encompassing a broader point of view.

MODEL

A model is a schema which, for a given family of questions, is considered as an abstract representation of a class of phenomena that an observer has more or less skillfully removed from their environment to help in an investigation and to facilitate communication.

A model is not necessarily a simplified description of reality. For the purposes of investigation or communication, it may propose a representation of the phenomena in question which relies on very unrealistic hypotheses. To the extent that it is contingent upon a family of questions, a model is more a caricature of real-life situations than an impoverished or approximate photograph of it.

MULTICRITERIA AGGREGATION PROCEDURE (MCAP)

A procedure which allows us to compare any two actions from a set of actions A by taking into account (in a comprehensive way) the performance levels of each action according to all the criteria of a given criterion family.

OPTIMUM (action)

An action a is an optimum action in a set A according to a criterion g if any other action in A is worse than or indifferent to a according to this criterion; in other words, if there is no other action a' in A whose performance level $g(a')$ is better than $g(a)$.

PERFORMANCE LEVEL

The performance level of an action according to a criterion is the degree on the scale associated with a criterion on which this action is positioned.

PERFORMANCE TABLEAU

A double-entry table showing actions in rows and criteria in columns. At the intersection of row i and column j we find the performance level $g_j(a_i)$ of action a_i according to criterion g_j.

It is important to indicate, for each criterion:

– the associated scale and the direction of preference variation;

– the values, if any, for the levels of rejection and aspiration, as well as for the discrimination thresholds.

POINT OF VIEW

A class of effects or attributes which share the same objective or the same type of concerns thought pertinent by at least one of the stakeholders for evaluating and comparing actions.

A point of view is verbally defined by a sentence or several key-words. It may encompass a class of concerns of varying scope: a more or less open angle from which an action is examined. The family of structure-providing points of view should constitute a clear framework allowing us to apprehend all effects and attributes thought to be pertinent. Those structure-providing points of view should be few in number. As a result, it is necessary to identify some more limited points of view in order to give birth to criteria (the same structure-providing point of view may be used to create several different criteria).

PROBLEMATIC

The way in which decision-aiding is envisaged. In other words, the manner in which a problem is formulated in order to arrive at *results* judged to be appropriate for illuminating decisions. These results can take on various forms, namely:

– selecting several actions (choice problematic);

– assigning each action to a category belonging to a set of pre-defined categories (sorting problematic);

– ordering actions along a complete weak order or a partial order (ranking problematic);

– providing only a performance tableau giving additional information (*aspiration levels*, rejection levels, discrimination thresholds, ...).

It is essentially the way in which DA is conceived of, with reference to integrating the team of analysis thoroughly into the decision-making process, that will orient the choice of the problematic.

RECOMMENDATION

An assertion (not necessarily devoid of subjectivity) derived from a sequence of results based on various sets of data and/or sets of working hypotheses.

REJECTION LEVEL

A degree on a criterion scale which marks a performance level which, if not attained by an action according to this criterion, justifies rejecting this action whatever its performance levels might be on other criteria.

RESULT

An output of a procedure when it is applied to a set of data within the framework of a specific working hypothesis.

SCALE (preference scale)

A set of elements, called degrees, ranked according to a complete order; each degree is characterized by either a number or a verbal statement; a degree is used to translate the evaluation of an action, taking into account clearly specified effects and attributes; relative to these, all other things being equal, the ranking of degrees reflects the direction of preference variation in relation to situations which the degrees are used to characterize. A scale can be:

a) Purely ordinal: the gap between two degrees does not have a clear meaning in terms of preference differences; this is especially true in cases with:

– a verbal scale, when nothing allows us to state that pairs of consecutive degrees reflect equal preference differences all along the scale;

– a numerical scale when nothing allows us to state that a difference fixed between two degrees reflects an unvarying preference difference when we move the pair of degrees considered along the scale.

This kind of scale is called a qualitative scale.

b) Quantitative: a numerical scale whose degrees are defined by reference to a clearly defined unit in a way that gives meaning, on the one hand, to the absence of quantity (degree 0) and, on the other hand, to the ratio between any two degrees as being equal to the ratio of the numbers which characterize them, each of them being interpretable as the addition of a given number (integer or fractional) of units of the quantity considered.

c) Intermediate scale between the two extreme cases given above; this is especially the case with:

– scales called interval scales: the ratio of numbers which characterize two degrees could not be significant but the ratio between differences in numbers associated with two pairs of distinct degrees is significant (example: evaluation of a temperature in Celsius or Fahrenheit from an energy point of view; to the contrary,

an action's immediate rate of return and, *a fortiori*, its internal rate of return can only with difficulty be regarded as being evaluated on an interval scale);

– scales on which we can define a complete weak order on the set of degree pairs.

We call this type of intermediate scale a non-quantitative numerical scale or a discrete scale (in the case of a finite number of degrees).

The difference between two degrees that are sufficiently close together may be judged non-significant for differentiating two actions. This stems from the fact that the procedure used to position these actions along the *criterion* considered on one degree and on the other appears insufficiently precise (with regard to the complexity of the reality in question) or reliable (especially taking into account uncertainty concerning the future) to show a significant difference between two actions. The concept of discrimination *threshold* allows us to model this state. It allows us to work with available information without seeking to impoverish it by rounding off procedures or by defining classes in such a way that non-significant differences do not appear. Such practice produce unfortunate side effects.

SIGNIFICANCE AXIS (of a criterion)

An underlying dimension to which a criterion refers; in other words, a dimension which gives meaning to the comparison of any two performance levels according to this criterion.

STAKEHOLDER

Any individual, entity or community likely to take part in (possibly through the offices of an appointed intermediary) in the unfolding of a decision-making process with the intention of influencing that process with regard to objectives s/he holds or according to his or her own stakes.

Certain authors [19] employ this term in a larger sense to designate any individual or group of individuals who have, consciously of unconsciously, an interest in the decision context. In other words, anyone holding a stake in a very broad sense. In this understanding of the term, future generations may also be stakeholders. The same is true of a team of analysts and many other actors.

THRESHOLDS (dispersion and discrimination thresholds)

A concept whose objective is, for each criterion, to take into account the imprecision of certain data concerned with past of present phenomena, the uncertainty of which affects our knowledge of the future, our difficulty in tackling very complex attributes and effects or, finally, the fact that certain precisions can be completely devoid of any informative value.

---

[19] See Banville et al. [8], Landry et al. [35].

Dispersion thresholds translate the plausible differences, due to over-estimating or under-estimating, which affect the evaluation of a *consequence* or of a performance level. They allow us to include in our reasoning not only a probable value but also an optimistic or pessimistic value.

Discrimination thresholds are used specifically to model the fact that the difference between performance levels associated with two actions may be (relative to the criterion considered and all other things being equal):

– probative of a preference in favor of one of the actions (preference threshold);

– compatible with indifference between the two actions (indifference threshold).

These threshold may be constant along a scale or, to the contrary, variable. If the latter is true, we must make a distinction between direct and inverse thresholds.

WEIGHT (of a criterion)

See Importance.

ACKNOWLEDGEMENTS

I would like to thank Daniel Vanderpooten who was kind enough to re-read the first draft of this chapter and whose remarks and suggestions have enabled me to improve upon it. I am equally indebted to Albert David, José Figueira, Jean-Marc Martel, Vincent Mousseau and Nikolas Stathopoulos whose comments on an original French version of this text made a significant contribution to consolidating and clarifying the ideas presented here.

I am likewise deeply indebted to Dominique François whose help has been invaluable for putting this chapter in its final form. As always, she has taken much care with both the typing and the lay-out. My grateful thanks also go out to Evelyn Perry for the great care she has taken with the translation so as to faithfully transmit the content of the original French-language version to a non-French-speaking audience. It was a great pleasure for me to work once more with her.

2 THEORY OF VECTOR MAXIMIZATION: VARIOUS CONCEPTS OF EFFICIENT SOLUTIONS

Johannes Jahn

CONTENTS

Abstract: This chapter introduces the basic concepts of vector optimization. After the discussion of a simple example from structural engineering partial orderings on $\mathbb{R}^m$ are defined and connections to convex cones are investigated. Then we present the definitions of several variants of the efficiency notion: weak, proper, strong and essential efficiency. Relationships between these different concepts are investigated and simple examples illustrate these notions The last section is devoted to the scalarization of vector optimization problems. Based on various concepts of monotonicity basic scalarization results are described and the weighted sum approach is investigated in detail.

T. Gal et al. (eds.), *Multicriteria Decision Making*

2.1 BASIC NOTIONS

Optimization problems with several criteria arise in economics, engineering, applied mathematics and physics. As a simple example we discuss a design problem from structural engineering.

Example 2.1: We consider the design of a beam with a rectangular cross-section and a given length l (see Fig. 2.1 and 2.2). The height x_1 and the width x_2 have to be determined.

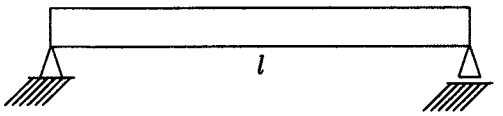


Figure 2.1: Longitudinal section.

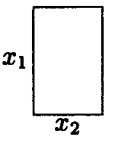


Figure 2.2: Cross-section.

The design variables x_1 and x_2 have to be chosen in an area which makes sense in practice. A certain stress condition must be satisfied, i.e. the arising stresses cannot exceed a feasible stress. This leads to the inequality

$$2000 \leq x_1^2 x_2.$$

Moreover, a certain stability of the beam must be guaranteed. In order to avoid a beam which is too slim we require

$$x_1 \leq 4x_2$$

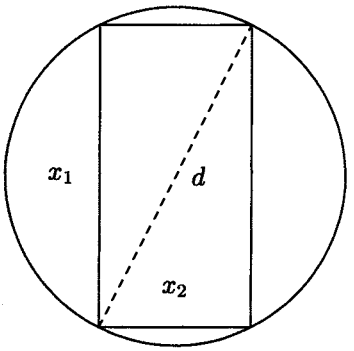
and

$$x_2 \leq x_1.$$

Finally, the design variables should be nonnegative which means

$$x_1 \geq 0, \quad x_2 \geq 0.$$

Among all feasible values for x_1 and x_2 we are interested in those which lead to a light *and* cheap construction. Instead of the weight we can also take the volume of the beam given as lx_1x_2 as a possible criterion (where we assume that the material is homogeneous). As a measure for the costs we take the sectional area of a trunk from which a beam of the height x_1 and the width x_2 can just be cut out. For simplicity this trunk is assumed to be a cylinder. The sectional area is given by $\frac{\pi}{4}(x_1^2 + x_2^2)$ (see Fig. 2.3).



$$d = \sqrt{x_1^2 + x_2^2}$$

Figure 2.3: Sectional area.

Hence, we obtain a vector optimization problem of the following form:

$$\text{"max"} \begin{pmatrix} -lx_1x_2 \\ -\frac{\pi}{4}(x_1^2 + x_2^2) \end{pmatrix}$$

subject to the constraints

$$\begin{aligned} 2000 - x_1^2 x_2 &\leq 0 \\ x_1 - 4x_2 &\leq 0 \\ -x_1 + x_2 &\leq 0 \\ -x_1 &\leq 0 \\ -x_2 &\leq 0. \end{aligned}$$

■

In this chapter we investigate vector optimization problems in finite dimensional spaces of the general form

$$\text{"}\max_{x \in S}\text{"} f(x) \tag{2.1}$$

Here we assume that S is a nonempty subset of $\mathbb{R}^n$ $(n \in \mathbb{N})$ and $f : S \to \mathbb{R}^m$ $(m \in \mathbb{N})$ is a given vector function. In the case of $m = 1$ this problem reduces to a standard optimization problem with a scalar-valued function f.

Since f is a vector-valued function one speaks of a so-called *vector optimization problem* or a *multi-objective optimization problem.* The set S is called

constraint set and the vector function f is called *objective function*. Actually it does not matter whether we investigate maximization or minimization problems. In this chapter we consider only vector maximization problems.

Maximization of a real-valued function f means that we look for the maximum value of all function values $f(x)$ with $x \in S$. In the vector-valued case we have to clarify in which sense we maximize among vectors $f(x) \in \mathbb{R}^m$ with $x \in S$ (see Fig. 2.4). The important question of how to order vectors in $\mathbb{R}^m$ is investigated in the next section.

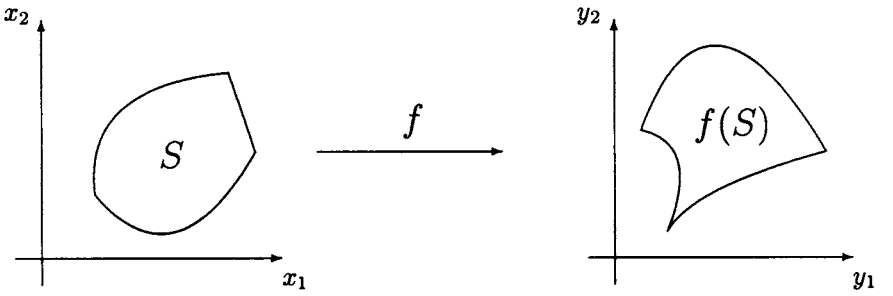


Figure 2.4: Preimage and image set of f.

2.2 ORDER RELATIONS

The mathematical theory of partially ordered sets provides the fundamental tool for the question of how to order vectors in $\mathbb{R}^m$. A partial ordering in $\mathbb{R}^m$ is defined as follows (e.g., see [12], [13]).

Definition 2.1:

(a) Every nonempty subset R of the product space $\mathbb{R}^m \times \mathbb{R}^m$ is called a *binary relation* R on $\mathbb{R}^m$ (one writes xRy for $(x, y) \in R$).

(b) Every binary relation $\leq$ on $\mathbb{R}^m$ is called a *partial ordering* on $\mathbb{R}^m$, if for arbitrary $w, x, y, z \in \mathbb{R}^m$:

(i) $x \leq x$ (reflexivity);

(ii) $x \leq y,\ y \leq z \ \Rightarrow\ x \leq z$ (transitivity);

(iii) $x \leq y,\ w \leq z \ \Rightarrow\ x+w \leq y+z$ (compatibility with the addition);

(iv) $x \leq y,\ \alpha \in \mathbb{R}_+ \ \Rightarrow\ \alpha x \leq \alpha y$ (compatibility with the scalar multiplication).

(c) A partial ordering $\leq$ on $\mathbb{R}^m$ is called *antisymmetric,* if for arbitrary $x, y \in \mathbb{R}^m$:

$$x \leq y, \quad y \leq x \quad \Rightarrow \quad x = y.$$

In the case that we equip the vector space $\mathbb{R}^m$ with a partial ordering we speak of a *partially ordered vector space.*

Example 2.2: If one defines the componentwise partial ordering $\leq_m$ on $\mathbb{R}^m$ by

$$\leq_m := \{(x, y) \in \mathbb{R}^m \times \mathbb{R}^m \mid x_i \leq y_i \quad \text{for all } i \in \{1, \ldots, m\}\},$$

then the vector space $\mathbb{R}^m$ becomes a partially ordered vector space. ∎

Notice that two arbitrary elements of a partially ordered vector space may not always be compared with each other with respect to the partial ordering. For instance, the vector (1,2) is neither larger nor smaller than the vector (2,1) with respect to the partial ordering given in Example 2.2.

The following definition shows that it is also possible to introduce a complete ordering on $\mathbb{R}^m$ allowing to compare arbitrary vectors.

Definition 2.2: A vector $x \in \mathbb{R}^m$ is called *lexicographically greater* than a vector $y \in \mathbb{R}^m$, if $x \neq y$ and the first component of $x - y$ being nonzero is positive.

It is obvious that for arbitrary vectors $x, y \in \mathbb{R}^m$ the following statement holds: Either $x = y$ or x is lexicographically greater than y or y is lexicographically greater than x. For instance, the vector (2,1) is lexicographically greater than (1,2).

Since we have a partial ordering on the vector space $\mathbb{R}^m$, it is not necessary to work with the algebraic structure in Definition 2.1 but we can use convex cones characterizing a partial ordering.

Definition 2.3: Let C be a nonempty subset of $\mathbb{R}^m$.

(a) The set C is called a *cone* if

$$x \in C, \quad \lambda \geq 0 \quad \Rightarrow \quad \lambda x \in C$$

(see Fig. 2.5).

(b) A cone C is called *pointed* if

$$x \in C, \quad -x \in C \quad \Rightarrow \quad x \stackrel{!}{=} 0_{\mathbb{R}^m}$$

(see Fig. 2.6).

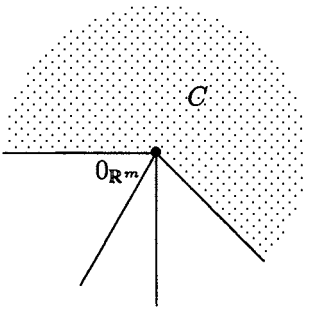


Figure 2.5: Cone.

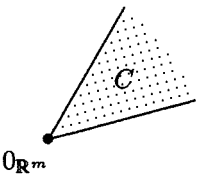


Figure 2.6: Pointed cone.

Example 2.3: The set

$$\mathbb{R}^m_+ := \{x \in \mathbb{R}^m \mid x_i \geq 0 \text{ for all } i \in \{1, \ldots, n\}\}$$

is a pointed cone. ■

Recall that a set $T \subset \mathbb{R}^m$ is called *convex* if for every $x, y \in T$

$$\lambda x + (1 - \lambda)y \in T \quad \text{for all } \lambda \in [0, 1]$$

(see Fig. 2.7 and 2.8).

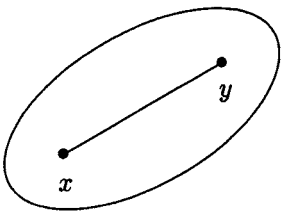


Figure 2.7: Convex set.

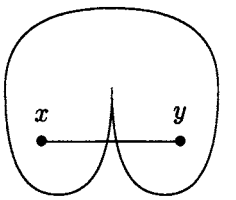


Figure 2.8: Non-convex set.

If T is a cone, then convexity is simply characterized.

Theorem 2.1: A cone $C \subset \mathbb{R}^m$ is convex if and only if

$$x + y \in C \quad \text{for all } x, y \in C. \tag{2.2}$$

Proof:

(a) Let C be a convex cone. Then it follows for all $x, y \in C$

$$\frac{1}{2}(x+y) = \frac{1}{2}x + \frac{1}{2}y \in C$$

which implies $x + y \in C$.

(b) For arbitrary $x, y \in C$ and $\lambda \in [0,1]$ we have $\lambda x \in C$ and $(1-\lambda)y \in C$. Then we get with the condition (2.2)

$$\lambda x + (1-\lambda)y \in C.$$

Consequently, the cone C is convex.

■

We now come to a central relationship between a partial ordering and a convex cone.

Theorem 2.2:

(a) If $\leq$ is a partial ordering on $\mathbb{R}^m$, then the set

$$C := \{x \in \mathbb{R}^m \mid 0_{\mathbb{R}^m} \leq x\}$$

is a convex cone. If, in addition, the partial ordering is antisymmetric, then C is pointed.

(b) If C is convex cone in $\mathbb{R}^m$, then the binary relation

$$\leq := \{(x,y) \in \mathbb{R}^m \times \mathbb{R}^m \mid y - x \in C\}$$

is a partial ordering on $\mathbb{R}^m$. If, in addition, C is pointed, then the partial ordering $\leq$ is antisymmetric.

Proof:

(a) Since the partial ordering $\leq$ is compatible with the scalar multiplication, it is evident that C is a cone. Now, take arbitrary $x, y \in C$. Then $0_{\mathbb{R}^m} \leq x$ and $0_{\mathbb{R}^m} \leq y$, and with property (iii) in Definition 2.1, (b) we obtain $0_{\mathbb{R}^m} \leq x + y$ which means $x + y \in C$. By Theorem 2.1 we conclude that C is convex. Next, assume that $x \in C$ and $-x \in C$ implying $0_{\mathbb{R}^m} \leq x$ and $x \leq 0_{\mathbb{R}^m}$. Then by the antisymmetry of $\leq$ we get $x = 0_{\mathbb{R}^m}$. Hence, C is pointed in this case.

(b) Since $0_{\mathbb{R}^m}$ is always contained in a cone, the binary relation $\leq$ is reflexive. For $x \leq y$ and $y \leq z$ $(x, y, z \in \mathbb{R}^m)$ we have $y - x \in C$ and $z - y \in C$. The convexity of C then implies $z - x \in C$ or $x \leq z$. This shows the transitivity of $\leq$. For arbitrary $x, y, w, z \in \mathbb{R}^m$ with $x \leq y$ and $w \leq z$ we have $y - x \in C$ and $z - w \in C$ and because of the convexity we conclude $(y + z) - (x + w) \in C$, i.e. $x + w \leq y + z$. So, the binary relation $\leq$ is compatible with the addition. For the proof of the compatibility with the scalar multiplication take arbitrary $x, y \in \mathbb{R}^m$ and $\alpha \in \mathbb{R}_+$ with $x \leq y$. Then $y - x \in C$, and because C is a cone we conclude $\alpha y - \alpha x \in C$ or $\alpha x \leq \alpha y$. Finally, assume that the cone C is pointed and take any $x, y \in \mathbb{R}^m$ with $x \leq y$ and $y \leq x$. Then $y - x \in C$ and $-(y - x) \in C$ implying $y - x = 0_{\mathbb{R}^m}$ or $x = y$. Consequently, the partial ordering $\leq$ is antisymmetric.

■

Definition 2.4: A convex cone characterizing the partial ordering on $\mathbb{R}^m$ is called an *ordering cone* (or also a *positive cone*).

Example 2.4:

(a) For the componentwise partial ordering given in Example 2.2 the ordering cone is given in Example 2.3.

(b) The ordering cone

$$C := \{x \in \mathbb{R}^2 \mid x_1 \geq 0 \text{ and } x_2 = 0\}$$

induces the partial ordering $\leq_C$ with

$$x \leq_C y \iff x_1 \leq y_1 \text{ and } x_2 = y_2.$$

■

In the following we denote the partial ordering induced by a given ordering cone $C \subset \mathbb{R}^m$ by $\leq_C$. And we assume that the ordering cone is pointed implying that the induced partial ordering $\leq_C$ is antisymmetric.

2.3 VARIOUS CONCEPTS OF SOLUTIONS IN VECTOR MAXIMIZATION

We come back to the vector optimization problem (2.1) with $\emptyset \neq S \subset \mathbb{R}^n$ and $f : S \to \mathbb{R}^m$. Now we assume that the vector space $\mathbb{R}^m$ is partially ordered by a binary relation $\leq_C$ induced by a pointed ordering cone C. This is the standard assumption for this section.

Definition 2.5: Let T be an arbitrary nonempty subset of $\mathbb{R}^m$.

(a) $\bar{y} \in T$ is called a *maximal* element of T, if there is no $y \in T$ with $y \neq \bar{y}$ and $\bar{y} \leq_C y$ (see Fig. 2.9).

(b) $\bar{y} \in T$ is called a *minimal* element of T, if there is no $y \in T$ with $y \neq \bar{y}$ and $y \leq_C \bar{y}$ (see Fig. 2.9).

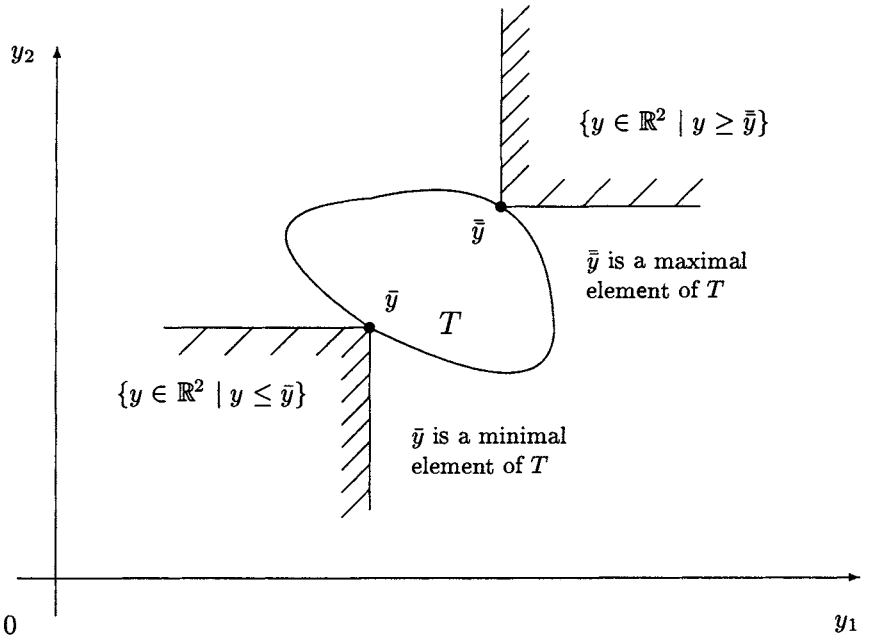


Figure 2.9: Maximal and minimal elements of T with respect to $C := \mathbb{R}^2_+$.

Example 2.5: We consider the unit circle in $\mathbb{R}^2$

$$T := \{(y_1, y_2) \in \mathbb{R}^2 \mid y_1^2 + y_2^2 \leq 1\},$$

and we assume that $C := \mathbb{R}^2_+$. The set of all maximal elements of T is given as

$$\left\{(y_1, y_2) \in \mathbb{R}^2 \mid y_1 \in [0,1] \text{ and } y_2 = \sqrt{1 - y_1^2}\right\},$$

and the set of all minimal elements of T reads

$$\left\{(y_1, y_2) \in \mathbb{R}^2 \mid y_1 \in [-1,0] \text{ and } y_2 = -\sqrt{1 - y_1^2}\right\}$$

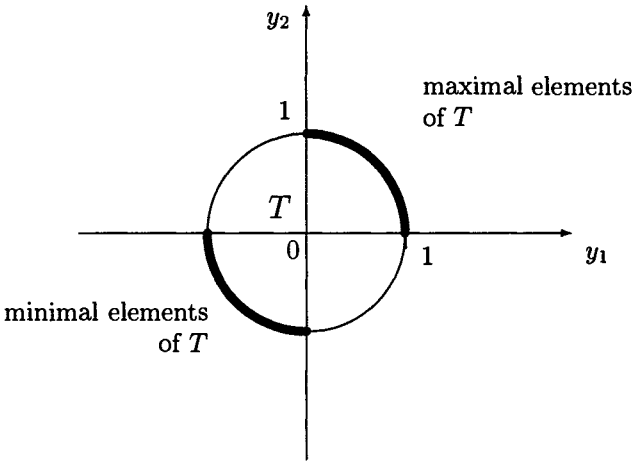


Figure 2.10: Maximal and minimal elements of the unit circle.

(see Fig. 2.10). ∎

In practice the maximal elements of $T := f(S)$ do not play the central role but their preimages.

Definition 2.6: $\bar{x} \in S$ is called an *efficient solution* (or an *Edgeworth-Pareto optimal point* or a *maximal solution* or a *nondominated point*) of problem (2.1), if $f(\bar{x})$ is a maximal element of the image set $f(S)$.

The notion of efficient solutions is often used in economics whereas the notion "Edgeworth-Pareto optimal" can be found in engineering, and in applied mathematics one speaks of maximal solutions. The efficiency concept has been introduced by Koopmans [15].

Example 2.6: Consider the constraint set

$$S := \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 - x_2 \leq 0, \quad x_1 + 2x_2 - 3 \leq 0\}$$

and the vector function $f : S \to \mathbb{R}^2$ with

$$f(x_1, x_2) = \begin{pmatrix} x_1 \\ -x_1 - x_2^2 \end{pmatrix} \quad \text{for all } (x_1, x_2) \in S.$$

Assume that $C := \mathbb{R}^2_+$. The point $(-\frac{3}{2}, -\frac{57}{16})$ is the only minimal element of $T := f(S)$, and the set of all maximal elements of T reads

$$\left\{(y_1, y_2) \in \mathbb{R}^2 \mid y_1 \in \left[-\frac{1}{2}\sqrt[3]{2}, 1\right] \text{ and } y_2 = y_1 - y_1^4\right\}.$$

The set of all efficient solutions is given as

$$\left\{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 \in \left[-\frac{1}{2}\sqrt[3]{2}, 1\right] \text{ and } x_2 = x_1^2\right\}$$

(see Fig. 2.11). ∎

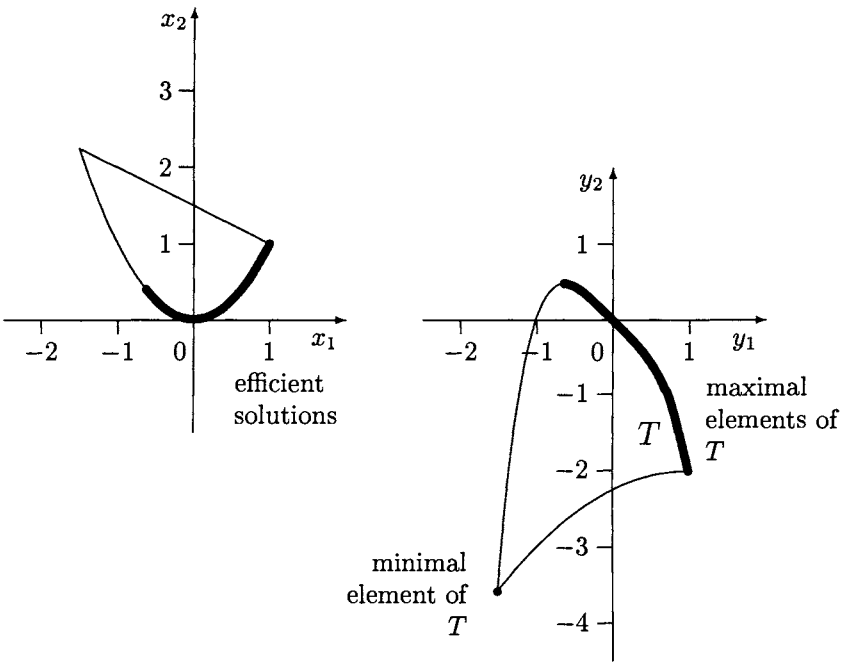


Figure 2.11: Maximal and minimal elements of T.

The efficiency concept is the main optimality notion used in vector optimization. But there are also other concepts being more weakly or more strongly formulated. First we present a weaker optimality notion.

Definition 2.7: Let the pointed ordering cone C have a nonempty interior int(C). $\bar{x} \in S$ is called a *weakly efficient solution* (or a *weakly Edgeworth-Pareto optimal point* or a *weakly maximal solution*) of problem (2.1), if there is no $x \in S$ with

$$f(x) - f(\bar{x}) \in \text{int}(C).$$

This weak efficiency notion is often only used if it is difficult to characterize theoretically efficient solutions or to determine them numerically. In general, in the applications one is not interested in weakly efficient solutions; this optimality notion is only of mathematical interest.

Example 2.7: We consider the vector optimization problem (2.1) with the set

$$S := \left\{(x_1, x_2) \in \mathbb{R}^2 \mid 0 \leq x_1 \leq 1,\ 0 \leq x_2 \leq 1\right\},$$

the identity $f : S \to \mathbb{R}^2$ with

$$f(x_1, x_2) = (x_1, x_2) \quad \text{for all} \quad (x_1, x_2) \in S,$$

and we assume $C := \mathbb{R}^2_+$. S describes a square in $\mathbb{R}^2$. Since f is the identity, the image set $f(S)$ equals S. The point (1,1) is the only efficient solution whereas the set

$$\left\{(x_1, x_2) \in S \mid x_1 = 1 \text{ or } x_2 = 1\right\}$$

is the set of all weakly efficient solutions (see Fig. 2.12). ∎

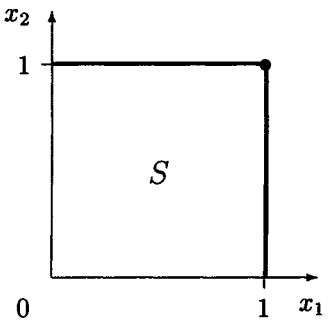


Figure 2.12: Weakly efficient solutions.

In the preceding example the set of efficient solutions is contained in the set of weakly efficient solutions. This fact holds in general (in the case of int(C) $\neq \emptyset$) and is proved in the next theorem.

Theorem 2.3: Let the pointed ordering cone C have a nonempty interior. Every efficient solution of problem (2.1) is a weakly efficient solution of problem (2.1).

Proof: If $\bar{x} \in S$ is an efficient solution of problem (2.1), then there is no $x \in S$ with $f(x) \neq f(\bar{x})$ and

$$f(\bar{x}) \leq_C f(x),$$

or equivalently

$$f(x) - f(\bar{x}) \in C.$$

Consequently, there is also no $x \in S$ with

$$f(x) - f(\bar{x}) \in \operatorname{int}(C).$$

This means that $\bar{x} \in S$ is a weakly efficient solution. ∎

Notice that the converse statement of Theorem 2.3 is not true in general (compare Example 2.7).

In the following we present a sharper optimality notion. It is the concept of properly efficient solutions introduced by Geoffrion [9] (but already earlier used by Kuhn and Tucker [16]).

Definition 2.8: Let $C := \mathbb{R}^m_+$ be given. $\bar{x} \in S$ is called a *properly efficient solution* (or a *properly Edgeworth-Pareto optimal point* or a *properly maximal solution*) of problem (2.1), if $\bar{x}$ is an efficient solution and there is a real number $\mu > 0$ so that for every $i \in \{1, \ldots, m\}$ and every $x \in S$ with $f_i(x) > f_i(\bar{x})$ at least one $j \in \{1, \ldots, m\}$ exists with $f_j(x) < f_j(\bar{x})$ and

$$\frac{f_i(x) - f_i(\bar{x})}{f_j(\bar{x}) - f_j(x)} \leq \mu.$$

An efficient solution which is not properly efficient is also called an *improperly efficient solution.*

In the applications improperly efficient solutions are not desired because a possible improvement of one component leads to a drastic deterioration of another component.

Example 2.8: For simplicity we investigate the vector optimization problem (2.1) with the unit circle

$$S := \left\{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1\right\},$$

the identity $f : S \to \mathbb{R}^2$ with

$$f(x_1, x_2) = (x_1, x_2) \quad \text{for all} \quad (x_1, x_2) \in S$$

and $C := \mathbb{R}^2_+$. By Example 2.5 the set of efficient solutions reads

$$\left\{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 \in [0, 1] \quad \text{and} \quad x_2 = \sqrt{1 - x_1^2}\right\}$$

(see Fig. 2.10). Except the points (0,1) and (1,0) all other efficient solutions are also properly efficient solutions. In the following we show that the point $\bar{x} := (0,1)$ is an improperly efficient solution. For an arbitrary $n \in \mathbb{N}$ consider the point $x(n) := \left(\frac{1}{n}\sqrt{2n-1}, 1-\frac{1}{n}\right)$ of the unit circle. For every $n \in \mathbb{N}$ we have $f_1\big(x(n)\big) > f_1(\bar{x})$ and $f_2\big(x(n)\big) < f_2(\bar{x})$, and we conclude

$$\frac{f_1(\bar{x}) - f_1\big(x(n)\big)}{f_2\big(x(n)\big) - f_2(\bar{x})} = \frac{\bar{x}_1 - x_1(n)}{x_2(n) - \bar{x}_2} = \frac{0 - \frac{1}{n}\sqrt{2n-1}}{1 - \frac{1}{n} - 1} = \sqrt{2n-1}$$

for all $n \in \mathbb{N}$.

It is obvious that an upper bound $\mu > 0$ of this term does not exist. Consequently, $\bar{x} = (0,1)$ is an improperly efficient solution. ∎

Example 2.9: It can be shown that one properly efficient solution of the design problem discussed in Example 2.1 is, for instance, the point $(10\sqrt[3]{4}, 5\sqrt[3]{4})$. This solution leads to a beam with the height $10\sqrt[3]{4} \approx 15.874$ and the width $5\sqrt[3]{4} \approx 7.937$. ∎

Until now there are more than a dozen variants of this proper efficiency concept (for instance, see [18]). Since Geoffrion's definition is based on the natural partial ordering in $\mathbb{R}^m$, extensions of this notion to more general partially ordered vector spaces are formally different from the original concept. Hartley [10] has formulated an extension of this concept to Hilbert spaces. Various generalizations use the tangential approximation of the image set $f(S)$ expressed by the contingent cone (see [13, p. 84]) or similar approximations (we refer to the papers of Borwein [2, 3], Vogel [19] and Benson [1]). Many generalizations define proper efficiency as efficiency with respect to a larger cone with suitable properties (we refer to the important contributions given by Wierzbicki [20–22] and Henig [11]). The notion of super efficiency introduced by Zhuang [23] in normed spaces is a possible generalization of the proper efficiency concept as well. It combines the idea of enlarging the ordering cone with the idea of using tangential approximations. It is shown in [18] that many of these generalized proper efficiency concepts coincide with Geoffrion's notion in the finite dimensional case with $C = \mathbb{R}^m_+$ and a convex image set $f(S)$. For further investigations of the proper efficiency conept we also refer to the book of Kaliszewski [14].

Next we come to a very strong optimality notion.

Definition 2.9: $\bar{x} \in S$ is called a *strongly efficient solution* (or a *strongly Edgeworth-Pareto optimal point* or a *strongly maximal solution*) of problem (2.1) if

$$f(x) \leq_C f(\bar{x}) \quad \text{for all} \quad x \in S.$$

This concept naturally generalizes the standard maximality notion used in scalar optimization. But it is clear that this concept is too strong for vector optimization problems.

Example 2.10: Consider the vector optimization problem in Example 2.7. Here the point (1,1) is a strongly efficient solution. The problem discussed in Example 2.8 has no strongly efficient solutions. ∎

Theorem 2.4: Every strongly efficient solution is an efficient solution.

Proof: Let $\bar{x} \in S$ be a strongly efficient solution, i.e.

$$f(x) \leq_C f(\bar{x}) \quad \text{for all} \quad x \in S.$$

Then there is no $x \in S$ with $f(x) \neq f(\bar{x})$ and $f(\bar{x}) \leq_C f(x)$. Hence, $\bar{x}$ is an efficient solution. ∎

Finally, we come to an optimality concept proposed by Brucker [4] for discrete problems.

Definition 2.10: $\bar{x} \in S$ is called an *essentially efficient solution* (or an *essentially Edgeworth-Pareto optimal point* or an *essentially maximal solution*) of problem (2.1), if $f(\bar{x})$ is a maximal element of the convex hull of the image set $f(S)$.

Since the image set $f(S)$ is contained in its convex hull it is evident that every essentially efficient solution $\bar{x} \in S$ is also an efficient solution. Morover, there is also a relationship to the strong efficiency concept.

Theorem 2.5: Every strongly efficient solution is an essentially efficient solution.

Proof: Let $\bar{x} \in S$ be a strongly efficient solution. Then we have

$$f(x) \leq_C f(\bar{x}) \quad \text{for all} \quad x \in S$$

or

$$f(S) \subset \{f(\bar{x})\} - C$$

("$-$" denotes the algebraic difference of sets). Since the set $\{f(\bar{x})\} - C$ is convex, we conclude for the convex hull $\text{co}(f(S))$ of $f(S)$ being the intersection of all convex subsets of $\mathbb{R}^m$ containing $f(S)$

$$\text{co}(f(S)) \subset \{f(\bar{x})\} - C.$$

Then there is no $y \in \text{co}(f(S))$ with $y \neq f(\bar{x})$ and $f(\bar{x}) \leq_C y$. Hence, $f(\bar{x})$ is a maximal element of the set $\text{co}(f(S))$, i.e. $\bar{x} \in S$ is an essentially efficient solution. ∎

Example 2.11: Consider the vector optimization problem (2.1) with the discrete constraint set

$$S := \{(0,3),\ (1,1),\ (3,0)\},$$

the identity as objective function f and $C := \mathbb{R}^2_+$ (see Fig. 2.13).

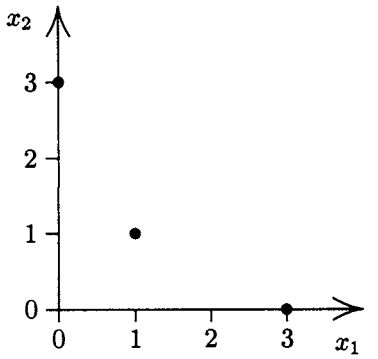


Figure 2.13: Constraint set S.

Every feasible point is an efficient solution, but only the points (3,0) and (0,3) are essentially efficient solutions. ∎

Summarizing the relationships between the presented optimality concepts we obtain the diagram in Table 2.1. Notice that the converse implications are not true in general.

strong efficiency
$\Downarrow$
essential efficiency
$\Downarrow$
proper efficiency $\overset{C=\mathbb{R}^m_+}{\Longrightarrow}$ efficiency $\overset{\text{int}(C)\neq\emptyset}{\Longrightarrow}$ weak efficiency

Table 2.1: Relationships between different solution concepts.

2.4 SCALARIZATION

In general, scalarization means the replacement of a vector optimization problem by a suitable scalar optimization problem which is an optimization problem with a real-valued objective function. This principle makes it possible that solutions of a vector optimimization problem can be characterized and also computed by using these scalar problems. In economics these problems are also called *auxiliary problems*, *auxiliary programs* or *compromise models* (for instance, see [6, 8]).

In this section we discuss appropriate scalar optimization problems which can be used for scalarization and we present the theoretical basics. These investigations enclose the efficiency and the weak and proper efficiency concept.

2.4.1 General Results

For the formulation of a general sufficient condition for efficient solutions of the vector optimization problem (2.1) we need appropriate monotonicity concepts. We use the same standard assumption as in Section 2.3.

Definition 2.11: Let M be a nonempty subset of $\mathbb{R}^m$.

(a) A function $\varphi : M \to \mathbb{R}$ is called *monotonically increasing on* M if

$$x, y \in M \text{ and } x \leq_C y \implies \varphi(x) \leq \varphi(y).$$

(b) A function $\varphi : M \to \mathbb{R}$ is called *strongly monotonically increasing on* M if

$$x, y \in M, x \neq y \text{ and } x \leq_C y \implies \varphi(x) < \varphi(y).$$

(c) Let C have a nonempty interior int(C). A function $\varphi : M \to \mathbb{R}$ is called *strictly monotonically increasing on* M if

$$x, y \in M \text{ and } y - x \in \text{int}(C) \implies \varphi(x) < \varphi(y).$$

It is evident in the case of int(C) $\neq \emptyset$ that every strongly monotonically increasing function is strictly monotonically increasing as well.

Example 2.12: Let $C = \mathbb{R}^m_+$.

(a) For arbitrary real numbers $t_1, \ldots, t_m \geq 0$ let the function $\varphi : \mathbb{R}^m \to \mathbb{R}$ be defined by

$$\varphi(y_1, \ldots, y_m) = \sum_{i=1}^{m} t_i y_i \text{ for all } (y_1, \ldots, y_m) \in \mathbb{R}^m.$$

Then φ is monotonically increasing on $\mathbb{R}^m$. In order to see this we fix arbitrary vectors $x, y \in \mathbb{R}^m$ with $x \leq_{\mathbb{R}^m_+} y$. Because of $t_1, \ldots, t_m \geq 0$ and

$$x_i \leq y_i \text{ for all } i \in \{1, \ldots, m\}$$

we conclude

$$\varphi(x) = \sum_{i=1}^{m} t_i x_i \leq \sum_{i=1}^{m} t_i y_i = \varphi(y).$$

This completes the proof. Moreover, one can show for $\sum_{i=1}^{m} t_i > 0$ that φ is strictly monotonically increasing on $\mathbb{R}^m$.

(b) For arbitrary real numbers $t_1, \ldots, t_m > 0$ let the function $\varphi : \mathbb{R}^m \to \mathbb{R}$ be defined by

$$\varphi(y_1, \ldots, y_m) = \sum_{i=1}^{m} t_i y_i \text{ for all } (y_1, \ldots, y_m) \in \mathbb{R}^m.$$

Then φ is strongly monotonically increasing on $\mathbb{R}^m$. For the proof of this assertion we choose arbitrary vectors $x, y \in \mathbb{R}^m$ with $x \neq y$ und $x \leq_{\mathbb{R}^m_+} y$. Then we have

$$t_i x_i \leq t_i y_i \text{ for all } i \in \{1, \ldots, m\}$$

where for at least one $i \in \{1, \ldots, m\}$ this inequality is strict. Consequently, we get

$$\varphi(x) = \sum_{i=1}^{m} t_i x_i < \sum_{i=1}^{m} t_i y_i = \varphi(y).$$

(c) In $\mathbb{R}^m$ norms can easily be given being strictly or strongly monotonically increasing on $M := \mathbb{R}^m_+$. Every weighted ℓ_p norm with $p \in [1, \infty)$ is strongly monotonically increasing on $\mathbb{R}^m_+$ (and, therefore, also strictly monotonically increasing on $\mathbb{R}^m_+$). For the proof take for an arbitrary $p \in [1, \infty)$ the weighted ℓ_p norm $\|\cdot\|_p$ with

$$\|y\|_p = \left(\sum_{i=1}^{m} w_i |y_i|^p \right)^{\frac{1}{p}} \text{ for all } y \in \mathbb{R}^m.$$

Let the weights $w_1, \ldots, w_m$ be positive real numbers (see Fig. 2.14). For arbitrary vectors $x, y \in \mathbb{R}^m_+$ with $x \neq y$ and $x \leq_{\mathbb{R}^m_+} y$ it follows

$$0 \leq x_i^p \leq y_i^p \text{ for all } i \in \{1, \ldots, m\}$$

where the right inequality is strict for at least one $i \in \{1, \ldots, m\}$. Then we conclude

$$\|x\|_p = \left(\sum_{i=1}^{m} w_i |x_i|^p \right)^{\frac{1}{p}} < \left(\sum_{i=1}^{m} w_i |y_i|^p \right)^{\frac{1}{p}} = \|y\|_p.$$

Hence, the weighted ℓ_p norm $\|\cdot\|_p$ with $p \in [1, \infty)$ is strongly monotonically increasing on $\mathbb{R}^m_+$.

On the other hand the weighted maximum norm $\|\cdot\|_\infty$ with

$$\|y\|_\infty = \max_{i \in \{1,\ldots,m\}} \{w_i |y_i|\} \text{ for all } y \in \mathbb{R}^m$$

$(w_1, \ldots, w_m > 0)$ is strictly monotonically increasing on $\mathbb{R}^m_+$. For the proof take arbitrary vectors $x, y \in \mathbb{R}^m_+$ with

$$x_i < y_i \text{ for all } i \in \{1, \ldots, m\}.$$

Then we obtain

$$\|x\|_\infty = \max_{i \in \{1,\ldots,m\}} \{w_i |x_i|\} < \max_{i \in \{1,\ldots,m\}} \{w_i |y_i|\} = \|y\|_\infty.$$

(d) Let the function $\varphi : \mathbb{R}^m \to \mathbb{R}$ be defined by

$$\varphi(y_1, \ldots, y_m) = \max_{i \in \{1,\ldots,m\}} \{y_i\} \text{ for all } (y_1, \ldots, y_m) \in \mathbb{R}^m.$$

Then φ is strictly monotonically increasing on $\mathbb{R}^m$. For the proof of this assertion choose arbitrary vectors $x, y \in \mathbb{R}^m$ with

$$x_i < y_i \text{ for all } i \in \{1, \ldots, m\}.$$

Then it follows

$$\varphi(x) = \max_{i \in \{1,\ldots,m\}} \{x_i\} < \max_{i \in \{1,\ldots,m\}} \{y_i\} = \varphi(y).$$

■

The following theorem gives a basic scalarization result.

Theorem 2.6:

(a) Let $\varphi : f(S) \to \mathbb{R}$ be a function being strongly monotonically increasing on $f(S)$. If there is an $\bar{x} \in S$ with

$$\varphi(f(\bar{x})) \geq \varphi(f(x)) \text{ for all } x \in S, \tag{2.3}$$

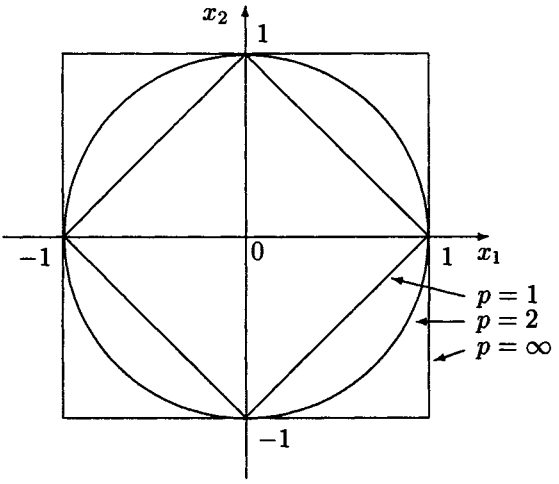


Figure 2.14: Unit spheres of the ℓ_p norms with weights $w_1 = w_2 = 1$ in $\mathbb{R}^2$.

then $\bar{x}$ is an efficient solution of problem (2.1).

(b) Let $\varphi : f(S) \to \mathbb{R}$ be a function being monotonically increasing on $f(S)$. If there is an $\bar{x} \in S$ with

(2.4) $$\varphi(f(\bar{x})) > \varphi(f(x)) \text{ for all } x \in S \text{ with } f(x) \neq f(\bar{x}),$$

then $\bar{x}$ is an efficient solution of problem (2.1).

(c) Let C have a nonempty interior $\operatorname{int}(C)$, and let $\varphi : f(S) \to \mathbb{R}$ be a function being strictly monotonically increasing on $f(S)$. If there is an $\bar{x} \in S$ with

(2.5) $$\varphi(f(\bar{x})) \geq \varphi(f(x)) \text{ for all } x \in S,$$

then $\bar{x}$ is a weakly efficient solution of problem (2.1).

(d) Let C equal $\mathbb{R}^m_+$, and let $t_1, \ldots, t_m > 0$ be given real numbers. If there is an $\bar{x} \in S$ with

(2.6) $$\sum_{i=1}^{m} t_i f_i(\bar{x}) \geq \sum_{i=1}^{m} t_i f_i(x) \text{ for all } x \in S,$$

then $\bar{x}$ is a properly efficient solution of problem (2.1).

Proof: In the first two cases we assume that $\bar{x}$ is no efficient solution. Then there is an $x \in S$ with $f(\bar{x}) \leq_C f(x)$ and $f(x) \neq f(\bar{x})$. In part (a) we obtain $\varphi(f(\bar{x})) < \varphi(f(x))$, a contradiction to the inequality (2.3). In part (b) it follows $\varphi(f(\bar{x})) \leq \varphi(f(x))$, a contradiction to the inequality (2.4).

For the proof of part (c) we assume that $\bar{x}$ is no weakly efficient solution of problem (2.1). Then there is an $x \in S$ with

$$f(x) - f(\bar{x}) \in \operatorname{int}(C).$$

Since φ is strictly monotonically increasing on $f(S)$, it follows

$$\varphi(f(\bar{x})) < \varphi(f(x)),$$

a contradiction to the inequality (2.5).

We now prove the assertion in part (d). Since the function $\varphi : f(S) \to \mathbb{R}$ with

$$\varphi(y_1, \ldots, y_m) = \sum_{i=1}^{m} t_i y_i \quad \text{for all} \quad y \in f(S)$$

is strongly monotonically increasing on $f(S)$ (compare Example 2.12, (b)), by part (a) of this theorem $\bar{x}$ is an efficient solution of problem (2.1). Assume that $\bar{x}$ is no properly efficient solution. Then we choose

$$\mu := (m-1) \max_{i,j \in \{1,\ldots,m\}} \left\{ \frac{t_j}{t_i} \right\} \quad \text{for} \quad m \geq 2,$$

and we obtain for some $i \in \{1, \ldots, m\}$ and some $x \in S$ with $f_i(x) > f_i(\bar{x})$

$$\frac{f_i(x) - f_i(\bar{x})}{f_j(\bar{x}) - f_j(x)} > \mu \quad \text{for all} \quad j \in \{1, \ldots, m\} \quad \text{with} \quad f_j(x) < f_j(\bar{x}).$$

This implies

$$f_i(x) - f_i(\bar{x}) > \mu\,(f_j(\bar{x}) - f_j(x)) \geq (m-1)\frac{t_j}{t_i}\,(f_j(\bar{x}) - f_j(x))$$

$$\text{for all} \quad j \in \{1, \ldots, m\} \setminus \{i\}.$$

Multiplication with $\frac{t_i}{m-1}$ and summation with respect to $j \neq i$ leads to

$$t_i\,(f_i(x) - f_i(\bar{x})) > \sum_{\substack{j=1 \\ j \neq i}}^{m} t_j (f_j(\bar{x}) - f_j(x))$$

and

$$0 > \sum_{j=1}^{m} t_j (f_j(\bar{x}) - f_j(x))$$

implying

$$\sum_{j=1}^{m} t_j f_j(\bar{x}) < \sum_{j=1}^{m} t_j f_j(x)$$

contradicting to the inequality (2.6). ∎

By Theorem 2.6, (a) - (c) efficient and weakly efficient solutions of problem (2.1) can be obtained as solutions of the scalar optimization problems

$$\max_{x \in S} (\varphi \circ f)(x) = \max_{x \in S} \varphi(f(x))$$

where in the case (b) the image uniqueness of the solution is additionally required. By Theorem 2.6, (d) for arbitrary positive numbers $t_1, \ldots, t_m$ every solution of the scalar optimization problem

$$\max_{x \in S} \sum_{i=1}^{m} t_i f_i(x)$$

is a properly efficient solution of problem (2.1).

2.4.2 Weighted Sum Approach

If one combines the assertions of Theorem 2.6, (b), (c) for $C := \mathbb{R}^m_+$ with the remarks in Example 2.12, (a), (b) (i.e., φ is chosen as a special linear function), then we obtain the scalarization results given in Table 2.2. The result in Theorem 2.6, (d) is also considered. This approach uses the weighted sum of the components of the objective vector function. Therefore, one speaks of a *weighted sum approach* (for instance, see [5, 6, 8]).

Example 2.13:

(a) In Example 2.6 we have already investigated the following vector optimization problem (see also Fig. 2.11):

$$\left\{ \begin{array}{c} \text{"max"} \begin{pmatrix} x_1 \\ -x_1 - x_2^2 \end{pmatrix} \\ \text{subject to the constraints} \\ x_1^2 - x_2 \leq 0 \\ x_1 + 2x_2 - 3 \leq 0. \end{array} \right. \tag{2.7}$$

Every solution of the scalar optimization problem $\max_{x \in S} \sum_{i=1}^{m} t_i f_i(x)$ with		
$t_1, \ldots, t_m > 0$	$t_1, \ldots, t_m \geq 0$, $t_i > 0$ for some $i \in \{1, \ldots, m\}$, where image uniqueness of the solution is given	$t_1, \ldots, t_m \geq 0$, $t_i > 0$ for some $i \in \{1, \ldots, m\}$,
is		
a properly efficient solution of problem (2.1).	an efficient solution of problem (2.1).	a weakly efficient solution of problem (2.1).

Table 2.2: Sufficient conditions for optimal solutions.

For the computation of a properly efficient solution of this problem one can choose, for instance, $t_1 = 1$ and $t_2 = 2$. Then one solves the scalar optimization problem

$$\left\{ \begin{array}{c} \max\ -x_1 - 2x_2^2 \\ \text{subject to the constraints} \\ x_1^2 - x_2 \leq 0 \\ x_1 + 2x_2 - 3 \leq 0. \end{array} \right. \tag{2.8}$$

$\bar{x} = (-\frac{1}{2}, \frac{1}{4})$ is the unique solution of the problem (2.8). By Theorem 2.6, (d) $\bar{x}$ is also a properly efficient solution of the vector optimization problem (2.7).

(b) The application of Theorem 2.6 to discrete problems allows a fast computation of efficient solutions. As a very simple example (see [7, p. 165]) all maximal elements of the discrete set

$$S := \{(16,9), (6,14), (11,13), (10,10)\} \subset \mathbb{R}^2$$

are determined. For this purpose we choose the vector function $f : S \to \mathbb{R}^2$ with

$$f(x_1, x_2) = (x_1, x_2) \text{ for all } (x_1, x_2) \in S.$$

The maximal elements of S are exactly the efficient solutions of the problem

$$\text{(2.9)} \qquad \text{"}\max_{x \in S}\text{"} \ f(x).$$

For the computation of these efficient solutions one can choose the weight vector $t = (\alpha, 1-\alpha)$ with $\alpha \in (0,1)$ and obtains the scalar optimization problem

$$\max_{(x_1,x_2) \in S} \alpha x_1 + (1-\alpha)x_2$$

for arbitrary $\alpha \in (0,1)$. The maximal elements of the set S are given in Table 2.3.

α	$\bar{x}$	$\alpha\bar{x}_1 + (1-\alpha)\bar{x}_2$
$0 < \alpha < \frac{1}{6}$	$(6,14)$	$6\alpha + 14(1-\alpha)$
$\alpha = \frac{1}{6}$	$(6,14)$ or $(11,13)$	$\frac{38}{3}$
$\frac{1}{6} < \alpha < \frac{4}{9}$	$(11,13)$	$11\alpha + 13(1-\alpha)$
$\alpha = \frac{4}{9}$	$(11,13)$ or $(16,9)$	$\frac{109}{9}$
$\frac{4}{9} < \alpha < 1$	$(16,9)$	$16\alpha + 9(1-\alpha)$

Table 2.3: Maximal elements and function values for different parameters (Example 2.13, (b)).

(c) The result of Theorem 2.6 can be well applied in linear vector optimization. As an example let us determine all efficient solutions of the following problem (see [5, pp. 155]):

$$\text{“max”} \begin{pmatrix} 4x_1 + 2x_2 \\ 8x_1 + 10x_2 \end{pmatrix}$$
subject to the constraints
$$x_1 + x_2 \leq 70$$
$$x_1 + 2x_2 \leq 100$$
$$x_1 \leq 60$$
$$x_2 \leq 40$$
$$x_1,\ x_2 \geq 0.$$

The constraint set of this example is illustrated in Fig. 2.15. Again, let

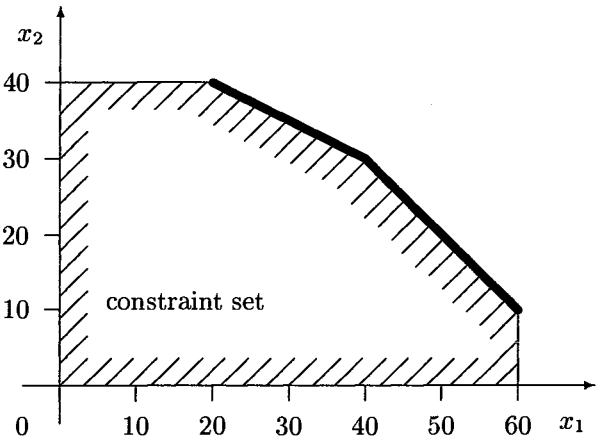


Figure 2.15: Constraint set in Example 2.13, (c).

the vector t of the weights be given as

$$t = \begin{pmatrix} \alpha \\ 1 - \alpha \end{pmatrix} \text{ with } \alpha \in (0,1).$$

Consequently, one obtains for $\alpha \in (0,1)$ the parametric optimization problem

$$\max\ (8 - 4\alpha)x_1 + (10 - 8\alpha)x_2$$
subject to the constraints
$$x_1 + x_2 \leq 70$$
$$x_1 + 2x_2 \leq 100$$
$$x_1 \leq 60$$
$$x_2 \leq 40$$
$$x_1, x_2 \geq 0.$$

All solutions of this problem are given in Table 2.4. These are also efficient solutions of the considered vector optimization problem.

■

α	$\bar{x}_1$	$\bar{x}_2$
$0 < \alpha < \frac{2}{5}$	20	40
$\frac{2}{5}$	$20\lambda + 40(1-\lambda)$	$40\lambda + 30(1-\lambda)$
$\frac{2}{5} < \alpha < \frac{6}{7}$	40	30
$\frac{6}{7}$	$40\lambda + 60(1-\lambda)$	$30\lambda + 10(1-\lambda)$
$\frac{6}{7} < \alpha < 1$	60	10

Table 2.4: Efficient solutions for different parameters (Example 2.13, (c)). $\lambda \in [0,1]$ can be arbitrarily chosen.

Notice that for general nonlinear vector optimization problems not *every* efficient solution can be determined using the weighted sum approach. For instance, Figure 2.16 shows that only two maximal points of the set T can be determined in such a way. Only these two points are supporting points of an appropriate supporting function.

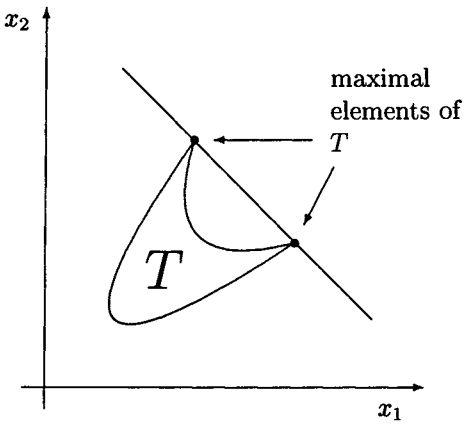


Figure 2.16: Weighted sum approach in the nonconvex case.

The weighted sum approach seems to be only suitable for convex problems,

like linear problems. In general, this approach cannot be used for vector optimization problems arising in engineering. For these problems other approaches, for instance, like the weighted maximum norm approach, are more suitable. Next, we answer the question for which special problems the weighted sum approach is appropriate.

Theorem 2.7: Let $C = \mathbb{R}^m_+$ be given, and let the set

$$f(S) - \mathbb{R}^m_+ := \{y \in \mathbb{R}^m \mid y \leq_{\mathbb{R}^m_+} f(x) \text{ for some } x \in S\}$$

be convex. If $\bar{x} \in S$ is a weakly efficient solution of problem (2.1), then there are real numbers $t_1, \ldots, t_m \geq 0$, with $t_i > 0$ for at least one $i \in \{1, \ldots, m\}$, so that

$$\sum_{i=1}^{m} t_i f_i(\bar{x}) \geq \sum_{i=1}^{m} t_i f_i(x) \text{ for all } x \in S.$$

Proof: Let $\bar{x} \in S$ be a weakly efficient solution of problem (2.1), then there is no $x \in S$ with

$$f_i(x) > f_i(\bar{x}) \text{ for all } i \in \{1, \ldots, m\}.$$

If we define the sets

$$A := \{y \in \mathbb{R}^m \mid y_i > f_i(\bar{x}) \text{ for all } i \in \{1, \ldots, m\}\}$$

and

$$B := \{y \in \mathbb{R}^m \mid y \leq_{\mathbb{R}^m_+} f(x) \text{ for some } x \in S\},$$

we conclude

$$A \cap B = \emptyset.$$

It is evident that the set A is convex and open. By assumption the set B is convex as well. By the Eidelheit separation theorem (see [13, Thm. C.2]) there are real numbers $t_1, \ldots, t_m$, with $t_i \neq 0$ for at least one $i \in \{1, \ldots, m\}$, and a real number α so that

$$\sum_{i=1}^{m} t_i a_i > \alpha \geq \sum_{i=1}^{m} t_i b_i \qquad \text{for all } (a_1, \ldots, a_m) \in A \text{ and all } (b_1, \ldots, b_m) \in B. \tag{2.10}$$

If $\mathrm{cl}(A)$ denotes the closure of A, i.e.

$$\mathrm{cl}(A) = \{y \in \mathbb{R}^m \mid y_i \geq f_i(\bar{x}) \text{ for all } i \in \{1, \ldots, m\}\},$$

then $f(\bar{x}) \in \mathrm{cl}(A) \cap B$, and one obtains by (2.10) $\alpha = \sum_{i=1}^{m} t_i f_i(\bar{x})$. Then we conclude

$$\sum_{i=1}^{m} t_i \left(f_i(\bar{x}) + z_i\right) \geq \sum_{i=1}^{m} t_i f_i(\bar{x}) \quad \text{for all } z_1, \ldots, z_m \geq 0$$

or

$$\sum_{i=1}^{m} t_i z_i \geq 0 \quad \text{for all } z_1, \ldots, z_m \geq 0.$$

This implies that $t_1, \ldots, t_m$ are nonnegative. Finally, we get from the right inequality in (2.10)

$$\sum_{i=1}^{m} t_i f_i(\bar{x}) \geq \sum_{i=1}^{m} t_i f_i(x) \quad \text{for all } x \in S.$$

This completes the proof. ∎

The preceding theorem can easily be extended to arbitrary ordering cones with nonempty interior. For cones with empty interior one has to apply a special separation theorem in $\mathbb{R}^m$.

In order to point out the importance of the weighted sum approach for convex vector optimization problems (i.e., for problems for which $f(S) - \mathbb{R}^m_+$ is a convex set), we combine the results of Theorem 2.6, (c) (in connection with Example 2.12, (a)) and Theorem 2.7 in the following corollary.

Corollary 2.1: Let $C = \mathbb{R}^m_+$ be given, and let the set

$$f(S) - \mathbb{R}^m_+ := \{y \in \mathbb{R}^m \mid y \leq_{\mathbb{R}^m_+} f(x) \text{ for some } x \in S\}$$

be convex. Then $\bar{x} \in S$ is a weakly efficient solution of problem (2.1) if and only if there are real numbers $t_1, \ldots, t_m \geq 0$, with $t_i > 0$ for at least one $i \in \{1, \ldots, m\}$, so that

$$\sum_{i=1}^{m} t_i f_i(\bar{x}) \geq \sum_{i=1}^{m} t_i f_i(x) \text{ for all } x \in S.$$

Although the concept of weak efficiency is not suitable in applications (see Example 2.7), the preceding corollary shows that a theoretically elegant characterization is possible under convexity assumptions. The following corollary illustrates that the efficiency concept cannot be treated in a theoretically smooth way.

Corollary 2.2: Let $C = \mathbb{R}^m_+$ be given.

(a) Let $t_1, \ldots, t_m > 0$ be given real numbers. If there is an $\bar{x} \in S$ with

$$\sum_{i=1}^{m} t_i f_i(\bar{x}) \geq \sum_{i=1}^{m} t_i f_i(x) \text{ for all } x \in S,$$

then $\bar{x}$ is an efficient solution of problem (2.1).

(b) Let the set

$$f(S) - \mathbb{R}^m_+ := \{y \in \mathbb{R}^m \mid y \leq_{\mathbb{R}^m_+} f(x) \text{ for some } x \in S\}$$

be convex. If $\bar{x} \in S$ is an efficient solution of problem (2.1), then there are real numbers $t_1, \ldots, t_m \geq 0$, with $t_i > 0$ for at least one $i \in \{1, \ldots, m\}$, so that

$$\sum_{i=1}^{m} t_i f_i(\bar{x}) \geq \sum_{i=1}^{m} t_i f_i(x) \text{ for all } x \in S.$$

Proof: Part (a) of the assertion follows from Theorem 2.6, (a) in connection with Example 2.12, (b). Part (b) immediately follows from Theorem 2.7 in connection with Theorem 2.3. ∎

In economics the result of Corollary 2.2 is also called *efficiency theorem.* Kuhn and Tucker [16] have already given a first formulation of this result (for instance, see also [5, pp. 160] and [17, pp. 57]).

We now complete these weighted sum approaches with an investigation of the proper efficiency concept.

Corollary 2.3: Let $C = \mathbb{R}^m_+$ be given, and let the set

$$f(S) - \mathbb{R}^m_+ := \{y \in \mathbb{R}^m \mid y \leq_{\mathbb{R}^m_+} f(x) \text{ for some } x \in S\}$$

be convex. Then $\bar{x} \in S$ is a properly efficient solution of problem (2.1) if and only if there are real numbers $t_1, \ldots, t_m > 0$ so that

$$\sum_{i=1}^{m} t_i f_i(\bar{x}) \geq \sum_{i=1}^{m} t_i f_i(x) \text{ for all } x \in S.$$

Proof: One part of the assertion is shown in Theorem 2.6, (d). For the converse part see [9]. ∎

The results concerning the weighted sum approach are summarized in Table 2.5. The corresponding mathematical results can be found in the Corollaries 2.1, 2.2 and 2.3.

<table>
<tr><td colspan="3">If the set $f(S) - \mathbb{R}^m_+$ is convex, then
a solution of the scalar optimization problem

$$\max_{x \in S} \sum_{i=1}^{m} t_i f_i(x)$$

is</td></tr>
<tr><td>a properly efficient solution of problem (2.1)</td><td>an efficient solution of problem (2.1)</td><td>a weakly efficient solution of problem (2.1)</td></tr>
<tr><td colspan="3">if and only if</td></tr>
<tr><td>$t_1, \ldots, t_m > 0$.</td><td>$t_1, \ldots, t_m > 0$ (suff. cond.).
$t_1, \ldots, t_m \geq 0$, $t_i > 0$ for some $i \in \{1, \ldots, m\}$, (necess. cond.).</td><td>$t_1, \ldots, t_m \geq 0$, $t_i > 0$ for some $i \in \{1, \ldots, m\}$.</td></tr>
</table>

Table 2.5: Necessary and sufficient conditions for optimal solutions.

In economics vector optimization problems are very often linear, i.e. they are of the form

$$\text{(2.11)} \qquad \left\{ \begin{array}{c} \text{"max"}\ Cx \\ \text{subject to the constraints} \\ Ax \leq b \\ x \in \mathbb{R}^n \end{array} \right.$$

where C and A are appropriate matrices (notice that $\mathbb{R}^m_+$ is the ordering cone) and b is a given vector (compare Example 2.13, (c)). For these problems the set $f(S) - \mathbb{R}^m_+$ is always convex and, therefore, the results in Table 2.5 can be applied. Moreover, it can be shown that efficient solutions and properly efficient solutions coincide in this case. This is the result of the following theorem.

Theorem 2.8: Let the linear vector optimization problem (2.11) be given where C is a real (m, n) matrix, A is a real (k, n) matrix and $b \in \mathbb{R}^k$ is a given vector. Let the constraint set

$$S := \{x \in \mathbb{R}^n \mid Ax \leq b\}$$

be nonempty, and let $\mathbb{R}^m_+$ be the ordering cone. Then $\bar{x} \in S$ is an efficient solution of the linear vector optimization problem (2.11) if and only if $\bar{x} \in S$ is a properly efficient solution of problem (2.11).

Proof: See Geoffrion [9, Theorem 2, p. 620]. ∎

REFERENCES

[1] Benson, H.P.: An Improved Definition of Proper Efficiency for Vector Maximization with Respect to Cones. J. Math. Anal. Appl. 71 (1979) 232–241

[2] Borwein, J.M.: Proper Efficient Points for Maximizations with Respect to Cones. SIAM J. Control Optim. 15 (1977) 57–63

[3] Borwein, J.M.: The Geometry of Pareto Efficiency over Cones. Math. Operationsforsch. Statist. Ser. Optim. 11 (1980) 235–248

[4] Brucker, P.: Diskrete parametrische Optimierungsprobleme und wesentlich effiziente Punkte. Z. Oper. Res. 16 (1972) 189–197

[5] Dinkelbach W.: *Sensitivitätsanalysen und parametrische Programmierung.* Springer, Berlin 1969

[6] Dinkelbach, W.: "Über einen Lösungsansatz zum Vektormaximumproblem". In: M. Beckmann (ed.): *Unternehmensforschung Heute.* Springer, Berlin 1971, pp. 1–13

[7] Dinkelbach W.: *Entscheidungsmodelle.* de Gruyter, Berlin 1982

[8] Fandel G.: *Optimale Entscheidung bei mehrfacher Zielsetzung.* Springer, Berlin 1972

[9] Geoffrion, A.M.: Proper Efficiency and the Theory of Vector Maximization. J. Math. Anal. Appl. 22 (1968) 618–630

[10] Hartley, R.: On Cone-Efficiency, Cone-Convexity and Cone-Compactness. SIAM J. Appl. Math. 34 (1978) 211–222

[11] Henig, M.I.: Proper Efficiency with Respect to Cones. J. Optim. Theory Appl. 36 (1982) 387–407

[12] Jahn, J.: *Mathematical Vector Optimization in Partially Ordered Linear Spaces.* Peter Lang, Frankfurt 1986

[13] Jahn, J.: *Introduction to the Theory of Nonlinear Optimization.* Springer, Berlin 1996

[14] Kaliszewski, I.: *Quantitative Pareto Analysis by Cone Separation Technique.* Kluwer, Boston 1994

[15] Koopmans, T.C.: "Analysis of Production as an Efficient Combination of Activities". In: T.C. Koopmans (ed.): *Activity Analysis of Production and Allocation.* Wiley, New York 1951, pp. 33–97

[16] Kuhn, H.W., A.W. Tucker: "Nonlinear Programming". In: J. Neyman (ed.): *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability.* University of California Press, Berkeley 1951, pp. 481–492

[17] Lommatzsch K.: *Anwendungen der linearen parametrischen Optimierung.* Birkhäuser, Basel 1979

[18] Reichenbach, G.: *Der Begriff der eigentlichen Effizienz in der Vektoroptimierung.* Diplom thesis, University of Erlangen-Nürnberg, Erlangen 1989

[19] Vogel, W.: *Vektoroptimierung in Produkträumen.* Anton Hain, Meisenheim am Glan 1977

[20] Wierzbicki, A.P.: Basic Properties of Scalarizing Functionals for Multiobjective Optimization. Math. Operationsforsch. Statist. Ser. Optim. 8 (1977) 55–60

[21] Wierzbicki, A.P.: "A Mathematical Basis for Satisficing Decision Making". In: J.N. Morse (ed.): *Organisations: Multiple Agents with Multiple Criteria.* Springer, Lecture Notes in Economics and Mathematical Systems No. 190, Berlin 1981, pp. 465–485

[22] Wierzbicki, A.P.: A Mathematical Basis for Satisficing Decision Making. Math. Modelling 3 (1982) 391–405

[23] Zhuang, D.: *Regularity and maximality properties of set-valued structures in optimization.* Dissertation, Dalhousie University, Halifax 1989

3 DUALITY IN MULTI-OBJECTIVE OPTIMIZATION

HIROTAKA NAKAYAMA

CONTENTS

Abstract: Duality is an attractive topic in multi-objective optimization as well as in usual mathematical programming with a single objective function. However, there seems to be no unified approach to dualization in multi-objective optimization. One of the difficulties is in the fact that the efficient solution to multi-objective optimization is not necessarily unique, but in general becomes a set. The definition of infimum (or supremum) of a set with a partial order plays a key role in development of duality theory in multi-objective optimization. In this chapter, these notions will be considered from some geometric viewpoint.

T. Gal et al. (eds.), *Multicriteria Decision Making*

3.1 REVIEW OF DUALITY IN SINGLE-OBJECTIVE PROGRAMMING AND RELATED PROPERTIES

In order to get a better understanding, we shall begin with a brief review of duality and related properties in single objective optimization, in particular, stability and the theorem of the alternative. As is well known, duality in mathematical programming is based on the property that any closed convex set can be also represented by the intersection of closed half spaces including it. Let X' be a subset of an n-dimensional Euclidean space R^n and let $f : X' \to R$ and $g : X' \to R^m$. Then for the following traditional scalar objective problem

$$\text{(SP)} \qquad \text{Minimize } \{f(x) |\ x \in X' \subset R^n,\ g(x) \leqq 0\},$$

an associated dual problem is given by

$$\text{(SD)} \qquad \text{Maximize } \{\phi(u) |\ u \geqq 0,\ \phi(u) = \inf\{L(x,u) |\ x \in X'\}\}.$$

Here the vector inequality $\leqq$ is, as usual, componentwise, and $L(x,u) := f(x) + u^T g(x)$ is the usual Lagrangean with the multiplier $u \in R^m$.

Now, set

$$X = \{x \in X' |\ g(x) \leqq 0\},$$

$$X(z) = \{x \in X' |\ g(x) \leqq z\},$$

$$w(z) = \inf\{f(x) |\ x \in X',\ g(x) \leqq z\}$$

and

$$\text{epi } w = \{(z,y) |\ y \geqq w(z),\ X(z) \neq \phi\}.$$

$w(z)$ is called a *perturbation function*. The primal problem can be embedded in $w(0)$. Under appropriate convexity conditions for f, g, and X', it is well known that the set epi w is convex.

For simplicity of notation, we denote the set of optimal solutions in the objective space by

$$\inf(\text{SP}) := \inf\{f(x) |\ x \in X' \subset R^n,\ g(x) \leqq 0\}.$$

For the dual problem, the set sup(SD) can be defined similarly:

$$\sup(\text{SD}) := \sup\{\phi(u) |\ u \geqq 0\}.$$

In cases in which there exists an optimal solution $x^* \in X'$ which attains inf(SP), we use the notation min(SP) instead of inf(SP). Similarly, max (SD) is used if there exists an optimal solution to the dual problem.

Definition 3.1 *Duality* between the problems (SP) and (SD) means that the property

$$\inf(\text{SP}) = \sup(\text{SD}) \text{ (or, } \max(\text{SD})).$$

holds.

In order to show a geometric meaning of duality, we shall use the following notation:

$$G = \{(z,y)|\ y \geqq f(x),\ z \geqq g(x),\ x \in X'\}$$
$$Y_G = \{y \in R^1|\ (0,y) \in G,\ 0 \in R^m\},$$
$$Y_{\mathrm{cl}\ G} = \{y \in R^1|\ (0,y) \in \mathrm{cl}\ G,\ 0 \in R^m\}$$

Here, 'cl' denotes the usual closure. Note that the primal problem (SP) is equivalent to finding inf Y_G. On the other hand, the dual problem is equivalent to finding the maximal intercept of linear support for cl G with y-axis. Therefore, it is readily seen that the duality holds if and only if the following normality condition holds (Van Slyke and Wets [19]):

Definition 3.2 (normality condition) The *normality condition* implies

$$\mathrm{cl}\ Y_G = Y_{\mathrm{cl}\ G}.$$

Remark 3.1 Let $\mathrm{clco}f$ denote the closed convex hull of function f on R^n defined by

$$\mathrm{clco}f(x) = \sup\{h(x)|\ h(z) \leqq f(z),\ \forall z \in R^n,\ h \text{ is an affine function}\}.$$

Then, the normality condition can be restated as

$$w(0) = \mathrm{clco}\ w(0).$$

Remark 3.2 Let f^* denote the *conjugate function* of f defined by

$$f^*(u) = \sup_{x \in R^n} \{< x,u > -f(x)\}.$$

where $< x,u >$ denotes the usual inner product of x and u, i.e., $u^T x$. Then it is known that $f^{**} = \mathrm{clco}f$ holds (Rockafellar [17]). Therefore, the above normality condition is equivalent to

$$w(0) = w^{**}(0),$$

which is known as the *conjugate duality* between (SP) and (SD).

The above statements do not mention the existence of an optimal solution to the dual problem. Note that the existence of an optimal solution to the dual problem is assured if and only if the optimal value of the primal problem is finite and there exists a nonvertical supporting hyperplane for epi$w$.

Definition 3.3 (stability) The problem (SP) is said to be *stable* if there exists a nonvertical supporting hyperplane for epi w at $(0,w(0))$.

Remark 3.3 The problem (SP) is stable if and only if $w(z)$ is subdifferentiable at $z = 0$.

Theorem 3.1 (Rockafellar [17])
inf(SP) = max(SD) holds if and only if inf(SP) is finite and (SP) is stable.

Remark 3.4 When inf(SP) is finite, one of well known sufficient conditions for stability in convex programming is the Slater's constraint qualification, i.e.,

$$\exists x^0 \in X' \text{ such that } g_i(x^0) < 0 \quad (i = 1, \ldots, m).$$

Now, we can consider the duality from another viewpoint.

Definition 3.4 (the condition of the alternative) The condition of the alternative involving the pairs (f, X) and (ϕ, R^m_+) means that for any $\alpha \in (-\infty, \infty)$ exactly one of the following (I_α) and (II_α) holds:

$$(\mathrm{I}_\alpha) \qquad \exists x \in X \quad \text{sucht that} \quad f(x) < \alpha$$

$$(\mathrm{II}_\alpha) \qquad \exists u \in R^m_+ \quad \text{such that} \quad \phi(u) \geqq \alpha.$$

Theorem 1.2 (MacLinden [10])
inf(SP)=max(SD) holds if and only if the condition of the alternative involving the pairs (f, X) and (ϕ, R^m_+) holds.

3.2 DUALITY IN MULTI-OBJECTIVE OPTIMIZATION

Many results in the following have been discussed in detail in the book by Sawaragi-Nakayama-Tanino [18], and hence will be described briefly without proofs here. New aspects for the duality theory, e.g., the relationship with the condition of the alternative, and relatively new results related with the duality for weak efficiency by Tanino [21] will be described in some detail in the subsequent sections.

Let X be a set of alternatives in an n-dimensional Euclidean space R^n, and let $f = (f_1, \ldots, f_p)$ be a vector-valued objective function from R^n into R^p. For two given vectors y^1 and y^2 and a pointed cone K, the following notations for cone-order will be used:

$$y^1 \leqq_K y^2 \Leftrightarrow y^2 - y^1 \in K,$$

$$y^1 \leq_K y^2 \Leftrightarrow y^2 - y^1 \in K \backslash \{0\},$$

$$y^1 <_K y^2 \Leftrightarrow y^2 - y^1 \in \operatorname{int} K.$$

Furthermore, the *K-minimal* and the *K-maximal solution* set of Y are defined, respectively, by

$$\operatorname{Min}_K Y := \{\overline{y} \in Y | \text{ no } y \in Y \text{ such that } y \leq_K \overline{y}\},$$

$$\text{Max}_K\ Y := \{\overline{y} \in Y |\ \text{no } y \in Y \ \text{ such that } \ y \geq_K \overline{y}\}.$$

K-minimal solutions and K-maximal solutions are also called *efficient*. On the other hand, the *weakly K-minimal* and the *weakly K-maximal solution* set of Y are defined, respectively, by

$$\text{w-Min}_K\ Y := \{\overline{y} \in Y |\ \text{no } y \in Y \ \text{ such that } \ y <_K \overline{y}\}$$

$$\text{w-Max}_K\ Y := \{\overline{y} \in Y |\ \text{no } y \in Y \ \text{ such that } \ y >_K \overline{y}\}.$$

For any cone K in R^p we denote the positive dual cone of K by K°, that is,

$$K^\circ := \{p \in R^P |\ \ < p, q > \geqq 0 \ \text{ for any } q \in K\}$$

where $< p, q >$ denotes the usual inner product of p and q. A set Y is said to be K-convex if $Y + K$ is convex [25]. In this chapter, a K-minimal solution $\overline{y}$ is said to be *proper*, if there exists $\mu \in \text{int } K^\circ$ such that

$$< \mu,\ y > \ \geqq \ < \mu,\ \overline{y} > \quad \text{for all } \ y \in Y.$$

Then, for a given cone order with a pointed closed cone D, a general type of multi-objective optimization problem may be formulated as follows:

$$\text{(P)} \qquad \text{Find} \quad \text{Min}_D\ \{f(x) |\ x \in X\},$$

where $f = (f_1, \ldots, f_p)$ and

$$X := \{x \in X' |\ \ g(x) \leqq_Q \ 0, \ \ X' \subset R^n\}.$$

Defining a dual problem (D) in some appropriate way associated with the problem (P), our aim is to show the property Min_D (P) = Max_D (D). Here Min_D(P) denotes the set of efficient points of the problem (P) in the objective function space R^p, and similarly Max_D (D) the one of the dual problem (D).

In contrast to the usual mathematical programming, the optimal value of the primal problem (and the dual problem) are not necessarily determined uniquely in multi-objective optimization. Hence, there have been developed several kinds of formulation of dual problem in order to get the desirable property Min_D (P) = Max_D (D). Regarding Lagrange duality, three typical dualizations can be seen in linear cases, nonlinear cases and geometric approaches.

3.2.1 Linear Cases

The first result on duality for multi-objective optimization seems to be the one given by Gale et al. [3] for linear cases. This is formulated as a matrix optimization including the vector optimization as a special case. Although there have been several related works, the probably most attractive one is given by Isermann [5] because it is formulated as a natural extension of traditional linear programming: Let A be an $m \times n$ matrix, C a $p \times n$ matrix, and b an m-vector. Furthermore, let D, Q and M be pointed convex polyhedral cones in R^p, R^m

and R^n, respectively. Then the primal problem (P) in linear cases is formulated as

$$(\mathrm{P_I}) \qquad \text{Find} \quad \mathrm{Min}_D\{Cx|\ x \in X\}$$

where

$$X = \{x \in R^n|\ Ax \geqq_Q b,\ \ x \geqq_M 0\}.$$

Associated with ($\mathrm{P_I}$), Isermann defined the dual problem as

$$(\mathrm{D_I}) \qquad \text{Find} \quad \mathrm{Max}_D\{Ub|\ U \in \mathcal{U}_0\}$$

where

$$\mathcal{U}_0 = \{U \in \mathcal{U}|\ \exists \mu \in \mathrm{int}D^\circ,\ A^T U^T \mu \geqq_{M^\circ} C^T \mu\}$$

Here, $\mathcal{U}$ is a set of $p \times m$ matrices U such that $UQ \subset D$. Such matrices are said to be positive [16], [2].

Isermann duality is then given by

(i) $Ub \not\geq_D Cx$ for all feasble x and U.

(ii) Suppose that $\overline{U}b = C\overline{x}$ for some feasible $\overline{x}$ and some feasible $\overline{U}$. Then $\overline{U}$ is an efficient solution to ($\mathrm{D_I}$) and $\overline{x}$ is an efficient solution to ($\mathrm{P_I}$).

(iii) $\mathrm{Min}_D(\mathrm{P_I})=\mathrm{Max}_D(\mathrm{D_I})$

3.2.2 Nonlinear cases

The probably first result on duality of nonlinear multi-objective optimization may be seen in Tanino-Sawaragi [22]. The result was obtained in parallel with the Lagrange duality of single-objective optimization. Unfortunately, however, it is not complete in a sense that $\mathrm{Min}_D(\mathrm{P})=\mathrm{Max}_D(\mathrm{D})$ does not hold for (strong) efficiency. Later, in order to obtain the more desirable duality, $\mathrm{Min}_D(\mathrm{P})=\mathrm{Max}_D(\mathrm{D})$, several authors suggested other dualizations (e.g., Jahn [6] and Nakayama [12]). In their formulation, however, there appears explicitly no perturbation map, and hence no dual map. In order to develop duality theory in connection with these notions, we have to define the notion of "inf" and "sup"" in some appropriate way. This has been accomplished for the weak efficiency by Kawasaki [8] and Tanino [21], which will be discussed in some detail later in this chapter. Many results in the subsequent sections have been discussed in detail in the book by Sawaragi-Nakayama-Tanino [18], and hence they will be described briefly without proof. New aspects for the duality theory, e.g., the relationship with the condition of the alternative, and realtively new results related to the duality for weak efficiency by Tanino [21] will be described in some details there.

3.3 LAGRANGE DUALITY IN MULTI-OBJECTIVE OPTIMIZATION

Restate our nonlinear multi-objective optimization problem under consideration:

$$\text{(P)} \qquad \text{Find} \quad \text{Min}_D \{f(x) | \ x \in X\},$$

where $f = (f_1, \ldots, f_p)$ and

$$X := \{x \in X' | \ g(x) \leqq_Q 0, \ \ X' \subset R^n\}.$$

Initially, we impose the following assumptions:

(i) X' is a nonempty compact convex set.

(ii) D and Q are pointed closed convex cones with nonempty interior respectively of R^p and R^m.

(iii) f is continuous and D-convex.

(iv) g is continuous and Q-convex.

Under these assumptions, it can be readily shown that for every $z \in R^m$, both sets $X(z) := \{x \in X' | \ g(x) \leqq z\}$ and $Y(z) =: f[X(z)] = \{y \in R^p | \ y = f(x), x \in X', g(x) \leqq z\}$ are compact and convex.

The primal problem (P) can be embedded as (P_0) in a family of perturbed problems (P_z) given by

$$(\text{P}_z) \qquad \text{Find} \quad \text{Min}_D \, Y(z).$$

Defining $\Gamma = \{z \in R^m | \ X(z) \neq \emptyset\}$, the set Γ is convex. Now in a similar fashion to the ordinary mathematical programming problem, the *perturbed map* can be defined by

$$W(z) = \text{Min}_D\{f(x) | \ x \in X', g(x) \leqq_Q z\}.$$

It is known that for every $z \in \Gamma$, $W(z) + D$ is convex and

$$W(z) + D = Y(z) + D.$$

In addition, the map W is monotone and convex on Γ.

A vector-valued Lagrangean function for the problem (P) is defined on $X' \times U$ by

$$L(x, U) = f(x) + Ug(x).$$

The set-valued map $\Phi : U \to \mathcal{P}(R^p)$ defined by

$$\Phi(U) = \text{Min}_D\{L(x, U) | \ x \in X'\}$$

is called a *dual map*, where $\mathcal{P}(R^P)$ denotes the power set of R^p. Using this teminology, Tanino-Sawaragi [22] gave the following dualization of (P):

$$(\mathrm{D}_{TS}) \qquad \text{Find} \quad \mathrm{Max}_D \bigcup_{U \in \mathcal{U}} \Phi(U).$$

It can be shown that Φ is a D-concave set-valued map on Γ, namely

$$\Phi(\alpha U^1 + (1-\alpha)U^2) \subset \alpha\Phi(U^1) + (1-\alpha)\Phi(U^2) + D$$

and $\Phi(U) + D$ is a convex set in R^p for each $U \in \mathcal{U}$.

Tanino-Sawaragi introduced the following as a duality in multi-objective optimization (for details, see Tanino-Sawaragi [22] and Sawaragi-Nakayama-Tanino [18]):

Proposition 3.1

(i) For any $x \in X$ and $y \in \Phi(U)$

$$y \not\geq_D f(x).$$

(ii) Suppose that $\hat{x} \in X$, $\hat{U} \in \mathcal{U}$ and $f(\hat{x}) \in \Phi(\hat{U})$. Then $\hat{y} = f(\hat{x})$ is an efficient point to the primal problem (P) and also to the dual problem $(\mathrm{D_{TS}})$.

(iii) Suppose that $\hat{x}$ is a properly efficient solution to (P) and that Slater's constraint qualification is satisfied (i.e.,$\exists \overline{x} \in X$ such that $g(\overline{x}) <_Q 0$). Then

$$f(\hat{x}) \in \mathrm{Max}_D \bigcup_{U \in \mathcal{U}} \Phi(U)$$

Remark 3.5 In the sense that the equality

$$\mathrm{Min}_D(\mathrm{P}) = \mathrm{Max}_D(\mathrm{D_{TS}}). \tag{3.1}$$

does not hold, the above dualization of vector optimization is not complete. In order to get $\mathrm{Min}_D(\mathrm{P}) = \mathrm{Max}_D(\mathrm{D_{TS}})$, several authors suggested other dualizations for multi-objective optimization. Some of them took a geometric approach to this. Before proceeding to the geometric duality, we shall discuss the condition of the alternative in multi-objective optimization in the next section.

3.4 CONDITION OF THE ALTERNATIVE IN MULTI-OBJECTIVE OPTIMIZATION

For two given sets $A \subset R^p$ and $B \subset R^p$, we define

$$A^+ := A + D,$$

$$B^- := B - D.$$

Throughout this section, we assume that A is closed.

Definition 3.5 The *condition of the alternative* (CA1) for multi-objective optimization means that for any $\alpha \in A^+ \cup B^-$ exactly one of the following (I_α) and (II_α) holds:

$$(\mathrm{I}_\alpha) \qquad \exists\, a \in A \text{ such that } a \leq_D \alpha,$$

$$(\mathrm{II}_\alpha) \qquad \exists\, b \in B \text{ such that } b \geqq_D \alpha.$$

Proposition 3.2 Suppose that $\mathrm{Min}_D\, A \neq \emptyset$. Then, the condition of the alternative (CA1) for multi-objective optimization holds if and only if

$$\mathrm{Min}_D\, A \subset \mathrm{Max}_D\, B. \tag{3.2}$$

A proof of this theorem, which was originally given by Luc [9] follows immediately via the following lemma:

Lemma 3.1 Define the condition (D1), (D2), (A1) and (A2) as follows:

$$\text{(D1)} \qquad \forall a \in A,\ \forall b \in B,\ a \not\leq_D b$$

$$\text{(D2)} \qquad \forall a \in \mathrm{Min}_D\, A,\ \exists b \in B,\ a \leqq_D b$$

$$\text{(A1)} \qquad \alpha \in A^+ \cup B^-,\ (\mathrm{II}_\alpha) \Rightarrow \text{not } (\mathrm{I}_\alpha)$$

$$\text{(A2)} \qquad \alpha \in A^+ \cup B^-,\ \text{not } (\mathrm{I}_\alpha) \Rightarrow (\mathrm{II}_*\alpha)$$

Then, (D1) is equivalent to (A1), and (D2) is equivalent to (A2).

Proof: (D1) $\Rightarrow$ (A1): From the condition (II_α), there exists some $b \in B$ such that $b \geqq_D \alpha$. Suppose to the contrary that the condition (I_α) holds, i.e., there exists some $a \in A$ such that $a \leq_D \alpha$. Then we have $a \leq_D b$, which contradicts (D1).
(A1) $\Rightarrow$ (D1): Putting $\alpha = b$, the condition (II_b) holds. Therefore, for any $b \in B$ we have *not* (I_b) due to (A1), i.e., there exists no $a \in A$ such that $a \leq_D b$, which is identical to (D1).
(D2) $\Rightarrow$ (A2): The negation of (I_α) for any $\alpha \in A^+ \cup B^-$ implies that for any $\alpha \in A^+ \cup B^-$ there exists no $a \in A$ such that $a \leq_D \alpha$. It follows then from the definition of $A^+ \cup B^-$ that $\alpha \in \mathrm{Min}_D A$ or $\alpha \in B^-$. $\alpha \in \mathrm{Min}_D A$ with (D2)

yields that there exists some $b \in B$ such that $\alpha \leqq_D b$, which is also obtained in case of $\alpha \in B^-$ from the definition of B^-.
(A2) $\Rightarrow$ (D2): For any $\alpha \in \mathrm{Min}_D A$, the condition ($\mathrm{I}_\alpha$) does not hold. Putting $\alpha = a$, it follows then from the condition of (A2) that there exists some $b \in B$ for any $a \in \mathrm{Min}_D A$ such that $a \leqq_D b$. ■

Remark 3.6 The condition (D1) is well known as the weak duality. It is easy to see that we have a kind of strong duality (3.2) from (D1) and (D2).

In the above discussion, it has been observed that we have to modify the condition of alternative (CA1) in order to get a more desirable strong duality $\mathrm{Min}_D A = \mathrm{Max}_D B$.

Definition 3.6 The *condition of the alternative* (CA2) for vector optimization means that for any $\alpha \in R^p$ exactly one of (I_α) and (II_α) holds.

The following lemma is crucial for understanding a geometric relationship between the condition of alternative (CA2) and the duality of multi-objective optimization:

Lemma 3.2 Denoting the weakly D-minimal solution set of A^+ by w-$\mathrm{Min}_D A^+$ and setting $S(A^+) = \text{w-}\mathrm{Min}_D A^+ \backslash \mathrm{Min}_D A^+$, then under the condition of the alternative (CA2), the following (i)-(iii) hold:

(i) $\mathrm{int}\, A^+ \cap \mathrm{int}\, B^- = \emptyset$

(ii) $A^+ \cup B^- = R^p$

(iii) $S(A^+) \cap B^- = \emptyset$

Proof: If (i) is false, then there exists a point $\alpha \in R^p$ such that both (I_α) and (II_α) hold. Furthermore, if (ii) is false, then there exists a point $\alpha \in R^p$ such that neither (I_α) nor (II_α) of the condition of alternatives (CA2) hold. Finally, if (iii) is false, there exists $\bar{b} \in S(A^+) \cap B^-$. Then, by setting $\alpha = \bar{b}$, both (I_α) and (II_α) hold. ■

Remark 3.7 Even though (i)-(iii) of Lemma 3.2 hold, the condition of the alternative (CA2) does not necessarily hold. To see this, for example, let $B^- = (A^+)^c$, where $(A^+)^c$ represents the complement of A^+. However, if $\mathrm{Min}_D A \subset \mathrm{Max}_D B$, then (i)-(iii) of Lemma 3.2 is equivalent to (CA2) by Lemma 3.2 and the following lemma.

Lemma 3.3 In cases in which $\mathrm{Min}_D A \subset \mathrm{Max}_D B$, the conditions (i)-(iii) in Lemma 3.2 are also sufficient for the condition of the alternative (CA2).

Proof: Note that (i)-(iii) implies $B^- \cap \mathrm{int}A^+ = \emptyset$. Therefore, if α belongs to $\mathrm{int}A^+$, then it is clear that there exists an $a \in A$ such that $a \leq_D \alpha$ but no $b \in B$ such that $b \geqq_D \alpha$. In addition, if $\alpha \in \mathrm{bd}(A^+)$, then either $\alpha \in S(A^+)$ or $\alpha \in \mathrm{Min}_D A$, where 'bd$(A^+)$' denotes the boundary of A^+. Here, we used the relation bd(cl(Y^+))=w-Min$_D$cl(Y^+) for any set $Y \in R^p$ as is seen in Lemma 3.4 below. In the case in which $\alpha \in S(A^+)$, there exists no $b \in B$ such that $b \geqq_D \alpha$ according to (iii), and it is clear that there is an $a \in A$ such that $a \leq_D \alpha$. On the other hand, in case in which $\alpha \in \mathrm{Min}_D A$, we have a $b \in B$ such that $b \geqq_D \alpha$ because $\mathrm{Min}_D A \subset \mathrm{Max}_D B$, and no $a \in A$ such that $a \leq_D \alpha$ by the definition of $\mathrm{Min}_D A$. This completes the proof. ■

The following lemma gives a basis of geometric property of weakly D-Minimal solution to a set Y.

Lemma 3.4 For any set $Y \in R^p$ with w-Min$_D$cl$(Y^+) \neq \emptyset$,

$$\mathrm{bd}(\mathrm{cl}(Y^+)) = \text{w-}\mathrm{Min}_D\mathrm{cl}(Y^+).$$

Similarly, if w-Max$_D$cl$(Y^-) \neq \emptyset$,

$$\mathrm{bd}(\mathrm{cl}(Y^-)) = \text{w-}\mathrm{Max}_D\mathrm{cl}(Y^-).$$

Proof: It is easy to show w-Min$_D$cl$(Y + D) \subset$ bd(cl$(Y + D)$). In order to prove the reverse inclusion, suppose that $\overline{y} \in \mathrm{bd}(Y + D)$ and that $\overline{y} \notin$ w-Min$_D$cl$(Y + D)$. From the latter assumption, there exists a point $y' \in \mathrm{cl}(Y + D)$ such that $\overline{y} \in y' + \mathrm{int}D$. Since $y' + \mathrm{int}D$ is an open set included by cl$(Y + D)$, $\overline{y}$ can never be a boundary point of $Y + D$, which leads to a contradiction.

The last half of the lemma can be obtained similarly. ■

Proposition 3.3 Suppose that $\mathrm{Min}_D A \neq \emptyset$. Then, $\mathrm{Min}_D A = \mathrm{Max}_D B$ holds under the condition of the alternative (CA2) for vector optimization.

Proof: $\mathrm{Min}_D A \subset \mathrm{Max}_D B$ follows in the same way as the proof of Proposition 3.2. Next, we shall show $\mathrm{Max}_D B \subset \mathrm{Min}_D A$. Suppose that $\overline{b} \in \mathrm{Max}_D B$. From Lemma 3.2, we have bd(A^+) = bd(B^-), where bd (A^+) denotes the boundary of A^+. From Lemma 3.4, therefore, we have

$$\text{w-}\mathrm{Min}_D\ A^+ \ = \ \text{w-}\mathrm{Max}_D\ B^-.$$

It follows then from (iii) of Lemma 3.2 that

$$\overline{b} \ \in \ \text{w-}\mathrm{Min}_D\ A^+ \setminus S(A^+) \ = \ \mathrm{Min}_D\ A.$$

This completes the proof. ■

Remark 3.8 Defining $\mathrm{Inf}_D A = \mathrm{Min}_D \mathrm{cl}A$ in cases in which A is not necessarily closed, the condition of the alternative (CA2) replacing A by cl$A$ implies $\mathrm{Inf}_D A = \mathrm{Max}_D B$.

In the following, we shall show the relationship between the condition of the alternative and the duality for weak efficiency.

Definition 3.7 Hereafter in this section, int $D \neq \emptyset$ is assumed. Define

$$\begin{aligned} A^{++} &:= A + \mathrm{int}D, \\ B^{--} &:= B - \mathrm{int}D. \end{aligned}$$

Then, the condition of the alternative (CA3) for multi-objective optimization means that for any $\alpha \in R^p$ exactly one of the following (I'$_\alpha$) and (II'$_\alpha$) holds:

$$\text{(I}'_\alpha\text{)} \qquad \exists\, a \in \mathrm{cl}(A^{++}) \text{ such that } a \leqq_D \alpha,$$

$$\text{(II}'_\alpha\text{)} \qquad \exists\, b \in B^{--} \text{ such that } b >_D \alpha,$$

Proposition 3.4 Suppose that $0^+(\mathrm{cl}A^{++}))\cap(-D) = \{0\}$, where $0^+(\mathrm{cl}(A^{++}))$ denotes the recession cone of $\mathrm{cl}(A^{++})$. Then, the condition of the alternative (CA3) holds if and only if the following (i) and (ii) hold:

$$\text{(i)} \qquad B^{--} \cap \mathrm{cl}(A^{++}) = \emptyset$$

$$\text{(ii)} \qquad B^{--} \cup \mathrm{cl}(A^{++}) = R^p$$

Proof: [only if part]: If (i) is false, there exists some $\alpha \in B^{--} \cap \mathrm{cl}(A^{++})$. Then, clearly both (I'$_\alpha$) and (II'$_\alpha$) hold. If (ii) is false, there exists some $\alpha \in R^p$ such that $\alpha \notin B^{--}$ and $\alpha \notin \mathrm{cl}(A^{++})$. When $\alpha \notin B^{--}$, suppose to the contrary that there exists some $b \in B^{--}$ such that $b >_D \alpha$. This implies that $\alpha \in b - \mathrm{int}D \subset B^{--}$ which leads to a contradiction. Hence (II'$_\alpha$) does not hold. On the other hand, when $\alpha \notin \mathrm{cl}(A^{++})$, suppose that $(\alpha - D) \cap \mathrm{cl}(A^{++}) \neq \emptyset$. Then, there exist some $d \in D$ and some $k \in \mathrm{cl}(A^{++})$ such that $\alpha - d = k$. This means that $\alpha = k + d \in \mathrm{cl}(A^{++}) + D = \mathrm{cl}(A^{++})$, which is a contradiction. Note here that the last equality follows from the assumption of $0^+(\mathrm{cl}(A^{++})) \cap (-D) = \{0\}$ (see, for example, Corollary 9.1.1 of Rockafeller [17]). Therefore, $(\alpha - D) \cap \mathrm{cl}(A^{++}) = \emptyset$, which implies that (I'$_\alpha$) does not hold.

[if part]: If both (i) and (ii) hold, for any $\alpha \in R^p$ exactly either $\alpha \in \mathrm{cl}(A^{++})$ or $\alpha \in B^{--}$ holds. If $\alpha \in \mathrm{cl}(A^{++})$, then (I'$_\alpha$) holds clearly. Also, since $\alpha \notin B^{--}$, it follows immediately from the proof of the *only if* part that (II'$_\alpha$) does not hold. On the other hand, if $\alpha \in B^{--}$, clearly (II'$_\alpha$) holds. In addition, since $\alpha \notin \mathrm{cl}(A^{++})$, it follows immediately from the proof of the *only if* part that (I'$_\alpha$) does not hold. This completes the proof. ■

Remark 3.9 Under the assumption that D is a pointed closed convex cone, the condition of $0^+(\mathrm{cl}(A^{++})) \cap (-D) = \{0\}$ assures the existence of $\mathrm{Min}_D\mathrm{cl}(A^{++})$, and hence also w-$\mathrm{Min}_D\mathrm{cl}(A^{++})$ (See, for example, Sawaragi-Nakayama-Tanino [18]).

Proposition 3.5 Suppose that w-Min$_D$cl$(A^{++}) \neq \emptyset$. If the condition of the alternative (CA3) holds, then

$$\text{w-Min}_D\text{cl}(A^{++}) = \text{w-Max}_D\text{cl}(B^{--}).$$

Proof: As stated in Lemma 3.4, for any subset Y in R^p, bd$(Y + D)$ = w-Min$_D$cl$(Y + D)$ and bd$(Y - D)$ = w-Max$_D$cl$(Y - D)$. Note first that

$$(\text{cl}(A + D))^c = (\text{cl}(A + D))^c - D. \tag{3.2}$$

In fact, for an arbitrary subset Y of R^p, $Y = Y + D$ implies $Y^c = Y^c - D$. For, if $y' \in Y^c - D$, there exist $\overline{y} \notin Y$ and $\overline{d} \in D$ such that $y' = \overline{y} - \overline{d}$. Suppose to the contrary that $y' \in Y$. Then, we have $\overline{y} = y' + \overline{d} \in Y + D = Y$, which leads to a contradiction. Hence, for a proof of the equation (3.2), it suffices to show that

$$\text{cl}(A + D) = \text{cl}(A + D) + D. \tag{3.3}$$

To this end, suppose that $\hat{y} \in \text{cl}(A + D) + D$. Then, there exists $\overline{y} \in \text{cl}(A + D)$ and $\overline{d} \in D$ such that $\hat{y} = \overline{y} + \overline{d}$. Further, $\overline{y} \in \text{cl}(A + D)$ implies that there exist sequences $\{y^i\}$ in A and $\{d^i\}$ in D such that $y^i + d^i$ converges to $\overline{y}$. Now, observe that $y^i + d^i + \overline{d} \in A + D$, because D is a convex cone. Finally, since $\hat{y}$ is a limit point of $y^i + d^i + \overline{y}$, we have that $\hat{y} \in \text{cl}(A + D)$, from which (3.3) follows.

Now, turn to the proof of the theorem. According to the above property, we have

$$\begin{aligned}
\text{w-Min}_D\text{cl}(A + D) &= \text{bd}(\text{cl}(A + D)) = \text{bd}(\text{cl}(\text{cl}(A + D))^c) \\
&= \text{bd}(\text{cl}[\text{cl}(A + D))^c - D]) \\
&= \text{w-Max}_D[\text{cl}((\text{cl}(A + D))^c - D)] \\
&= \text{w-Max}_D\text{cl}(\text{cl}(A + D))^c.
\end{aligned}$$

Note that $B^{--} = (\text{cl}(A^{++}))^c$, because (CA3) yields both (i) and (ii) in Proposition 3.4 hold. This leads to the conclusion of the proposition immediately. ■

Remark 3.10 Note that cl(A^{++}) = w-Min$_D$cl$(A^{++}) \cup$ (w-Min$_D$cl(A^{++}) + int$D$) and two sets in the right-hand side are disjoint with each other. It follows then that the above Proposition is equivalent to the following Proposition 3.6 which was originally given by Nieuwenhuis [14]. This proposition plays an important role in the conjugate duality for weak efficiency in multi-objective optimization as will be seen later.

Proposition 3.6 Let w-Min$_D A \neq \emptyset$. Then

$$R^p = \text{w-Min}_D\text{cl}(A^{++}) \cup (\text{w-Min}_D\text{cl}(A^{++}))^{++} \cup (\text{w-Min}_D\text{cl}(A^{++}))^{--}$$

and three sets in the right-hand side are disjoint with each other.

3.5 GEOMETRIC DUALITY IN CONVEX MULTI-OBJECTIVE OPTIMIZATION

3.5.1 Geometric Duality

In the preceeding section, it has been observed that we need another dualization from Tanino-Sawaragi [22] so that the condition of the alternative (CA2), or almost equivalently Lemma 3.2 may hold.

A geometric approach to duality in multiobjective optimization has been given by Jahn [6] and Nakayama [12]. There, some devices for dualization were made in such a manner that the condition of the alternative (CA2) holds (note Proposition 3.3 and Lemma 3.2). We shall review them briefly without proof. For a more detailed discussion, refer to Sawaragi-Nakayama-Tanino [18] or their original papers.

As in the previous section, the convexity assumption of f and g will be also imposed here, but X' is not necessarily compact.

Define

$$G := \{(z,y) \in R^m \times R^p |\ y \geqq_D f(x),\ z \geqq_Q g(x),\ x \in X'\},$$

$$Y_G := \{y \in R^p |\ (0,y) \in G,\ 0 \in R^m,\ y \in R^p\}.$$

We restate the primal problem as

$$\text{(P)} \qquad \text{Find} \quad \text{Min}_D\{f(x) |\ x \in X\},$$

where

$$X := \{x \in X' |\ g(x) \leqq_Q\ 0,\ X' \subset R^n\}.$$

Associated with this primal problem, the dual problem formulated by Nakayama [12] is as follows:

$$\text{(D}_\text{N}\text{)} \qquad \text{Find} \quad \text{Max}_D \bigcup_{U \in \mathcal{U}} Y_{S(U)},$$

where

$$Y_{S(U)} := \{y \in R^p |\ f(x) + Ug(x) \nleqq_D y, \text{ for all } x \in X'\}.$$

An alternative dual problem is given by Jahn [6]:

$$\text{(D}_\text{J}\text{)} \qquad \text{Find} \quad \text{Max}_D \bigcup_{\substack{\mu \in \text{int}D^\circ \\ \lambda \in Q^\circ}} Y_{H^-(\lambda,\mu)}$$

where

$$Y_{H^-(\lambda,\mu)} = \{y \in R^p | < \mu, f(x) > + < \lambda, g(x) > \geqq < \mu, y >,\ \forall x \in X'\}.$$

Proposition 3.7 Suppose that G is closed, and that there is at least a properly efficient solution to the primal problem. Then, under the condition of Slater's constraint qualification,

$$(Y_G)^c \subset \bigcup_{\substack{\mu \in \mathrm{int}D^\circ \\ \lambda \in Q^\circ}} Y_{H^-(\lambda,\mu)} \subset \bigcup_{U \in \mathcal{U}} Y_{S(U)} \subset \mathrm{cl}\,(Y_G)^c.$$

where $(Y_G)^c$ denotes the complement of set Y_G.

Lemma 3.5 The following holds:

$$\mathrm{Min}_D\,(P) = \mathrm{Min}_D\, Y_G.$$

Proposition 3.8 (i) For any feasible x in (P) and for any feasible y in ($\mathrm{D_N}$) or ($\mathrm{D_J}$),

$$y \not\geq_D f(x).$$

(ii) Assume that G is closed, that there exists at least one efficient solution to the primal problem, and that these solutions are all proper. Then, under the condition of Slater's constraint qualification, the following holds:

$$\mathrm{Min}_D(\mathrm{P}) = \mathrm{Max}_D(\mathrm{D_N}) = \mathrm{Max}_D(\mathrm{D_J}).$$

In some cases, one might not assume that G is closed. In this situation, we can invoke some appropriate normality condition in order to derive the duality, which will be stated briefly here. For more details, see for example, Jahn [6], Borwein-Nieuwenhuis [1], and Sawaragi-Nakayama-Tanino [18].

Define

$$\begin{aligned} A_{G(\mu)} &= \{\alpha |\ (0,\alpha) \in G(\mu), 0 \in R^m, \alpha \in R^1\}, \\ Y_G &= \{y |\ (0,y) \in G, 0 \in R^m, y \in R^m\}. \end{aligned}$$

Definition 3.8 The primal problem (P) is said to be *J-normal*, if for every $\mu \in \mathrm{int}D^\circ$,

$$\mathrm{cl}(A_{G(\mu)}) = A_{\mathrm{cl}\ G(\mu)}.$$

The primal problem (P) is said to be *J-stable*, if it is J-normal and for an arbitrary $\mu \in \mathrm{int}D^\circ$ the problem

$$\sup_{\lambda \in Q^\circ} \inf_{x \in X} < \mu, f(x) > + < \lambda, g(x) >$$

has at least one solution.

On the other hand, Nieuwenhuis [14] suggested another normality condition:

Definition 3.9 The primal problem (P) is said to be *N-normal*, if

$$\text{cl } Y_G = Y_{\text{cl } G}.$$

Lemma 3.6 Under the given condition of convexity of the problem, Slater's constraint qualification ($\exists \hat{x}$, $g(\hat{x}) <_Q 0$) yields J-stability and N-normality.

Theorem 3.3 Suppose that Y_G is closed, $\text{Min}_D(\text{P}) \neq \emptyset$, and the D-minimal solutions to (P) are all proper. Then, if Slater's constraint qualification holds

$$\text{Min}_D(\text{P}) = \text{Max}_D(\text{D}_\text{N}) = \text{Max}_D(\text{D}_\text{J}).$$

3.5.2 Duality of Multi-objective Linear Programming via Geometric Duality

In linear cases, fortunately, it is readily seen that the set G is closed and that each efficient solution is proper. In addition, we have $G = \text{epi } W$ if there exists no $x \in X$ such that $(C - UA)x \leq_D 0$, as will be seen later. Therefore, we can derive Isermann's duality in linear cases via the stated geometric duality.

Lemma 3.7 There exists some $\mu \in \text{int}D^\circ$ such that

$$(C - UA)^T \mu \geqq_{M^\circ} 0$$

if and only if there exists no $x \in M$ such that

$$(C - UA)x \leq_D 0.$$

Proof: See Nakayama [13]. ■

Proposition 3.9 For linear cases with $b \neq 0$,

$$\bigcup_{U \in \mathcal{U}_0} \{Ub\} = \bigcup_{U \in \mathcal{U}_0} \Phi(U) = \bigcup_{\substack{\mu \in \text{int}D^\circ \\ \lambda \in Q^\circ}} Y_{H^-(\lambda,\mu)}.$$

Proof: See Nakayama [13]. ■

Now we can obtain Isermann duality via Propositions 3.7, 3.8 and 3.9. Namely,

Theorem 3.4 For $b \neq 0$, we have

$$\text{Min}_D(\text{P}_\text{I}) = \text{Max}_D(\text{D}_\text{I}).$$

3.5.3 *Geometric Duality for Weak Efficiency*

As in the preceding sections, the convexity of f, g, X' is assumed here, but X' is not necessarily compact nor closed. Define

$$Y_{S'(U)} = \{y \in R^p |\ f(x) + Ug(x) \not\leq_D y, \forall x \in X'\}.$$

Then the following properties are known (Nakayama [12]).

Proposition 3.10 Suppose that Y_G is a nonempty D-bounded subset in R^p. Then under the condition of N-normality, we have

$$\begin{aligned} \mathrm{cl}(\mathrm{cl}Y_G)^c &= \mathrm{cl}\left(\bigcup_{\substack{\mu \in D^\circ \setminus \{0\} \\ \lambda \in Q^\circ}} Y_{H^-(\lambda,\mu)} \right) \\ &= \mathrm{cl}\left(\bigcup_{U \in \mathcal{U}} Y_{S'(U)} \right). \end{aligned}$$

Note from Proposition 3.5 that

$$\begin{aligned} \text{w-Min}_D\ \mathrm{cl}Y_G &= \text{w-Max}_D\ \mathrm{cl}\left(\bigcup_{\substack{\mu \in D^\circ \setminus \{0\} \\ \lambda \in Q^\circ}} Y_{H^-(\lambda,\mu)} \right) \\ &= \text{w-Max}_D\ \mathrm{cl}\left(\bigcup_{U \in \mathcal{U}} Y_{S'(U)} \right). \end{aligned}$$

This leads immediately to the following duality.

Theorem 3.5 Suppose that Y_G is a nonempty D-bounded subset in R^p. Then under the condition of N-normality

$$\text{w-Min}_D(\mathrm{P}) = \text{w-Max}_D(\mathrm{D_{TS}}) = \text{w-Max}_D(\mathrm{D_N}) = \text{w-Max}_D(\mathrm{D_J}).$$

Remark 3.11 As can be readily seen, by defining w-Inf$_D Y$ for a set $Y \in R^p$ essentially as w-Min$_D$cl(Y^{++}) and similarly w-Sup$_D Y$ essentially as w-Max$_D$cl(Y^{--}), we can have inf(P)=sup(D$_{\mathrm{TS}}$)=sup(D$_{\mathrm{N}}$)=sup(D$_{\mathrm{J}}$) under some appropriate stability condition (Tanino [21]).

3.6 CONJUGATE DUALITY IN MULTI-0BJECTIVE OPTIMIZATION

3.6.1 Conjugate Map and Conjugate Duality

In parallel to the Lagrange duality, Tanino-Sawaragi [23] developed the conjugate duality corresponding to efficient solutions in vector optimization. Their defintion of conjugate map for a set-valued map is as follows:

Definition 3.10 Let F be a set-valued map from R^n into R^p. The point-to-set map $F^* : R^{p\times n} \to R^p$ defined by

$$F^*(T) = \mathrm{Max}_D \bigcup_{x\in R^n} [Tx - F(x)] \quad \text{for } T \in R^{p\times n}$$

is called the *conjugate map* of F. The conjugate map of F^* is called the *biconjugate* of F and given by

$$F^{**}(T) = \mathrm{Max}_D \bigcup_{T\in R^{p\times n}} [Tx - F^*(T)].$$

For a vector-valued function f from R^n to $R^p \cup \{+\infty\}$, letting $\mathrm{dom}f = \{x \in R^n |\ f(x) \neq +\infty\}$, the conjugate map f^* of f is defined by

$$f^*(T) = \mathrm{Max}_D\{Tx - f(x) |\ x \in \mathrm{dom}f\}.$$

Here $+\infty$ is an imaginary point whose every component is $+\infty$. We identify the function f as the set-valued map that is equal to $\{f(x)\}$ for $x \in \mathrm{dom}f$ and is empty otherwise.

Definition 3.11 Let f be a function from R^n to $R^p \cup \{+\infty\}$. A $p \times n$ matrix T is called a *subgradient* of f at $\hat{x} \in \mathrm{dom}f$ if

$$f(x) \not\leq_D f(\hat{x}) + T(x - \hat{x}) \quad \text{for any } x \in R^n,$$

in other words, if

$$f(\hat{x}) - T\hat{x} \in \mathrm{Min}_D\{f(x) - Tx \in R^p |\ x \in R^n\} = \mathrm{Min}_D\{f(x) - Tx |\ x \in \mathrm{dom}f\}.$$

The set of all subgradients of f at $\hat{x}$ is called the *subdifferential* of f at $\hat{x}$ and is denoted by $\partial f(\hat{x})$. If $\partial f(\hat{x})$ is not empty, then f is said to be *subdifferentiable* at $\hat{x}$.

Similarly, a set-valued map F: $R^n \to R^p$ is said to be *subdifferentiable* at $\hat{x} \in R^n$ if there exists a $p \times n$ matrix T such that

$$\hat{y} - T\hat{x} \in \mathrm{Min}_D \bigcup_{x\in R^n} [F(x) - Tx].$$

for every $\hat{y} \in F(\hat{x})$. The set of all subgradients of F at $(\hat{x};\hat{y})$ is called the *subdifferential* of F at $(\hat{x};\hat{y})$ and is denoted by $\partial F(\hat{x};\hat{y})$. If $\partial F(\hat{x};\hat{y})$ is not empty for every $\hat{y} \in F(\hat{x})$, then F is said to be *subdifferentiable* at $\hat{x}$.

The following property is essential to the conjugate duality for set-valued map.

Proposition 3.11 A set-valued map $F : R^n \to R^p$ is subdifferentiable at $\hat{x}$ if and only if $F(\hat{x}) \subset F^{**}(\hat{x})$.

Proof: For $\hat{y} \in F(\hat{x})$, it follows from the definition of subgradient of set-valued map that $\partial F(\hat{x}; \hat{y}) \neq \emptyset$ if and only if there exists $T \in R^{p \times n}$ such that

$$T\hat{x} - \hat{y} \in F^*(T).$$

Hence, if $\hat{y} \in F^{**}(\hat{x})$, then it is clear that $\partial F(\hat{x}; \hat{y}) \neq \emptyset$. Conversely, suppose that $\partial F(\hat{x}; \hat{y}) \neq \emptyset$, namely that $\hat{y} \in T\hat{x} - F^*(T)$ for some T. Hence, we have $\hat{y} \not\leq y$ for any $y \in T'\hat{x} - F^*(T')$ with any T'. Therefore, $\hat{y} \in F^{**}(\hat{x})$, as was to be proved. ■

Now embed the original problem (P) into a family of perturbed problems

$$(\mathrm{P}_z) \qquad \text{Find} \quad \mathrm{Min}_D\{\, \phi(x,z) |\, x \in R^n,\ z \in R^m\}$$

by considering the function $\phi : R^n \times R^m \to R^p \cup \{+\infty\}$ such that

$$\phi(x,z) = \begin{cases} f(x) & \text{for } x \in X(z), \\ +\infty & \text{otherwise.} \end{cases}$$

where $X(z) = \{x |\, g(x) \leqq_Q z,\ x \in X' \subset R^n\}$. Clearly, (P)=($\mathrm{P}_0$). Let ϕ^* be the conjugate map of ϕ, namely

$$\phi^*(T,U) = \mathrm{Max}_D\{Tx + Uz - \phi(x,z) \in R^p |\, x \in R^m,\ z \in R^m\}.$$

Now consider the dual problem ($\mathrm{D_C}$) to (P) as follows:

$$(\mathrm{D_C}) \qquad \text{Find} \quad \mathrm{Max}_D \bigcup_{U \in \mathcal{U}} -\phi^*(0,U).$$

Let us consider now the perturbation map $W(z) : R^m \to R^p$ defined by

$$W(z) = \mathrm{Min}_D\{\phi(x,z) \in R^p |\, x \in R^p\}.$$

Clearly,

$$\mathrm{Min}_D(\mathrm{P}) = W(0).$$

The following properties can be seen in Sawaragi-Nakayama-Tanino [18] which were originally given by Tanino-Sawaragi [23].

Lemm 3.8 For every $U \in \mathcal{U}$,

$$W^*(U) \supset \phi^*(0,U).$$

If every $W(z)$ is externally stable, i.e., if $\{\phi(x,z) \in R^p |\, x \in R^n\} \subset W(z) + D$ for any $z \in R^m$, then

$$W^*(U) = \phi^*(0,U).$$

Lemma 3.9

$$\mathrm{Max}_D(\mathrm{D_C}) = \mathrm{Max}_D \bigcup_{U \in \mathcal{U}} [-W^*(U)] = W^{**}(0).$$

Definition 3.12 The multi-objective optimization problem (P) is said to be *stable* if the perturbation map W is subdifferentiable at 0.

In view of Proposition 3.11, the problem (P) is stable if and only if

$$\mathrm{Min}_D(\mathrm{P}) = W(0) \subset W^{**}(0) = \mathrm{Max}_D(\mathrm{D_C}),$$

which yields immediately the following duality property similar to the one in Lagrange duality of Tanino-Sawaragi:

Proposition 3.12 (i) The problem (P) is stable if and only if, for each solution $\hat{x}$ of (P), there exists a solution $\hat{U}$ of the dual problem ($\mathrm{D_C}$) such that

$$\phi(\hat{x}, 0) \in -\phi^*(0, U).$$

(ii) Conversely, if $\hat{x} \in R^n$ and $\hat{U} \in R^{p \times m}$ satisfy the above relationship, then $\hat{x}$ is a solution of (P) and $\hat{U}$ is a solution of ($\mathrm{D_C}$).

Note that the above duality is essentially identical to $W(0) \subset W^{**}(0)$. Unfortunately, a more desirable relation $W(0) = W^{**}(0)$ has not yet been established by using the definition of conjugate map on the basis of strong efficiency. However, it can be easily shown that this stronger duality holds for the weak efficiency just similarly to the Lagrange duality as stated in the previous section. This fact will be discussed in the following subsection.

3.6.2 Conjugate Duality for Weak Efficiency

In Lagrange Duality, we suggested to define the infimum of a set A in terms of ordering induced by a closed pointed cone D as follows:

$$\mathrm{Inf}_D A := \mathrm{Min}_D(\mathrm{cl}A) \quad \text{if} \quad A \neq \emptyset$$

As is discussed in Sawarai-Nakayama-Tanino [18], however, this definition is not adequate for definition of conjuate map for a vector-valued function f because $f(0) \notin f^{**}(0)$ in general.

Example 3.1 (Sawaragi-Nakayama-Tanino [18])
Let

$$f(x) = \begin{bmatrix} x \\ -x \end{bmatrix} \qquad \text{for} \quad x \in R.$$

Then, for

$$T = \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} \in R^{2 \times 1},$$

$$-f^*(T) = \mathrm{Inf}_D \left\{ \begin{bmatrix} (1-t_1)x \\ -(1+t_2)x \end{bmatrix} : x \in R \right\}.$$

It is readily seen that

$$\begin{aligned} \bigcup_T (-f^*(T)) &= (R^2 \backslash (R^2_+ \cup R^2_-)) \\ &\cup \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -\infty \end{bmatrix}, \begin{bmatrix} -\infty \\ 0 \end{bmatrix}, \begin{bmatrix} -\infty \\ -\infty \end{bmatrix}, \begin{bmatrix} -\infty \\ +\infty \end{bmatrix}, \begin{bmatrix} +\infty \\ -\infty \end{bmatrix} \right\} \end{aligned}$$

Hence,

$$f^{**}(0) = \left\{ \begin{bmatrix} 0 \\ +\infty \end{bmatrix}, \begin{bmatrix} +\infty \\ 0 \end{bmatrix} \right\}$$

which does not contain

$$f(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Now, we introduce another notion of infimum based on weak efficiency which is seen in Nieuwenhuis [14], Kawasaki [7],[8], and Tanino [21]. The following discussion is based on Tanino [20], [21].

In order to develop the conjugate duality, it is convenient to extend the space R^p into $\overline{R}^p := R^p \cup \{+\infty\} \cup \{-\infty\}$. Here, $\pm\infty$ are imaginary points such that for any $y \in R^p$, $-\infty <_D y <_D +\infty$ and $y + \infty = +\infty$ and $y - \infty = -\infty$. Clearly, $-(+\infty) = -\infty$, and the sum $+\infty - \infty$ is not considered.

Given a set $Y \subset \overline{R}^p$, we define the set of all points above Y, $A(Y)$, by

$$A(Y) := \{y \in \overline{R}^p \mid y >_D y' \text{ for some } y' \in Y\}.$$

Similarly, the set of all points below Y, $B(Y)$, by

$$B(Y) := \{y \in \overline{R}^p \mid y <_D y' \text{ for some } y' \in Y\}.$$

Note that $B(Y) = \emptyset$ if and only if $Y = \emptyset$ or $Y = \{-\infty\}$. Also, $B(+\infty) = R^p \cup \{-\infty\}$.

The definition of weakly maximal points of $Y \subset \overline{R}^p$ can be extended directly from the one in Section 3.2. Namely, a point $\hat{y} \in \overline{R}^p$ is called a *weakly D-maximal point* of Y, if $\hat{y} \in Y$ and there is no $y' \in Y$ such that $\hat{y} >_D y'$. The set of all weakly D-maximal points of Y is denoted by w-Max$_D Y$. The definition of weakly D-minimal points of $Y \subset \overline{R}^p$ is similar.

The closure of $B(Y)$ in $\overline{R}^p$ is defined by

$$\overline{\mathrm{cl}} B(Y) = \begin{cases} \{-\infty\} & \text{if } B(Y) = \emptyset \\ \overline{R}^p & \text{if } B(Y) = R^p \cup \{-\infty\} \\ \mathrm{cl}[B(Y) \cap R^p] \cup \{-\infty\} & \text{otherwise} \end{cases}$$

Here, the notation of "cl" on the right-hand side means the usual closure in R^p.

Now, we define the set of weakly supremal points of the set Y with respect to the cone ordering $\geqq_D$, w-Sup$_D Y$, as follows:

$$\text{w-Sup}_D Y := \text{w-Max}_D\, \overline{\text{cl}} B(Y).$$

It can be shown that a point $\hat{y} \in \overline{R}^p$ is a weakly supremal point of $Y \subset \overline{R}^p$ if and only if there is no $y \in Y$ such that $\hat{y} <_D y$ and in addition $y' <_D \hat{y}$ implies the existence of some $y \in Y$ such that $y' <_D y$. The closure of $A(Y)$ and the weak D-infimum of $Y \subset \overline{R}^p$ can be defined similarly.

Note that w-Max$_D\emptyset = \emptyset$ and w-Sup$_D\emptyset = \{-\infty\}$. Also, w-Max$_D\overline{R}^p = \{+\infty\}$. Furthermore, it follows that w-Sup$_D Y = \{-\infty\}$ if and only if $B(Y) = \emptyset$. Also, w-Sup$_D Y = \{+\infty\}$ if and only if $B(Y) = R^p \cup \{-\infty\}$.

The following lemma is an extension of Lemma 3.4.

Lemma 3.10

$$\text{w-Sup}_D Y = [\overline{\text{cl}} B(Y)] \backslash B(Y).$$

Proof: When $Y = \emptyset$ or $Y = \{-\infty\}$, then $B(Y) = \emptyset$. Therefore, $[\overline{\text{cl}}B(Y)] \backslash B(Y) = \{-\infty\}$. Clearly, w-Max$_D\{-\infty\} = \{-\infty\}$, which yields the property of the theorem. Next, when $Y = R^p$ or $Y = \{+\infty\}$, we have $B(Y) = R^p \cup \{-\infty\}$. Hence, $[\overline{\text{cl}}B(Y)]\backslash B(Y) = \{+\infty\}$, from which we have the property of the theorem because w-Max$_D\overline{R}^p = \{+\infty\}$. Thirdly, if Y is a nonempty subset of R^p, then $B(Y) = Y^{--} \cup \{-\infty\}$, where $Y^{--} = Y - \text{int}D$. In this case, $\overline{\text{cl}}B(Y) = \text{cl}[B(Y) \cap R^p] \cup \{-\infty\}$. Clearly, w-Max$_D\overline{\text{cl}}B(Y)$=w-Max$_Dcl[B(Y) \cap R^p] \cup \{-\infty\})$=w-Max$_Dcl[B(Y) \cap R^p]$. Furthermore, in this case, cl$[B(Y) \cap R^p]$ = cl(Y^{--}) and $B(Y) \cap R^p = Y^{--}$. Since w-Max$_D$cl(Y^{--}) = cl$(Y^{--}) \backslash Y^{--}$ from Lemma 3.4, we have the property of the theorem. Lastly, other cases are some combination of the above. This completes the proof. ■

Lemma 3.11

$$B(Y) = B(\text{w-Sup}_D Y).$$

Proof: Since w-Max$_D\overline{\text{cl}}B(Y) \subset \overline{\text{cl}}B(Y)$, we have $B(\text{w-Max}_D\overline{\text{cl}}B(Y)) \subset B(\overline{\text{cl}}B(Y)) = B(Y)$. Here, the last equality follows from Lemma 4.1 of Tanino [21]. Therefore, we have $B(\text{w-Sup}_D Y) \subset B(Y)$ from the definition of w-Sup$_D Y$ immediately. In the following, we shall prove $B(Y) \subset B(\text{w-Sup}_D Y)$. If w-Sup$_D Y = \{+\infty\}$ or $\{-\infty\}$, the relation is obvious. In the other case, $-\infty$ is contained in both sets. Let $\hat{y} \in B(Y)$ and $\hat{y} \neq -\infty$. Then there exists $y \in R^p \cap Y$ such that $\hat{y} <_D y$. Take an arbitrary $d \in \text{int}D$. Then there exists a positive number α_0 such that $y + \alpha d \notin \text{cl}B(Y)$ for all $\alpha > \alpha_0$, since otherwise cl$B(Y) \supset R^p$. Thus we can define a finite nonnegative number $\overline{\alpha}$ by

$$\overline{\alpha} = \sup\{\alpha |\ y + \alpha d \in \text{cl}B(Y)\}.$$

Then it is clear that $y + \overline{\alpha} d \in \text{w-Sup}_D Y = \text{w-Max}_D[\overline{\text{cl}}B(Y)]$. Since $\overline{y} <_D y \leqq_D y + \overline{\alpha} d$, we have finally $B(Y) \subset B(\text{w-Sup}_D Y)$. This completes the proof. ■

Now we can prove the following property which is an extension of Proposition 3.6:

Proposition 3.13

$$\overline{R}^p = (\text{w-sup}_D Y) \cup A(\text{w-sup}_D Y) \cup B(\text{w-sup}_D Y),$$

and the above three sets in the right-hand side are disjoint.

Proof: It should be noted from Lemmas 3.10 and 3.11 that $(\text{w-Sup}_D Y) \cup B(\text{w-Sup}_D Y) = \overline{\text{cl}} B(Y)$. Therefore, it suffices to show that $y \notin \overline{\text{cl}} B(Y)$ implies $y \in A(\text{w-Sup}_D Y)$. When $\text{w-Sup}_D Y = \{-\infty\}$ or $\{+\infty\}$, the property of the theorem is obviously true. Therefore, we shall consider the remaining ordinary case. Since $+\infty \in A(\text{w-Sup}_D Y)$ in this case, it suffices to show $y \in A(\text{w-Sup}_D Y)$ for $y \notin \overline{\text{cl}} B(Y)$ such that $y \neq +\infty$. Fix an arbitrary $d \in \text{int} D$. Since $Y \cap R^p \neq \emptyset$ in this case, $y - \alpha d \in B(Y)$ for sufficiently large $\alpha > 0$. Let

$$\overline{\alpha} = \inf\{\alpha > 0 |\ y - \alpha d \in B(Y)\}$$

and $\overline{y} = y - \overline{\alpha} d$. Showing that $\overline{y} \in \text{w-Sup}_D Y$ completes the proof. Since $\overline{y} \in \overline{\text{cl}} B(Y) = (\text{w-Sup}_D Y) \cup B(Y)$, it suffices to show that $\overline{y} \notin B(Y)$. If we suppose to the contrary that $\overline{y} \in B(Y)$, then $y - \alpha d \in B(Y)$ for some α smaller than $\overline{\alpha}$, which contradicts the definition of $\overline{\alpha}$. Therefore, $\overline{y} \notin B(Y)$ as was to be proved. ■

Now we can define the conjugate map of a set-valued map from n-dimensional Euclidean space R^n to the extended p-dimensional Euclidean space $\overline{R}^p$. The results below can be easily extended to a set-valued map from a linear topological space to an extended partially ordered linear topological space.

Let $\mathcal{T}$ be a set of all $p \times n$ matrices, and let F be a set-valued map from R^n to $\overline{R}^p$.

Defintion 3.13 A set-valued map from $\mathcal{T}$ to $\overline{R}^p$ defined by

$$F^*(T) = \text{w-sup}_D \bigcup_{x \in R^n} [Tx - F(x)] \quad \text{for } T \in \mathcal{T}.$$

is called the *conjugate map* of F. Moreover, a set-valued map F^{**} from R^n to $\overline{R}^p$ defined by

$$F^{**}(x) = \text{w-sup}_D \bigcup_{T \in \mathcal{T}} [Tx - F^*(T)] \quad \text{for } x \in R^n.$$

is called the *biconjugate map* of F. When f is a vector-valued function, its conjugate map can be defined by identifying it with the set-valued map $x \to \{f(x)\}$.

Lemma 3.12 For any $x \in R^n$ and any $T \in \mathcal{T}$,

$$[F(x) - Tx] \cap B(-F^*(T)) = \emptyset.$$

Proof: Since $F^*(T) = \text{w-Sup}_D \bigcup_{x \in R^n} [Tx - F(x)]$, it is clear from Lemma 3.10 and Proposition 3.13 that $[Tx - F(x)] \cap A(F^*(T)) = \emptyset$, as was to be proved. ■

Corollary 3.1 If $y \in F(0)$ and $y' \in -F^*(T)$, then $y \not\leq_D y'$.

Corollary 3.2 If $y \in F(x)$ and $y' \in F^{**}(x)$, then $y \not\leq_D y'$. In other words,

$$F(x) \subset F^{**}(x) \cup A(F^{**}(x)).$$

Now we define w-subgradients for a set-valued maps from R^n to $\overline{R}^p$.

Definition 3.14 Let $\hat{x} \in R^n$ and $F(\hat{x})$. An element $T \in \mathcal{T}$ is said to be a *w-subgradient* of F at $(\hat{x}; \hat{y})$ if

$$T\hat{x} - \hat{y} \in \text{w-Max}_D \bigcup_{x \in R^n} [Tx - F(x)].$$

The set of all w-subgradients of F at $(\hat{x}; \hat{y})$ is called the *w-subdifferential* of F at $(\hat{x}; \hat{y})$ and is denoted by $\partial' F(\hat{x}; \hat{y})$. When $\partial' F(\hat{x}; \hat{y}) \neq \emptyset$ for $\forall \hat{y} \in F(\hat{x})$, F is said to be *w-subdifferentiable* at $\hat{x}$.

Proposition 3.14 If F is w-subdifferentiable at $\hat{x}$, then $F(\hat{x}) \subset F^{**}(\hat{x})$. Moreover, if $F(\hat{x}) = \inf_D F(\hat{x})$ in addition, then $F(\hat{x}) = F^{**}(\hat{x})$.

Proof: It is readily shown in general (Tanino [21], Proposition 3.1) that for a point $\overline{x} \in R^n$, a set-valued map G from R^n to $\overline{R}^p$ defined by $G(x) = F(x + \overline{x}), \forall\, x \in R^n$ has the property that (i) $G^*(T) = F^*(T) - T\overline{x} \quad \forall T \in \mathcal{T}$ and (ii) $G^{**}(x) = F^{**}(x + \overline{x}) \quad \forall x \in R^n$. Therefore, it suffices to prove the case $\hat{x} = 0$. First let $\hat{x} \in F(0)$. Since F is w-subdifferentiable at 0, there exists $\hat{T} \in \mathcal{T}$ such that $\hat{y} \in -F^*(\hat{T})$. Then, from Corollary 3.1,

$$\hat{y} \in \text{w-Min}_D \bigcup_{T \in \mathcal{T}} [-F^*(T)] \subset \text{w-Sup}_D \bigcup_{T \in \mathcal{T}} [-F^*(T)] = F^{**}(0).$$

Thus we have proved that $F(0) \subset F^{**}(0)$. Next we assume that $F(0) = \inf_D F(0)$ and take an arbitrary $\hat{y} \in F^{**}(0)$. From Proposition 3.13,

$$\overline{R}^p = F(0) \cup A(F(0)) \cup B(F(0)).$$

In view of Corollary 3.2, $\hat{y} \notin A(F(0))$. If we suppose that $\hat{y} \in B(F(0))$, there exists $y' \in F(0)$ such that $\hat{Y} < y'$. Then there exists $T' \in \mathcal{T}$ such that $y' \in -F^*(T')$ since F is assumed to be subdifferentiable at 0. However, this implies that $\hat{y} \in B(-F^*(T'))$ and hence contradicts the assumption $\hat{y} \in F^{**}(0) = \sup_D \bigcup_{T \in \mathcal{T}} [-F^*(T)]$. Therefore, $\hat{y} \in F(0)$ and we have proved that $F^{**}(\hat{x}) \subset F(\hat{x})$. This completes the proof. ■

Now we can show the conjugate duality based on weak efficiency in vector optimization. The primal problem is

$$(\mathrm{P}') \qquad \text{Find} \quad \text{w-Min}_D\{f(x)|\ x \in X\}$$

where $X = \{x \in R^n|\ g(x) \leqq_Q 0\}$. Let ϕ be a function from $R^n \times R^m$ into $R^p \cup \{+\infty\}$ such that

$$\phi(x,z) = \begin{cases} f(x) & \text{if } x \in X \\ +\infty & \text{otherwise.} \end{cases}$$

Then the perturbed problem is

$$(\mathrm{P}'_z) \qquad \text{Find} \quad \text{w-Min}_D\{\phi(x,z)|\ x \in R^n\}.$$

Definition 3.15 The set-valued map W from R^m to $\overline{R}^p$ defined by

$$W(z) = \text{w-Inf}_D(\mathrm{P}'_z) = \text{w-Inf}_D\{\phi(x,z)|\ x \in R^n\}$$

is called the perturbation map for Problem (P').

Consider the conjugate map of ϕ:

$$\phi^*(T,U) = \text{w-Sup}_D\{Tx + Uz - \phi(x,z)|\ x \in R^n,\ z \in R^m\}.$$

Then

$$\begin{aligned} -\phi^*(0,U) &= -\text{w-Sup}_D\{Uz - \phi(x,z)|\ x \in R^n,\ z \in R^m\} \\ &= \text{w-Inf}_D\{\phi(x,z) - Uz|\ x \in R^n,\ z \in R^m\}. \end{aligned}$$

The dual problem to (P') is given by

$$(\mathrm{D}'_{\mathrm{C}}) \qquad \text{Find} \quad \text{w-Max}_D\{-\phi^*(0,U)|\ U \in \mathcal{U}\},$$

where $\mathcal{U}$ is a set of $p \times m$ matrices such that $UQ \subset D$.

Since $-\phi^*(0,\cdot)$ is a set-valued map, the problem (D') is regarded as a problem of finding

$$\text{w-Sup}_D(\mathrm{D}'_{\mathrm{C}}) = \text{w-Sup}_D \bigcup_{U \in \mathcal{U}} [-\phi^*(0,U)].$$

Lemma 3.13

$$W^*(U) = \phi^*(0,U).$$

Proof:

$$\begin{aligned} W^{**}(U) &= \text{w-Sup}_D \bigcup_{z \in R^m} [Uz - W(z)] \\ &= \text{w-Sup}_D \bigcup_{z \in R^m} [Uz - \text{w-Inf}_D\{\phi(x,z)|\ x \in R^n\}] \\ &= \text{w-Sup}_D \bigcup_{z \in R^m} [Uz + \text{w-Sup}_D\{-\phi(x,z)|\ x \in R^n\}] \end{aligned}$$

$$
\begin{aligned}
&= \text{w-Sup}_D \bigcup_{z \in R^m} [\text{w-Sup}_D \{Uz - \phi(x,z) |\ x \in R^n\}] \\
&= \text{w-Sup}_D \bigcup_{z \in R^m} \{Uz - \phi(x,z) |\ x \in R^n\} \\
&= \text{w-Sup}_D \{Uz - \phi(x,z) |\ x \in R^n,\ z \in R^m\} \\
&= \phi^*(0,U).
\end{aligned}
$$

This completes the proof. ■

In view of the above Lemma 3.13, we can rewrite $\text{w-Sup}_D(\text{D}_C^{'})$ as

$$\text{w-Sup}_D(\text{D}_C^{'}) = \text{w-Sup}_D \bigcup_{U \in \mathcal{U}} [-W^*(U)] = W^{**}(0).$$

Since $\text{w-Inf}_D(\text{D}_C^{'}) = W(0)$, the desirable duality property is identical with $W(0) = W^{**}(0)$.

Definition 3.16 The problem $(\text{P}^{'})$ is said to be *w-stable* if the perturbation map W is w-subdifferentiable at 0.

Now, immediately from Proposition 3.14, we have the following duality property.

Theorem 3.6 If the problem $(\text{P}^{'})$ is w-stable, then

$$\text{w-inf}_D(\text{P}^{'}) = \text{w-Sup}_D(\text{D}_C^{'}).$$

It should be noted that the convexity is essentially sufficient for $(\text{P}^{'})$ being w-stable, as will be seen in the following.

Lemma 3.14 If the function $\phi : R^n \times R^m \to \overline{R}^p$ is convex, then the perturbation map W is a convex set-valued map from R^m to $\overline{R}^p$.

Proof: Let

$$Y(z) = \{\phi(x,z) |\ x \in R^n\} \subset R^p \cup \{+\infty\}$$

for each $z \in R^m$. Then, from Lemma 3.10,

$$W(z) = \text{w-Inf}_D Y(z) = [\text{cl}A(Y(z))] \backslash A(Y(z)).$$

Suppose that $(z^1, y^1),\ (z^2, y^2) \in \text{epi}\ W$. Then

$$
\begin{aligned}
y^i \in W(u^i) + D \ &\subset\ A(W(z^i)) \\
&=\ A(ZY(z^i)) \\
&\subset\ \text{cl}A(Y(z^i)) \quad \text{for}\ \ i = 1,2.
\end{aligned}
$$

For each α such that $0 \leq \alpha \leq 1$,

$$\begin{aligned} \alpha y^1 + (1-\alpha)y^2 &\in \alpha \mathrm{cl}A(Y(z^1)) + (1-\alpha)\mathrm{cl}A(Y(z^2)) \\ &\subset \mathrm{cl}\{\alpha A(Y(z^1)) + (1-\alpha)A(Y(z^2))\}, \end{aligned}$$

where we put $0 \cdot (+\infty) = 0$. Since ϕ is convex, it is readily shown that

$$\alpha A(Y(z^1)) + (1-\alpha)A(Y(z^2)) \subset A(Y(\alpha z^1 + (1-\alpha)z^2)).$$

Therefore,

$$\begin{aligned} \alpha y^1 + (1-\alpha)y^2 &\in \mathrm{cl}A(Y(\alpha z^1 + (1-\alpha)z^2)) \\ &= W(\alpha z^1 + (1-\alpha)z^2) \cup A(W(\alpha z^1 + (1-\alpha)z^2)), \end{aligned}$$

which implies that

$$\alpha(z^1, y^1) + (1-\alpha)(z^2, y^2) \in \text{epi } W.$$

Hence, epi W is a convex set in $R^m \times R^p$, which means W is a convex set-valued map. This completes the proof. ■

Corollary 3.3 If the function ϕ is convex and if $0 \in \text{int dom } \phi(x, \cdot)$ for some x, then the problem (P$'$) is w-stable.

Proof. The perturbation map W is convex by Lemma 3.14. Hence, if $W(0) = \{-\infty\}$, then $W(z) = \{-\infty\}$ for all z, which implies that W is w-subdifferentiable at 0. On the other hand, since $0 \in \text{int dom } \phi(x, \cdot)$ for some x, $W(z) \neq \{+\infty\}$ for all z in some neighborhood of 0. This follows in general because for any set $Y \subset \overline{R}^p$, (i) w-Sup$_D Y = \{-\infty\}$ if and only if $B(Y) = \emptyset$, (ii) w-Sup$_D Y = \{+\infty\}$ if and only if $B(Y) = R^p \cup \{-\infty\}$, and (iii) w-Sup$_D Y \subset R^p$ except for the cases of (i) and (ii). In general, if a set-valued map F from R^n to $R^p \cup \{+\infty\}$ is convex, if $\hat{x} \in \text{int dom } F$, and if $F(\hat{x}) \subset$ w-Inf$_D F(\hat{x})$, then F is w-subdifferentiable at $\hat{x}$ (Tanino [21], Proposition 4.3). This property and $W(0) =$ w-Inf$_D W(0)$ yields the conclusion of the corollary. ■

Remark 3.12 If the problem (P$'$) is convex and if Slater's constraint qualification holds, then (P$'$) is w-stable.

3.7 CONCLUSION

Many results in this chapter are due to Jahn [6], Nakayama [12], Nieuwenhuis [14], Tanino-Sawaragi [22], [23], and Tanino [20], [21]. Some of them were already discussed in detail in Sawaragi-Nakayama-Tanino [18]. This chapter is devoted with as much a unified approach to duality in multi-objective optimization as possible. Although there are several other kinds of duality in multi-objective optimization, e.g., Fenchel's duality by Gros [4] and Mond-Weir type of duality by Mond-Weir [11] are not described here due to the page limitation.

Acknowledgment The author would like to express sincere gratitude to Prof. T. Tanino of Osaka University in Japan for his valuable suggestions for preparing this chapter.

References

[1] Borwein, J.M. and J.W. Nieuwenhuis: Two Kinds of Normality in Vcector Optimization. *Mathematical Programming* 28(1984)185-191

[2] Corley, H.W.: Duality for Maximization with respect to Cones. *J. Mathematical Analysis and Applications.* 84(1981)560-568

[3] Gale, D., Kuhn, H. W. and A. W. Tucker: "Linear Programming and the Theory of Games". In: T. C. Koopmans (ed.): *The Theory of Linear Economics Models*, Wiley, New York 1951, pp.317-329

[4] Gros, C.: Generalization of Fenchel's Duality Theorem for Convex Vector Optimization. *European J. Operaionl Research* 2(1978)368-376

[5] Isermann, H.: On Some Relations between a Dual Pair of Multiple Objective Linear Programs. *Zeitschrift fur Oper. Res* 22(1978)33-41

[6] Jahn, J.: Duality in Vector Optimization. *Mathematical Programming* 25(1983)343-353

[7] Kawasaki, H.: Conjugate Relations in Weak Subdifferentials of Relations. *Math. of Oper. Res.* 6(1981)593-607

[8] Kawasaki, H.: A Duality Theorem in Multi-objective Nonlinear Programming. *Math. Oper. Res.* 7(1982)95-110

[9] Luc, D.T.: About Duality and Alternative in Multiobjective Optimization. *J. Optim. Theory Appl.* 53(1987)303-307

[10] MacLinden, L.: Duality Theorems and Theorems of the Alternative. *Proc. of AMS* 53(1975)172-175

[11] Mond, B. and T. Weir: "Symmetric Duality for Nonlinear Multiobjective Programming". In: Kumar, S. (ed.): *Recent Developments in Mathematical Programming*, Gordon and Breach Science, Philadelphia 1991, pp.137-153

[12] Nakayama H.: Geometric Consideration on Duality in Vector Optimization. *J. Optimi. Theory Appl.* 44(1984)625-655

[13] Nakayama H.: Some Remarks on Dualization in Vecotr Optimization. *J. Multi-criteria Decision Analysis* 5(1996)218-225

[14] Nieuwenhuis, J.W.: Supremal Points and Generalized Duality. *Math. Oper. Stat.*, Ser. Optim. 11-1(1980)41-59

[15] Ponstein, J.: *Approaches to the Theory of Optimization*, Cambridge Univ. Press, London 1980

[16] Ritter, K.: Optimization Theory in Linear Spaces III. *Math. Annal.* 184(1970)133-154

[17] Rockafellar, T.: *Convex Analysis*, Princeton Univ. Press, Princeton 1968

[18] Sawaragi Y., H. Nakayama and T. Tanino: *Theory of Multiobjective Optimization*, Academic Press, Orlando 1985

[19] Van Slyke, R.M. and R.J.B. Wets: A Duality Theory for Abstract Mathematical Programs with Applications to Optimal Control Theory. *J. Math. Anal. and Appl.* 22(1968)679-706

[20] Tanino, T.: On Supremum of a Set in a Multi-dimensional Space. *J. Math. Analysis and Applications* 130(1988)386-397

[21] Tanino, T.: Conjugate Duality in Vector Optimization. *J. Math. Analysis and Applications* 167(1992)84-97

[22] Tanino, T. and Y. Sawaragi: Duality Theory in Multiobjective Programming. *J. Optim. Theory Appl.* 27(1979)509-529

[23] Tanino, T. and Y. Sawaragi: Conjugate Maps and Duality in Multi-objective Optimization. *J. Optim. Theory Appl.* 31(1980)473-499

[24] Weir, T. and B. Mond: Symmetric and Self Duality for Multiple Objective Programming. *Asia-Pacific J. Oper. Res.* 5(1988)124-133

[25] Yu, P.L.: Cone convexity, Cone Extreme Points and Nodominates Solutions. *J. Optimi. theory and Appl.* 14(1974)319-377

4 PREFERENCE RELATIONS AND MCDM

Carlos A. Bana e Costa
Jean-Claude Vansnick

CONTENTS

Abstract: Multicriteria decision aid is above all a human activity in which value judgements of involved actors play a crucial role. Therefore, "how to represent such judgements?" is a key question in MCDM. This chapter is devoted to this subject. Depending on the particular paradigm adopted for preference modelling, different questioning procedures can be conceived which lead to different preference structures. We present a few questioning procedures related to three basic paradigms, together with some preference structures that are useful for MCDM. First, the classical preference-indifference structure is discussed, followed by the introduction of the ideas of "incomparability" and "hesitation". Finally, we present some complementary questioning procedures particularly relevant for cardinal modelling of value judgements.

T. Gal et al. (eds.), *Multicriteria Decision Making*

4.1 INTRODUCTION

Multicriteria decision aid is above all a human activity in which value judgements of involved actors play a crucial role. Therefore, "how to represent such judgements?" is a key question in MCDM.

Let us first note that the representation of value judgements is part of the more general problem of the representation of judgements of a person J about the degree to which the elements of a set **X** possess a certain property $\wp$. (For instance, **X** could be a set of cars and $\wp$ their comfort, or **X** could be a set of offences and $\wp$ the seriousness of these offences, etc.) In the multicriteria decision aid framework:

- ✓ **X** can be a set of actions (real or fictitious) and $\wp$ the (partial) attractiveness of the actions in regard of a particular point of view.
- ✓ **X** can be a set of actions (real or fictitious) and $\wp$ the (overall) attractiveness of the actions in regard of several points of view simultaneously.
- ✓ **X** can be a set of sentences (for example: "action a is at least as attractive as action b in regard of a particular point of view") and $\wp$ the credibility of such sentences.

... etc.

Mathematically, the notion of a binary relation plays a fundamental role in the representation of value judgements. Section 4.2 offers a general overview of binary relations, which constitutes the basic background for the next sections. In order to represent value judgements, a first step consists in assessing them. The *questioning procedure* used for this purpose plays a fundamental role. Indeed, depending on the particular paradigm adopted for preference modelling, different questioning procedures can be conceived which lead to different preference structures. Sections 4.3 to 4.5 are devoted to the presentation of a few questioning procedures related to three basic *paradigms*, together with some preference structures that are useful for MCDM. The classical preference-indifference structure is discussed in section 4.3, in section 4.4 the notion of "incomparability" is introduced, and the idea of "hesitation" is addressed in section 4.5. Finally, we present some complementary questioning procedures particularly relevant for cardinal modelling of value judgements.

4.2 RELEVANT BACKGROUND ON BINARY RELATIONS

4.2.1 Notion of a binary relation

A binary relation defined on a set **X** is a subset of the cartesian product **X**×**X**, that is, a set of ordered pairs (x, y) such that x and y are in **X**. In other words, R is a binary relation on **X** if and only if (iff) $R \subseteq \mathbf{X}\times\mathbf{X} = \{(x, y) \mid x \in \mathbf{X} \text{ and } y \in \mathbf{X}\}$. We use either $(x, y) \in R$ or xRy to denote that the pair (x, y) belongs to the binary relation R. We use either $(x, y) \notin R$ or xR̸y to denote that (x, y) does not belong to R.

If R and T are two binary relations on **X**, $R \cup T$ [resp. $R \cap T$] denotes the union [resp. the intersection] of the two subsets R and T of **X**×**X**. We have, therefore:

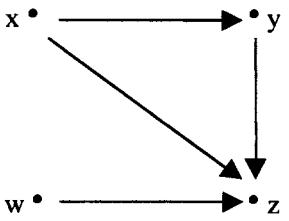
$\forall x, y \in \mathbf{X}$: $x(R \cup T)y$ iff [xRy or xTy],
$\forall x, y \in \mathbf{X}$: $x(R \cap T)y$ iff [xRy and xTy].

4.2.2 Graph of a binary relation

A binary relation R defined on a finite set **X** can be represented by a directed graph such that:

- each element of **X** is represented by a point (node of the graph)
- there exists an arc from node x to node y iff xRy.

Example 4.1 The graph representation of the binary relation R = {(w, z), (x, y), (y, z), (x, z)}defined on **X** = {w, x, y, z} is:



■

4.2.3 Basic properties of binary relations

A binary relation R on the set **X** is

✓	reflexive	iff	$\forall x \in \mathbf{X}$: xRx
✓	irreflexive	iff	$\forall x \in \mathbf{X}$: x$\not$Rx
✓	symmetric	iff	$\forall x, y \in \mathbf{X}$: xRy $\Rightarrow$ yRx
✓	antisymmetric	iff	$\forall x, y \in \mathbf{X}$ with $x \neq y$: xRy $\Rightarrow$ y$\not$Rx
✓	asymmetric	iff	$\forall x, y \in \mathbf{X}$: xRy $\Rightarrow$ y$\not$Rx
✓	complete	iff	$\forall x, y \in \mathbf{X}$ with $x \neq y$: xRy or yRx
✓	strongly complete	iff	$\forall x, y \in \mathbf{X}$: xRy or yRx
✓	transitive	iff	$\forall x, y, z \in \mathbf{X}$: [xRy and yRz] $\Rightarrow$ xRz
✓	negatively transitive	iff	$\forall x, y, z \in \mathbf{X}$: [x$\not$Ry and y$\not$Rz] $\Rightarrow$ x$\not$Rz
✓	semitransitive	iff	$\forall x, y, z, w \in \mathbf{X}$: [xRy and yRz] $\Rightarrow$ [xRw or wRz]
✓	Ferrers	iff	$\forall x, y, z, w \in \mathbf{X}$: [xRy and wRz] $\Rightarrow$ [xRz or wRy]
✓	acyclic	iff	$\forall n \in \{1, 2, 3, \ldots\}$ and $\forall x_1, x_2, \ldots, x_n \in \mathbf{X}$: $[x_1 R x_2, \ldots, x_{n-1} R x_n] \Rightarrow x_n \not R x_1$.

4.2.4 Relationships between basic properties of binary relations

The following are some of the relationships that can be established between basic properties of binary relations - see [8, 16]:

- R acyclic $\Rightarrow$ R asymmetric
- R asymmetric and transitive $\Rightarrow$ R acyclic
- R asymmetric and semitransitive $\Rightarrow$ R transitive
- R asymmetric and Ferrers $\Rightarrow$ R transitive
- R asymmetric and negatively transitive $\Rightarrow$ R semitransitive
- R asymmetric and negatively transitive $\Rightarrow$ R Ferrers
- R asymmetric and negatively transitive $\Rightarrow$ R transitive
- R acyclic and complete $\Rightarrow$ R negatively transitive.

Remarks

1) [R asymmetric and transitive] does not imply [R negatively transitive] as proved by example 4.1, in which R is asymmetric and transitive but R is not negatively transitive: xRw, $w\not Ry$ but xRy.
2) When R is asymmetric

   [R negatively transitive] $\Leftrightarrow$ [R transitive and R^N transitive]

   being R^N (N for negation) the symmetric binary relation defined on **X** by $\forall x, y \in \mathbf{X}$: xR^Ny iff [$x\not Ry$ and $y\not Rx$]. In example 4.1, $R^N = \{(w, x), (x, w), (y, w), (w, y), (x, x), (y, y), (z, z), (w, w)\}$. This relation is not transitive, because yR^Nw, wR^Nx but $y\not R^Nx$.

4.2.5 Dual relation and symmetric and asymmetric parts of a relation

Let R be a binary relation defined on **X**.

We call the *dual relation* R^d of R the binary relation defined on **X** by xR^dy iff $y\not Rx$.

We call the *symmetric part* R^S of R the symmetric binary relation defined on **X** by xR^Sy iff [xRy and yRx].

We call the *asymmetric part* R^A of R the asymmetric binary relation defined on **X** by xR^Ay iff [xRy and $y\not Rx$].

It follows immediately that $R = R^S \cup R^A$, $R^S \cap R^A = \varnothing$ and that R transitive $\Rightarrow$ [R^S transitive and R^A transitive]. If R is strictly complete than: $(R^A)^N = R^S$, $(R^A)^d = R$ and it can be proved that [R^S transitive and R^A transitive] $\Rightarrow$ R transitive.

Concerning duality, it is easy to verify that $(R^d)^d = R$. On the other hand, it can be proved that the following relationships among properties exist - see [12]:

R asymmetric	$\Leftrightarrow$	R^d strongly complete
R antisymmetric	$\Leftrightarrow$	R^d complete
R transitive	$\Leftrightarrow$	R^d negatively transitive
R semitransitive	$\Leftrightarrow$	R^d semitransitive
R Ferrers	$\Leftrightarrow$	R^d Ferrers.

4.3 THE CLASSICAL QUESTIONING PROCEDURE (QP1)

4.3.1 Presentation

Let **X** be a finite set of actions in a decision problem, $\wp$ the property of (partial or overall) attractiveness and J the person who will express judgements about $\wp$ for the elements of **X**.

Two actions (denoted x and y) are involved in the formulation of any questioning procedure, and the questions must be asked for all pairs $\{x, y\} \subset \mathbf{X}$. The only answers accepted are those indicated next to the questions. Moreover, one (and only one) possible answer can be chosen.

Questioning procedure 1 (QP1)

Is one of the actions (x or y) more attractive than the other? YES NO

If YES, which action (x or y) is more attractive than the other? x y

4.3.2 Paradigm 1 and related binary relations

The questioning procedure **QP1** is associated with the following paradigm:

Paradigm 1 Two (and only two) basic preference situations (distinct and exclusive) should be considered for the expression of value judgements: the *situation of preference* and the *situation of indifference.*

The answers obtained with **QP1** allow the construction of two binary relations on **X**: the relation P defined by xPy iff J judged x more attractive than y

and

the relation I defined by xIy iff J did not judge any of the two actions more attractive than the other one

(i.e., iff J answered NO to the first question).

The questioning procedure **QP1** automatically leads to [P asymmetric] and [I symmetric]; moreover, it will always be supposed that J's answers are such that [I reflexive]. P is called a *"strict" preference relation* and I an *indifference relation.*

With paradigm 1, one has xIy iff [x$\not P$y and y$\not P$x] that is $P^N = I$ (see remark 2 in section 4.2.4). The dual relation P^d of P also has an interesting interpretation in this context. Indeed, $P^d = P \cup I$ which means that xP^dy iff "x is at least as attractive as y".

In general, the relation S semantically defined by [xSy iff x is at least as attractive as y] is called a *"large" preference relation* or an *outranking relation.* The following relationships between S, P and I always exist: $I = S^S$ (symmetric part of S) and $P = S^A$ (asymmetric part of S). It is only with the questioning procedure **QP1** (paradigm 1) that one has $S = P^d$; this is due to the fact that paradigm 1 implies that S is strictly complete (recall from section 4.2.5 that, when R is strictly complete, $(R^A)^d = R$).

Many particular binary relations ("particular" in the sense of verifying one or several basic properties) have been introduced and studied in the framework of paradigm 1. The six relations (3 plus the corresponding dual ones) in Table 4.1 are those that have been more extensively used in MCDM, in the framework of paradigm 1

Table 4.1 Some particular binary relations.

Relation	Properties	Dual	Properties	Relation
Strict simple order	Asymmetric Complete Negatively transitive	d ↔	Strictly complete Antisymmetric Transitive	*Simple order* or *Total order*
Strict weak order	Asymmetric Negatively transitive	d ↔	Strictly complete Transitive	*Weak order* or *Total preorder*
Strict semiorder	Asymmetric Semitransitive Ferrers	d ↔	Strictly complete Semitransitive Ferrers	*Semiorder*

Taking into account the relationships that exist among basic properties of binary relations, it can be verified that:

- R is a strict simple order ⇔ R is asymmetric, complete and transitive
- R is a strict simple order ⇔ R is acyclic and complete
- R strict simple order ⇒ R strict weak order ⇒ R strict semiorder
- R total order ⇒ R total preorder ⇒ R semiorder.

4.3.3 Some theoretical results

The relevance of the six particular relations in Table 4.1 is due to the fact that for any one relation denoted by ●, there exists a homomorphism (sometimes injective) from $(\mathbf{X}, \bullet)$ to $(\Re, \blacktriangle)$, where $\Re$ is the set of real numbers and, depending on the case, ▲ denotes $>$, $>_1$, $\geq$ or $\geq_1$ ($\forall r, s \in \Re$, $r >_1 s$ iff $r > s + 1$, and $r \geq_1 s$ iff $r \geq s + 1$).

In a more precise way, for the case of the three particular relations on the left side of Table 4.1, the following theorems can be proved (the respective theorems for the three corresponding dual relations can be obtained by replacing strict inequality ($>$) by non-strict ($\geq$) inequality) - see, for instance, [8, 16, 17].

Let **X** be a finite set and R a binary relation defined on **X**.

Theorem 4.1 There exists an injective mapping μ from **X** to $\Re$ such that

$$\forall x, y \in \mathbf{X} \ [xRy \text{ iff } \mu(x) > \mu(y)]$$

iff R is a strict simple order.

Moreover, μ is an *ordinal scale*, that is, μ is unique up to a strictly monotone increasing transformation.

Theorem 4.2 There exists a mapping μ from **X** to $\Re$ such that

$$\forall x, y \in \mathbf{X} \ [xRy \text{ iff } \mu(x) > \mu(y)]$$

iff R is a strict weak order.

Moreover, μ is an *ordinal scale*, that is, μ is unique up to a strictly monotone increasing transformation.

Theorem 4.3 There exists a mapping μ from **X** to $\Re$ such that

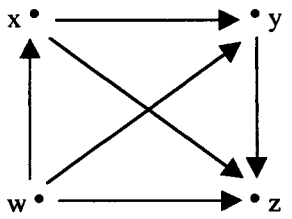
$$\forall x, y \in \mathbf{X} \ [xRy \text{ iff } \mu(x) > \mu(y) + 1]$$

iff R is a strict semiorder.

Note With an appropriate change of μ, any positive number could be used in Theorem 4.3 in place of 1.

In substantive terms, theorem 4.1 essentially means that the mathematical structure that models the notion of *ranking without ex-aequo* is the (strict) simple order, and theorem 4.2 means that the mathematical structure that models the notion of *ranking with possibility of ex-aequo* is the (strict) weak order.

Example 4.2

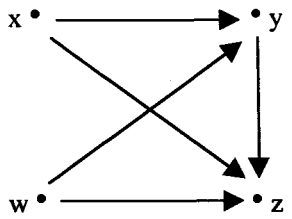


This relation is asymmetric, complete and transitive. It is a strict simple order, to which corresponds the following ranking without ex-aequo:

w
↓
x
↓
y
↓
z

■

Example 4.3



This relation is asymmetric and negatively transitive. It is a strict weak order, to which corresponds the following ranking with ex-aequo (w and x):

w , x
↓
y
↓
z

■

Theorem 4.3 shows that the notion of semiorder is linked to the notion of *threshold*. Even if we use precise measurement tools, a threshold exists below which

differences are not noticeable, leaving people sometimes to declare that "x is equal to y", "y is equal to z", but "x is not equal to z".

As stated by Pirlot and Vincke [14], *this fact has been acknowledged in the past by scientists like G. Fechner as early as in 1860, by H. Poincaré in 1905 and by others. In decision-aid, it is D. Luce who pointed out, in 1956, the phenomenon of intransitivity of "equality".*

Luce is the author of the famous example of the cup of tea, which shows that, even for a "totally" rational person, the indifference relation can be non-transitive for she is physically unable to perceive little differences. Pirlot and Vincke [14] present the example as follows: *Let t_i a cup of tea containing i milligrams of sugar. Any human being, comparing cups of tea, will generally consider that there is no difference between t_i and t_{i+1} (nobody is able to perceive a difference of 1 milligram of sugar), and this, for every i. We say that a person is indifferent between t_i and t_{i+1}. However, she may have a preference for t_N over t_0 (or the contrary) when N is large enough.*

Concerning the relations P and I obtained with questioning procedure **QP1** (P asymmetric, $I = P^N$), the results above allow to state that:

1) The answers of J give rise to a ranking without ex-aequo of the elements of **X** (this ranking being linked to the attractiveness of the elements of **X** for J) iff P is transitive and I is reduced to $\{(x, x) \mid x \in \mathbf{X}\}$.
2) The answers of J give rise to a ranking with possibility of ex-aequo of the elements of **X** iff P is transitive and I is transitive.
3) If P is transitive but I is not transitive, we will be in presence of a threshold structure of the semiorder type iff P (or $S = P \cup I$) is semitransitive and Ferrers.

4.3.4 Practical issues

How can it be verified in practice that we are facing one of the three cases above?

Letting n be the number of elements of **X**, a simple practical test consists of:

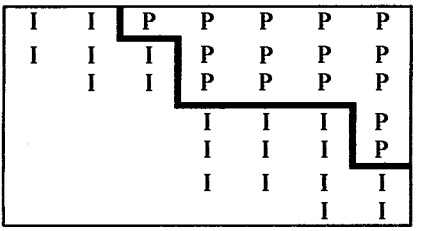
- ✓ The construction of a n×n tableau in which the entry x, y (line x, column y) is "P" iff xPy, or "I" iff xIy, or is empty iff yPx.
- ✓ The association, with each element x of **X**, of a number n(x) (called the *score* of x) equal to the difference between the number of P's in the line x and the number of P's in column x.
- ✓ The definition of a ranking without ex-aequo *Rn* of the elements of **X** such that, $\forall x, y \in \mathbf{X}$: xPy $\Rightarrow$ x is ranked before y in *Rn*.
- ✓ The construction of an n×n tableau of the relations P and I, adapted to *Rn* in the sense that, $\forall i \in \{1, 2, ..., n\}$, line i and column i correspond to the element ranked i in *Rn*.

It can easily be proved that:

A) (P, I) has a semiorder structure iff in the tableau adapted to *Rn*:
- *i* any entry in the main diagonal of the tableau is "I" and all "I" entries are symmetric in regard of the main diagonal
- *ii* any entry at the right or above a "P" entry is a "P" entry.

Remark Condition ***ii*** can be formulated in the following more original way: the broken line separating the "P" entries from the entries with no "P" (empty or "I" entries) is not interrupted and represents "steps" which are above the main diagonal.

Example 4.4



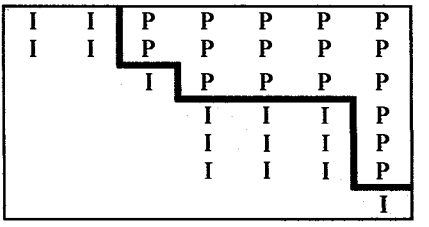
■

B) (P, I) corresponds to a ranking with possibility of ex-aequo iff in the tableau adapted to *Rn*:
i and ***ii*** are both verified, and
iii in any situation of the type

the "I" entries are situated in the main diagonal of the tableau.

Remark Conditions ***ii*** and ***iii*** can be formulated in the following more original way: the broken line separating the "P" entries from the entries with no "P" (empty or "I" entries) is not interrupted and represents "steps" which are above the main diagonal and such that each one is supported (at least in part) by the main diagonal.

Example 4.5

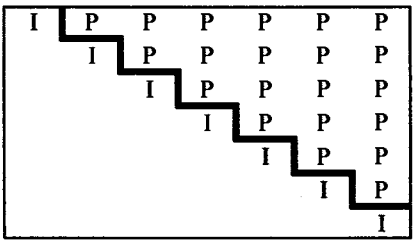


■

C) (P, I) corresponds to a ranking without ex-aequo iff in the tableau adapted to *Rn*: ***i***, ***ii*** and ***iii*** are all verified, and
iv all the "I" entries are situated in the main diagonal of the tableau.

Remark Conditions ***ii***, ***iii*** and ***iv*** can be formulated in the following more original way: the broken line separating the "P" entries from the entries with no "P" (empty or "I" entries) is not interrupted and represents "steps" formed by only an entry and supported by the main diagonal.

Example 4.6



■

Example 4.7 (illustrating the entire procedure)

Let $\mathbf{X} = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}\}$ and suppose that the judgements of J are:

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
x_1	I	P		P	P	I	P	P	P	P
x_2		I		P	I		P		I	I
x_3	P	P	I	P	P	P	P	P	P	
x_4				I			I			
x_5		I		P	I		P		I	
x_6	I	P		P	P	I	P	P	P	P
x_7				I			I			
x_8		P		P	P		P	I	P	P
x_9		I		P	I		P		I	
x_{10}		I	P	P	P		P		P	I

Scores:

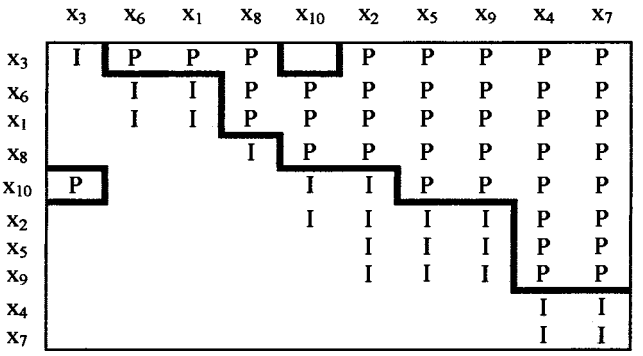
$n(x_1) = 7 - 1 = 6$
$n(x_2) = 2 - 4 = -2$
$n(x_3) = 8 - 1 = 7$
$n(x_4) = 0 - 8 = -8$
$n(x_5) = 2 - 5 = -3$
$n(x_6) = 7 - 1 = 6$
$n(x_7) = 0 - 8 = -8$
$n(x_8) = 6 - 3 = 3$
$n(x_9) = 2 - 5 = -3$
$n(x_{10}) = 5 - 3 = 2$

Ranking *Rn*:

x_3 ↓ x_1 ↓ x_6 ↓ x_8 ↓ x_{10} ↓ x_2 ↓ x_5 ↓ x_9 ↓ x_4 ↓ x_7

Note Other different *Rn* rankings could be taken. For all of them the procedure would give rise to the same conclusion stated hereafter.

The tableau adapted to *Rn* is



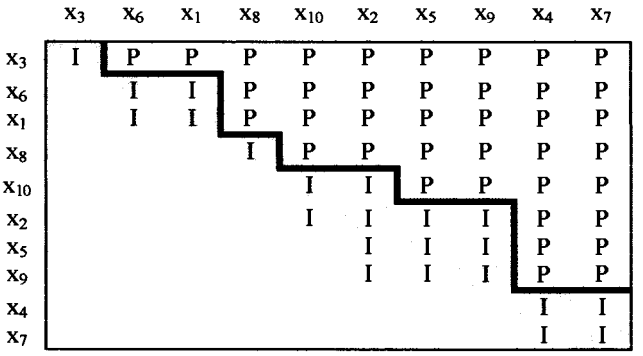
Conclusion The broken line separating the "P" entries from the entries with no "P" (empty or "I" entries) is interrupted, which reveals that the judgements of J correspond neither to a ranking nor to a semiorder.

An advantage of this type of tableau for decision-aiding is that it highlights the judgements that generates the absence of ranking. Confronted with such situation, J can reflect more deeply on her own judgements and eventually revise them. Helping people to learn about their own preferences is indeed a fundamental goal of the decision-aid activity.

In the tableau above, the judgement $x_{10}Px_3$ seems very peculiar. It appears in the following three cyclic situations:

$$x_3Px_6,\ x_6Px_{10},\ x_{10}Px_3$$
$$x_3Px_1,\ x_1Px_{10},\ x_{10}Px_3$$
$$x_3Px_8,\ x_8Px_{10},\ x_{10}Px_3.$$

Suppose that, after reflection, J reviews her initial judgement between x_{10} and x_3, and judges now that x_3 is more attractive than x_{10}. The new scores of the actions are such that one can always take as *Rn* ranking the same ranking as before. The tableau adapted to *Rn* will then become:



The judgements of J have now a semiorder structure. It is possible to associate a real number $\mu(x)$ to each element of **X** in such a way that, $\forall x, y \in \mathbf{X}$ [xPy iff $\mu(x) > \mu(y) + 1$]. For instance, one can take

$$\mu(x_3) = 10$$
$$\mu(x_6) = \mu(x_1) = 8$$
$$\mu(x_8) = 5$$
$$\mu(x_{10}) = 3$$
$$\mu(x_2) = 2.4$$
$$\mu(x_5) = \mu(x_9) = 1.7$$
$$\mu(x_4) = \mu(x_7) = 0.$$

One can also easily see why there is no ranking. J has judged:

$x_{10}Ix_2$, x_2Ix_5, $x_{10}Px_5$ and $x_{10}Ix_2$, x_2Ix_9, $x_{10}Px_9$

This is a situation which - although considered as normal by Luce - is often a matter of reflection by a person who makes the value judgements. If, confronted with the situation, J says that she felt the difference of attractiveness between x_{10} and x_2 very weak, and that, for a matter of precision, she would judge x_{10} more attractive than x_2, one would have the following tableau:

	x_3	x_6	x_1	x_8	x_{10}	x_2	x_5	x_9	x_4	x_7
x_3	I	P	P	P	P	P	P	P	P	P
x_6		I	I	P	P	P	P	P	P	P
x_1		I	I	P	P	P	P	P	P	P
x_8				I	P	P	P	P	P	P
x_{10}					I	P	P	P	P	P
x_2					I	I	I	I	P	P
x_5						I	I	I	P	P
x_9						I	I	I	P	P
x_4									I	I
x_7									I	I

Now, the judgements of J give rise to a ranking which can be obtained by extending the "steps" horizontally or vertically.

	x_3	x_6	x_1	x_8	x_{10}	X_2	x_5	x_9	x_4	x_7
x_3	I	P	P	P	P	P	P	P	P	P
x_6		I	I	P	P	P	P	P	P	P
x_1		I	I	P	P	P	P	P	P	P
x_8				I	P	P	P	P	P	P
x_{10}					I	P	P	P	P	P
x_2					I	I	I	I	P	P
x_5						I	I	I	P	P
x_9						I	I	I	P	P
x_4									I	I
x_7									I	I

The ranking is

$$x_3$$
$$\downarrow$$
$$x_6 , x_1$$
$$\downarrow$$
$$x_8$$
$$\downarrow$$
$$x_{10}$$
$$\downarrow$$
$$x_2 , x_5, x_9$$
$$\downarrow$$
$$x_4 , x_7$$

■

The learning process exemplified above for aiding J to construct a ranking of the elements of **X** (linked to their attractiveness for J) can be seen as the first phase of the MACBETH approach (**M**easuring **A**ttractiveness by a **C**ategorical **B**ased **E**valuation **Tec**H**nique**) developed by Bana e Costa and Vansnick (see [2] and [4]). Applied to a tableau of judgements, the MACBETH software computes the scores of the different elements, establishes a ranking *Rn* and displays on the screen the tableau adapted to *Rn* (with all "P" entries coloured (light grey in Fig. 4.1)). The software also indicates if the judgements of J give rise to a ranking and, if not, it verifies if at least a semiorder structure is present. Moreover, as the "P" entries are coloured, an examination of the tableau allows to easily identify which judgements are responsible for the absence of ranking. This is a friendly visual interactive platform for discussing with J.

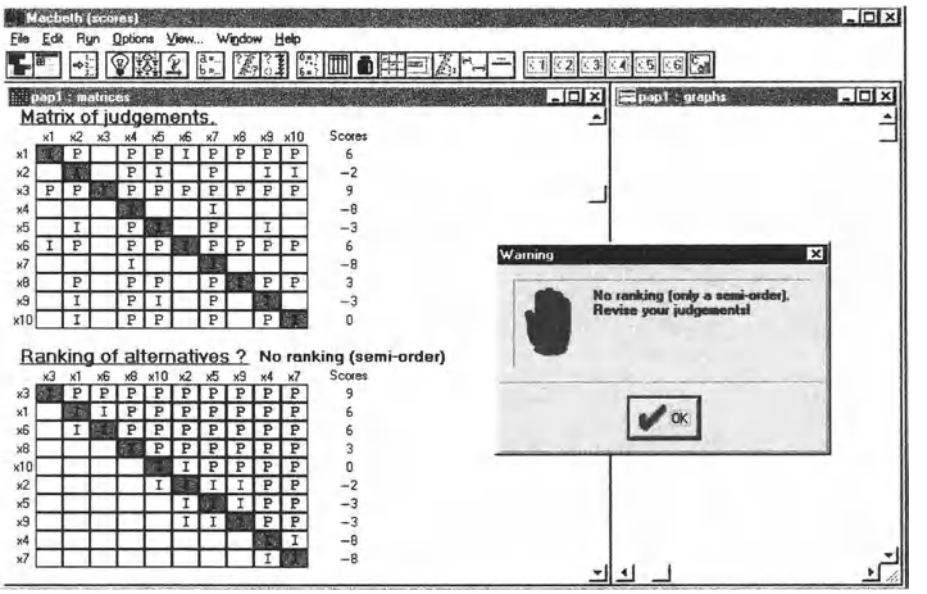


Figure 4.1 First phase of the MACBETH approach.

Fig. 4.1 shows the MACBETH screen for the case discussed in example 4.7, where the initial judgement "x_{10} is more attractive than x_3" has already been modified to "x_3 is more attractive than x_{10}."

4.4 INTRODUCING THE IDEA OF INCOMPARABILITY

4.4.1 Questioning procedures QP2 and QP2^{bis} and paradigm 2

Questioning procedure 2 (QP2)			
Is one of the actions (x or y) more attractive than the other?	YES	NO	I DON'T KNOW
If YES, which action (x or y) is more attractive than the other?	x	y	

The questioning procedure **QP2** is associated with the following paradigm:

Paradigm 2 Three (and only three) basic preference situations (distinct and exclusive) should be considered for the expression of value judgements: the *situation of preference*, the *situation of indifference* and the *situation of incomparability*.

The answers obtained with **QP2** allow the construction of three binary relations on **X**:

the relation P defined by xPy iff J judged x more attractive than y,
the relation I defined by xIy iff J answered NO to the first question
and
the relation ? defined by x?y iff J answered
I DON'T KNOW to the first question.

The questioning procedure **QP2** automatically leads to [P asymmetric], [I symmetric] and [? symmetric]; moreover, it will always be supposed that J's answers are such that [I reflexive] and [? irreflexive]. P is called *"strict" preference relation*, I *indifference relation* and ? *incomparability relation*.

It is worthwhile noting that **QP2** is equivalent to the following questioning procedure (**QP2^{bis}**), provided that an answer NO to both of its questions is equivalent to answer I DON'T KNOW to the first question of **QP2**:

Questioning procedure 2bis (QP2^{bis})		
Is x at least as attractive as y?	YES	NO
Is y at least as attractive as x?	YES	NO

The answers obtained with **QP2^{bis}** allow the construction on **X** of a binary relation S defined by xSy iff "x is at least as attractive as y". S is called a *"large" preference relation* or an *outranking relation*. When **QP2^{bis}** is used, the rational hypothesis required in J's answers is that they should be such that S is reflexive.

From S, one can derive the three relations P, I and ? as follows:

$$\begin{array}{lll} xPy & \text{iff} & xSy \text{ and } y\not Sx \quad (P = S^A) \\ xIy & \text{iff} & xSy \text{ and } ySx \quad (I = S^S) \\ x?y & \text{iff} & x\not Sy \text{ and } y\not Sx \quad (? = S^N). \end{array}$$

Note that under paradigm 2 - contrary to what happens under paradigm 1 - it is no more possible to derive I and S from P, except when ? is empty; indeed, one has $P^N = I \cup ?$ and $P^d = S \cup ?$. However, one always has $S = P \cup I$.

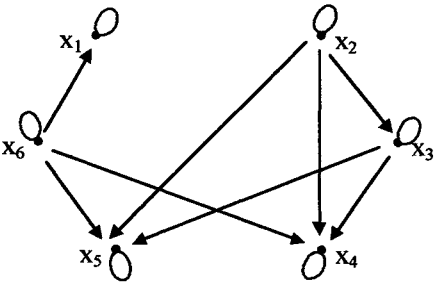
Theoretically, the difference between paradigm 1 and paradigm 2 is that, contrary to the latter, the former implies that $S = P \cup I$ is strictly complete.

4.4.2 Related preference structures

The *total order*, *total preorder*, and *semiorder* structures, introduced in the framework of paradigm 1 (see Table 4.1) have been also studied in the context of paradigm 2. *Partial order* and *partial preorder* are classical structures but *partial semiorder* is a more recent one.

The definitions of *partial order* and *partial preorder* can be obtained from the definitions of *total order* and *total preorder* by replacing strictly complete by reflexive. Intuitively, a partial order corresponds to the possibility of ranking without ex-aequo the elements of some subsets of **X**, while the partial preorder corresponds to the possibility of ranking, eventually with ex-aequo, the elements of some subsets of **X**.

Example 4.8 The following is a representation of a partial order on $\mathbf{X} = \{x_1, x_2, x_3, x_4, x_5, x_6\}$.



It corresponds to the following rankings without ex-aequo of the subsets of **X**:

$\{x_6, x_1\}$	$\{x_6, x_4\}$	$\{x_6, x_5\}$	$\{x_2, x_3, x_4\}$	$\{x_2, x_3, x_5\}$
x_6	x_6	x_6	x_2	x_2
$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$
x_1	x_4	x_5	x_3	x_3
			$\downarrow$	$\downarrow$
			x_4	x_5

■

It is easy to verify that, if $R_1, R_2, \ldots, R_k$ are k total orders (resp. total preorders) on **X**, then $R = R_1 \cap R_2 \cap \ldots \cap R_k$ is a partial order (resp. partial preorder). It is often in this framework that these structures appear in MCDM.

Given a partial order (or preorder), the problem of replacing incomparability by preference situations in order to obtain a total order (or preorder) can also be interesting in MCDM. For the case of partial orders, this problem has been studied by Szpilrajn [21] since 1930 and it is the basis for the concept of *dimension* of a partial order.

For a deep study of partial structures, we refer to [17, 10].

In the frameworks of both questioning procedures **QP1** and **QP2**, the relation P is transitive in all the particular preference structures introduced in sections 4.3 and 4.4. Let us point out that P transitive implies P acyclic, which is very interesting for MCDM. Indeed, in such a case, for each non-empty subset Y of **X**, $C(Y, P) \equiv \{y \in Y \mid zPy \text{ for no } z \in Y\} \neq \emptyset$, that is, C(. , P) is a choice function on **X**.

4.5 INTRODUCING THE IDEA OF HESITATION

4.5.1 Questioning procedure QP3 and paradigm 3

Questioning procedure 3 (QP3)

Is one of the actions (x or y) more attractive than the other?
YES NO I HESITATE I DON'T KNOW

If YES, which action (x or y) is more attractive than the other? x y

If you HESITATE, is your hesitation:
1) between "x and y are equally attractive" and "x is more attractive than y"
or (exclusive)
2) between "x and y are equally attractive" and "y is more attractive than x"?

The questioning procedure **QP3** is associated with the following paradigm:

Paradigm 3 Four (and only four) basic preference situations (distinct and exclusive) should be considered for the expression of value judgements: the *situation of preference*, the *situation of indifference*, the *situation of hesitation between indifference and preference* and the *situation of incomparability*.

The answers obtained with **QP3** allow to construct four binary relations on **X**:
the relation P defined by xPy iff J judged x more attractive than y,
the relation I defined by xIy iff J answered NO to the first question,
the relation Q defined by xQy iff J hesitated between
"x and y are equally attractive" and "x is more attractive than y"
and
the relation ? defined by x?y iff J answered I DON'T KNOW to the first question.

The questioning procedure **QP3** automatically leads to [P asymmetric], [Q asymmetric], [I symmetric] and [? symmetric]; moreover, it will always be supposed that J's answers are such that [I reflexive] and [? irreflexive].

Roy has termed the relations P *strict preference*, I *indifference*, Q *"weak" preference*, and ? *incomparability* - see, for instance, [18, 19]. However, it is worthwhile noting that the use of the term "weak" preference for Q has unfortunately been prone to misunderstanding. Indeed, many people think that Roy's work was developed in a *preference intensity* framework, but the truth is that Roy works in the framework of *credibility of preference*. Note also that the incomparability relation is often represented by "R" instead of "?" (see, for example, [19]).

The main theoretical work concerning (P, Q, I, ?)-structures deals essentially with the three relations P, Q, and I, i.e. is developed for the case in which the incomparability relation ? is empty. The (P, Q, I)-structure does represent an important case, particularly when it implies the existence of a mapping μ from $\mathbf{X}$ to $\mathfrak{R}$ and two real constants $c_2 > c_1 \geq 0$ such that, $\forall x, y \in \mathbf{X}$,

$$\begin{aligned} xIy \quad & \text{iff} \quad |\mu(x) - \mu(y)| \leq c_1 \\ xQy \quad & \text{iff} \quad c_1 < \mu(x) - \mu(y) \leq c_2 \\ xPy \quad & \text{iff} \quad \mu(x) - \mu(y) > c_2. \end{aligned}$$

In such a particular case, Cozzens and Roberts [5] speak about "double semiorder" and Roy and Vincke [20] speak about "pseudo-order with constants thresholds". For a deep study of the (P, Q, I)-type structures see [5, 20, 24].

4.5.2 Questioning procedures QP3⁺

In section 4.4, we have shown that it is possible to obtain a questioning procedure equivalent to **QP2** based on questions of the type "x is at least as attractive as y". The same idea can be used to define another questioning procedure (**QP3⁺**) which is not equivalent to **QP3** but is very interesting as it easily permits the taking into account of missing, ambiguous, contradictory or conflicting information.

Questioning procedure 3⁺ (QP3⁺)

Is x at least as attractive as y?

YES NO I HESITATE* I DON'T KNOW**

Is y at least as attractive as x?

YES NO I HESITATE* I DON'T KNOW**

* I HESITATE between YES and NO because certain aspects invite me to answer YES while other aspects invite me to answer NO

** I DON'T KNOW because I have no information that makes me answer YES or NO.

The answers obtained with **QP3⁺** allow the construction on $\mathbf{X}$ of ten binary relations, six of them being asymmetric (answers {YES, NO}, {YES, I HESITATE}, {YES, I DON'T KNOW}, {NO, I HESITATE}, {NO, I DON'T KNOW}, {I HESITATE, I DON'T KNOW}) and the other four being symmetric (answers {YES, YES}, {NO, NO}, {I HESITATE, I HESITATE}, {I DON'T KNOW, I DON'T KNOW}).

QP3$^+$ seems particularly well adapted to the case in which the questions concern the overall attractiveness of the elements of **X**. For the few available results concerning this new type of approach, see [22, 23].

4.5.3 Questioning procedures QP3^F

The following is another questioning procedure (**QP3^F**) linked with paradigm 3, that is, the paradigm that introduces the hypothesis of hesitation in the expression of value judgements:

Questioning procedure 3^F (QP3^F)

To each of the following statements, assign a real number between 0 and 1 quantifying the credibility of the statement being true

x is at least as attractive as y	Credibility: ...
y is at least as attractive as x	Credibility: ...

The answers obtained with **QP3^F** allow the construction of a mapping S^F from $\mathbf{X}\times\mathbf{X}$ to [0, 1] which associates with each ordered pair (x, y) a number $S^F(x, y)$ quantifying the *credibility* of the statement "x is at least as attractive as y". This mapping is often called a *fuzzy outranking relation* on **X**.

In general, a mapping R^F: $\mathbf{X}\times\mathbf{X} \rightarrow [0, 1]$ is called a *fuzzy binary relation* defined on **X**. Such mapping can be seen as an extension of the classic notion of a (crisp) binary relation R on **X**. Indeed, if $\forall x, y \in \mathbf{X}$, we state that

$$R(x, y) = 1 \text{ iff } (x, y) \in R \quad \text{and} \quad R(x, y) = 0 \text{ iff } (x, y) \notin R,$$

the binary relation R on **X** appears as a mapping from $\mathbf{X}\times\mathbf{X}$ to $\{0, 1\}$.

Given the fuzzy relation R^F defined on **X** and a real number $\lambda \in [0, 1]$, we denote by R_λ the classic binary relation defined on **X** by: $\forall x, y \in \mathbf{X}$, $xR_\lambda y$ iff $R^F(x, y) \geq \lambda$. R_λ is called *λ-cut* of R^F.

A possible way of studying fuzzy binary relations consists in extending the basic properties of (crisp) binary relations to the fuzzy context. There are several possible ways of making such an extension. The following definitions of fuzzy properties are of particular interest given the fact that all λ-cuts of a fuzzy relation R^F verifying one (or several) of these properties will verify the corresponding crisp property(ies).

A fuzzy relation R^F on **X** is

✓	reflexive	iff	$\forall x \in \mathbf{X}$: $R^F(x, x) = 1$
✓	symmetric	iff	$\forall x, y \in \mathbf{X}$: $R^F(x, y) = R^F(y, x)$
✓	asymmetric	iff	$\forall x, y \in \mathbf{X}$: $\min(R^F(x, y), R^F(y, x)) = 0$
✓	strongly complete	iff	$\forall x, y \in \mathbf{X}$: $\max(R^F(x, y), R^F(y, x)) = 1$
✓	transitive	iff	$\forall x, y, z \in \mathbf{X}$: $\min(R^F(x, y), R^F(y, z)) \leq R^F(x, z)$
✓	semitransitive	iff	$\forall x, y, z, w \in \mathbf{X}$: $\min(R^F(x, y), R^F(y, z)) \leq \max(R^F(x, w), R^F(w, z))$
✓	Ferrers	iff	$\forall x, y, z, w \in \mathbf{X}$: $\min(R^F(x, y), R^F(w, z)) \leq \max(R^F(x, z), R^F(w, y))$.

On the other hand, it is also interesting to study a fuzzy binary relation based on its λ-cuts. It is clear that it is equivalent to work with a fuzzy relation R^F or with all the λ-cuts of R^F for $\lambda \in [0, 1]$, or even only with the λ-cuts of R^F corresponding to λ values equal to $R^F(x, y)$ for at least a pair (x, y). In this last case, one will be in the presence of a nested family of relations (because $\lambda_1 < \lambda_2 \Rightarrow R_{\lambda 1} \supset R_{\lambda 2}$) and this family is of course finite when **X** is finite.

The notion of a *family of relations* has given rise to wide research, from the pioneering work of Roberts [15] and Fishburn [9] to the more recent article by Doignon, Monjardet, Roubens and Vincke [7]. These references cover the topic in considerable depth.

Still in the framework of questioning procedure **QP3**^{F}, let us finally mention the hard problem of, based on the fuzzy outranking relation S^F on **X**, defining a fuzzy preference relation P^F on **X**, a fuzzy indifference relation I^F on **X** and a fuzzy incomparability relation $?^F$ on **X**. For the study of this problem, together with the representation of fuzzy relations in preference modelling, see [13].

4.6 OTHER COMPLEMENTARY QUESTIONING PROCEDURES

In going from paradigm 1 to paradigms 2 and 3, we have introduced the possibility of explicitly taking into account the situations of "hesitation" and of "inability to choose one of the proposed preference situations" when expressing value judgements (be it because of uncertainty, ambiguous or contradictory information, or due to lack of sufficiently rich information). The possibility of modelling situations of poor information offered by questioning procedures **QP2**, **QP2**^{bis}, **QP3**, **QP3**$^+$ and **QP3**^{F} is undoubtedly interesting.

On the other hand, it is also very interesting to complement **QP1** with a richer and more precise questioning procedure enabling one to go from ordinal to cardinal preference modelling, in particular when one wants to assess value judgements about the relative attractiveness of actions with regard to a certain point of view. As a matter of fact, assessing ordinal value judgements is not too difficult but, unfortunately, this information is not enough in most practical applications of MCDM, in which one needs to know, not only that *x* is more attractive than *y*, but also by how much. It is precisely to address this problem that the MACBETH approach has been developed.

Based on semantic judgements of a person J about the attractiveness of the elements of a set **X**, MACBETH is an interactive approach to aid J to quantify the attractiveness of each element, in such a way that the measurement scale constructed is a cardinal scale on **X**. As a matter of fact, once a ranking (with possibility of ex-aequo) is obtained with questioning procedure **QP1** (that is the relations P and I are transitive) the following complementary questioning procedure is used in MACBETH to enter into the cardinal domain.

Complementary questioning procedure (QP1^C)

$\forall (x, y) \in P$, the difference of attractiveness between x and y is

VERY WEAK WEAK MODERATE STRONG VERY STRONG EXTREME

Other than the relations I and P previously obtained with questioning procedure **QP1**, **QP1**^{C} gives rise to a partition of P into six binary relations C_1, C_2, C_3, C_4, C_5 and C_6 defined as follows: $\forall$ (x, y) $\in$ P,

(x, y) $\in C_1$	iff	J judged the difference of attractiveness (dif. att.) between x and y to be VERY WEAK
(x, y) $\in C_2$	iff	J judged the dif. att. between x and y to be WEAK
(x, y) $\in C_3$	iff	J judged the dif. att. between x and y to be MODERATE
(x, y) $\in C_4$	iff	J judged the dif. att. between x and y to be STRONG
(x, y) $\in C_5$	iff	J judged the dif. att. between x and y to be VERY STRONG
(x, y) $\in C_6$	iff	J judged the dif. att. between x and y to be EXTREME.

Taking into account the goal of constructing a cardinal scale on X, the minimum rationality required in J's answers to questioning procedure **QP1**^{C} is that they should be such that there exists a mapping μ: X $\rightarrow \Re$ satisfying the following conditions:

Condition 4.1 $\forall x, y \in \mathbf{X}$: $xPy \Leftrightarrow \mu(x) > \mu(y)$

Condition 4.2 $\forall k, k' \in \{1, 2, 3, 4, 5, 6\}$,
$\forall x, y, w, z \in \mathbf{X}$ with $(x, y) \in C_k$ and $(w, z) \in C_{k'}$
$k > k' \Rightarrow \mu(x) - \mu(y) > \mu(w) - \mu(z)$.

Certainly, there exists μ: $\mathbf{X} \rightarrow \Re$ satisfying condition 4.1, because J's answers to **QP1** were supposed to give rise to a ranking of the elements of **X**. This is to say that an ordinal scale "measuring" the attractiveness of the elements of **X** can be defined on **X**. But, to construct with J a cardinal measurement scale on **X**, conditions 4.1 and 4.2 should be simultaneously satisfied. J's answers to **QP1**^{C} are said to be *consistent* (with regard to the main goal of constructing a cardinal scale) when it is possible to satisfy conditions 4.1 and 4.2 simultaneously.

The MACBETH software tests the consistency of J's judgements by linear programming. In case of inconsistency, the software immediately finds the possible sources of inconsistency, in order to enable a discussion with J and thereby to aid her to eventually revise her initial judgements. Moreover, in case of consistency, a first scale μ is obtained by linear programming. This scale can be used as the starting point for a discussion with J aiming at constructing a cardinal scale on **X** measuring the attractiveness, for J, of the elements of **X**. The MACBETH software was conceived to facilitate the perception by J of the first scale μ, the discussion of that scale and the evolution towards the cardinal domain (see [4] for details).

Theoretically, **QP1**^{C} can be linked to the two important research domains of *families of relations* and of *difference measurement*.

1) The binary relations C_1, C_2, C_3, C_4, C_5 and C_6 allow the definition of the following six relations, which form a nested family of binary relations on **X**:

$P_1 = C_1 \cup C_2 \cup C_3 \cup C_4 \cup C_5 \cup C_6$ (= P)
$P_2 = C_2 \cup C_3 \cup C_4 \cup C_5 \cup C_6$
$P_3 = C_3 \cup C_4 \cup C_5 \cup C_6$
$P_4 = C_4 \cup C_5 \cup C_6$
$P_5 = C_5 \cup C_6$
$P_6 = C_6$.

As said before, there are many works devoted to this subject. Let us detach the work of Doignon [6] that studies and solves the following general problem:
Given m relations $P_1, P_2, \ldots, P_m$ on the finite set **X**, when do there exist a real-valued mapping f on **X** and nonnegative real numbers $\sigma_1, \sigma_2, \ldots, \sigma_m$ such that for all $j \in \{1, 2, \ldots, m\}$ and $x, y \in \mathbf{X}$, $xP_jy \Leftrightarrow f(x) > f(y) + \sigma_j$?
Doignon's results generalize the results of Cozzens, Roberts, Roy and Vincke mentioned in the last paragraph of section 4.5.1. Inspired by Doignon's work, Bana e Costa and Vansnick [3] give also necessary and sufficient conditions for the consistency of J's answers to **QP1**^{C}.

2) The binary relations I (here noted C_0), C_1, C_2, C_3, C_4, C_5 and C_6 enable one to define a binary relation $\succ^*$ on $\mathbf{X}\times\mathbf{X}$ by making, $\forall(x, y), (w, z) \in \mathbf{X}\times\mathbf{X}$,
either $(x, y) \succ^* (w, z)$ iff
$[(x, y) \in C_k, (w, z) \in C_{k'}, k, k' \in \{0, 1, 2, 3,4, 5, 6\}$ and $k > k']$
(*positive-difference structure*)
or $(x, y) \succ^* (w, z)$ iff
$[(x, y) \in C_k, (w, z) \in C_{k'}, k, k' \in \{0, 1, 2, 3, 4, 5, 6\}$ and $k > k']$
or $[(y, x) \in C_k, (z, w) \in C_{k'}, k, k' \in \{0, 1, 2, 3, 4, 5, 6\}$ and $k < k']$
or $[(x, y), (z, w) \in P = C_1 \cup C_2 \cup C_3 \cup C_4 \cup C_5 \cup C_6]$.
(*algebraic-difference structure*)

For the study of such structures see [8, 11, 16]. Relating more closely to the basic framework of the MACBETH approach, Adams [1] also gives a theoretical answer to the problem of the consistency of J's judgements.

Let us finally mention that other complementary questioning procedures than **QP1**^{C} can be added to **QP1**. A particular interesting one can be obtained from **QP1**^{C} by introducing the possibility of hesitation between semantic categories of difference of attractiveness. This option exists in MACBETH, in which J can choose to answer with several (consecutive) categories, for instance "weak to moderate" as illustrated in Fig. 4.2, if she hesitates in answering with only one category.

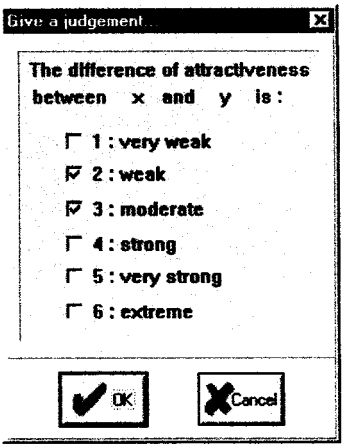


Figure 4.2 Hesitation between semantic categories.

REFERENCES

[1] Adams, E.W.: Elements of a theory of inexact measurement. Philosophy of Science 32 (1965) 205-228.

[2] Bana e Costa, C.A., J.C. Vansnick: MACBETH - An interactive path towards the construction of cardinal value functions. International Transactions in Operational Research 1, 4 (1994) 489-500.

[3] Bana e Costa, C.A., J.C. Vansnick: "A theoretical framework for measuring attractiveness by a categorical based evaluation technique (MACBETH)." In: J. Clímaco (ed): *Multicriteria Analysis*. Springer, Berlin 1997, pp. 15-24.

[4] Bana e Costa, C.A., J.C. Vansnick: "The MACBETH approach: Basic ideas, software and an application." In: N. Meskens, M. Roubens (eds): *Advances in Decision Analysis*. Kluwer Academic Publishers, Dordrecht (forthcoming).

[5] Cozzens, M.B., F.S. Roberts: Double semiorders and double indifference graphs. SIAM Journal of Algebraic Discrete Methods 3, 4 (1982) 566-583.

[6] Doignon, J.P.: Thresholds representations of multiple semiorders. SIAM Journal of Algebraic and Discrete Methods 8 (1987) 77-84.

[7] Doignon, J.P., B. Monjardet, M. Roubens, Ph. Vincke: Biorder families, valued relations and preference modelling. Journal of Mathematical Psychology 30, 4 (1986) 435-480.

[8] Fishburn, P.C.: *Utility Theory for Decision Making*. John Wiley & Sons, New York 1970.

[9] Fishburn, P.C.: Binary choice probabilities: On the varieties of stochastic transitivity. Journal of Mathematical Psychology 10 (1973) 327-352.

[10] Kacprzyk, J., M. Roubens (eds): *Non-Conventional Preference Relations in Decision Making*, Springer, Berlin 1988.

[11] Krantz, D.H., R.D. Luce, P. Suppes, A. Tversky: *Foundations of Measurement, Volume 1, Additive and Polynomial Representations*, Academic Press, New York 1971.

[12] Monjardet, B.: Axiomatiques et propriétés des quasi-ordres. Mathématiques et Sciences Humaines 63 (1978) 51-82.

[13] Perny, P., M. Roubens: "Fuzzy preference modelling." In: R.Slowinski (ed): *Fuzzy Sets in Decision Analysis, Operations Research and Statistics*. Kluwer Academic Publishers, Dordrecht 1998, pp. 3-30.

[14] Pirlot, M., Ph. Vincke: *Semiorders: Properties, representations, applications Statistics*. Kluwer Academic Publishers, Dordrecht 1997.

[15] Roberts, F.S.: Homogeneous families of semiorders and the theory of probabilistic consistency. Journal of Mathematical Psychology 8 (1971) 248-263.

[16] Roberts, F.S.: *Measurement Theory with Applications to Decisionmaking, Utility, and the Social Sciences*. Addison-Wesley, London 1979.

[17] Roubens, M., Ph. Vincke: *Preference Modelling*. Springer, Berlin 1985.

[18] Roy, B.: "The outranking approach and the foundation of ELECTRE methods." In: C.A. Bana e Costa (ed): *Readings in Multiple Criteria Decision Aid*. Springer, Berlin 1990, pp. 155-183.

[19] Roy, B.: *Multicriteria Methodology for Decision Aiding*. Kluwer Academic Publishers, Dordrecht 1996.

[20] Roy, B., Ph. Vincke: Pseudo-orders: Definition, properties and numerical representation. Mathematical Social Sciences 14 (1987) 263-274.

[21] Szpilrajn, E.: Sur l'extension de l'ordre partiel. Fundamenta Mathematicae 16 (1930) 386-389.

[22] Tsoukiàs, A., Ph. Vincke: A new axiomatic foundation of the partial comparability theory. Cahiers du LAMSADE 123 (1994) Université de Paris-Dauphine, Paris.

[23] Tsoukiàs, A., Ph. Vincke: "Extended preference structures in multi criteria decision aid." In: J. Clímaco (ed): *Multicriteria Analysis*. Springer, Berlin 1997, pp. 37-50.

[24] Vincke, Ph.: "(P,Q,I)-preference structures." In: J. Kacprzyk, M. Roubens (eds): *Non-Conventional Preference Relations in Decision Making,* Springer, Berlin 1988, pp. 72-80.

5 NORMATIVE AND DESCRIPTIVE ASPECTS OF DECISION MAKING

Oleg I. Larichev

CONTENTS

T. Gal et al. (eds.), *Multicriteria Decision Making*

Abstract: The problems of human behavior in decision processes are central in this chapter. The gaps between the requirements of decision methods and the possibilities of human information processing systems are analyzed. The qualitative model describing the decision maker's behavior is proposed. The model defines the guidelines for the construction of decision methods justified from behavior point of view.

5.1 SPECIFIC FEATURES OF DECISION MAKING AS A BRANCH OF RESEARCH

Decision Making, as a branch of research, has two main features distinguishing it from other research disciplines such as Economics and Operations Research:

1. For typical Operations Research problems, the information needed to solve a problem is given in the problem's statement. Contrary to the latter, the initial statement of any decision making problem has elements of uncertainty connected with a lack of information on general criterion of the solution's quality and/or the consequences of the decision's variants. This is why information from a Decision Maker (DM) and experts is required for the solution of a decision making problem. It is possible to say that a statement of any decision making problem includes a priori unknown preferences of the DM. This is why inside any decision making method there are some procedures of information elicitation from the DM (or a group of DM).

2. The primary step in many economical studies and in Operations Research is the construction of models representing the *reality,* small pieces of the real world having a mathematical description. In contrast to the latter, typical Decision Making problems imply the construction of a *subjective* model representing the personal perception of such a decision problem by the DM as the primary step. This subjective model reflects the DM's policy in the situation of a choice.

The distinguishing features of Decision Making as a research branch make the DM the central figure of the decision making process. Therefore, behavioral aspects become the central features in this line of research.

The goal of this chapter is to look at the Decision Making field of research from this point of view, to summarize the existing knowledge about human behavior in the decision processes, to analyze the existing gap between descriptive and normative approaches in decision making, and to draw some criteria for the construction of decision aiding tools and to demonstrate the importance of behavior aspects.

5.2 THE GAP BETWEEN DESCRIPTIVE AND PRESCRIPTIVE

The source of the widely accepted gap between the requirements of decision methods to human beings and the possibilities of human information processing systems lies in the historical development of Decision Making as a research field. Decision Making has two roots:

- economical utility theory;
- operations research.

5.2.1 Utilities and probabilities

Each portion of a purchased commodity (e.g., bread or tea) has its consumer utility. The law of marginal utility reads as follows: the marginal utility decreases, that is, subsequent portions of a commodity are less valuable to the consumer than initially, which is quite understandable from our everyday experience. If there exists a need for several commodities, the consumer attempts to allocate money to support a constant ratio of the utility of a commodity to the general measurement unit (dollars, rubles, etc.).

Stated differently, more money must be invested into commodities of higher utilities. The same human behavior is inherent to the problem of investments - more money is put into more useful areas of activity. Economists believe this is the only correct behavior and refer to the person making such a choice as *a rational person.* It is assumed that rational persons are intrinsically consistent and that transitivity of choice is appropriate to them.

Second, it is assumed that, when making a decision, a rational person maximizes his utility.

To conclude, what does the rational person do? First of all, they list all possible decisions and their consequences for which utilities (or money values) are determined. For each variant of a decision, the probabilities of all its outcomes are determined (no matter how). Next, the expected utility of each variant is calculated by summing the products of utilities by corresponding probabilities. The best variant is that which has the maximum expected utility.

J. von Neumann and O. Morgenstern laid the scientific foundation for the utility theory in their well-known "Theory of Games and Economic Behavior" [48]. The utility theory as presented in this book is axiomatic. The originators of the utility theory made use of so-called lotteries, where two results (outcomes) with respective probabilities exist, as simple problems of choice and demonstrated that if human preferences for simple problems (lotteries) satisfy some axioms, then human behavior can be regarded as maximizing expected utility.

The axioms used by the authors of [48] assert, for example, that a person can compare all outcomes and he/she is transitive, due to possibility of determining the probabilities under which lotteries constructed on pairs of outcomes (out of three) are equivalent, etc. The axioms are required to infer the theorem of existence of the utility function for a person that agrees with the axioms.

The internal utility function of the DM is used to measure the utility of any outcome. The theory presented in the classic book by J. von Neumann and O. Morgenstern needs a quantitative measurement of all utilities and probabilities.

Von Neumann and Morgenstern's theory assumes that probabilities are given as objectively known magnitudes. D.Savage [39] developed an axiomatic theory enabling one to measure simultaneously utility and subjective probability which gave rise to the model of Subjective Expected Utility (SEU) where the probability is defined as the degree of confidence in fulfillment of one or another event.

Together with the development of utility theory and SEU some findings appeared concerning human behavior in the lotteries choice. Well known is the so-called Allais' paradox that was the object of hot disputes for several years [37]. People repeatedly demonstrated contradictory numerical evaluations of utility in the

tasks of choice.

Inconsistent human behavior in lotteries choice was demonstrated [5]. Furthermore, it was demonstrated that people don't believe in Savage axioms [43]. The difficulties of checking axioms in the applications became evident.

The entire research on Decision Making theory was greatly influenced by the psychological studies of P.Slovic, A.Tversky, B. Fischhoff [14], et al., who demonstrated the existence of human errors made when evaluating event probabilities. The main causes of these errors can be represented as follows [14]:

- Judgement from representativeness: people judge about the membership of an object A to the class B only from its similarity to the typical representative of B disregarding the a priori probabilities.
- Judgement from availability: people often evaluate the probabilities of events on the basis of their own meeting with such events.
- Judgement from the anchoring point: if initial information is used as a reference point for determining probabilities, then it exerts significant influence on the result.
- Superconfidence: people place too much confidence in their evaluations of event probabilities.
- Tendency to eliminate risk: people try to eliminate risky situations as much as possible.

These work may bring into question the possibility of practical application of utility theory and SEU theory.

Clearly, it was the first demonstration of a gap between descriptive and normative.

5.2.2 Prospect theory

Attempts were made to update utility theory so as to eliminate the most salient discrepancies between theory and real human behavior. The Theory of Prospects [4, 15] is the most conspicuous attempt of this kind. By prospect we mean a game with probabilistic outcomes.

Prospect theory allows for three behavioral effects:

- certainty effect, that is, the tendency to give greater weights to determinate outcomes,
- reflection effect, that is, the tendency to change preferences upon passing from gains to losses, and
- isolation effect, that is, the tendency to simplify choice by eliminating the common components of decision variants.

All these effects being taken into consideration, the value of a lottery to gain outcomes x and y with respective probabilities p and q is defined by multiplying the utilities of the outcomes by the subjective importance of the probabilities of these outcomes. The function of the subjective importance of the probabilities has some specific features that allow one to avoid the Allais' paradox and give some explanations to other disagreement between the theory and human behavior.

The theory of prospects recommends to "edit" prospects before comparing them - for example, to eliminate identical outcomes with identical probabilities, to merge in one prospects with identical outcome, and so forth.

Despite the fact that the theory of prospects eliminates some paradoxes of

choice stemming from the utility theory, it does not eliminate all problems and paradoxes appearing upon studying human behavior in the problems of choice. The possible paradoxes appear when editing the prospects. A solution of the problem depends very much on the way in which one "frames" it. Furthemore, the quite natural desire to round the probabilities and utilities leads to different results of the prospect comparison depending on the rounding [53].

The prospect theory, as well as the utility theory relies on an axiomatic basis. A common problem with all axiomatic theories is the validation of the axioms allowing one to use one or another form of the function of utility (value) of the theory.

The common feature for the utility theory, SEU and the prospect theory is the same representation (model) of the decision problem: in the form of the holistic parameters of the utility (value) and probability. In the middle of the 70-ies a different model became more popular and promising: the multicriteria description of the positive and negative factors influencing the choice. The reason is that utilities and probabilities manifest itself in the multiple criteria of alternatives' evaluation.

5.2.3 Multiattribute utility theory

The next step in the evolution of the utility theory was marked by the transition to the multicriteria or Multiattribute Utility Theory - "MAUT" [16]. The construction of a strict and harmonic mathematical theory of utility under multiple criteria was a great merit of R. Keeney and H. Raiffa. The theory is constructed axiomatically, where the general axioms of connectivity and transitivity on a set of alternatives, etc., are complemented by the axioms (conditions) of independence. There exist many conditions [12] which conceptually define the possibility of comparing alternatives in one criteria, while the estimates in other criteria are being fixed (at different levels). For example, the condition of the preference independence states that comparisons of alternatives in two criteria are valid if their estimates in other criteria are fixed at any level. If the conditions of such a kind are met for all pairs of criteria, then the existence of a utility function in different forms is proved. We note that the Multicriteria Utility Theory is directed to the problems where existence of many alternatives justifies great efforts that are required to construct a utility (value) function.

After the development of "MAUT", critical comments were made about possibilities of validating all axioms and conditions needed for the existence of a multicriteria utility function in one or other form. For example, the sum of the importance coefficients of the criteria is to be equal 1 for the existence of a utility function in the additive form [17]. The question is: if we take into account the possible small errors in the measurement, which value of the sum is close to 1? [51].

In the construction of one dimensional utility functions the lotteries were used as the preferences elicitation tool. But human behavior in a lottery choice is inconsistent [5].

Again there is an evidence of a gap between the requirements of the decision methods (normative) and the possibilities of the human information

processing system (descriptive).

5.2.4 Multicriteria counterparts of the well-known problems of operations research

The second root of Decision Making as a branch of research is Operations Research.

The introduction of multiple quality criteria enables one to obtain multicriteria counterparts of the well-known problems of the operations research. For example, additional criteria are readily built into the generalized transportation problem [49] which can be formalized as a multicriteria problem of linear programming for which multitude of methods were developed [46]. Also, there are multicriteria assignment problems [22], multicriteria bin packing problems [24] which are counterparts of the well-known Operations Research problems.

There exist a great deal of the man-computer procedures enabling DM to examine the domain of the admissible decisions and at the same time to establish a compromise between the criteria [46].

The man-computer procedure consists of alternating phases of analysis (performed by the DM) and optimization (performed by the computer). Each phase can consist of more than one step.

Optimization phase (computer):

- using the information received from the DM at the preceding step, a new decision is computed and auxiliary information for the DM is generated.

Analysis phase (DM):

- the presented decision (or decisions) is estimated and its admissibility is determined. If the answer is positive, then the procedure terminates; otherwise, auxiliary information is considered;
- additional information is communicated to enable computation of a new decision.

The man-computer procedures differ in content and execution of the above steps. Their efficiency depends mostly on the nature of the DM-computer interaction that is represented in terms of the quality and quantity of the information.

Together with the development of many man-computer procedures, there appeared papers with the evaluation of such procedures from the behavioral point of view [18, 25]. The analysis demonstrated that many operations required from people in the framework of the man-computer procedures are difficult for the human information processing system. People show intransitivity in the process of choice, show suspiciously fast convergence to the solution and so on.

Again we witness here the evidence of a gap between descriptive and normative.

All the gaps mentioned above lead us to the question: what could be said about a human being as a DM?

5.3 THE QUALITATIVE MODEL OF THE HUMAN DECISION MAKER

On the level of the existing knowledge it is possible to summarize the evidence about human behavior in the decision processes in the following way.

5.3.1 The features of the human information processing system

A. Limited span of the working memory.

According to cognitive psychology [45], human beings have a limited span of the short -term memory. In repeated tasks the span of the working memory could be enlarged [9] but it takes both time and efforts.

That is why the DM cannot simultaneously pay attention to many factors (or evaluations of alternatives upon criteria) in the new decision tasks. As a matter of fact, for the new tasks DM has no possibility to create the internal structure of the necessary knowledge.

This limitation manifests itself in such known facts as:
The DM is trying to simplify the description of the decision situation by replacing some of the criteria by limitations, by eliminating some of the criteria, by grouping the alternatives and so on [25]. Such behavior is the unconscious desire to decrease the load on the short-term memory.

Experienced DMs have usually the skill of simplifying the decision situation in the best possible way. For inexperienced DMs a significant increase in the number of contradictions for more complex decision tasks is typical [25]

B. Limited exactness in quantitative measurements.

According to the existing knowledge, a human being is not an exact measurement device producing quantitative measurements. The famous experiment of A.Tversky [47] demonstrated that people neglect small differences in the evaluations. It is the reason for the intransitive behavior in some problems of choice. Inability to take into account small differences in the evaluation leads to the elimination of the dominating alternatives by the conservation of the dominated ones [19].

The experiments demonstrated that people can poorly measure the probabilities in the quantitative way (see above). The change in the method of measurement, the transfer from the quantitative to the verbal probability allows one to decrease significantly the number of the preference reversals [13].

It was demonstrated in the experiments [28] that slightly different procedures of the quantitative measurements for the same variables give quite different results.

C. Human errors and contradictions.

It has been known since the time of antiquity that "To err is human". People err when processing information. There could be different reasons for such a behavior: weariness, lack of attention, habitual heuristics and so on.

5.3.2 The features of human behavior in the decision processes

A. Absence of preconceived decision rules in new decision tasks.

As many researchers supposed, the DM has no preconceived decision rules. As noted in [52] it can be hardly expected that the utilities and numbers expressing the subjective estimates of the objects and situations are just stored in our minds until elicited.

To develop a decision rule the DM needs time and some learning procedures. Usually people use some kind of a "trial and error " approach in such procedures.

B. Search of the dominance structure.

At every step of the decision making procedures people pay attention to a limited number of objects. This is a possible explanation to the psychological theory of human behavior in the decision tasks-the search of the dominance structure [32]. According to the theory, in the case of the limited number of alternatives people make a preliminary selection of the potentially best alternative and compare it pair-wise with other alternatives , trying to check the fact of dominance.

In the case of a larger number of alternatives, people use initially the strategy of eliminating by aspects and after that utilize a more elaborated process (like the search of dominance) for a smaller number of alternatives.

C. Minimization of human efforts.

J. Payne et al. [35] suggested and substantiated another theory of human behavior upon choosing the best multicriteria alternative(s) that can be called the theory of the constructive processes.

When comparing multicriteria alternatives, people can use various strategies. The studies of J. Payne et al. [35] have demonstrated that in the process of the decision making subjects often choose a strategy depending on the specific features of the alternatives under consideration (their evaluations by criteria). Here, the human preferences of the alternatives and criteria are very unstable. At the local stages of the comparison, rules (or their parts) can vary depending on the relation between the required human effort and the accuracy of choice.

As J. Payne et al. notes, such a behavior is a characteristic of the untrained subjects. People experienced in the decision making, as well as regular decision makers have their preferable strategies for solving problems.

5.3.3 The features of human behavior in organizations

A. Satisfactory decisions.

The studies of economists and psychologists provided an insight into the human decision making in large organizations.

Ch. Lindblom [31] notes the officers organizations try to make as small changes in the existing policy as possible to be able to adjust to the environmental changes. It is not only easier to work out such changes, but also to coordinate them within an organization. The sequence of changes is mostly the means for forming the current policy. Lindblom also believes that this way of solving problems is more realistic because it requires less effort and is more customary for the managers. On the other hand, this approach is more conservative and is not adjusted to dramatic changes in the policy.

Similar discoveries were made by H.Simon [41] who introduced notion of satisfactory decisions as a counter to the optimal ones. In organizations, the life itself brings people to seek satisfactory decisions - the environment is too complicated to be described by a model, the multiple criteria are defined incompletely, there are many active groups influencing the choice, etc. This natural behavior of the personnel resulted in the loss of the strategic objectives amid the petty, everyday routine.

B. Taking the power in the hands.

The desire to have the decision situation under control is typical for the behavior of a DM in organizations. It means that the DM is trying to control all stages of the decision making, all transformations of the information influencing a decision.

Speaking differently, the DM is trying to have the power in the hands. In the case when it is necessary for him/her to take into account the interests of different active groups, the DM is looking for a mutually satisfying decision. [27] but he/she is always trying to implement the principal components of own policy.

5.4 HOW TO REDUCE THE GAP BETWEEN DESCRIPTIVE AND PRESCRIPTIVE

The above features of human behavior define the numerous gaps between the requirements of the different normative methods and the possibilities of human beings to meet such requirements. The discrepancy manifests itself in human errors and contradictions badly influencing the results of an analysis, as well as in mistrust in the results of the DM to the results presented by an analyst and so on.

There are several remedies to save the situation. First, it was the idea to improve the human performance in the process of choice: to teach people how to use the axiomatic methods or to train them to make the quantitative measurements.

Unfortunately, we do not have any evident confirmation of success for this approach. Even more, it became clear that many features of human behavior could be explained by the basic organization of the human information processing system [42].

The second reaction to the gaps is the following: human behavior is not important factor in the decision processes. One could take evaluations in the qualitative form but transform them quickly in the qualitative form appropriate for many decision methods. On the final stage of the decision process it is possible to use the so-called sensitivity analysis to check the influence of the different factors on the output of an the analysis.

Unfortunately, the task of sensitivity analysis is very complex. It is an independent difficult problem and only the skill of an analyst could shed light on the influence of the different factors.

The approach we have taken [26, 30] differs from the others. From our point of view it is necessary to adapt the decision methods to human behavior.

A possible way to close the gaps consists of taking the behavior finding concerning human behavior as constrains for the normative decision aiding methods [26]. By going on such a way it is possible to use the qualitative DM model described above as a base for the construction of the decision aiding tools and the decision support systems.

In other words, on the basis of the behavior findings it is possible to formulate special requirements to the characteristics of the decision aiding methods [30].

5.5 BEHAVIORAL REQUIREMENTS ON THE METHODS OF DECISION MAKING

The knowledge about human behavior in the decision processes allows us to define

the requirements for the methods for the decision making [26].

5.5.1 Measurements

The methods must be adjusted to the language of the problem description that is natural to the DM and their environment. To be socially acceptable, the decision method must be readily adjustable to the accepted way of discussing problems in a particular organization.

The kind of a "natural language" depends on the type of the problem. For so -called ill-structured [40] problems the combination of the quantitative and qualitative variables is typical. It is true, for example, for the multicriteria counterparts of the well-known problems of operations research. One has usually an objective quantitative model for such problems. The criteria for the evaluation of the decision's quality are some functions from the quantitative variables. Therefore, for such problems the quantitative language of measurement is natural.

There are many ill-structured and some unstructured problems where main variables (or criteria) have an objective quantitative nature like distance, money, a number of residents and so on [16]. Here also the quantitative language is natural and widely accepted.

For typical unstructured [30] problems, this usually means that the estimates of the criteria and, consequently, the estimates of the variants by the criteria are presented in a verbal form. The verbal estimates are located usually on the ordinal scales of the criteria. Such estimates create an adequate language for describing unstructured problems.

The decision method must be adjusted to such a description. Hence, by defining one or another form of verbal scales the DM defines the "measurer" for the experts estimating alternatives on these scales. The same verbal estimates are used by the DM to define the requirements to alternatives, that is, the decision rule.

Additionally to the kind of a problem, the methods of measurement are to be defined by an uncertainty level of evaluations upon criteria. The uncertainty level demonstrates itself in the ability of having an exact measure of the variable. In the case of exact measurements made by a measurement device one has negligible level of uncertainty. In case of the human measurements, levels of uncertainty are defined by the possibilities of a human being to give the information in a different form and with different reliability.

5.5.1.1 Quantitative human measurements

For the situation where the quantitative language of measurement is accepted it is necessary to take into account the characteristics of a human being as a measurement device.

The estimates on the criteria scales must reflect changes in the value (utility, preference, importance, distinctness, etc.) of an alternative with a corresponding change of the estimate by a given criterion. It is known that people can poorly estimate and compare objects of close utilities. With the continuous scales, slight distinctions in the estimates can result in different comparisons of the alternatives. Indeed, all other estimates being equal, the preferableness of an alternative will be defined by one insignificant difference.

The experiment of A. Tversky [47] that demonstrated the stable intransitivity of choice was based precisely on this property. The subjects were given successive pairs of alternatives where the gain slightly increased with a slight increase in the payment. And they persistently preferred to get a higher gain for a slight increase in the payment. Yet, when given the alternatives from the first and the last pairs, they persistently preferred the first one, because they could not admit such a great increase in the payment even for a corresponding increase in the gain. In our view, the continuous scale of the estimates prevented the subjects from seeing the gradual transition from quantity to quality. If the same subjects were given the same task but with the ordinal (qualitative or verbal) estimates of the payment and the gain, then a transitive relationship could easily result.

Some of the results of our experiments in choosing a summer country-house [33] are indicative for the inconvenience of the continuous scales. Two continuous-scale criteria, the cost and the size of the territory, were used in these experiments. It was noted that for insignificant discrepancies in the evaluation of the country-house the subjects do not necessarily find the alternatives that dominate the remaining ones. The subjects sometimes eliminated from the subset of the best alternatives (even if their number is only in the range from four to seven) an alternative dominating one of the remaining alternatives. This observation can also be attributed to the fact that insignificant (5-7%) variations in the cost do not affect appreciably the values of the alternatives. Though the subjects assert that `the cheaper the better' if this difference is pointed out, on the whole they agree that both variants have the same utility.

The experiments suggest that the quantitative measurements are the most sensitive to small errors and differences in the DM answers, which gives rise to the question of the accuracy of human measurements, especially under indefiniteness. It is well known that in physics the accuracy of measurements depends on the precision of the instrumentation. The same applies to human measurements. The available results of the experiments are indicative for the fact that man cannot make precise quantitative measurements.

It means that while performing the quantitative measurements it is much better to replace the continuous scales by the ones with discrete evaluations. Such evaluations could represent some intervals on the continuous scales which have a meaningful interpretation for the DM. Sometimes it is preferable to use for such intervals verbal labels like "expensive", "cheap" and so on for the cost evaluation.

5.5.1.2 Transition from qualitative notions to numbers

For the situations where qualitative language of a problem's description is natural (unstructured problems) let us discuss the attempts to combine the qualitative measurement scales and the quantitative representation of the results. First of all, we should mention the simple means of establishing a mutual correspondence between the primary qualitative measurement scale and the quantitative scale of scores where the primary measurements are carried out in a qualitative form and (independently of the expert's will) are assigned certain numbers which are then used to estimate the variants of the decisions. This method of measurement is not reliable because no logical basis underlies the assignment of one or another

numerical value to the primary estimates. The worst of it is that the numbers are further treated as the results of the objective physical measurements. For example, when estimating the quality of the objects by multiple criteria, the scores by criteria are regarded as the results of the quantitative measurements and are often multiplied by the weights of the criteria and summed up.

When considering the problems of a political choice, Dror [6] drew attention to the fact that people assign different numerical estimates to the same verbal definitions. We do not think that this necessarily means that one person believes that this event occurs with 70% probability and that 70% refers to a highly probable event, whereas another person believes that this event occurs with 90% probability and that 90% refers to a highly probable event. Both experts are, possibly, sure that this event is `very probable,' but when they are asked to evaluate this probability numerically (for example, in terms of a percent or somehow else) they replace their ignorance of this number by some (rather arbitrary) number. Human estimates corresponding to the same verbal definition on the scale were experimentally shown to have a rather great dispersion [50] which is especially great for the estimates representing the mean `neutral' level of the quality.

The second popular approach is that of the theory of fuzzy sets where measurements are carried out in terms of the descriptive qualitative values which are then transformed to the quantitative form by the means of a given membership function assigning numbers to any word.

To what extent is this transformation reliable? To what extent is man error-free? It is obvious that a person constructing the membership function performs approximately the same operation as when establishing the correspondence between the qualitative and the quantitative scales where the DM cannot evaluate the effect of small deviations in the estimates on the resulting comparison of the alternatives. The references to the check for sensitivity after quantitative measurements are of no avail. Indeed, in the presence of the multiple quantitative parameters the sensitivity check becomes an independent involved problem that can be solved only by eliciting from the DMs information that they hardly can provide.

5.5.1.3 Comparative verbal probabilities

Some experiments focus on the relationship between the language of measurement and the degree of indefiniteness of the events [8]. For example, the subjects were asked to estimate the chances of basketball teams to win in games between them. The experimenters noticed that in the case of an unknown team (higher indefiniteness) the experts were able to discriminate only two levels of verbal probabilities in comparative forms — for example, 'it is believed that the host always plays better than the guests'.. It is stated [8] that compelling people to quantify the probability estimates in the situations where only a few levels of indefiniteness can be discriminated can result in erroneous estimates. This example shows that some measurements can be carried out only in a verbal form with the use of the `more probable than' relationships. Methodical studies of the comparative probabilistic estimates [11] demonstrated that the comparative probabilities are much more frequently used by the common people (both adults and children) than

quantitative estimates of the probabilities of events. The experiments used tasks such as the estimation of the probabilities of hitting the sectors of a rotating disk and estimating the winners in competitions and games. The authors of this work formulated six mathematical principles for comparative probabilities in the form of axioms representing the mathematical concept of the qualitative probabilities. The main experimental result obtained with adults and children above five is as follows: human comparisons follow completely the principles of the mathematical theory of the qualitative probabilities. The authors of this work conclude that the six principles provide a more reliable foundation for describing human behavior than the laws of the quantitative probability.

5.5.1.4 Qualitative measurements

We regard decision making in the unstructured problems as the domain of the human activity where quantitative (the more so, objective) means of measurement are not developed, and it is unlikely that they will appear in the future. Therefore, it is required to estimate the possibility of doing reliable qualitative measurements. Following R. Carnap, we turn to the methods of measuring physical magnitudes that were used before the advent of the reliable quantitative measurements. Before the invention of balances, for example, objects were compared in weight using two relationships — equivalence (E) and superiority (L), that is, people determined whether the objects are equal in weight or one is heavier than the other. There are four conditions to be satisfied by E and L [3]:

1.E is the equivalence relationship,
2.E and L must be mutually exclusive,
3.L is transitive, and
4.For two objects a and b either (i) a E b, or (ii) a L b, or (iii) b L a.

One can easily see that the above scheme enables one to carry out relatively simple comparisons of the objects in one quality (weight). It is required here that all objects be accessible to the measurement maker (expert).

Two more remarks are due. It is obvious that the thus-constructed absolute ordinal scale cannot have many values; otherwise, they will be poorly distinguishable by the measurement makers. To come to terms easier, it is required to identify commonly understandable and identically perceived points on the scale and explain their meaning in detail. Therefore, these scales must have detailed verbal definitions of the estimates (grades of quality). Moreover, these definitions focus on those estimates on the measurement scale that were emphasized by the persons constructing the scale (for example, they could be interested only in very heavy and very light objects). Thus, the estimates on the ordinal scale are defined both by the persons interested in one or another kind of measurement (in our case, it is the DM) and by the distinguishability of estimates, that is, the possibility of describing them verbally in a form understandable to the experts and the DMs.

There is no reason to question the fact that before the coming of the reliable methods of quantitative measurement of the physical magnitudes, they were already measured qualitatively. Today, these methods could seem primitive because we have much more reliable quantitative methods. Yet, there is no doubt that the pre-quantitative (qualitative) methods of measuring physical magnitudes did exist.

When they were superseded by the quantitative methods, they were treated with negligence as something `unscientific' and obsolete. The progress of physics gave rise to the well-known statement that the science appears wherever the number (quantity) occurs. To our mind, these declarations refer mostly to the natural sciences, but in the sciences dealing with human behavior qualitative measurements were and will be the most reliable.

5.5.1.5 How to measure

For the conclusion, we could put the following requirements to human measurements in decision processes.

1. The measurements must be made in a language that is natural to DMs and their environment.
2. In the case of quantitative variables (criteria) it is preferable to use discrete scales with the evaluations representing some intervals meaningful for "measurement makers".
3. In the case of qualitative measurements the ordinal scales with verbal evaluations are the best way of measurement.
4. For the cases with big uncertainty the comparative verbal measurement (better, worse and so on) are the most correct way of receiving information from human beings.

In the general situation, one could take as the output of measurement process the discrete evaluations on criteria scales. Very often such evaluations have verbal labels or verbal descriptions.

5.5.2 Information elicitation for the construction of a decision rule

The next problem after the measurements is the construction of the decision rules for the evaluation of the alternatives. The problem for the DM is to construct the decision rules using the kind of criteria evaluations described above.

The operations performed by the DM in the process of constructing the DM's decision rules are to be psychologically correct. We shall differentiate between two types of measurements. We discussed above measurements of the main factors influencing the decision. We shall refer to them as the primary measurements. In some normative methods, the primary measurements suffice for reaching the final decision. In the method of the subjective expected utility, for example, the quantitative measurement of the utility and subjective probability allows one to calculate the expected utility of every alternative.

Yet, for a large majority of the normative methods this is insufficient, and some cognitive operations of the information elicitation are needed to construct a decision rule. We will call them the secondary measurements. For example, one needs to measure weights of criteria to decide whether the utility function is additive or multiplicative [16].

Analysis of the different normative techniques enables one to distinguish three groups of the information processing operations such as operations with criteria, operations with the estimates of the alternatives by criteria, and operations with the alternatives. Let us refer to an operation as elementary if it is not decomposable into simpler operations over to the objects of the same group, that is,

to criteria, alternatives, and alternative estimates by criteria.

In the survey [26] the results of psychological experiments demonstrating the validity of different cognitive operations used in the decision methods were collected. Below is a group of the information processing operations which are admissible from the psychological point of view for the construction of a decision rule [26, 30].

1. Ordering criteria by importance.
2. Qualitative comparison of two estimates taken from two criteria scales.
3. Qualitative comparison of the probabilities of two alternatives.
4. Attribution of alternatives to decision classes.
5. Comparison of two alternatives viewed as a set of estimates by criteria and selection of the best one.
6. Comparison of two alternatives viewed as something whole and selection of the best one.
7. Determination of a satisfactory level by one criterion.

Let us note that the operations 4 and 5 are admissible in some limits defined by the parameters of a problem. For the operations 6 and 7 we do not have enough of the psychological research demonstrating the validity of the operations, it is a preliminary conclusion.

The admissible operations, reduce to qualitative comparisons (of the type "better", "worse", "approximately equal") of criteria, pairs of estimates on two criteria scales, holistic images of alternatives. Also, we may assign satisfactory values, exercise a simple decomposition of criteria, alternatives. Given a relatively small number of the criteria, we may compare two alternatives. With a not too big number of criteria, decision classes, and estimates on scales we may assign alternatives to the decision classes. All this together seems to be an essential constraint for a researcher working on the prescriptive [2] techniques. But psychologically valid methods give the reliable output. To avoid the gaps between normative and descriptive only a psychologically correct operation of the information elicitation are to be used.

5.5.3 Consistency test

One of the inherent characteristics of human behavior is proneness to error. In transmitting and processing information, people make errors. They make less and sometimes considerably less errors when using the psychologically valid information elicitation procedures, but all the same they do make errors. The latter may be caused by the distraction of human attention, a person's fatigue, or other reasons.

Errors are observed both in practice and in psychological experiments. They differ essentially from the human errors in psychometric experiments which are known to follow the Gauss law and have the greater probability for the greater deviations from the true value. The human errors in the procedures of the information processing are of a different nature. For example, our studies of the multicriteria classification demonstrated that in the problems of small dimension (which are simple for man) gross errors leading to many contradictions are rare — 1 or 2 out of 50 cases [25]. These errors are obvious. Errors of the same kind are

met when comparing pairs of the estimates by the criteria, ranking criteria, etc. Stated differently, man can once and again commit essential errors. Therefore, the information elicited from man must be validated and not used uncontrollably.

In other words, an individual can make unavoidable errors from time to time. Hence, information obtained from a person must be subject to verification, rather than to be used uncontrollably.

How to check the information for consistency?

The efficient methods are so-called closed procedures [20, 30] under which the earlier collected information is subject to an indirect rather than direct test. The questioning procedure is built so that the questions are duplicated, but the duplication is exercised implicitly, through other questions logically associated with the former.

The following closed procedure was first suggested for the method ZAPROS [20]. Let there be Q criteria with ordinal scales and a small (2-5) number of estimates. It is required to order the estimates of all criteria, that is, to arrange them on the joint ordinal scale. To this purpose, it was suggested to perform pair-wise comparisons of the criteria scales.

All 0.5Q(Q-1) pairs of criteria were pair-wise compared, which enabled a rather reliable validation of the DMs' information. We note that as the number of criteria (hence, the complexity of the problem) increases, the potential amount of the redundant information generated by this comparison increases as well. A closed procedure of this type has been employed to advantage in the ZAPROS method.

Note that pinpointing a logical inconsistency should not lead, in general, to the automatic exclusion of an error but to the creation of the premises for a logical analysis.

The decision methods must incorporate means for checking the DM's information for consistency. No matter what method is used to elicit information from DMs, one must be aware of the possibility of occasional errors and of the stages of the DM training. In this connection, the procedures for checking the elicited information for consistency are required, as well as the methods for detecting and eliminating contradictions in the DM's information.

The need for consistency checks is not eliminated by the psychologically correct methods of the information elicitation from the DM's. This checking is extremely important because it improves the efficiency of the training and compels the DMs to recognize their errors and work out a reasonable compromise.

5.5.4 Learning procedures

As was noted above, learning is a part of human behavior. It is one of the inherent properties of human behavior, and the trial-and-error approach is the most characteristic human feature. Learning involves the study of a multicriteria problem and gradual working out of the DM's policy (decision rule).

One can hardly expect that the needed information is just stored in human minds until elicited [52]. Despite the fact that such expectations were not made explicitly, they were implied. Indeed, in many decision methods people are required to give immediately all parameters of the decision rules. It can hardly be expected that at the initial stages of the decision making an individual can define sensibly

and consistently the decision rule. It can be assumed that an experienced DM (especially, that who dealt previously with such a problem) has some elements of the policy such as a (possibly incomplete) list of criteria, comparative importance of some criteria and estimates, etc., but usually all this is specified in the course of the decision making where all tradeoffs are defined.

To allow the human ability of learning to manifest itself, the decision method must comprise special procedures for a gradual, rather than instantaneous, working out of the DM's policy. These procedures must allow the individuals to err and correct themselves, to work out partial compromises, and go on to the next ones. This process must allow the individuals to challenge their own decisions and return to the beginning.

5.5.5 Possibility to receive explanations

From a behavioral point of view, one of the requirements for any method is explainability of its results. The DM making a responsible decision would like to know why alternative A is superior to B and why both are superior to C. This requirement is quite legitimate. The stages of the information elicitation from the DM (measurements) and presentation of the final results are separated by the information's transformation. Understandably, the DMs want to make sure that the assessments of alternatives are based, without any distortion, precisely on their own preferences. To meet this requirement, the decision method must be `transparent,' that is, allow one to find the one-to-one correspondence between the DM's information and the final evaluations of the alternatives.

The DMs must have an opportunity to check whether there is a correspondence between the resulting estimates of the alternatives, on the one hand, and their own preferences, on the other hand. This check allows the DMs to make sure that it is precisely their preferences that uniquely define the results of using this method. Consequently, the DM must get explanations from the method in an understandable language. Only after that DM can receive the feeling of power in the hands and would like to use the results of the analysis.

5.5.6 New decision methods adapted to human behavior

The requirements formulated above create the possibility to develop new decision methods adapted to known features of human behavior.

First, the statement of a multicriteria decision problem in an organization gives the chance to make a step forward from the usual satisfactory behavior. It was demonstrated in many practical cases of the multicriteria decision methods application: the choice of a pipeline route [34], the location of an airport [16].

Second, the utilization of the ways of measurements adapted to the human information processing system gives the possibility to justify the decision methods from the psychological point of view. Using the natural language strongly increases the chances of a successful practical implementation.

Third, new methods of the decision rules' construction reduce the load on the human short-term memory and give the chances to reduce significantly the number of human errors, contradictions, biases.

Fourth, the special procedures for checking the information and

eliminating the contradictions give the decision methods the new quality of the reliable tools.

Fifth, the process of a gradual development of a decision rule gives the DM the time for learning, for careful development of a compromise between the criteria.

Finally, the possibility to get an explanation increases the chances for a successful implementation of the multicriteria decision analysis.

5.6 PRACTICAL IMPORTANCE OF BEHAVIOR ISSUES

How important are the requirements given above for the practice? How big is the influence of the incorrect measurements and human errors on the possibility to get a practically valid output of a decision method? Are behavior issues important only for theoretical reasons or do they define the practical value of the decision analysis?

The partial answers to these questions are provided by the results of the comparison of three decision aiding methods implemented as decision support systems. One of them was the method of the Verbal Decision Analysis -ZAPROS, satisfying the requirements given above [21, 29].

5.6.1 The decision problem

The experimental study was done to compare three methods of the decision making [28]. The subjects were college students nearing graduation, which were in a job search process, facing opportunities similar to those given in the study.

Let us suppose that a college graduate has several offers (after interviews) and he (or she) is to make a decision. These jobs are very similar in quality (that is, every variant is acceptable, but of course, one variant is better upon one aspect and the other - on the other). So, the student has to present this task as a multicriteria problem and try to solve it with the help of an appropriate multicriteria method.

Four criteria are used as the focus of the study: salary, job location, job position (type of work involved), and prospects (career development and promotion opportunities). The following alternatives were used:

FIRM	SALARY	JOB LOCATION	POSITION	PROSPECTS
a1	$30 000	Very attractive	Good enough	Moderate
a2	$35 000	Unattractive	Almost ideal	Moderate
a3	$40 000	Adequate	Good enough	Almost none
a4	$35 000	Adequate	Not appropriate	Good
a5	$40 000	Unattractive	Good enough	Moderate

It is easy to note that in this case there are three possible estimates upon the scale of each criterion. The greater the salary the more attractive an alternative would be to a rational subject. Thus, we have four criteria with three possible values each and the values of each criterion are rank-ordered from the most to the least preferable one.

It is evident, that there are no dominated alternatives. Therefore, the comparison of these alternatives required some value function, which would take into account the advantages and disadvantages of each alternative upon each criterion.

5.6.2 Two decision support systems based on numerical measurements

Two decision support systems based on the Multiattribute Utility Theory ("MAUT") [16, 17] were used for the solution of the problem given above. These systems are LOGICAL DECISION [44] and DECAID [36]. The third DSS was one based on the Verbal Decision Analysis (see below).

Both decision support systems LOGICAL DECISION and DECAID were used to solve this task. Both systems implement ideas from multiattribute utility theory, providing possibilities for the construction of an additive utility function for the case of the risky decisions, and an additive value function for a decision making under certainty. In our study, we used only additive value functions.

The value function obtained from both systems would therefore have the linear additive form of the weighted sum of the criteria estimates. The coefficients of importance for the criteria (the weights of importance) are used.

Both systems are easy to use, have a flexible dialogue and graphical tools to elicit the decision maker's preferences.

The main difference in the systems (besides interface) is the way of the determination of the numerical values for the evaluations upon separate criteria. In DECAID a pure graphical (direct) estimation is used (a point on the line of the size 1). In LOGICAL DECISION there is a possibility to use a special function for the criterion values. To determine the parameters of this function it is enough to mark the "middle" value for the criterion (sure thing for a lottery with 50% possibility for the best and the worst estimates).

The criteria weights are also defined in a different manner in these two systems. In LOGICAL DECISION criteria weights are defined on the basis of the trade-offs in a rather traditional way [17]. In DECAID weights are elicited directly (in a graphical way - point on a line), though the system provides also the possibility to make trade-offs, but after that the result is presented as points on the lines.

Taking into account the commonness of the approach implemented in both systems and also the similarity of the information, received from the DM in the process of the task solution, one could suppose that the attempt to solve the above described task with the help of these systems must lead to very close results.

5.6.3 Decision support system ZAPROS

The third DSS is one from the family of Verbal Decision Analysis [29]. Only verbal measurements are used on all stages of this method. ZAPROS uses ranking rather than rating information, but the additive overall value rule is correct if there is an additive value function. In ZAPROS the additive rule does not provide the summation of the verbal estimates, but rather the means of obtaining a pair-wise compensation between the components of the two alternatives.

For the preference elicitation from the subjects the following procedure was used.

Subjects were asked to compare several specially formed alternatives by pairs. For each pair two alternatives differed on evaluation according to two criteria only (one evaluation was best for each alternative) and had equal evaluations (best or worst) on other criteria.

For the task presented above it was necessary for the subjects to compare the pairs of the alternatives different on each pair of the four criteria. The example of a typical question is:

"What do you prefer: the firm giving salary $ 40 000 with an adequate location or the firm giving salary $35 000 with a very attractive location? Please, take into account that on the criteria "Position" and "Prospects" both firms are good".

Comparing these alternatives, subjects were to choose one of the following responses:

1. alternative 1 is more preferable than alternative 2;
2. alternative 2 is more preferable than alternative 1;
3. alternatives 1 and 2 are equally preferable.

The implementation of a such simple system for the comparison of the pairs of the alternatives gives us a possibility for a simple check of the received comparisons on the basis of transitivity.

The method provides a verification of the received comparisons for the transitivity and allows to change some of the responses on the request of the user to eliminate the intransitivity. It also guarantees that the comparison of each pair of the alternatives from this set is supported by at least two responses of the user.

Let us note that such a way of the preference elicitation is psychologically valid (see above). The received information allows one to build joint ordinal scale combining all evaluations on the separate criteria scales. The joint ordinal scale provides the possibility for the construction of a partial ranking for every given set of the alternatives.

Thus, this rank-ordering may be used for the comparison of the initial five alternatives because in our task the additive value function is supposed to be the right one and the criteria were formed to be preferentially independent. This algorithm does not guarantee the comparison of all alternatives because for some pairs of the alternatives ZAPROS gives only the incomparability relation.

5.6.4 The comparison of three decision support systems

Each subject from the group used all three DSS for the solution of the problem presented above. The difference in the outputs of the methods consisted in following: some pairs of the alternatives have not been compared with the ZAPROS method. Simple method of the preferences elicitation used by ZAPROS gave no possibility (in general case) to compare all given alternatives. ZAPROS gave only a partial ranking of the alternatives.

In contrast to it, two other methods give the complete ranking for the given alternatives. Also, LOGICAL DECISION and DECAID gave numerical values of the utility for all alternatives.

The results of the experiment were analyzed in a different form: the ranking of the given alternatives, the ranking of the specially formed alternatives used in ZAPROS, the ranking of the criteria weights and so on.

First of all, it was found that the correlation between the outputs of LOGICAL DECISION and DECAID was very low. The ANOVA test demonstrated that for the group of subjects the outputs of LOGICAL DECISION and DECAID

have not been statistically significant in the measurements of the criteria weights and the ranking of the alternatives.

The following results were very interesting: the outputs of the pairs LOGICAL DECISION-ZAPROS and DECAID-ZAPROS were correlated and were significantly correlated. It means that only for the alternatives compared by ZAPROS the relations were essentially the same.

It is possible to give the following explanation of the results.

The alternatives that could be ordered by ZAPROS are in the relations closed to the ordinal dominance. Such relations are more stable. Moreover, they were constructed in a very reliable way: verbal measurements psychologically correct way of preference elicitation, a possibility to check information and eliminate contradictions.

Two complete orders constructed by LOGICAL DECISION and DECAID were based on the numerical measurements and the weighted sum of the alternatives estimations by criteria. The difference in the utility (even small) defined the final order of the alternatives. The errors (even small) made by people while performing numerical (primary and secondary) measurements resulted in quite different orders of the alternatives.

5.7 CONCLUSIONS

The fact that a DM can not perform some cognitive operations in a reliable way is very important. It is the starting point for the development of a new decision theory where behavioral issues play at least the same role as the mathematical ones.

This new theory is needed not only for the decision analysis. The economists still believe in the myth of the existence of a *rational person* who behaves like a robot in the mathematical models of economics. But the facts demonstrate that consumers' behavior is far from rational [10]. The same is true for investor's decisions.

Behavioral issues are very important for political science [7]. We witness many errors made by leading political figures in different countries. Some such errors are difficult to explain. But a lot of them could be explained by the behavior factors. To describe how politicians make decisions, to help them choose the better strategies one must take into account knowledge about human behavior.

The reason of gaps between prescriptive and descriptive is basically the lack of the joint work between psychologists and sociologists engaged in the behavior research and mathematicians developing normative methods. The members of different research communities have different goals that generally do not coincide. Only joint multidisciplinary work could change the situation.

The understanding of the importance of research directed to the elimination of the gap between normative and descriptive is constantly increasing.

It is a good sign, the sign of the formation of a new image for the decision theory. New multidisciplinary research will give the important theoretical and practical results.

REFERENCES

[1] Bettman, I., P. Kakkar: Effects of information presentation format on consumer information acquisition strategies. Journal of Consumer Research, 3(1977).

[2] Brown, R., A. Vary: "Towards a research agenda for prescriptive decision science: the normative tempered by the descriptive." In: O. Huber, J. Mumpower, J. van der Pligt, P. Koele (Eds.) *Current Themes in Psychological Decision Research*, North-Holland Amsterdam 1992.

[3] Carnap, R.: *Philisophical Foundation Of Physics.*, Basic Books Inc. Publishers London 1969.

[4] Currim, I. S., R. K. Sarin: Prospect versus utility, Management Science, 35, 1(1989)22-41.

[5] Dolbear, F. I., L. B. Lave: Inconsistent behavior in lottery choice experiments, Behavior Science, 12,1(1967).

[6] Dror, Y.: *Public Policy Making Reexamined*, Transaction Publishers, Oxford. 1989.

[7] Dror, Y.: Basic Consepts in Advanced Policy Sciences., In: S. Nagel (Ed.) *Encyclopedia of Policy Studies,* Marcel Dekker, New York 1994.

[8] Erev, I., B. Cohen: Verbal versus numerical probabilities: efficiency, biases, and the preference paradox, Organizational Behavior and Human Decision Processes, 45(1990)1-18.

[9] Ericsson, K.A., W. Kintsch: Long-term working memory, Psychological Review, 102(1995)211-245.

[10] Foxall, G.R.: *Consumer Choice.* Kent, Maidstone 1983.

[11] Huber, B., O. Huber: Development of the concept of comparative subjective probability, Journal of Experimental Child Psychology, 44 (1987)304-316.

[12] Humphreys, P. C.: "Application of multiattribute utility theory", In: H. Jungermann, G. De Zeeuw (eds.), *Decision making and change in human affairs,* Reidel, Dordrech 1977.

[13] Gonzalez-Vallejo, C., T. Wallsten: The effects of communication mode on preference reversal and decision quality, Journal of Experimental Psychology: learning, memory and cognition, 18 (1992)855-864.

[14] Kahneman, D., P. Slovic, A. Tversky (eds.): *Judgement Under Uncertainty: Heuristics And Biases.*, Cambridge University Press, Cambridge 1982.

[15] Kahneman, D., A. Tversky: Prospect theory: an analysis of decisions under risk., Econometrica, 47(1979)263-291.

[16] Keeney, R. L., H. Raiffa: *Decisions With Multiple Objectives: Preferences And Value Tradeoffs*, Wiley, New York 1976.

[17] Keeney, R. L.: *Siting energy facilities*: Academic Press, New York 1980.

[18] Korhonen, P., J. Wallenius, "Behavior Issues in MCDM: Neglected Research Questions", In: J. Climaco (eds.) *Multicriteria Analysis*, Springer, Berlin 1997, pp. 412-422.

[19] Korhonen, P., O. Larichev, H. Moshkovich, A. Mechitov, J. Wallenius: Choice behavior in a computer-aided multiattribute decision task, Journal of Multi-criteria Decision Analysis, 6(1997)233-246.

[20] Larichev, O. I., Yu. A. Zuev, L. S. Gnedenko: "Method ZAPROS (CLosed Procedures near Reference Situations) for the analysis of variants of complex decisions", In: S. V. Emelyanov (eds.), *Multicriteria Choice for the Solution of Ill-structured Problems*, Moscow: VNIISI Proceedings, N 5 (in Russian), 1978, pp. 83-97.

[21] Larichev, O. I.: *A method for evaluating R&D proposals in large research organizations*, Collaborative paper CP-82-75, Laxenburg, Austria: IIASA 1982.

[22] Larichev, O. I.: *Objective models and Subjective Decisions*: Nauka (in Russian), Moscow 1987.

[23] Larichev, O. I., O. A. Polyakov, A. D. Nikiforov: Multicriteria linear programming problems (Analytical Survey), Journal of Economical Psychology, 8(1987)387-407.

[24] Larichev, O. I., E. M. Furems, "Multicriteria Packing Problems.", In: Y. Sawaragi, K. Inoue, H. Nakayama (eds.), *Toward Interactive and Intelligent Decision Suppport Systems*, Springer, Berlin, 1987, pp. 218-226.

[25] Larichev, O. I., H. M. Moshkovich, S. B. Rebrik: Systematic research into human behavior in multiattribute object classification problems, Acta Psychologica, 68(1988)171-182.

[26] Larichev, O. I.: Cognitive validity in design of decision-aiding techniques, Journal of Multi-criteria Decision Analysis, 1, 3(1992)127-138.

[27] Larichev, O. I., R. Brown, L. Andreeva, and N. Flanders: "Categorical Decision Analysis for Environmental Management: a Siberian Gas Distribution Case.", In: Caverni J. P., Bar-Hillel M., Barron F. N. and Jungermann H. (eds.), *Contributions to Decision Research*. North-Holland, Amsterdam, 1995, pp. 255-279.

[28] Larichev, O. I., D. L. Olson, H. M. Moshkovich, A. I. Mechitov: Numerical vs. cardinal measurements in multiattribute decision making: how exact is exact enough? Organizational Behavior and Human Decision Processes, 64, 1 (1995)9-21.

[29] Larichev, O. I., H. M. Moshkovich: ZAPROS-LM-A method and system for rank-ordering of multiattribute alternatives.., European Journal of Operations Research, 82 (1995)503-521.

[30] Larichev, O. I., H. M. Moshkovich: *Verbal Decision Analysis For Unstructured Problems.*, Kluwer Academic Publishers, Boston 1997.

[31] Lindblom, C. H.: The science of mudding through, Public Administration Review, 19(1959)155-169.

[32] Montgomery, H., O. Svenson: "A think-aloud study of dominance structuring in decision processes", In: H. Montgomery, O. Svenson (eds.), *Process and Structure on Human Decision Making*, J. Wiley and Sons, Chichester 1989, pp. 135-150.

[33] Nikiforov, A. D., S. B. Rebrik, L. P. Sheptalova: "Experimental study of stability of preferences in some decision making tasks", In: S.V.Emeljanov, O.I.Larichev (eds.) *Procedures for evaluation of multicriteria objects..*, VNIISI Proceedings, N 9 (in Russian), 1984, pp. 69-80.

[34] Oseredko, Y. S., O. I. Larichev, and A. I. Mechitov: "Main Gas Pipeline Route Selection Problem.", In: H. Kunreuther and E. Ley (*Eds.*) *Risk Analysis Controversy: An Institutional Perspective*, Springer, Berlin 1982, pp. 91-101.

[35] Payne, J. W., J. R. Bettman, E. Coupey, E. J. Johnson: "A constructive process view of decision making :multiple strategies in judgment and choice", In: O. Huber, J. Mumpower, J van der Pligt, P. Koele (*Eds.*), *Current Themes in Psycological Decision Research,* North Holland, Amsterdam 1993, pp. 107-142.

[36] Pitz, G.F.: *DECAID Computer Program.* Carbondale, IL, Univ. Of Southern Illinois 1987.

[37] Raiffa, H.: *Decision analysis. Readings*, Addison –Wesley, Menlo Park, California 1968.

[38] Russo, J. E., L. D. Rosen: An eye fixation analysis of multiattribute choice, Memory and Cognition., 3(1975)267-276.

[39] Savage, L.J. *The foundation of statistics.* Wiley, New York 1954.

[40] Simon, H., A. Newell: Heuristic problem solving: the next advance in operations research., Operations Research, 6(1958).

[41] Simon, H. A.: *The New Science of Management Decision*, Harper and Row Publ., New York 1960.

[42] Simon, H.A.: "Computational Theories Of Cognition.", In: W.O'Donohue, R.F. Kitchener (eds.), *The Philosophy Of Psychology.*, England, Sage Publications, London 1996.

[43] Slovic, P., A. Tversky: Who accepts Savage's axiom? Behavioral Science, 19(1974)368-373.

[44] Smith, G. R., F. Speiser: *Logical Decision: Multi-Measure Decision Analysis Software.* Golden, CO: PDQ Printing, 1991.

[45] Solso, R. L.: *Cognitive Psychology*, Allyn and Bacon, Inc., Boston 1988.

[46] Steuer, R. E.: *Multiple Criteria Optimization: Theory, Computation, and Application*, J. Wiley and Sons, New York 1986.

[47] Tversky, A., Intransitivity of preferences, Psychological Review, 76 (1969)31-48.

[48] Von Neumann, J., O. Morgenstern: *Theory Of Games And Economic Behavior.* Princeton University Press, Priceton 1947.

[49] Wagner, H. M.: *Principles of Operations Research*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey 1969.

[50] Wallsten, T., D. Budescu, R. Zwick: Comparing the calibration and coherence numerical and verbal probability judgments, Management Science, 39(1993)176-190.

[51] von Winterfeldt, D., G.W. Fischer: "Multiattribute utility theory: Models and assessment procedures", In: D. Wendt, C. Vlek (eds.) *Utility, probability and human decision making,*, Reidel, Dordrecht 1975.

[52] von Winterfeldt, D., W. Edwards: *Decision Analysis and Behavioral Research,* Cambridge University Press, Cambridge 1986.

[53] Wu, G.: Editing and Prospect Theory: Ordinal Independence and Outcome Cancellation, *Working Paper of Harvard Business School*, 1993.

6 META DECISION PROBLEMS IN MULTIPLE CRITERIA DECISION MAKING

Thomas Hanne

Contents

T. Gal et al. (eds.), *Multicriteria Decision Making*

Abstract: The problem of selecting a method for solving an MCDM problem is dicussed. This problem called meta decision problem can be formulated and solved in different ways. The most common approach is to define it as a problem of choosing one method from a finite set of methods which are evaluated according to several criteria. This leads to a formalization of the meta decision problem as an MCDM problem itself. Scalar evaluations of methods can help to avoid the meta decision problem becoming too complex. The problem of assessing the parameter(s) of an MCDM problem is similar to the meta decision problem above and can be interpreted as the problem to design an MCDM method. This leads to a formalization of the meta decision problem as a scalar parameter optimization problem. Information for solving meta decision problems can be submitted by a decision maker or is given as a data file originating, e.g., from prior decision making processes. Thus a meta decision problem can be solved in an interactive framework or through machine learning.

6.1 INTRODUCTION

The rapid growth of multiple criteria decision making (MCDM) since the first international conference in 1972 was accompanied by the evolution of different schools for solving complex decision problems with several objectives, goals, or criteria. Many methods based on different methodological assumptions have been proposed during this time. Chapters 8 – 12 of this volume survey the main schools or "philosophies" of MCDM.

Before MCDM became an established field of research, a decision maker (DM) may have felt quite helpless being confronted with an MCDM problem because of a lack of methodological support. Today, he/she may have similar feelings as a result of there beeing too many competing approaches for solving his/her problem. These approaches are usually based on different assumptions about the decision maker and the decision problem, require different information for their application, and, of course, usually lead to different results. The DM might consider such an indefiniteness of the results as most disturbing.

Although there have been long-standing debates concerning the "right" or "best" approach to MCDM which, especially, led to more or less relevant criticism on various methods (see below for details), there is no generally accepted foundation of a special school of MCDM or a specific method. On the other hand, there is an implicit universality claim of MCDM methods which means that these methods are usually proposed independently of the specific problem situation, without restrictions of their domain of application (other than technical assumptions like the formal problem definition). It is then assumed that the method is able to handle any MCDM problem as long as it fulfils the formal requirements (e.g. finite set of alternatives, or: polyhedral set of alternatives) of such a problem.

Let us mention just some examples of methodological criticism on MCDM approaches given in the literature: A valuable critical survey about several schools of MCDM is done by Stewart [60]. One of the most popular methods for finite MCDM problems, the analytical hierarchy process, has provoked an intensive methodological discussion: See, e.g., [4, 6, 12, 13, 30, 44, 60]. The

most central aspect in this discussion is based on the phenomenon of rank-reversal which means that the inclusion of "irrelevant" additional alternatives may change the ranking order of the other alternatives.

Goal programming, likely the oldest school of MCDM approaches developed in the 50s as an extension of linear programming [7], has also been critically discussed by several authors: See, e.g., [18, 52, 53, 72]. Some of the criticism concerns, for instance, special variants of goal programming like the lexicographic version.

Utility theory for MCDM is criticized because of its "strict" assumptions which are usually not empirically valid (see, for example, [23, 63]) although this would not necessarily imply that they are not useful for a normative context as in MCDM.

Outranking methods which are especially popular in France and Belgium are criticized, for instance, by Alley [1] and Stewart [60]. Alley ([1], see also [17]) points out the possibility of obtaining dominated solutions with the ELECTRE approach and considers the ranking process to be "a mystery to the DM." Stewart [60] supposes some outranking procedures to be "difficult to verify empirically as models of human preferences." Interactive approaches which seem to be the most popular way of dealing with MCDM problems cannot be criticized easily without considering the special concept they rely on (e.g., utility theory or a reference point approach). Larichev [34] (see also chapter 5) in his work took into account the human limitations in interactivity.

6.2 WHAT IS A META DECISION PROBLEM IN MCDM?

There is neither a strong reason to reject a particular school of MCDM nor a convincing argument to give general preference to one of the many methods. Therefore, it seems to be reasonable to accept a certain methodological pluralism within MCDM.[1] The problem of method choice should instead be solved in the context of a given decision situation consisting of decision problem, decision maker, and various restrictions on the decision process, e.g., time limits or information costs. This problem of method choice is called a *meta decision problem.*

The question of method choice can be formulated as a problem to select a method to solve an MCDM problem from a finite set of methods. Such a problem is treated, for instance, in the following studies: Buchanan and Daellenbach [5], Cohon and Marks [8], Gershon and Duckstein [16], Hobbs [25], Khairullah and Zionts [31], Ozernoy [45, 46], and Wallenius [65]. In Sections 6.3 and 6.4 we will discuss the applicability of such an approach. Let us note that many papers dealing with the problem of plurality of MCDM methods do not formulate it as a method selection problem formally. Instead, judgements on different MCDM methods are compiled or comparisons of several MCDM

[1]For a general discussion on methodological pluralism within OR see Mingers [39] and Mingers and Brocklesby [40, 41].

approaches are provided (see, e.g., [2, 11, 19, 21, 38, 42, 58, 59, 62]). For instance, characteristic features of methods are pointed out, advantages and disadvantages of specific approaches are summarized, different assumptions of application are discussed, or characteristic features of solutions obtained with different methods are analyzed.

Another possibility to formulate and analyze the meta decision problem is to consider the problem of method design, i.e., to choose an MCDM method from an infinite set of methods characterised by a parameter. From this more general point of view it is possible to analyze the question of method choice and the question of parameter assessment in an integrated approach. This approach is discussed in section 6.5.

6.3 CRITERIA FOR METHOD EVALUATION AND SELECTION

A straightforward approach for solving a meta decision problem is to formulate this question as a problem of selecting an MCDM method from a given set of methods. A first step in structuring this problem is to look for a criterion or several criteria for evaluating a method. The following subsections summarize possible criteria as discussed in the literature.

6.3.1 Suitability for the problem type

The chosen method must be able to handle problems of the given type. Often, MCDM is subdivided into MODM (multiple objective decision making) and MADM (multiple attribute decision making). MODM considers problems with a continuous set of alternatives whereas MADM treats problems with a finite, usually small set of alternatives. Many methods can handle only one of these problem types. For instance, interactive approaches usually assume that an MODM problem is given whereas the AHP or outranking approaches are constructed for MADM. Several MODM methods assume that the set of alternatives is convex and that the objective functions are concave (in case of maximization). An important special case of MODM is MOLP (multiple objective linear programming) with a polyhedral set of alternatives and linear objective functions.

Other problem types are such with discrete, integer or binary, stochastic or fuzzy decision variables. Extensions to dynamic problems, control problems, problems with uncertainty, incomplete information or qualitative criteria etc. have been proposed. For such special problem types usually only a few methods have been discussed in the literature such that the method selection problem is almost "solved" when a formally suitable method has been found at all. However, for the problem types discussed most often, like MADM, MODM, and MOLP, several formally suitable methods have been proposed such that considerations based on the problem type do not solve the method selection problem.

However, let us mention that the differentiation between problem definition and method selection is an idealization. In practice the modelling of a problem

as, for instance, a fuzzy problem is frequently the consequence of the availability of an appropriate method for handling such problems. If an appropriate method would not be available then the problem might be defined as a conventional (non-fuzzy) MCDM problem. In a similar manner convex MODM problems are occasionally linearized such that a theoretically and methodologically well-established MOLP problem can be analyzed.

6.3.2 Criteria based on solution concepts

If the DM has clear preferences concerning possible solution concepts which an MCDM method should apply, then the considered set of methods can be reduced to the feasible ones which fit his/her preferences. Some examples for such preferences:

If the DM prefers the alternatives to be compared with a reference point, then, of course, a reference point approach would be appropriate; if the DM wishes to work with trade-offs between criteria (see, e.g., Cohon and Marks [8]), then a utility-based method might be better than, for instance, a reference point approach or an outranking method.

Desired solution concepts may be aspiration levels, pairwise comparisons (of criteria and/or alternatives), interactivity or non-interactivity, the calculation of a complete ranking of alternatives, a classification of alternatives into groups, or the selection of a most preferred alternative, etc.

Hwang and Yoon ([29], p. 211) propose a solution to the method selection problem for MADM based on using decision trees. Based on a hierarchy of questions the DM is asked whether he/she wishes certain method features.[2] Then a proposal for method choice is made. Let us note that such an approach usually does not solve the method selection problem uniquely when several methods belong to the same leaf of the decision tree. The features in question often correspond to method features such that a good methodological knowledge is required for the application of this approach. In practice, a DM might have difficulties to decide a priori whether, for instance, he/she prefers the articulation of pairwise preference to the articulation of pairwise proximity. However, if such a DM with well-structured preferences on methodological features exists it would not make much difference to ask him/her directly for the most preferred method.

If the choice of an MCDM method is considered there may be given a requirement for the method to fulfil certain assumptions of "*rationality*". The most important of such assumptions is that an MCDM method yields a non-dominated alternative as the most preferred solution.[3] Such assumptions about the results calculated by an MCDM method usually help to reduce the set of feasible methods but they do not solve the method selection problem uniquely. In a more general context these and other properties can be used to measure

[2]For instance, the DM is asked whether optimizing or satisficing is desired, or whether pairwise preference is preferred to pairwise proximity.

[3]Although there may be reasons why this is not always necessary, see, e.g., [73].

the "*validity*" of a method. The validity of a method is often defined by the (non-) respecting of transformations allowed by the criteria scale (see, e.g., [51]). Validity can also be interpreted in a descriptive context of application as correspondence between results of the method and observable phenomena (of a physical or socio-economic reality).

This leads to another interesting criterion for evaluating a method: the possibility to *predict preferences.* The idea behind this concept is that preferences on alternatives are assumed to be given a priori. An MCDM method is then applied to the problem. The differences between the method results interpreted as predicted preferences and the actual preferences are measured (see also section 6.5.2. for a formal elaboration of a similar approach). Frequently, this criterion oriented towards a descriptive application of MCDM is used as a single quality measure of methods (see, e.g., [9, 47, 55]). Khairullah and Zionts [31, 32] propose the prediction of the best alternative, and the prediction of a (complete) ranking order as criteria for method evaluation.

Another scalar and quantitatively measurable criterion taking into consideration the solution concept is based on approximate solutions. For problems where it is not possible or too difficult to find an exact (efficient) solution[4] the distance to an optimal solution can be measured to judge the methods. Daniels [10] uses such an approach where the distance is measured on the basis of an additive utility function.

Usually, data about the decision problem, e.g., the objective functions or criteria evaluations of the alternatives, are not given precisely. The DM may also feel unsure about the preference information he/she has to assess. Therefore, the sensitivity of the solution calculated by the method should be quite robust.

6.3.3 Implementation-oriented criteria

In this group of criteria we consider characteristics of the method and its computer implementation. A typical software-oriented feature to judge an optimization method is its *computing time.* This can either be measured using concepts from complexity theory (see, e.g., [26, 71]) or performing practical tests where the actual computing time for typical problems is measured for different problem sizes. However, it should be noted that MADM problems as usually treated in literature consist of quite few criteria and alternatives, and the discussion of MODM avoids hard-to-solve nonlinear problems or degenerate linear problems which can cause computing difficulties (see Kruse [33], esp. p. 36). Considering the power of modern computers as well, the criterion of the computing time does not seem to be relevant in most cases to the discussion of MCDM method selection.

Instead the *ease of use* of a method is an important criterion for evaluating its application especially when non-experts have to deal with the method (see

[4]This is the case, for instance, in scheduling problems where "multicriteria heuristics" are used; see [10].

Stewart [60]). This is particularly relevant for methods which are at least in some respect interactive and which are developed for non-expert use. The design of the man-machine dialogue is essential to the success of an interactive approach. Today, it is generally important for software to be user-friendly. User-friendliness can be achieved, for instance, by the usage of a graphical user interface (as Windows), by easy data input (e.g. via mouse), and by providing help functions.

Interactivity causes opportunity costs because it takes time and efforts for the DM to formulate the information he/she provides for the computer algorithm and to understand the information provided by the program. This aspect is measured frequently by the time needed for applying a method or the number of iterations (see, e.g., [35]). The time required for interactions is often significantly longer than the time for the computation itself (see Buchanan and Daellenbach [5], p. 355). The costs of method application also depend on the difficulty for the DM to understand the method and what kind of information he/she has to provide. The logic of a method should therefore be transparent and the information required from the DM should be free from ambiguity (see Stewart [60]).

It takes time for the DM to improve his/her understanding, or the quality of the results suffers or he/she gets frustrated when the transparency of a method is not good enough. The same holds when the information output by the method is not explicit enough, when it is unclear, incomplete, difficult to grasp, or simply when it does not seem to be useful. For instance, the DM may prefer an indirect assessment of weights by pairwise comparisons of criteria (as, for instance, in the AHP) rather than a direct assessment because he/she feels quite unsure about their meaning or specific values. Or it is the other way round if the DM feels quite sure about weight values and wishes to avoid the information overhead by pairwise comparisons.

Another relevant aspect concerns the costs of implementing a method into an (institutionalized) decision process. This comprises the costs for self-implementing the computer software or for buying a ready-made or custom-made software, the cost for training the DM(s) or decision analyst(s) (DA) who are working with the software. The knowledge and the abilities of the DM have to be considered here as well as other aspects which are related to the decision environment when information costs for integrating a method into an institutionalized decision process have to be analyzed.

6.3.4 Criteria depending on the specific decision situation

In some of the studies on the meta decision problem it is, however, not clear whether a method evaluation takes place in the context of a given decision situation or not. A solution obtained without considering a specific decision situation would be obligatory for any DM and decision problem as long as the evaluations of the criteria in sections 6.3.1.-6.3.3. do not change.

However, we believe that the *appropriateness* of an MCDM method *for a specific decision situation* is a key issue for its selection. A decision situation

consists of (1.) the decision problem, (2.) the decision maker, and (3.) various aspects of the decision process such as time limits, knowledge of the DM about different methods for MCDM, institutional restrictions, etc. Concerning the given decision problem it has to be judged whether the method can *actually* handle the problem. Problems might arise when the number of criteria, the number of alternatives, the number of model variables, or the number of restrictions is too large. Considering typical problem sizes[5] this does not seem to be a problem which really matters.

The DM's or DA's understanding of MCDM methods is a very important issue. As Balesta and Tsoukiàs ([3], p. 422) note it is often possible that a method selection problem does not become evident because only one method is known and *therefore* preferred by the DM. Such ignorance can be thoroughly rational considering the costs of implementing methods into an institutionalized decision process, esp. the costs for learning to understand several methods.

If the DM/DA understand more or less different methods then his/her preferences for a method may depend on the specific problem. For instance, the DM's preferences for interactions change from problem to problem depending on the available time and the importance of the problem. The availability and confidence in a prior preference structure may influence the method choice, or the plausibility of preliminary results of a method can influence the DM's preference for that approach. The DM's confidence in a method can change from problem to problem and is subject to a learning process.

Restrictions resulting from the specific decision process can be given by the available interaction time. Another resource restriction is possibly based on the costs of the decision process. The usability of a method for group decisions and other aspects of embedding it in a multi-person decision environment may be an essential feature for its choice. Institutional restrictions may require a method which adequately supports group decision making or provides a logic intelligible to all. Political decision processes, for instance, may be affected by hidden preferences or strategic considerations in preference articulation such that formal MCDM methods are avoided or solutions inconsistent with method results are chosen (see, e.g., [6]). It is also possible that a simple and non-iteractive method, e.g. additive weighting, is used to justify a decision process by a seemingly transparent decision rule. All these considerations complicate the finding of a solution to the meta decision problem in multiple criteria decision making.

6.4 SCALAR VS. MULTICRITERIA META DECISION PROBLEMS

6.4.1 Scalar evaluations of MCDM methods

A simple and easy-to-handle approach for solving the meta decision problem is to reduce it to a problem with a scalar evaluation criterion. If it is possible to define such a single criterion for comparing the methods the one with the

[5] As discussed in published application studies of MCDM. See Vincke ([64], p. 1-4) for a discussion of the set of alternatives.

best evaluation would be chosen. The formerly mentioned approaches based on a prediction of preferences usually use such a single criterion defined as an average error of prediction. The error definition is based on the aggregated differences between the method results interpreted as predicted preferences and the actual preferences.[6] Let us note that the prediction of preferences is a questionable criterion because it relies on a basically descriptive approach which is not necessarily a good foundation of prescriptive MCDM [9].

Another single criterion well-known from OR methods in areas where an exact solution cannot be calculated[7] is based on approximate solutions. Following this concept Daniels [10] evaluates "multicriteria heuristics" for discrete MCDM problems like scheduling. This analysis is based on an additive utility function which serves to measure the distance of an approximate solution to an efficient solution.

In section 6.5.2. we discuss the application of a scalar criterion for solving the method design problem using machine learning. A general problem related to scalar definitions of the meta decision problem is, however, the acquisition of the necessary information, e.g. data of "true" preferences or data on exact solutions. These difficulties are discussed in more details in section 6.6. However, the approach of modelling the meta decision problem as an MADM problem discussed in the next section, leads to similar problems of information assessment.

6.4.2 Method selection as an MADM problem

Considering the alleged ubiquity of multiple criteria in decision problems (see, e.g., Zeleny [73], p. 1-11), it is not surprising that in many studies related to the MCDM method selection problem it is proposed that this problem should be formulated as a multicriteria problem itself. Different criteria which express features or qualities of methods relevant for the selection of one of them have to be considered. This approach is quite common and has already been applied in one of the earliest papers on method selection in MCDM, a study by Cohon and Marks [8]. Other papers on MCDM method evaluation or selection in which the problem is defined explicitly as an MCDM problem are, e.g., [16, 25, 45, 46, 49].

The multicriteria method selection problem can be formally defined as

$$opt f(a) \tag{6.1}$$

with

$$a \in A = \{a_1, ..., a_l\}, f : A \to R^q, \tag{6.2}$$

where A is a discrete set of MCDM methods, $f_k, k \in \{1, ..., q\}$, are criterion evaluation functions and *opt* means that the single criteria are either to be

[6] See section 6.5.2. for a formal notation of such an approach.

[7] E. g. for the travelling salesman problem and other problems in combinatorics whose exact solution would be too time consuming.

maximized or minimized. The problem can also be represented by a decision matrix $Z \in R^{l \times q}$ with $z_{ij} = f_j(a_i), j \in \{1, ..., q\}, i \in \{1, ..., l\}$.

Example 6.1:

Gershon and Duckstein [16] present one of the most elaborate approaches to the problem of selecting an MCDM method from a finite set of methods. First, 28 criteria for method evaluation are proposed. The criteria which are derived from characteristics of the specific decision situation (problem, method, DM) can be grouped into the following categories:

- mandatory binary criteria
- non-mandatory binary criteria
- technique-dependent criteria
- application-dependent criteria

With the help of the binary criteria MCDM methods can be classified. The mandatory binary criteria, e.g. the property to choose among discrete and/or continuous alternatives, also serve to distinguish between "feasible" and "infeasible" methods for a given MCDM problem. Among the non-mandatory binary criteria we find characteristics of specific MCDM methods, e.g. the property to use a reference point, or to apply pairwise comparisons of alternatives. The technique-dependent criteria comprehend the following:

- computer time,
- interaction time,
- implementation time,
- level of DM's sophistication required
- consistency of results with those of other techniques
- robustness to parameter value changes
- applicability to group decision making.

The application-depended criteria are, for instance, the number of criteria and alternatives, or the DM's understanding of MCDM methods.

In an illustrative application Gershon and Duckstein propose first to reduce the set of these 28 criteria to the relevant ones which are 16 criteria in their example. For these criteria preference information, e.g. weights, has to be assessed. Then a set of alternative MCDM methods is determined. 13 methods are considered in their approach: $A = \{$sequential optimization, weighting, ϵ-constraint, compromise programming, cooperative game theory, multiattribute utility theory, surrogate worth trade-off, ELECTRE, Q-analysis, dynamic compromise programming, PROTRADE, STEP$\}$. Infeasible alternatives are eliminated by consideration of mandatory criteria. For the remaining alternatives, $a_i \in A' = \{$compromise programming, cooperative game theory, multiattribute utility theory, ELECTRE, Q-analysis, STEP$\}$, the quantitative evaluations, $z_{ij}, j \in \{1, ..., 16\}$, have to be assessed which might cause problems because of the non-quantitative and subjective character of some of the criteria. Finally, the problem of method selection is formally defined as an MADM problem. ■

If the problem of MCDM method selection is formalized as an MCDM problem then it is straightforward to solve it with an MCDM method itself. The

question then, of course, is: Which of the many MCDM methods should be used for solving the meta decision problem? Ford, Keeney and Kirkwood [14], for instance, propose to use multiple attribute utility theory for method evaluation. Gershon and Duckstein [16] solve the MCDM problem of MCDM method evaluation with compromise programming. Depending on the preference assumptions different solutions of the method choice problem are obtained in their study. Ozernoy ([45]; see also [46]) proposes firstly to eliminate infeasible methods (e.g. because of criterion scales) and, secondly, to evaluate the remaining approaches with "screening". A similar approach is proposed by Cohon and Marks [8] who first exclude dominated alternatives (e.g. ELECTRE I in their example) and then evaluate the remaining alternatives verbally.

The problem of choosing a method for method selection, a meta meta decision problem, demonstrates the recursive character of the question concerning method choice. Analogously, meta meta meta ... decision problems can be analyzed. Since the solution of a meta meta decision problem influences the solution of a meta decision problem just as the solution of a meta decision problem influences the solution of the original decision problem such questions have practical relevance.[8] Let us note that, usually, even in several of the papers where the meta decision problem is structured as a formal MCDM problem it is treated quite informally, i.e. without applying an MCDM method. The goal is usually to find an overall evaluation of methods which is, for instance, called "operational usefulness" by Stewart [60].

6.5 METHOD DESIGN

A more general problem compared to the method selection problem is that of "optimal" design or construction of an MCDM method. If it is questionable whether an arbitrary method proposed in the literature (or given as software) is appropriate for a particular MCDM problem, then there is also no guarantee that a satisfying method can be found within a small finite and fixed set of MCDM approaches compiled from the literature. One might instead ask for a custom-made method designed or constructed for the particular MCDM situation.

6.5.1 Parameter assessment in MCDM methods

In most cases an MCDM method does not compute a solution automatically like a "black box" but requires additional information to solve a problem. Such information is used to determine one or several parameters of the method. The choice of a parameter serves to adapt an MCDM method to the decision problem being analyzed. Typical examples of parameters utilized within MCDM methods are weights, achievement levels, threshold values, trade-offs, parameters for value scaling, ideal and anti-ideal points, parameters of utility

[8] Although a practical solution, i.e. with involvement of the DM, does not seem to be realistic.

functions, of preference functions, and other parameters. The problem of parameter choice can be interpreted as an attempt to design an MCDM method appropriate to the decision problem, i.e. as a meta decision problem.

This leads to a problem mostly not discussed in the context of method selection: How should the parameter(s) for a given method be determined? Usually, it is assumed that parameters are assessed directly or indirectly by the DM such that they reflect his/her preferences. Parameters are determined a priori or can be changed during the process of method application (interactive approach).

Some methods or decision support systems (DSSs) for MCDM offer additional help for assessing parameters. The analytic hierarchy process (AHP) [56] uses the idea of structuring the objectives hierarchically and comparing objectives of the same level and group pairwise. For q criteria of a given group of criteria $(q^2 - q)/2$ relative importance values have to be assessed by the DM. They serve to calculate weights for the criteria of the group such that they can be aggregated additively to a single criterion assuming ratio scale measurement.[9] Since the $(q^2 - q)/2$ relative importance values include redundancy (assuming consistent preferences) deviations of the DM's inputs to the ideal values corresponding to the obtained weights can be used to calculate a consistency index. A high degree of inconsistency can lead the DM to revise the relative importance values articulated formerly or to perform more fundamental changes of the problem definition such that improved results can be obtained by repeating the weight calculation.

Several DSSs for MCDM offer a sensitivity analysis [15] for the parameter values of the implemented MCDM methods. These are, for instance, the SENSATO software developed by Rios Insua [50] and the PROMCALC software [37] implementing the PROMETHEE outranking method. SENSATO offers a sensitivity analysis of weights within a utility-based MCDM approach. The PROMCALC software also offers a sensitivity analysis of weights but within the PROMETHEE method where they have a completely different meaning. The other parameters to be assessed for PROMETHEE are not considered during the sensitivity analysis. Some available software packages for MCDM do not include any forms of sensitivity analysis at all.

Within some interactive methods it is possible to revise the given parameter information. These methods usually work along the following scheme (see chapter 10): Depending on the DM's input a preliminary solution is calculated. If this result does not sufficiently correspond to the DM's preferences then he/she is encouraged and supported to change the parameter values such that a new preliminary solution can be calculated. This process continues until the DM is satisfied with the results. Some other interactive methods work by reducing the set of possible parameters (e.g. weights) and thus by reducing the set of reachable solutions.

[9] Interval scale measurement is assumed in some modifications of the AHP.

6.5.2 A parameter optimization model

Let us define a parameterized method with parameter p as follows: A method M_p with input in calculates an output $out = f_{M_p}(in) = f_M(in, p)$, e.g. a scalar value for each alternative, where M is an ordinary, nonparameterized algorithm which calculates its result with two inputs, the problem and a parameter p. Below we distinguish two cases with different assumptions on the information processing of an MCDM method and thus with different interpretations of inputs in, outputs out, and reference outputs out^*.

The meta decision problem can be formulated as a scalar parameter optimization problem:

$$(6.3) \qquad argmin_{p \in P} \sum_{(in, out^*) \in \Omega} dist(f_M(in, p), out^*)$$

where P is an appropriate set of parameters[10], Ω is a finite training set consisting of pairs of example inputs in and corresponding reference outputs out^*. $dist$ is, for instance, based on quadratic differences or on a metrics defined on problem data types which fulfils $dist(x, x) = 0, dist(x, y) > 0$ for $x \neq y$ and $dist(x, y) + dist(y, z) \geq dist(x, z)$ for all x, y, z.

If the MCDM method is used for evaluating each alternative separately then Ω consists of example alternatives a represented by their objective vectors, $in = f(a)$, together with their (scalar or vector-valued) reference evaluations $out^* = y^* \in R^u$.[11] The MCDM method calculates for each in (with $(in, out^*) \in \Omega$) an evaluation $f_M(in, p) = y \in R^u$ assumed to be scaled as the reference evaluations. We can use the square of the difference of the reference and calculated alternative evaluations such that the problem (6.3) leads to a minimization of the quadratic error:

$$(6.4) \qquad dist(y, y^*) = \sum_{i=1}^{u} (y_i^* - y_i)^2.$$

Similarly a distance function $dist$ can easily be defined by an l_p metrics as

$$(6.5) \qquad dist(y, y^*) = \sqrt[p]{\sum_{i=1}^{u} (y_i^* - y_i)^p}.$$

The assumption that each alternative is evaluated separately can be dropped. In several MCDM methods, e.g., outranking methods, the AHP, or some reference point approaches, the treatment of an alternative may depend on properties of other alternatives. If the alternatives cannot be evaluated without the context of other alternatives then the axiom of independence of irrelevant

[10] Usually, the set of parameters is an infinite subset of a vector space.

[11] If scalar alternative evaluations are considered then $u = 1$.

alternatives does not hold. In our parameter optimization model the set of alternatives has to be evaluated as a whole.

The input of the method then consists of an MADM problem, i.e. *in* is an MCDM problem (A, f) with a finite, feasible set of alternatives A and a vector-valued objective function $f : A \to R^q$. The output can also be interpreted as an MADM problem with one[12] or several objective values for each alternative. We assume that the output is represented by a decision matrix $out = Z' \in R^{l \times u}$. The quadratic difference to a reference output $out^* = Z^* \in R^{l \times u}$ can be obtained as

$$dist(out, out^*) = \sum_{h=1}^{l} \sum_{k=1}^{u} (z^*_{hk} - z'_{hk})^2. \tag{6.6}$$

The parameter $p \in P$ in (6.3) can be interpreted as an index which determines a specific method, or as a parameter of a given method or as both, the type and parameter(s) of a method. Let us consider an example: Ramesh, Zionts and Karwan [49] who analyze the meta decision problem for interactive branch and bound methods in integer multiobjective linear programming introduce an additional parameter for combining two parameterized families of methods. In this way the meta decision problem is expressed as a method design problem or a problem of choosing parameters from a continuous set. Two objectives for evaluating the methods are proposed: the minimization of questions for the decision maker and the minimization of solution time. Using these criteria, different efficient alternatives (methods) are obtained as solutions of the meta decision problem.

Another, more flexible framework allowing MCDM method design via parameter optimization is based on neural networks.[13] A function calculated by such a neural network is controlled by parameters which are assessed by solving an optimization problem similar to (6.3), (6.4) through machine learning (see section 6.6.4.). Parameters of neural networks are weights of the connections between single neurons and neuron-specific parameters like threshold values. Also the structure or size of a neural network is occasionally expressed by parameters. One of the most popular approaches are feedforward neural networks which allow approximation of arbitrary continuous functions.[14] Therefore, a neural network can learn to work like different MCDM methods depending on the training set Ω.

Example 6.2:

Wang [69] analyzes the capabilities of neural networks to approximate utility functions for solving MCDM problems. Apriori it is usually not clear what kind

[12] This is the standard case of scalar evaluations.

[13] See, e.g., [24, 43] for an introduction to neural networks.

[14] A proof of this property based on a theorem by Kolmogorov is given by Hornik, Stinchcombe and White [27].

of utility aggregation model (additive, multiplicative, etc.) and what criterion-specific utility functions fit well the DM's preferences. The main idea with using a neural network is that a general structure is applied which can work like different utility models (or other MCDM methods). The specific design of the neural network is obtained by a learning process. For this a utility evalution of a small set of alternatives (training set) is required.

Wang performs tests of the neural network model (feedforward network) using learning data obtained from a simulated DM who can perform utility evaluation of alternatives. The MCDM problem under consideration possesses 3 objective functions. Four different utility functions are analyzed as models of the DM's preferences:

1. additive utility function:

$$U(f_1(a), f_2(a), f_3(a)) = 0.5f_1(a) + 0.3f_2(a) + 0.2f_3(a) + \epsilon$$

2. multilinear utility function:

$$U(f_1(a), f_2(a), f_3(a)) = 0.4f_1(a) + 0.3f_2(a) + 0.2f_3(a) + 0.04f_1(a)f_2(a)$$
$$+0.03f_1(a)f_3(a) + 0.02f_2(a)f_3(a) + 0.01f_1(a)f_2(a)f_3(a) + \epsilon$$

3. quadratic utility function:

$$U(f_1(a), f_2(a), f_3(a)) = f_1(a) + 0.6f_2(a) + 0.4f_3(a) - 0.5f_1^2(a)$$
$$- 0.3f_2^2(a) - 0.2f_3^2(a) + \epsilon$$

4. polynomial-exponential utility function:

$$U(f_1(a), f_2(a), f_3(a)) = 0.5f_2(a)e^{3f_1(a)-3} + 0.3f_3^2(a)e^{2f_2(a)-2}$$
$$+ 0.2f_1^3(a)e^{f_3(a)-1} + \epsilon$$

ϵ is a (0,0.0001)-normally distributed random variable.

For the learning process a set Ω consisting of randomly generated alternatives together with their evaluations according to the respective utility function is used. The resultant 10 test alternatives, together with their evaluations, are used for measuring the quality of the neural network after the training process. The learning process is based on a minimization of a training error of the neurral network according to (6.3) and (6.4). The parameters p to be optimized are weights of connections between neurons and neuron-specific parameters, i.e. threshold values.

The results of Wang's experiments are the following: Neural networks are capable to approximate all analyzed utility functions quite well. Training and test errors are between $1.14 \cdot 10^{-4}$ and $4.77 \cdot 10^{-4}$. The best alternative is always predicted correctly. The obtained neural networks are small and also the training time is very good (except for the quadratic utility model). ■

The results of such test applications of neural networks were in general quite positive. Of course, the function of other MCDM approaches than utility-based

concepts can be realized by a neural network computing architecture as well. The meta decision problem of method design can be approached by neural networks which are a well-established tool in areas like pattern recognition or data prediction.

6.6 INFORMATION HANDLING

If the meta decision problem for a given MCDM problem has to be solved, additional information is required. It is possible to distinguish two different sources of such information: It is either "hidden" in the knowledge and preferences of the DM, or it is given in some explicit and articulated form, i.e. as a data file. Let us discuss these two sources of information in some detail.

Several of the criteria as discussed in section 6.3. are not easy to measure. Some criteria like computation time can be objectively measured. Other criteria, e.g. the DM's understanding of a method, the ease of use or the usefulness of information provided by a method can be assessed only subjectively by interviewing the DM. Some criteria can be measured cardinally while others have an ordinal character.

6.6.1 Internal information

Frequently, the DM has some idea how a good solution of a decision problem might look like or how an appropriate method should work. This intuitive knowledge has to be articulated explicitly, e.g., in the form of preferences between MCDM methods. Such implicitly available information can then be used for the choice of a method and/or its parameters. In a simple approach the decision maker chooses a method directly, usually from a given set of methods. Such an approach to method selection corresponds to an a priori articulation of method parameters (see, e.g., [28], p. 8).

A more elaborate way of obtaining information relevant for decision making is to use the possibilities of a man-machine-dialogue.[15] The interactive approach allows to give the DM additional help to express his/her knowledge and preferences about the decision problem. So far no interactive DSS has been developped which specifically supports the acquisition and utilization of information for solving the meta decision problem in MCDM. If some software implementing several MCDM methods is available then it is easy for a DM to follow a trial and error approach to get to a method selection. The DM can "play" with different MCDM methods and find out which one might be most suitable to solve his/her MCDM problem in practice. With such a heuristic approach it is not necessary to assess criteria evaluations of the considered MCDM methods and to formalize the meta decision problem as in (6.1)-(6.2) or (6.3)-(6.6).

[15] See chapter 10 of this volume.

6.6.2 Laboratory and computer experiments

There are several studies based on experiments to get information on criteria, i.e. the values for z_{ij}, for evaluating MCDM methods. Some of them are based on computer experiments simulating a DM's behavior when dealing with an MCDM method, e.g. assuming a given utility function [31, 32, 35]. Other experiments compare and evaluate different methods used by "dummy" decision makers [5, 65]. While computer experiments can consider objectively measurable information only, the experiments with "dummy" DMs can also be used for assessing subjective information (based on DM's attitude) although it is in general not clear whether such information is intersubjectively valid.

Usually, in both types of experiments the approach discussed in 6.4.2. is applied: A discrete set of methods is evaluated according to several criteria. In these studies the focus lies on interactive methods and the human attitude towards the man-machine dialogue. For instance, Wallenius [65] compares the GDF method, STEM, and an unstructured interactive approach according to six criteria. Four of them are ordinal and subjectively assessed by the method users. Buchanan and Daellenbach [5] perform laboratory experiments with four interactive multiobjective decision making techniques. Six criteria are used for evaluating the methods. Four of them are subjectively assessed by the DM while the other two are objectively measurable. Occasionally, the DM states seemingly irrational judgements on the methods. For instance, one of the "dummy" DMs remarks on the Zionts-Wallenius method: "I understand the logic; it's easy to use, but I don't like it." ([5], p. 357). This shows that the confidence in or preference for a method cannot easily be expressed even by usage of complex subjectively measurable criteria like the ease of use or the transparency of a method.

6.6.3 External information

An alternative to the information acquisition from a DM during the decision making process (for solving the meta decision problem) is to use information already articulated in an explicit way, i.e. given as a data file. Let us discuss some examples where such information might originate from. External data can be assessed objectively or subjectively.

Three cases of such data can be identified:

a) The data are based on the DM's preferences. The intention is to find a solution which corresponds best to the preferences of the DM. A finite, relatively small subset of the original set of alternatives or coming from a similar decision problem is evaluated by the decision maker. For instance, a utility function for this subset of alternatives is assessed. The DSS receives the data of the alternatives (criteria evaluations) together with their "scores" as input for learning. With this information the DSS learns the functional relationship (e.g. the utility function) between the alternative criteria and scores and, thus, the preferences of the DM. The complete set of alternatives can then be evaluated with the learned function. Such an approach has been discussed, e.g.,

by Wang [66, 67, 68, 69] and Wang and Malakooti [70] who use a feedforward neural network approach for learning.

b) Data are taken from previous decisions for similar problems. If we consider a repetitive multicriteria decision problem which has formerly been solved without applying a formal MCDM approach then we can use records of these historical decisions for the learning process. For instance, from the historical decision problems we have obtained the data of the alternative sets together with evaluations of the single alternatives, or, at least, information about the most preferred (selected) alternative. In this case learning is oriented towards an adaptation to real, historical decision behavior, no matter whether it was "rational" or not.

c) Objective data are defined by ex post results of previous decisions. In this case we consider a type of decision problem for which alternatives are characterized by multiple criteria a priori (prior to a decision). Expost (some time after a decision) objective knowledge about the quality of the alternatives or, at least, about the realized alternative(s) becomes evident. The functional relationship between an alternative's criteria evaluations and a measure of quality constituted later on is not known (and possibly stochastically influenced). In such a decision framework the application of MCDM has a mainly descriptive character (see [57], p. 328).

Examples of such decision problems are, for instance, the choice of a stock market investment where a quality measure is given by future stock prices. At present measurable criteria, e.g. the price earning ratio or chart-related ratios, are in practice used to give a vague indication of the possibly highly stochastically influenced future success of an investment. Other examples (see [57]) include, for instance, prediction of political election results, the prediction of currency exchange rates or oil prices, the forecasting of economic indicators or sports results.

In these cases of application, data of alternatives together with an ex post measure can be used for learning method parameters. An adaptation to simulate systems behavior is intended. The purpose of an MCDM method is to approximate a "hidden" relationship between measurable multiple criteria and a nonmeasurable ("not-at-this-time measurable") scalar criterion. MCDM can here be interpreted as a search for a descriptive theory or as a tool for system identification. In a similar manner, MCDM approaches sometimes fulfil prediction tasks or serve for simulation applications (see, e.g., [48, 61]).

No matter where the external information comes from, it constitutes a kind of "reference solution" for a method which can be exploited by machine learning as discussed in section 6.6.4.

6.6.4 Machine learning

In considering information costs (esp. costs of interactivity) in multiple criteria decision making, it seems worthwhile to examine the usage of information already available and given as a data file. Such data can be accumulated, for instance, in the context of repetitive decision situations. The utilization of infor-

mation for adapting a method (or its parameters respectively) or DSS software to a given decision problem can be called machine learning. Machine learning[16] can be used to automate the decision process at least partially. Thus information usage can be regarded as an alternative concept to interactive knowledge acquisition and utilization discussed above. Learning can be interpreted as a way of approaching the meta decision problem in MCDM.

The learning process (or training process) then takes place prior to the application process of the method. The meta decision problem is solved prior to the original decision problem. In [22] an object-oriented DSS framework is proposed which treats MCDM problems, methods, and meta (=learning) methods as coupled objects and solves the meta and the original decision problem by a recursive solution process.

Learning methods can generally be used for assessing the parameters(s) of an MCDM method. Machine learning minimizes the "difference" between a solution calculated by a method and a "best" or reference solution. This "difference" can be measured by a distance function, e.g. an l_p metrics, defined on appropriate method outputs based on sample inputs (of the problem data type) and matching reference solutions. From a mathematical point of view this usually leads to a scalar nonlinear optimization problem as expressed in (6.1). Problems of this type can be solved with different algorithms, e.g. gradient approaches or evolutionary algorithms.

The approach of utilizing learning for multiple criteria decision making is used in applications of neural networks to MCDM. A discussion of the significance of neural networks (feedforward layered neural networks) to MCDM can, for instance, be found in papers by Wang [66, 67, 68, 69], Wang and Malakooti [70], Malakooti and Zhou [36], and Hanne [20, 22]. Neural networks can be considered as parameterized MCDM methods. The parameters are threshold values of the single neurons and weights of connections between them. Also the structure of the network can be expressed as a parameter. In the literature different learning algorithms for parameter assessment have been proposed. The most popular learning method for feedforward neural networks is the backpropagation algorithm (see [54]) which is based on a gradient approach. Other approaches are, for instance, based on evolutionary algorithms [22].

6.7 CONCLUSIONS

The meta decision problem in MCDM can be formulated as a method selection problem or as a method design problem. The method evaluation can be performed with respect to a single criterion or several criteria. The meta decision problem can be solved using information articulated by the DM or information given as a data file. The main alternatives of approaching the meta decision problem are to use an interactive framework or to apply machine learning.

[16] Aspects of machine learning in expert systems are not discussed here. For details concerning the application of artificial intelligence in MCDM see chapter 15 of this volume.

Let us now reconsider some main questions related to these different approaches for dealing with meta decision problems. In the literature, the meta decision problem is mostly treated (1.) as an MCDM problem and (2.) as a method selection (instead of a method design) problem. Both of these two assumptions involve some difficulties:

1. The approach of formulating the meta decision problem as a method selection problem is mainly based on a small set of existing methods and neglects effective possibilities of adapting a method to a specific problem situation (e.g. by parameter assessment).

2. The formalization of the meta decision problem as an MCDM problem demands the articulation of selection criteria which causes some problems in detail: Some of the evaluation criteria proposed in the literature have a classificatory character which allows, for instance, to distinguish suitable and unsuitable methods but not to find a unique (meta) solution. Further, it is possible to distinguish objective and subjective criteria. Some of the objective criteria like computation time are usually not relevant considering today's computer power or costs and typical problem sizes.

The subjective criteria which are more or less based on the use of interactive methods apply subjective judgements of the DM. These criteria are often questionable, e.g. in terms of measurability (scales) or operationality. The problem of prejudices and knowledge of the method user also has to be considered.

Some criteria are partially based on "descriptivity" or the adaptation to possibly irrational "preferences". Similarly those approaches which are based on the prediction of preferences have to be regarded as questionable for prescriptive or normative approaches as in MCDM. Behind several subjective criteria there is the idea that the DM can evaluate MCDM methods. This means that it is assumed that a DM has some intuitive knowledge as to what a "good" method should look like or, more specifically, what the results of a "good" method should be. Interactivity is a way of utilizing this implicit knowledge.

Last but not least, let us mention that occasionally prejudices on methods (predecisions for a method) may have a significant influence in the studies on the meta decision problem. Such influences can, for instance, be discovered in the considered set of methods, in the proposed selection criteria, and in the way of structuring the evaluation process, e.g. the applied (meta) MCDM method.

The approach of formulating the meta decision problem as a scalar parameter optimization problem (6.3) requires that training information consisting of example problems (alternatives) and corresponding reference solutions be given. This seems to be the main difficulty for applying this concept since often such information is simply not available or cannot be considered as a "good" standard for future decision making.

References

[1] Alley, W. M.: Comment on "Multiobjective river basin planning with qualitative criteria" by M. Gershon, L. Duckstein, and R. McAniff. Water Resources Research 19:1(1983)293-294

[2] Anderson, E. E.: Choice models for the evaluation and selection of software packages. Journal of Management Information Systems 6:4(1990)123-138

[3] Balestra, G., A. Tsoukiàs: Multicriteria analysis represented by artificial intelligence techniques. Journal of the Operational Research Society 40:5(1990)419-430

[4] Belton, V., T. Gear: On a short-coming of Saaty's method of analytic hierarchies. OMEGA International Journal of Management Science 11:3(1983)228-230

[5] Buchanan, J. T., H. G. Daellenbach: A comparative evaluation of interactive solution methods for multiple objective decision models. European Journal of Operational Research 29(1987)353-359

[6] Carlsson, C., P. Walden: AHP in political group decisions: A study in the art of possibilities. Interfaces 25:4(1995)14-29

[7] Charnes, A., W. W. Cooper: *Management models and industrial applications of linear programming.* Volume I. Wiley, New York 1961

[8] Cohon, J. L., D. H. Marks: A review and evaluation of multiobjective programming techniques. Water Resources Research 11:2(1975)208-220

[9] Currim, I. S., R. K. Sarin: A comparative evaluation of multiattribute consumer preference models. Management Science 30:5(1984)543-561

[10] Daniels, R. L.: Analytical evaluation of multi-criteria heuristics. Management Science 38:4(1992)501-513

[11] Despontin, M., F. Lehert, M. Roubens: Multi-attribute decision making by consumers associations. European Journal of Operational Research 23(1986)194-201

[12] Dyer, J. S.: Remarks on the analytic hierarchy process. Management Science 36:3(1990)249-258

[13] Dyer, J. S.: A clarification of "remarks on the analytic hierarchy process". Management Science 36:3(1990)274-275

[14] Ford, C. K., R. L. Keeney, C. W. Kirkwood: Evaluating methodologies: a procedure and application to nuclear power plant siting methodologies. Management Science 25:1(1979)1-10

[15] Gal, T.: *Postoptimal analyses, parametric programming, and related topics. Degeneracy, multicriteria decision making, redundancy.* De Gruyter, Berlin 1994

[16] Gershon, M., L. Duckstein: "An algorithm for choosing of a multiobjective technique." In: P. Hansen (ed.): *Proceedings on the Fifth International Conference on Multiple Criteria Decision Making*, Mons, August 9-13, 1982. Springer, Berlin 1983, pp. 53-62

[17] Gershon, M., L. Duckstein: Reply. Water Resources Research 19:1(1983) 295-296

[18] Hannan, E. L.: An assessment of some criticisms of goal programming. Computers Operational Research 12:6(1985)525-541

[19] Hanne, T.: *Bonitätsorientierte Bankbilanzanalyse mit Hilfe von Multikriteria-Methoden, insbesondere PROMETHEE.* Diplomarbeit. Ruhr-Universität Bochum, Fakultät für Wirtschaftswissenschaft 1989

[20] Hanne, T.: Die Integration von Multikriteria-Verfahren insbesondere mittels neuronaler Netze. OR Spektrum 16(1994)277-283

[21] Hanne, T.: "An application of different MCDM methods to bank balance sheet analysis." In: U. Derigs, A. Bachem, A. Drexl (eds): *Operations Research Proceedings 1994.* Springer, Berlin 1995, pp. 506-511

[22] Hanne, T.: "Decision support for MCDM that is neural network-based and can learn." In: J. Climaco (ed): *Multicriteria analysis.* Proceedings of the XIth International Conference on MCDM, Coimbra, Aug. 1-6, 1994. Springer, Berlin 1997, pp. 401-410

[23] Hemming, T.: On the validity of multi-attribute utility models. Working Paper, 3rd International Conference on MCDM, Königswinter 1979

[24] Hertz, J., A. Krogh, R. Palmer: *Introduction to the theory of neural computation.* Morgan Kaufmann, San Mateo 1991

[25] Hobbs, B. F.: What can we learn from experiments in multiobjective decision analysis? IEEE Transactions on Systems, Man and Cybernetics 16:3(1986)384-394

[26] Hopcroft, J. E., J. D. Ullman: *Introducton to automata theory, languages and computation.* Addison-Wesley, Reading 1979

[27] Hornik, K., M. Stinchcombe, H. White: Multilayer feedforward networks are universal approximators. Neural Networks 2(1989)359-366

[28] Hwang, C.-L., A. S. Md. Masud: *Multiple Objective Decision Making - Methods and Applications.* Springer, Berlin 1979

[29] Hwang, C.-L., K. Yoon: *Multiple Attribute Decision Making - Methods and Applications.* Springer, Berlin 1981

[30] Islei, G., A. G. Lockett: Judgemental modelling based on geometric least square. European Journal of Operational Research 36(1988)27-35

[31] Khairullah, Z. Y., S. Zionts: "An experiment with some algorithms for multiple criteria decision making." In: G. Fandel, T. Gal (eds): *Multiple criteria decision making theory and application.* Proceedings Hagen/Königswinter. Lecture Notes 177, Springer, Berlin 1979, pp. 150-159

[32] Khairullah, Z. Y., S. Zionts: An approach for preference ranking of alternatives. European Journal of Operational Research 28(1987)329-342

[33] Kruse, H.-J.: *Degeneracy graphs and the neighbourhood problem.* Springer, Berlin 1986.

[34] Larichev, O. I.: "Problem of man-machine interaction in decision support systems." In: L. B. Methlie, R. H. Sprague (eds): *Knowledge representation for decision support systems.* Elsevier Scientific Publisher, Amsterdam 1985, pp. 27-39

[35] Lotfi, V., Y. S. Yoon, S. Zionts: Aspiration-based search algorithm (ABSALG) for multiple objective linear programming problems: Theory and comparative tests. Management Science 43:8(1997)1047-1059

[36] Malakooti, B., Y. Q. Zhou: Feedforward artificial neural networks for solving discrete multiple criteria decision making problems. Management Science 40:11(1994)1542-1561

[37] Mareschal, B.: Weight stability intervals in multicriteria decision aid. European Journal of Operational Research 33(1988)54-64

[38] Michalowski, W.: Evaluation of a multiple criteria interactive programming approach: An experiment. INFOR Information Systems and Operations Research 25(1987)165-173

[39] Mingers, J.: "Towards critical pluralism." In: J. Mingers, A. Gill (eds): *Multimethodology: Theory & practice of combining management sciences methodologies.* Wiley, Chichester 1997

[40] Mingers, J., J. Brocklesby: Multimethodology: Towards a framework for critical pluralism. Systemist 18:3(1996)101-131

[41] Mingers, J., J. Brocklesby: Multimethodology: Towards a framework for mixing methodologies. OMEGA International Journal of Management Science 25:5(1997)489-509

[42] Mote, J., D. L. Olson, M. A. Venkataramanan: A comparative multiobjective programming study. Mathematical Computer Modelling 10:10(1988)719-729

[43] Müller, B., J. Reinhardt: *Neural networks. An introduction.* Springer, Berlin 1990

[44] Murphy, C. K.: Limits on the analytic hierarchy process from its consistency index. European Journal of Operational Research 65(1993)138-139

[45] Ozernoy, V. M.: "A framework of choosing the most appropriate discrete alternative multiple criteria decision-making method in decision support systems and expert systems." In: Y. Sawaragi, et al. (eds): *Toward interactive and intelligent decision support systems*, Springer, Berlin 1986, Vol. 2, pp. 56-64

[46] Ozernoy, V. M.: Some issues in designing an expert system for multiple criteria decision making. Acta Psychologica 68(1988)237-253

[47] Park, C. W.: A conflict resolution choice model. Journal of Consumer Research 5:2(1978)124-137

[48] Peschel, M., F. Breitenecker: "Interactive structure design and simulation of nonlinear systems from a multiobjective viewpoint using the Lotka-Volterra approach." In: Grauer, M., A. P. Wierzbicki (eds): *Interactive decision analysis.* Proceedings, Sept. 20-23, 1983. Springer, Berlin 1984, pp. 63-76

[49] Ramesh, R., S. Zionts, M. H. Karwan: A class of practical interactive branch and bound algorithms for multicriteria integer programming. European Journal of Operational Research 26(1986)161-172

[50] Rios Insua, D.: *Sensitivity analysis in multiobjective decision making.* Springer, Berlin 1990

[51] Roberts, F. S.: *Measurement theory with applications to decisionmaking, utility, and the social sciences.* Addison-Wesley, Reading 1979

[52] Romero, C.: *Handbook of critical issues in goal programming.* Pergamon Press, Oxford 1991

[53] Rosenthal, R. E.: Goal programming - A critique. New Zealand Operational Research 11:1(1983)1-7

[54] Rumelhart, D. E., G. E. Hinton, R. J. Williams: "Learning internal representations by error propagation." In: D. E. Rumelhart, et al. (eds): *Parallel Distributed Processing. Explorations in the Microstructure of Cognition. Vol. 1: Foundations.* MIT Press, Cambridge, London 1986, pp. 318-364

[55] Russ, F. A.: Evaluation process models and the prediction of preference. Proc. of the Association for Consumer Research, 1971, pp. 256-261

[56] Saaty, T. L.: *The analytic hierarchy process. Planning, priority setting, ressource allocation.* McGraw-Hill, New York 1980

[57] Saaty, T. L.: That is not the analytic hierarchy process: What the AHP is and what it is not. Journal of Multi-Criteria Decision Analysis 6(1997)324-335

[58] Srinivasan, V., A. D. Shocker: Estimating the weights for multiple attributes in a composite criterion using pairwise judgements. Psychometrika 38:4(1973)473-493

[59] Stewart, T. J.: Experience with prototype multicriteria decision support systems for pelagic fish quota determination. Naval Research Logistics 35(1988)719-731

[60] Stewart, T. J.: A critical survey on the status of multiple criteria decision making. Theory and practice. OMEGA International Journal of Management Science 20:5/6(1992)569-586

[61] Teleb, R., F. Azadivar: A methodology for solving multi-objective simulation-optimization problems. European Journal of Operational Research 72(1994)135-145

[62] Terry, H.: Comparative evaluation of performance using multiple criteria. Management Science 9:3(1963)B431-B442

[63] Tsoukiàs, A., P. Vincke: A survey on non conventional preference modeling. Ricera Operativa 61(1992)5-49

[64] Vincke, P.: *Multicriteria decision-aid.* Wiley, Chichester 1992

[65] Wallenius, J.: Comparative evaluation of some interactive approaches to multicriterion optimization. Management Science 21:12(1975)1387-1396

[66] Wang, J.: A neural network approach to multiple-objective cutting parameter optimization based on fuzzy preference information. Computers and Industrial Engineering 25(1993)389-392

[67] Wang, J.: Multiple-objective optimisation of machining operations based on neural networks. International Journal of Advanced Manufacturing Technology 8(1993)235-243

[68] Wang, J.: A neural network approach to modeling fuzzy preference relations for multiple criteria decision making. Computers & Operations Research 21:9(1994)991-1000

[69] Wang, J.: Artificial neural networks versus natural neural networks. A connectionist paradigm for preference assessment. Decision Support Systems 11(1994)415-429

[70] Wang, J., B. Malakooti: A feedforward neural network for multiple criteria decision making. Computers & Operations Research 19:2(1992)151-167

[71] Wegener, I.: *Effiziente Algorithmen für grundlegende Funktionen.* B. G. Teubner, Stuttgart 1989

[72] Zeleny, M.: The pros and cons of goal programming. Computers & Operations Research 8:4(1981)357-359

[73] Zeleny, M.: *Multiple criteria decision making.* McGraw-Hill, New York 1982

7 SENSITIVITY ANALYSIS IN MCDM

Tetsuzo Tanino

CONTENTS

Abstract: Stability and sensitivity analysis is both theoretically and practically interesting and important in optimization and decision making. In this chapter we will explain several approaches, though limited, to stability and sensitivity analysis in MCDM.

T. Gal et al. (eds.), *Multicriteria Decision Making*

7.1 INTRODUCTION

In this chapter we deal with stability and sensitivity analysis in multiple criteria decision making. It aims to analyze qualitative and quantitative behavior of the optimal solution or the optimal value according to changes of parameter values included in the original optimization or decision making problem.

We consider a family of parametrized multiple criteria decision making problems

$$\begin{array}{ll} \text{minimize} & f(x,u) = (f_1(x,u), f_2(x,u), \ldots, f_p(x,u))^T \\ \text{subject to} & x \in X(u), \end{array} \tag{7.1}$$

where x is an n-dimensional decision variable, u is an m-dimensional parameter vector, $f_i (i = 1, 2, \ldots, p)$ is a real-valued objective function on $R^n \times R^m$, X is a set-valued map from R^m to R^n which specifies a feasible decision set depending on u, and ^{T} denotes a transposed vector or matrix. Let Y be a set-valued map from R^m to R^p defined by

$$Y(u) = \{y \in R^p : y = f(x,u) \text{ for some } x \in X(u)\} \tag{7.2}$$

for each $u \in R^m$. Y is regarded as the feasible set map in the objective space.

In order to define a solution of the above problem we consider a partial order in the objective space R^p induced by a pointed closed convex cone K with a nonempty interior in R^p, where K is said to be pointed if $K \cap (-K) = \{0\}$. Though we may consider more general ordering (domination structure), we confine our analysis within the cone-induced ordering for simplicity. In the following we use the following inequality notations for vectors $y, z \in R^p$:

$$\begin{array}{ll} y \leqq_K z & \text{if and only if} \quad z - y \in K \\ y \leq_K z & \text{if and only if} \quad z - y \in K \setminus \{0\} \\ y <_K z & \text{if and only if} \quad z - y \in \text{int}K. \end{array} \tag{7.3}$$

Generally $K \supset R^p_+$, which denotes the nonnegative orthant of R^p, and in case of $K = R^p_+$, we omit the suffix K for simplicity.

We will define three kinds of cone minimal points based on the above inequality relations.

Definition 7.1: Let A be a set in R^p. Then
1) $\hat{y} \in A$ is said to be a *K-minimal point* of A if there exists no $y \in A$ such that $y \leq_K \hat{y}$.
2) $\hat{y} \in A$ is said to be a *properly K-minimal point* of A if there exists a cone C such that $\hat{y}$ is a C-minimal point of A, where C is a convex cone with $C \neq R^p$ and $K \setminus \{0\} \subset \text{int}C$.
3) $\hat{y} \in A$ is said to be a *weakly K-minimal point* of A if there exists no $y \in A$ such that $y <_K \hat{y}$.

These points are also often called efficient, properly efficient and weakly efficient, respectively, with respect to K. The set of all K-minimal, properly K-minimal and weakly K-minimal points of A are denoted by $\text{Min}_K A$, $\text{PrMin}_K A$

and $\mathrm{WMin}_K A$, respectively. Clearly

$$\mathrm{PrMin}_K A \subset \mathrm{Min}_K A \subset \mathrm{WMin}_K A \tag{7.4}$$

and if A is a polyhedral convex set

$$\mathrm{PrMin}_K A = \mathrm{Min}_K A \tag{7.5}$$

(see, e.g. Sawaragi et al. [33] Theorem 3.1.7).

According to these solution concepts we can define the following three set-valued maps W, G and S from R^m to R^p by

$$W(u) = \mathrm{Min}_K Y(u) \tag{7.6}$$

$$G(u) = \mathrm{PrMin}_K Y(u) \tag{7.7}$$

and

$$S(u) = \mathrm{WMin}_K Y(u) \tag{7.8}$$

for each $u \in R^m$, respectively. These set-valued maps W, G and S are called the perturbation map, the proper perturbation map, and the weak perturbation map, respectively. They are extensions of the well-known perturbation function (marginal function) in usual scalar optimization.

Moreover we may consider the set-valued map from R^m to R^n defined by

$$V(u) = \{x \in X(u) : f(x,u) \in W(u)\} \tag{7.9}$$

by considering the set of all K-minimal solutions of the perturbed problem in the decision space. If the original problem is a multiobjective linear programming problem, we may focus on the extreme points (basic solutions) of the feasible polyhedral set. Hence we may deal with the set of all K-minimal basic solutions $V_E(u)$.

Even in multiple criteria decision making problems, if the preference of the decision maker (DM) is represented by a scalar utility (value, or evaluation) function and if the DM is interested only in the best alternative which maximizes this utility function, we may directly apply existing methods of stability and sensitivity analysis in ordinary scalar optimization (see e.g. Fiacco [8]) to this case. On the contrary, if we deal with the whole set of K-minimal points (efficient solutions), the aim of stability and sensitivity analysis becomes to investigate the qualitative and/or quantitative behavior of the set-valued maps defined above. This chapter is mainly devoted to the analysis in the latter more complicated case and several approaches will be introduced. Since the available space for this chapter is limited, some interesting results must be omitted. The author owes quite much to the very nice survey paper by Gal and Wolf [13] in completing this chapter as well as several referred papers.

The contents of this chapter are as follows. First in Section 7-2, sensitivity in multiobjective linear programming is dealt with. Section 7-3 provides fundamental concepts of continuity and differentiation of set-valued maps as mathematical preliminaries. Section 7-4 is devoted to stability analysis, i.e. investigation of continuity of the perturbation maps. We consider derivatives of the perturbation maps in Section 7-5. Sensitivity analysis using duality theory will be dealt with in Section 7-6. In Section 7-7 we concentrate on sensitivity analysis in the discrete multicriteria decision making. The last section provides some concluding remarks.

Because of page limitation, all the proofs of the theorems, propositions and lemmas are omitted in this chapter and the reader is requested to see the references.

7.2 SENSITIVITY ANALYSIS IN MULTIOBJECTIVE LINEAR PROGRAMMING

This second section will be devoted to sensitivity analysis in multiobjective linear programming. Sensitivity analysis in ordinary linear programming is quite popular and useful for postoptimal analysis (see, e.g., Gal [11]). Since we may concentrate on the basic solutions in the linear case, and the number of those solutions is finite, our main interest lies in knowing whether a nominal optimal basic solution remains optimal when parameters included in the problems change. Some results related to duality theory will be explained later in Section 6.

In this section we introduce the results given by Gal and Leberling [12,10]. Other interesting results can be found in Deshpande and Zionts [7], Hansen et al. [16], Antunes and Clímaco [1] etc. (see the survey by Dauer and Liu [6]). We consider the following nominal multiobjective linear programming problem (the value of the parameter u is fixed at a certain nominal value):

$$\begin{array}{ll} \text{maximize} & Cx \\ \text{subject to} & x \in X = \{x \in R^n : Ax \leqq b,\ x \geqq 0\} \end{array} \tag{7.10}$$

where $A \in R^{m \times n}$ (which denotes the set of all $m \times n$ matrices) and $C \in R^{p \times n}$. Since "maximization" is considered in usual linear programming we follow this tradition in this section and hence the ordering cone is fixed as the nonpositive orthant. A feasible solution x is called efficient if there is no feasible x' such that $Cx' \geq Cx$. The set of all efficient basic solutions (i.e. efficient extreme points) of the problem is denoted by V_E (or $V_E(u)$ when there exists a parameter u in the problem). Let $\hat{x}$ be an efficient extreme point of the above problem which is already determined. We assume in the following that no degeneration occurs and that X is compact. The basis associated with $\hat{x}$ is denoted by B and the set of the uniquely characterized corresponding basis indices is denoted by ρ. Let P be the set of all $j \in \bar{\rho}$ such that for every $j \in P$, there exists at least one $y_{kj} > 0$, where $\bar{\rho}$ is the complement of ρ and y_{kj} are elements of $B^{-1}A$. Denote

by y^i_{0j} the current reduced costs associated with the ith objective function, i.e.

$$y^i_{0j} = c^i_B B^{-1} a^j - c^i_j \text{ for all } j \in \bar{\rho},\ i = 1, \ldots, p \tag{7.11}$$

where c^i_B is the part corresponding to the basis of the ith row of C and a^j is the jth column of A.

Gal and Leberling [12] called their sensitivity investigation postefficient analysis or relaxation analysis and divided it into two parts: with respect to the objective functions and with respect to the right-hand side.

7.2.1 *Sensitivity analysis with respect to the objective function*

Now we assume that the objective function is perturbed, i.e. let

$$C(u) = C + \bar{C}U \tag{7.12}$$

where $\bar{C}$ is a constant $p \times n$ matrix and

$$U = \begin{pmatrix} u_1 & \cdots & 0 \\ \cdots & \cdots & \cdots \\ 0 & \cdots & u_p \end{pmatrix} \tag{7.13}$$

is a diagonal matrix in which the main elements define a parameter vector $u = (u_1, \ldots, u_p)^T \in R^p$. Denote by $V(u)$ the set of all efficient solutions to the problem

$$\begin{array}{ll} \text{maximize} & C(u)x \\ \text{subject to} & x \in X = \{x \in R^n : Ax \leqq b,\ x \geqq 0\}. \end{array} \tag{7.14}$$

The task is then "Determine a region $\Omega \subset R^p$ such that for all $u \in \Omega$ the set $V(u)$ remains equal to $V(0)$".

To determine Ω, i.e. to maintain B optimal, the dual feasibility condition must be satisfied, i.e.

$$-\sum_{i=1}^{p} y^i_{0j} \lambda_i \leqq 0 \text{ for all } j \in \bar{\rho} \tag{7.15}$$

with a weighting vector $\lambda \in R^p_+$ (the nonnegative orthant). Adding slack variables s_j, $j \in \bar{\rho}$ and normalizing λ by $\sum_{i=1}^{p} \lambda_i = 1$, we have the following E-test (efficiency test, Gal [10]):
for each $j \in P$ solve

$$\begin{array}{ll} \text{minimize} & s_j \\ \text{subject to} & \sum_{i=1}^{p} y^i_{0j}(u)\lambda_i + s_j = 0 \text{ for all } j \in \bar{\rho} \\ & \sum_{i=1}^{p} \lambda_i = 1 \\ & \lambda_i \geqq 0 \text{ for all } i = 1, \ldots, p \\ & s_j \geqq 0 \text{ for all } j \in \bar{\rho}. \end{array} \tag{7.16}$$

Solving the above linear programming problem for each $j \in P$ and considering some additional conditions (for details see Gal [10] or Gal and Leberling [12]) the critical region Ω can be obtained.

7.2.2 Sensitivity analysis with respect to the right-Hand side

Next consider the following multiobjective linear programming problem with the perturbation in the right-hand side of the inequality constraint:

$$\begin{array}{ll} \text{maximize} & Cx \\ \text{subject to} & x \in X(u) = \{x \in R^n : Ax \leqq b(u),\ x \geqq 0\} \end{array} \tag{7.17}$$

where $b(u) = b + Lu$, where L is a constant $m \times q$ matrix and $u \in R^q$ is a parameter vector. As before, the set of all efficient solutions to the above problem is denoted by $V(u)$.

In order to proceed analysis Gal and Leberling [12] introduced an undirected graph generated by multiobjective linear programming problem ((MOLP) for short).

Definition 7.2: An undirected graph $\mathcal{G} = (U, \Gamma)$ is said to be *generated* by (MOLP) if
1) $\rho \in U$ if and only if B is an efficient basis (i.e. a basis associated with an efficient basic solution),
2) between two nodes $\rho,\ \rho' \in U$ there exists an arc if and only if the corresponding efficient points $x, x' \in X$ are efficient neighboring extreme points and the edge

$$[x, x'] = \{x \in R^n : x = \alpha x + (1 - \alpha)x',\ 0 \leqq \alpha \leqq 1\}$$

is included in the set of all efficient points.
3) to each node $\rho \in U$ there is assigned at least one positive weighting vector λ.

Based on this definition, the task can be formulated as follows: "Determine a region $\Omega \in R^q$ such that for all $u \in \Omega$ the graph $\mathcal{G}(u) = (U(u), \Gamma(u))$ generated by the perturbed (MOLP) with u remains the same as the graph generated by the nominal (MOLP), i.e. $\mathcal{G}(u) = \mathcal{G}(0)$ for all $u \in \Omega$".

They proved the following theorem.

Theorem 7.1: $\mathcal{G}(u) = \mathcal{G}(0)$ if and only if

$$u \in \bigcap \{u \in R^q : -B^{-1}Lu < B^{-1}b\} \tag{7.18}$$

where the intersection is taken over the set of all extreme points of X.

7.3 CONTINUITY AND DIFFERENTIATION OF SET-VALUED MAPS

Throughout this section let F be a set-valued map (point-to-set map) from the m-dimensional Euclidean space R^m to the p-dimensional Euclidean space R^p, namely a map from R^m to the power set 2^{R^p}. Let

$$\text{dom } F = \{x \in R^m : F(x) \neq \emptyset\}. \tag{7.19}$$

7.3.1 *Continuity of set-valued maps*

In this chapter we use the most simple concept of continuity of set-valued maps which is defined in terms of the convergence of sequences (e.g. Aubin and Frankowska [4], Maeda [24] or Hogan [17], though the terms "closed" and "open" are used instead of "u.s.c." and "l.s.c.", respectively, in the last). Recently Rockafellar and Wets [32] are using the terms *outer* and *inner* semicontinuity instead.

Definition 7.3: 1) F is said to be *lower semicontinuous (l.s.c.)* at a point $\bar{x} \in R^m$ if $\{x^k\} \subset R^m$, $x^k \to \bar{x}$, $\bar{y} \in F(\bar{x})$ imply the existence of an integer $\bar{k}$ and a sequence $\{y^k\} \subset R^p$ such that $y^k \in F(x^k)$ for $k \geq \bar{K}$ and $y^k \to \bar{y}$.
2) F is said to be *upper semicontinuous (u.s.c.)* at a point $\bar{x} \in R^m$ if $\{x^k\} \subset R^m$, $x^k \to \bar{x}$, $y^k \in F(x^k)$, $y^k \to \bar{y}$ imply that $\bar{y} \in F(\bar{x})$.
3) F is said to be *continuous* at a point $\bar{x}$ if it is both l.s.c. and u.s.c. at $\bar{x}$.
4) We say that F is l.s.c., u.s.c. or continuous on $X \subset R^m$ if it has the respective property at every $x \in X$.
5) F is said to be *locally bounded* (or uniformly compact) near $\bar{x}$ if there is a neighborhood N of $\bar{x}$ such that the set $\cup_{x \in N} F(x)$ is bounded.

These definitions can be rewritten by using the limit of a set-valued map. First define the distance function

$$d(y, C) = \inf\{\| z - y \| : z \in C\} \tag{7.20}$$

for a set C in R^m and a point $y \in R^m$, where $\| y \|$ denotes the Euclidean norm of a vector y.

Definition 7.4: Let $\bar{x} \in \text{dom } F$.
1) The set defined by

$$\limsup_{x \to \bar{x}} F(x) = \{y \in R^p : \liminf_{x \to \bar{x}} d(y, F(x)) = 0\} \tag{7.21}$$

is called the *upper limit* of F when $x \to \bar{x}$.
2) The set defined by

$$\liminf_{x \to \bar{x}} F(x) = \{y \in R^p : \lim_{x \to \bar{x}} d(y, F(x)) = 0\} \tag{7.22}$$

is called the *lower limit* of F when $x \to \bar{x}$.

It is clear that

$$\liminf_{x \to \bar{x}} F(x) \subset \text{cl } F(\bar{x}) \subset \limsup_{x \to \bar{x}} F(x). \tag{7.23}$$

Similarly we can define the upper limit and the lower limit of a sequence of sets $\{C_k\} \subset R^p$, i.e.,

$$\limsup_{k \to \infty} C^k = \{y \in R^p : \liminf_{k \to \infty} d(y, C^k) = 0\} \tag{7.24}$$

$$\liminf_{k\to\infty} C^k = \{y \in R^p : \lim_{k\to\infty} d(y, C^k) = 0\}. \tag{7.25}$$

Proposition 7.1: Let $\bar{x} \in \text{dom } F$.
1) F is u.s.c. at $\bar{x}$ if and only if

$$\limsup_{k\to\infty} F(x^k) \subset F(\bar{x}) \tag{7.26}$$

for any sequence $\{x^k\} \subset \text{dom } F$ converging to $\bar{x}$.
2) F is l.s.c. at $\bar{x}$ if and only if

$$F(\bar{x}) \subset \liminf_{k\to\infty} F(x^k) \tag{7.27}$$

for any sequence $\{x^k\} \subset \text{dom } F$ converging to $\bar{x}$.
3) F is continuous at $\bar{x}$ if and only if

$$\liminf_{k\to\infty} F(x^k) = F(\bar{x}) = \limsup_{k\to\infty} F(x^k) \tag{7.28}$$

for any sequence $\{x^k\} \subset \text{dom } F$ converging to $\bar{x}$.

7.3.2 Derivatives of set-valued maps

Next we will consider derivatives of set-valued maps. For that purpose we first introduce the concept of the contingent cone ([4], or tangent cone [32]) to a set, which is commonly used in optimization and mathematical programming.

Definition 7.5: Let A be a nonempty subset of R^m and $\hat{v} \in \text{cl } A$. The set $T_A(\hat{v}) \subset R^m$, defined by

$$T_A(\hat{v}) = \{v \in R^m : \liminf_{h\to 0+} \frac{d(\hat{v} + hv, A)}{h} = 0\} \tag{7.29}$$

is called the *contingent cone* to A at $\hat{v}$.

It is very convenient to note that $v \in T_A(\hat{v})$ if and only if there exist sequences $\{h_k\} \subset \text{int } R_+$(the set of positive numbers) and $\{v^k\} \subset R^m$ such that $v^k \to v$, $h_k \to 0$ and $\hat{v} + h_k v^k \in A$ for any k. This conditon is also equivalent to the following condition: there exist sequences $\{h_k\} \subset \text{int } R_+$ and $\{v^k\} \subset A$ such that $v^k \to \hat{v}$ and $h_k(v^k - \hat{v}) \to v$.

It is well known that $T_A(\hat{v})$ is a closed (but not always convex) cone.

The graph of a set-valued map F from R^m to R^p is defined by

$$\text{graph } F = \{(x, y) \in R^m \times R^p : y \in F(x)\}. \tag{7.30}$$

We define the contingent derivative of F in terms of the contingent cone to the graph of F.

Definition 7.6: Let $(\hat{x}, \hat{y}) \in \text{graph } F$. The set-valued map $DF(\hat{x}, \hat{y})$ from R^m to R^p defined by

$$\text{graph } DF(\hat{x}, \hat{y}) = T_{\text{graph}F}(\hat{x}, \hat{y}) \tag{7.31}$$

is called the *contingent derivative* ([4] or *graphical derivative* [32]) of F at $(\hat{x}, \hat{y})$.

The contingent derivative can be characterized in terms of sequences as follows: $y \in DF(\hat{x}, \hat{y})(x)$ if and only if there exist sequences $\{h_k\} \subset \text{int}R_+$ and $\{(x^k, y^k)\} \subset R^m \times R^p$ such that $h^k \to 0$, $(x^k, y^k) \to (x, y)$ and

$$\hat{y} + h_k y^k \in F(\hat{x} + h_k x^k) \text{ for all } k \tag{7.32}$$

or, equivalently, there exist sequences $\{h_k\} \subset \text{int}R_+$ and $\{(x^k, y^k)\} \subset \text{graph } F$ such that $(x^k, y^k) \to (\hat{x}, \hat{y})$ and

$$h_k((x^k, y^k) - (\hat{x}, \hat{y})) \to (x, y). \tag{7.33}$$

Proposition 7.2: Let $(\hat{x}, \hat{y}) \in \text{graph } F$. Then $y \in DF(\hat{x}, \hat{y})(x)$ if and only if

$$y \in \limsup_{(h,x') \to (0+,x)} \frac{F(\hat{x} + hx') - \hat{y}}{h}. \tag{7.34}$$

Therefore the contingent derivative is the same as the Dini upper derivative (Penot [26]). We can also define the Dini lower derivative by replacing limsup by liminf in the above proposition as in the following definition.

Definition 7.7: Let $(\hat{x}, \hat{y}) \in \text{graph } F$. The set-valued map defined by

$$D^L F(\hat{x}, \hat{y})(x) = \liminf_{(h,x') \to (0+,x)} \frac{F(\hat{x} + hx') - \hat{y}}{h} \tag{7.35}$$

is called the *Dini lower derivative* of F at $(\hat{x}, \hat{y})$.

Hence $y \in D^L F(\hat{x}, \hat{y})(x)$ if and only if, for any sequences $\{h_k\} \subset \text{int}R_+$ and $\{x^k\} \subset R^m$ satisfying $h^k \to 0, x^k \to x$, there exists a sequence $\{y^k\} \subset R^p$ such that $y^k \to y$ and $\hat{y} + h_k y^k \in F(\hat{x} + h_k x^k)$ for all k.

It is obvious from the definitions

$$D^L F(\hat{x}, \hat{y})(x) \subset DF(\hat{x}, \hat{y})(x) \tag{7.36}$$

for any $x \in R^m$. When the equality holds in the above relation, F is said to be Dini differentiable at $(\hat{x}, \hat{y})$.

In addition to the contingent derivative, we may also define the following two derivatives. The former was introduced by Shi [34] as the TP-derivative.

Definition 7.8: Let $(\hat{x}, \hat{y}) \in \text{graph } F$. The set-valued map defined by

$$PF(\hat{x},\hat{y})(x) = \{y \in R^p : \text{there exist sequences } \{h_k\} \subset \text{int}R_+ \text{ and } \{(x^k, y^k)\} \subset \text{graph } F \text{ such that } h^k \to 0,\ x^k \to x,\ h_k(x^k - \hat{x}, y^k - \hat{y}) \to (x, y)\} \tag{7.37}$$

is called the *upper semi-derivative* of F at $(\hat{x}, \hat{y})$. On the other hand, the set-valued map defined by

(7.38)

$$P^L F(\hat{x},\hat{y})(x) = \{y \in R^p : \text{for any sequences } \{h_k\} \subset \text{int}R_+ \text{ and } \{x^k\} \subset R^m \text{ satisfying } x^k \to \hat{x} \text{ and } h_k(x^k - \hat{x}) \to x, \text{ there exists a sequence } \{(x^{k_n}, y^{k_n})\} \subset \text{graph } F \text{ such that } h_{k_n}(y^{k_n} - \hat{y}) = y\}$$

is called the *lower semi-derivative* of F at $(\hat{x}, \hat{y})$. Moreover, when $P^L F(\hat{x}, \hat{y})(x) \neq \emptyset$ for any $x \in R^m$, F is said to be *lower semi-differentiable* at $(\hat{x}, \hat{y})$.

Before closing this section we provide a sufficient condition for the lower semi-differentiability.

Definition 7.9: 1) F is said to be *locally Lipschitz* at $\hat{x}$ if there exist a neighborhood N of $\hat{x}$ and a positive constant M such that

$$F(x') \subset F(x) + M \parallel x - x' \parallel B \quad \text{for all } x, x' \in N \tag{7.39}$$

where B is the unit ball.
2) F is said to be *upper locally Lipschitz* at $\hat{x}$ if there exist a neighborhood N of $\hat{x}$ and a positive constant M such that

$$F(x) \subset F(\hat{x}) + M \parallel x - \hat{x} \parallel B \quad \text{for all } x \in N. \tag{7.40}$$

Of course, if F is locally Lipschitz, then it is upper locally Lipschitz.

Proposition 7.3: If F is locally Lipschitz at $\hat{x}$, then it is lower semi-differentiable at $\hat{x}$.

7.4 CONTINUITY OF THE PERTURBATION MAPS

Stability analysis in MCDM has been mainly conducted by discussing continuity of the perturbation map (Naccache [25], Tanino and Sawaragi [41], Luc [23], Penot and Sterna-Karwart [27], Todorov [42] etc.). In this section we consider continuity of the perturbation maps W, G and S mainly based on the author's results. The following sufficient conditions for the upper semicontinuity of W were obtained by Tanino and Sawaragi [38].

Theorem 7.2: The perturbation map W is u.s.c. at u if the following conditions are satisfied:

1) the map Y is continuous at u,
2) $W(u) = S(u)$.

In order to guarantee the lower semicontinuity, we need the following definition.

Definition 7.10: The set-valued map Y is said to be K-*dominated* by W near u if there exists a neighborhood N of u such that

$$Y(u') \subset W(u') + K \tag{7.41}$$

for any $u' \in N$.

Theorem 7.3: The perturbation map W is l.s.c. at u if the following conditions are satisfied:
1) the map Y is continuous at u,
2) Y is locally bounded near u,
3) Y is K-dominated by W near u.

We also obtain the following results concerning the continuity of the proper perturbation map and the weak perturbation map.

Theorem 7.4: The proper perturbation map G is u.s.c. at u if the following conditions are satisfied:
1) the map Y is continuous at u,
2) $G(u) = S(u)$.

Theorem 7.5: The proper perturbation map G is l.s.c. at u if the following conditions are satisfied:
1) the map Y is continuous at u,
2) Y is locally bounded near u,
3) Y is K-closed (i.e. $Y(u) + K$ is a closed set) near u.

Theorem 7.6: The weak perturbation map S is u.s.c. at u if the following conditions are satisfied:
1) the map Y is continuous at u.

Theorem 7.7: The weak perturbation map S is continuous at u if the following conditions are satisfied:
1) the map Y is continuous at u,
2) the map Y is locally bounded and closed-valued near u,
3) W(u)=S(u).

The continuity of Y is, of course, closely related to the continuity of X and f.

Proposition 7.4: 1) If the map X is u.s.c. at u and locally bounded at u and if the function f is continuous on $X(u) \times \{u\}$, then the map Y is u.s.c. at u.
2) If the map X is l.s.c. at u and if the function f is continuous on $X(u) \times \{u\}$, then the map Y is l.s.c. at u.

Moreover, if $X(u)$ is defined by inequality constraints as

$$X(u) = \{x \in R^n : g_j(x,u) \leqq 0,\ j = 1,\ldots,l\} \tag{7.42}$$

sufficient conditions for its continuity are given as follows (Hogan [17]).

Proposition 7.5: 1) If each $g_j(j = 1,\ldots,l)$ is lower semicontinuous on $R^n \times \{u\}$, then X is u.s.c. at u.
2) Let

$$\bar{X}(u) = \{x \in R^n : g_j(x,u) < 0,\ j = 1,\ldots,l\}. \tag{7.43}$$

If each $g_j(j = 1,\ldots,l)$ is upper semicontinuous on $\bar{X}(u) \times \{u\}$ and

$$X(u) \subset \mathrm{cl}\bar{X}(u) \tag{7.44}$$

then X is u.s.c. at u.

Though we dealt with the preference ordering specified by a cone only, we may consider a more general preference structure (domination structure by Yu [43]) of the DM. We can also investigate stability according to the change of preference structure. See Sawaragi et al. [33], chapter 4 for details.

7.5 CONTINGENT DERIVATIVES OF THE PERTURBATION MAPS

7.5.1 General case

In contrast with stability analysis, sensitivity analysis usually requires quantitative investigation on the behavior of the perturbation map. Several papers have been published along this approach (e.g. Tanino [38,39], Shi [34,35], Klose [19], Kuk et al. [21,22]). In this section we investigate the contingent derivatives of the perturbation maps mainly based on the author's results. We start with the following propositions.

Proposition 7.6: Let F be a set-valued map from R^m to R^p and $\hat{y} \in F(\hat{u})$. Then

$$DF(\hat{u},\hat{y})(u) + K \subset D(F+K)(\hat{u},\hat{y})(u) \tag{7.45}$$

for all $u \in R^m$, where $F + K$ is a set-valued map from R^m to R^p defined by

$$(F+K)(u) = F(u) + K \quad \text{for all } u \in R^m. \tag{7.46}$$

The converse inclusion does not generally hold in the above proposition. A sufficient condition for the converse inclusion is given in the following proposition.

Proposition 7.7: Let F be a set-valued map from R^m to R^p and $\hat{y} \in F(\hat{u})$. If

$$PF(\hat{u},\hat{y})(0) \cap (-K) = \{0\} \tag{7.47}$$

then

$$DF(\hat{u},\hat{y})(u) + K = D(F+K)(\hat{u},\hat{y})(u) \tag{7.48}$$

$$\mathrm{Min}_K DF(\hat{u},\hat{y})(u) = \mathrm{Min}_K D(F+K)(\hat{u},\hat{y})(u) \tag{7.49}$$

$$\mathrm{PrMin}_K DF(\hat{u},\hat{y})(u) = \mathrm{PrMin}_K D(F+K)(\hat{u},\hat{y})(u) \tag{7.50}$$

$$\mathrm{WMin}_K DF(\hat{u},\hat{y})(u) \subset \mathrm{WMin}_K D(F+K)(\hat{u},\hat{y})(u) \tag{7.51}$$

for all $u \in R^m$. Moreover, if $\tilde{K}$ is a closed convex cone with $\tilde{K} \subset \mathrm{int}K \cup \{0\}$, then

$$\mathrm{WMin}_K DF(\hat{u},\hat{y})(u) \subset \mathrm{WMin}_K D(F+\tilde{K})(\hat{u},\hat{y})(u) \tag{7.52}$$

for all $u \in R^m$.

If the above condition

$$PF(\hat{u},\hat{y})(0) \cap (-K) = \{0\}$$

is satisfied, then $\hat{y} \in \mathrm{Min}_K\, F(\hat{u})$. In fact if otherwise, there exists $\bar{y} \in F(\hat{u})$ such that $\bar{y} \leq_K \hat{y}$. Putting $t_k \equiv 1$, $(u^k, y^k) \equiv (\hat{u}, \bar{y})$ implies that $\bar{y} - \hat{y} \in PF(\hat{u},\hat{y})(0)$. On the contrary the following result holds.

Lemma 7.1: Let F be a set-valued map from R^m to R^p and $\hat{y} \in F(\hat{u})$. If F is locally upper Lipschitz at $\hat{u}$ and $\hat{y} \in \mathrm{PrMin}_K\, F(\hat{u})$, then

$$PF(\hat{u},\hat{y})(0) \cap (-K) = \{0\}.$$

The following proposition is fundamental in deriving our sensitivity results.

Proposition 7.8: Let F be a set-valued map from R^m to R^p and $\hat{y} \in F(\hat{u})$. Then

$$\mathrm{Min}_K D(F+K)(\hat{u},\hat{y})(u) \subset DF(\hat{u},\hat{y})(u) \tag{7.53}$$

for all $u \in R^m$.

Now we are ready to analyze the contingent derivatives of the perturbation map W and the weak perturbation map S. Since $W(u) \subset Y(u)$, if Y is K-dominated by W near $\hat{u}$, then

$$Y(u) + K = W(u) + K \quad \text{for all } u \in N. \tag{7.54}$$

Hence, in this case,

$$D(Y+K)(\hat{u},\hat{y}) = D(W+K)(\hat{u},\hat{y}) \quad \text{for all } \hat{y} \in W(\hat{u}). \tag{7.55}$$

Similarly, if Y is K-dominated by S near $\hat{u}$,

$$D(Y+K)(\hat{u},\hat{y}) = D(S+K)(\hat{u},\hat{y}) \quad \text{for all } \hat{y} \in W(\hat{u}). \tag{7.56}$$

On the other hand, if the assumptions in Proposition 7.7 are satisfied for $F = Y$, we have the following relationships:

$$\mathrm{Min}_K DY(\hat{u},\hat{y})(u) = \mathrm{Min}_K D(Y+K)(\hat{u},\hat{y})(u), \tag{7.57}$$

$$\mathrm{WMin}_K DY(\hat{u},\hat{y})(u) = \mathrm{WMin}_K\ D(Y+\tilde{K})(\hat{u},\hat{y})(u). \tag{7.58}$$

Therefore we have the following theorem.

Theorem 7.8: Assume that

$$PY(\hat{u},\hat{y})(0) \cap (-K) = \{0\}. \tag{7.59}$$

1) If Y is K-dominated by W near $\hat{u}$, then

$$\mathrm{Min}_K DY(\hat{u},\hat{y})(u) \subset DW(\hat{u},\hat{y})(u) \quad \text{for any } u \in R^m. \tag{7.60}$$

2) If Y is $\tilde{K}$-dominated by S near $\hat{u}$, then

$$\mathrm{WMin}_K DY(\hat{u},\hat{y})(u) \subset DS(\hat{u},\hat{y})(u) \quad \text{for any } u \in R^m, \tag{7.61}$$

where $\tilde{K}$ is a closed convex cone with $\tilde{K} \subset \mathrm{int}K \cup \{0\}$.

Corollary 7.1: Assume that

$$PY(\hat{u},\hat{y})(0) \cap (-K) = \{0\}.$$

If, for each $u \in R^m$,

$$\mathrm{Min}_K DY(\hat{u},\hat{y})(u) = DY(\hat{u},\hat{y})(u) \tag{7.62}$$

then, for each $u \in R^m$,

$$\mathrm{Min}_K DY(\hat{u},\hat{y})(u) = DW(\hat{u},\hat{y})(u). \tag{7.63}$$

In order to obtain the relationship between DY and DG, we introduce the concepts of cone closedness and cone boundedness.

Definition 7.11: Let C be a nonempty set in R^p and D be a cone in R^p. Then C is said to be
1) *D-closed* if $C + D$ is closed,

2) *D-bounded* if $C^+ \cap (-D) = \{0\}$
where

$$C^+ = \{y \in R^p : \text{ there exist } \{h_k\} \subset \text{int}R_+ \text{ and } \{y^k\} \subset C \text{ such that } h_k \to 0,\ h_k y^k \to y\}. \tag{7.64}$$

Lemma 7.2: Suppose that $Y(u)$ be a K-bounded, K-closed set for any $u \in N$, where N is a neighborhood of $\hat{u}$. Then, for any $u \in R^m$

$$DG(\hat{u}, \hat{y})(u) = DW(\hat{u}, \hat{y})(u). \tag{7.65}$$

Theorem 7.9: Assume that

$$PY(\hat{u}, \hat{y})(0) \cap (-K) = \{0\}.$$

If $Y(u)$ be a K-bounded, K-closed set for any $u \in N$, then

$$\text{PrMin}_K DY(\hat{u}, \hat{y})(u) = DG(\hat{u}, \hat{y})(u) \tag{7.66}$$

for any $u \in R^m$.

Finally we obtain a relationship between DW and WMin DY as follows.

Theorem 7.10: If Y is Dini differentiable at $(\hat{u}, \hat{y})$, then

$$DW(\hat{u}, \hat{y})(u) \subset \text{WMin}_K DY(\hat{u}, \hat{y})(u) \tag{7.67}$$

for any $u \in R^m$.

Thus we have obtained the following relationships under some conditions:

(7.68)

$$\begin{array}{ccccc} \text{PrMin}_K DY(\hat{u}, \hat{y})(u) & \subset & \text{Min}_K DY(\hat{u}, \hat{y})(u) & \subset & \text{WMin}_K DY(\hat{u}, \hat{y})(u) \\ \cap & & \cap & & \cap \\ DG(\hat{u}, \hat{y})(u) & = & DW(\hat{u}, \hat{y})(u) & \subset & DS(\hat{u}, \hat{y})(u) \end{array}$$

Next we consider how to obtain the contingent derivative of the feasible set map Y from that of X and the derivatives of F, because

$$Y(u) = \{y \in R^p : y = f(x, u) \text{ for some } x \in X(u)\}. \tag{7.69}$$

For that purpose we define the following set-valued map $\tilde{X}$ from $R^m \times R^p$ to R^n:

$$\tilde{X}(u, y) = \{x \in X(u) : f(x, u) = y\}. \tag{7.70}$$

Proposition 7.9: Let $\hat{y} = f(\hat{x}, \hat{u})$ for $\hat{x} \in X(\hat{u})$. Then, for any $u \in R^m$

$$\nabla_x f(\hat{x}, \hat{u}) DX(\hat{u}, \hat{x})(u) + \nabla_u f(\hat{x}, \hat{u})u \subset DY(\hat{u}, \hat{y})(u). \tag{7.71}$$

Moreover, if $\tilde{X}$ is lower semi-differentiable at $(\hat{u}, \hat{y}, \hat{x})$, then the equality holds in the above relation, i.e., for any $u \in R^m$

$$\nabla_x f(\hat{x}, \hat{u}) DX(\hat{u}, \hat{x})(u) + \nabla_u f(\hat{x}, \hat{u}) u = DY(\hat{u}, \hat{y})(u). \tag{7.72}$$

The following lemma provides a suffcient condition for lower semi-differentiability of $\tilde{X}$.

Lemma 7.3: Let $\hat{y} = f(\hat{x}, \hat{u})$ for $\hat{x} \in X(\hat{u})$. If $\tilde{X}$ is upper locally Lipschitz at $(\hat{u}, \hat{y})$ and $\tilde{X}(\hat{u}, \hat{y}) = \{\hat{x}\}$ (i.e. singleton), then $\tilde{X}$ is lower semi-differentiable at $(\hat{u}, \hat{y}, \hat{x})$.

Moreover, if $X(u)$ is defined by inequality constraints as

$$X(u) = \{x \in R^n : g_j(x, u) \leqq 0,\ j = 1, \ldots, l\} \tag{7.73}$$

i.e.

$$\text{graph } X = \{(u, x) \in R^m \times R^n : g_j(x, u) \leqq 0,\ j = 1, \ldots, l\} \tag{7.74}$$

then its contingent derivatives is given as follows:

Proposition 7.10: Suppose that $X(u)$ is given in (7.68), that $\hat{x} \in X(\hat{u})$ and that there exists $(\bar{x}, \bar{u})$ such that

$$\nabla_x g_j(\hat{x}, \hat{u})\bar{x} + \nabla_u g_j(\hat{x}, \hat{u})\bar{u} < 0 \text{ for all } j \in J(\hat{x}, \hat{u}) \tag{7.75}$$

where

$$J(\hat{x}, \hat{u}) = \{j :\ g_j(\hat{x}, \hat{u}) = 0\}. \tag{7.76}$$

Then $x \in DX(\hat{x}, \hat{u})$ if and only if

$$\nabla_x g_j(\hat{x}, \hat{u})x + \nabla_u g_j(\hat{x}, \hat{u})u \leqq 0 \text{ for all } j \in J(\hat{x}, \hat{u}). \tag{7.77}$$

7.5.2 Convex case

If we assume appropriate convexity conditions on the original problem, we can refine the above results. First we consider convexity conditions.

Definition 7.12: A set-valued map F is said to be *K-convex* if, for any $u, u' \in R^m$ and any $\alpha \in [0, 1]$

$$\alpha F(u) + (1 - \alpha)F(u') \subset F(\alpha u + (1 - \alpha)u') + K. \tag{7.78}$$

F is simply said to be *convex* if it is $\{0\}$-convex, i.e. if graph F is a convex set.

Hereafter throughout this section we assume the following convexity conditions (CA):

Convexity Assumption (CA)

1) The set-valued map X is convex.

2) The function f is K-convex, i.e.

$$\alpha f(x,u) + (1-\alpha) f(x',u') \in f(\alpha x + (1-\alpha)x', \alpha u + (1-\alpha)u') + K \tag{7.79}$$

for any $u, u' \in R^m$ and any $\alpha \in [0,1]$.

Proposition 7.11: Under the convexity assumption (CA), the set-valued map Y defined by

$$Y(u) = \{y \in R^p : y = f(x,u),\ x \in X(u)\}$$

is K-convex.

We can omit the condition

$$PY(\hat{u},\hat{y})(0) \cap (-K) = \{0\}$$

under the convexity assumption (CA).

Theorem 7.11: 1) If Y is K-dominated by W near $\hat{u}$, then

$$\mathrm{Min}_K DY(\hat{u},\hat{y})(u) \subset DW(\hat{u},\hat{y})(u) \tag{7.80}$$

for any $u \in R^m$.

2) If Y is $\tilde{K}$-dominated by S near $\hat{u}$, then

$$\mathrm{WMin}_K DY(\hat{u},\hat{y})(u) \subset DS(\hat{u},\hat{y})(u) \tag{7.81}$$

for any $u \in R^m$.

Next we consider sufficient conditions for the converse inclusion of the above theorem. For a cone $C \subset R^p$, we denote its negative polar cone by C°, i.e.,

$$C^\circ = \{z \in R^p :\ z^T y \leqq 0 \text{ for all } y \in C\} \tag{7.82}$$

Definition 7.13: Let A be a nonempty K-convex set in R^p. If a point $\hat{y} \in \mathrm{Min}_K A$ satisfies the condition

$$[T_{A+K}(\hat{y})]^\circ \subset \text{int } K^\circ \cup \{0\} \tag{7.83}$$

or equivalently

$$\text{int } T_{A+K}(\hat{y}) \cup \{0\} \supset K, \tag{7.84}$$

then $\hat{y}$ is called the *normally K-minimal point* of A.

Remark 7.1: A point $\hat{y} \in A$ is said to be a properly K-minimal point of a K-convex set A if

$$T_{A+K}(\hat{y}) \cap (-K) = \{0\}. \tag{7.85}$$

In this case there exists a vector $\lambda \in [T_{A+K}(\hat{y})]^\circ \cap \text{int } K^\circ$. Thus, the normal K-minimality is a stronger concept than the proper K-minimality. From the geometric point of view, the latter implies the existence of the supporting hyperplane to A at $\hat{y}$ with the normal vector $\lambda \in \text{int } K^\circ$ and, on the other hand, the former implies that all the normal vectors of the supporting hyperplanes to A at $\hat{y}$ belong to int K°.

Theorem 7.12: If $\hat{u} \in \text{int}(\text{dom } Y)$ and $\hat{y}$ is a normally K-minimal point of $Y(\hat{u})$, then

$$DW(\hat{u}, \hat{y})(u) \subset \text{Min}_K DY(\hat{u}, \hat{y})(u) \tag{7.86}$$

for any $u \in R^m$.

Moreover we can obtain the following theorem concerning the contingent derivative of the weak perturbation map S.

Theorem 7.13: 1) If $\hat{u} \in \text{int}(\text{dom } Y)$, then

$$DS(\hat{u}, \hat{y})(u) \subset \text{WMin}_K DY(\hat{u}, \hat{y})(u) \tag{7.87}$$

for any $u \in R^m$. 2) Moreover, if Y is $\tilde{K}$-dominated by S near $\hat{u}$ and $\hat{y} \in G(\hat{u})$, then

$$DS(\hat{u}, \hat{y})(u) = \text{WMin}_K DY(\hat{u}, \hat{y})(u) \tag{7.88}$$

for any $u \in R^m$, where $\tilde{K}$ is a closed convex cone such that $\tilde{K} \subset \text{int } K \cup \{0\}$.

Therefore the following relationships hold under some appropriate conditions.

(7.89)

$$\begin{array}{ccccc} \text{PrMin}_K DY(\hat{u}, \hat{y})(u) & \subset & \text{Min}_K DY(\hat{u}, \hat{y})(u) & = & \text{WMin}_K DY(\hat{u}, \hat{y})(u) \\ \cap & & \| & & \| \\ DG(\hat{u}, \hat{y})(u) & = & DW(\hat{u}, \hat{y})(u) & = & DS(\hat{u}, \hat{y})(u) \end{array}$$

Example 7.1: As an illustrated example we consider the following convex multiobjective programming problem with $K = R^2_+$. Let $u \in [0, +\infty) \subset R$.

$$\begin{array}{ll} \text{minimize} & f(x) = (x_1 + x_2, x_1 - x_2)^T \\ \text{subject to} & g(x, u) = x_1^2 + 2x_2^2 - u \leqq 0 \end{array}$$

Then

$$X(u) = \{x \in R^2 : x_1^2 + 2x_2^2 \leqq u\}$$

$$Y(u) = \{y \in R^2 : 3y_1^2 - 2y_1y_2 + 3y_2^2 \leqq 4u\}$$

Hence

$$\begin{aligned} W(u) = S(u) = \{y \in R^2 : \quad & 3y_1^2 - 2y_1y_2 + 3y_2^2 \leqq 4u, \\ & -(\tfrac{3}{2}u)^{1/2} \leqq y_1 \leqq -(\tfrac{1}{6}u)^{1/2}, \\ & -(\tfrac{3}{2}u)^{1/2} \leqq y_2 \leqq -(\tfrac{1}{6}u)^{1/2}\} \end{aligned}$$

If we put

$$\hat{u} = 1, \ \hat{x} = (-1, 0)^T, \ \hat{y} = (-1, -1)^T$$

then, since

$$\nabla_x g(\hat{x}, \hat{u}) = (-2, 0), \ \nabla_u g(\hat{x}, \hat{u}) = -1, \ \nabla_x f(\hat{x}) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

we have

$$DX(\hat{u}, \hat{x})(u) = \{x : -2x_1 - u \leqq 0\}$$

$$\begin{aligned} DY(\hat{u}, \hat{y})(u) &= \nabla_x f(\hat{x}) DX(\hat{u}, \hat{x})(u) \\ &= \{y \in R^2 : y_1 + y_2 + u \geqq 0\} \end{aligned}$$

and

$$\mathrm{Min} DY(\hat{u}, \hat{y})(u) = \{y \in R^2 : y_1 + y_2 + u = 0\}$$

which clearly coincides with

$$DW(\hat{u}, \hat{y})(u) = \{y \in R^2 : y_1 + y_2 + u = 0\}$$

We should also note that the fact

$$DW(\hat{u}, \hat{y})(0) = \{y \in R^2 : y_1 + y_2 = 0\}$$

provides the trade-off ratio 1 between the two objectives f_1 and f_2 at $\hat{y} = (-1, -1)^T$. ■

7.6 SENSITIVITY ANALYSIS USING DUALITY THEORY

Duality theory in MCDM has been investigated by several authors (see, for example, Chapter 3 dealing with duality written by Nakayama in this monograph). Some sensitivity results can be obtained as a by-product of the duality theory. In this section we introduce two of them, one in the linear case and the other in the convex case.

7.6.1 *Case of linear programming*

As was noted before, in the linear case, we can concentrate on a finite number of extreme points (basic solutions) in a constrained polyhedral set in a multiobjective programming problem. Kornbluth [20] analyzed the sensitivity of an efficient solution by determining the set of possible weights, so-called indifference region, associated with each efficient solution.

The primal multiobjective linear programming problem which was considered by Kornbluth is as follows:

$$\begin{array}{ll} \text{maximize} & Cx \\ \text{subject to} & Ax \leqq Q\mu \\ & x \geqq 0 \end{array} \tag{7.90}$$

where $x \in R^n, \mu \in R^q, A \in R^{m \times n}, Q \in R^{m \times q}, C \in R^{p \times n}$. Its dual problem is given by

$$\begin{array}{ll} \text{minimize} & Q^T\pi \\ \text{subject to} & A^T\pi \geqq C^T\lambda \\ & \pi \geqq 0 \end{array} \tag{7.91}$$

where $\pi \in R^m$. Throughout this section we assume that the weighting vector satisfies $\mu \in S^q$ (q-dimensional fundamental simplex) and $\lambda \in S^p$, i.e.

$$\sum_{j=1}^{q} \mu_j = 1, \ \mu \geqq 0; \ \sum_{i=1}^{p} \lambda_i = 1, \ \lambda \geqq 0. \tag{7.92}$$

The ordering cone K is supposed to be R^p_+ (nonnegative orthant of R^p). Namely we seek for usual Pareto maximal and minimal points in the above problems, respectively. We call them efficient solutions as Kornbluth did hereafter in this subsection.

Proposition 7.12: Given the weighting vector $\hat{\mu}$, $\hat{x}$ is properly efficient for the primal problem with $\mu = \hat{\mu}$ if and only if there exist a weighting vector $\hat{\lambda}$ and a dual variable $\hat{\pi}$ such that $\hat{\pi}$ is properly efficient for the dual problem with $\lambda = \hat{\lambda}$.

Lemma 7.4: Given the matrices C, A and Q, a vector $\hat{x}$ is efficient for the primal problem if and only if the Kuhn-Tucker conditions are satisfied, namely there exist $\hat{\pi}, \hat{\lambda}, \hat{\mu}$ such that

$$\begin{array}{c} A^T\hat{\pi} \geqq C^T\hat{\lambda} \\ (\hat{\pi}^T A - \hat{\lambda}^T C)\hat{x} = 0 \\ A\hat{x} \leqq Q\hat{\mu} \\ \hat{\pi}^T(A\hat{x} - Q\hat{\mu}) = 0 \\ \hat{x}, \hat{\pi}, \hat{\lambda}, \hat{\mu} \geqq 0 \end{array} \tag{7.93}$$

Corollary 7.2: Assuming that the efficient solutions to the primal problem are nondegenerate, the Kuhn-Tucker conditions imply the following relationships:

$$\begin{aligned} \hat{x} &= B^{-1}Q\hat{\mu} \\ \hat{\pi} &= B^{-T}\hat{C}^T\hat{\lambda} \end{aligned} \tag{7.94}$$

where B is the associated basic matrix, $\hat{x}$ is the basic vector (including slacks where necessary), and $\hat{C}$ is the matrix of coefficients of C corresponding to B and $\hat{x}$.

Thus, for given $\hat{\mu}$ and given $\hat{x}$, there exists a space $\Lambda(\hat{\mu}, \hat{x})$ such that any $\lambda \in \Lambda(\hat{\mu}, \hat{x})$ satisfies the above relationships. This space is called the *indifference region* for the weights associated with an efficient solution $\hat{x}$. Of course

$$\bigcup_{\hat{x}} \Lambda(\hat{\mu}, \hat{x}) = S^p = \{\lambda \in R^p : \sum_{i=1}^{p} \lambda_i = 1,\ \lambda \geqq 0\}. \tag{7.95}$$

If we assume that the decision maker can recognize that his preference is within a paticular region $\Lambda(\hat{\mu}, \hat{x})$, we can deduce the problem of sensitivity analysis to that of estimating the effect of changing data on the indifference region $\Lambda(\hat{\mu}, \hat{x})$.

Example 7.2 (Kornbluth [20] and Hannan [15]): Let

$$C = \begin{pmatrix} 3 & 1 & 1 \\ 1 & -1 & 2 \\ 1 & 2 & 0 \end{pmatrix}, \quad A = \begin{pmatrix} 4 & 2 & 3 \\ 1 & 3 & 2 \\ 0 & 0 & 1 \end{pmatrix},$$

$$Q = \begin{pmatrix} 20 & 0 \\ 20 & -4 \\ 6 & 4 \end{pmatrix}, \quad \mu = \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix}.$$

There are six decision variables (including the three slack variables) and four efficient bases:

$$(x_1, x_5, x_6),\ (x_1, x_2, x_6),\ (x_2, x_3, x_6) \text{ and } (x_3, x_5, x_6)$$

Looking at the efficient solution $\hat{x} = (x_1, x_5, x_6) = (\frac{5}{2}, \frac{11}{2}, 5)$,

$$\hat{C} = \begin{pmatrix} 3 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad B^{-1} = \begin{pmatrix} \frac{1}{4} & 0 & 0 \\ -\frac{1}{4} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Hence

$$\hat{\pi} = B^{-T}\hat{C}^T\lambda = (\frac{3}{4}\lambda_1 + \frac{1}{4}\lambda_2 + \frac{1}{4}\lambda_3, 0, 0)^T.$$

Thus from the constraints $A^T\hat{\pi} \geqq C\lambda$ and $\hat{\pi} \geqq 0$, we can obtain the indifference region $\Lambda(\mu, \hat{x})$ as

$$\begin{aligned} &\tfrac{1}{2}\lambda_1 + \tfrac{3}{2}\lambda_2 - \tfrac{3}{2}\lambda_3 \geqq 0 \\ &\tfrac{5}{4}\lambda_1 - \tfrac{5}{4}\lambda_2 + \tfrac{3}{4}\lambda_3 \geqq 0 \\ &\tfrac{3}{4}\lambda_1 + \tfrac{1}{4}\lambda_2 + \tfrac{1}{4}\lambda_3 \geqq 0 \\ &\lambda_1 + \lambda_2 + \lambda_3 = 1,\ \lambda_1, \lambda_2, \lambda_3 \geqq 0 \end{aligned}$$

These inequalities specify the values of λ for which (x_1, x_5, x_6) is an efficient basic solution. ■

7.6.2 *Case of convex programming*

Next we consider sensitivity with respect to the right-hand side perturbation in convex multiobjective programming problem. The following results are essentially due to Balbás and Jiménez Guerra [5]. Though they considered a convex programming problem with right hand side perturbation in more general spaces, we focus on the case of finite dimensional Euclidean spaces. Thus we consider the following primal problem:

$$(\mathrm{P}_b) \qquad \begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in X,\ g(x) \leqq b \end{array} \tag{7.96}$$

where $f : R^n \to R^p$, $g : R^n \to R^m$, $X \subset R^n$. We assume the convexity i.e. every $f_i (i = 1, \ldots, p)$ and every $g_j (j = 1, \ldots, m)$ are convex functions and X is a convex set.

Let

$$X(b) = \{x \in X : g(x) \leqq b\} \tag{7.97}$$

and

$$\lambda \in S^p = \{\lambda \in R^p : \sum_{i=1}^{p} \lambda_i = 1,\ \lambda_i \geqq 0\ (i = 1, \ldots, p)\}. \tag{7.98}$$

Definition 7.14: $\hat{x} \in X(b)$ is said to be a λ-*optimal* solution of (P_b) if

$$\lambda^T f(\hat{x}) \leqq \lambda^T f(x) \quad \text{for any } x \in X(b). \tag{7.99}$$

Definition 7.15: A positive vector $\mu \in R^m_+$ satisfying

$$\inf_{x \in X(b)} \lambda^T f(x) = \inf_{x \in X} \{\lambda^T f(x) + \mu^T (g(x) - b)\} \tag{7.100}$$

is said to be a *Lagrangian* λ-*multiplier* for (P_b).

Proposition 7.13: Let $x_0 \in X(0)$ and $x_b \in X(b)$ be λ-optimal solutions for (P_0) and (P_b) respectively, and μ_0 and μ_b be corresponding Lagrangian multiplier vectors respectively. Then

$$-\mu_b^T b \geqq \lambda^T (f(x_b) - f(x_0)) \geqq -\mu_0^T b. \tag{7.101}$$

Now let

$$M_\lambda = \{\mu \in R_+^m : \{\lambda^T f(x) + \mu^T g(x) : x \in X\} \text{ is bounded below}\}, \tag{7.102}$$

$$\phi(\lambda, \mu) := \inf\{\lambda^T f(x) + \mu^T g(x) : x \in X\}, \quad \lambda \in S^p, \ \mu \in M_\lambda, \tag{7.103}$$

and

$$\psi(\lambda, \mu) := \frac{1}{\sum_{1=1}^p \lambda_i^2} \phi(\lambda, \mu)\lambda \tag{7.104}$$

For every $\lambda \in S^p$, the λ-dual program of (P_b) will be

$$\begin{array}{ll} \text{maximize} & \phi(\lambda, \mu) - \mu^T b \\ \text{subject to} & \mu \in M_\lambda. \end{array} \tag{7.105}$$

Moreover, the dual program of (P_b) will be

$$(D_b) \quad \begin{array}{ll} \text{maximize} & \psi(\lambda, G^T\lambda) - Gb \\ \text{subject to} & \lambda \in S^p, \ G \in R^{p\times m}, \ G^T\lambda \in M_\lambda. \end{array} \tag{7.106}$$

Proposition 7.14 (weak duality theorem): If $x \in X(b), G \in R^{p\times m}, \lambda \in S^p, G^T\lambda \in M_\lambda$, then

$$\psi(\lambda, G^T\lambda) - Gb \geq f(x) \tag{7.107}$$

never holds.

Corollary 7.3: If $x \in X(b), \lambda \in S^p, G^T\lambda \in M_\lambda$ and

$$\psi(\lambda, G^T\lambda) - Gb = f(x), \tag{7.108}$$

then x is an optimal solution of (P_b) and (λ, G) is an optimal solution of (D).

Definition 7.16: Under the notations of the above corollary, x and (λ, G) are called *associated* solutions.

Theorem 7.14: Let $\lambda \in S^p$ and $x \in X(b)$ be an λ-optimal solution of (P_b). If $b \neq 0$, the following assertions are equivalent.
1) There exists $\mu \in R_+^m$ which is a Lagrange λ-multiplier for (P_b).
2) There exists a $p \times m$ matrix G such that (λ, G) is an optimal solution of (D_b) and x and (λ, G) are associated solutions.

Based on these facts, the following sensitivity theorem can be obtained.

Theorem 7.15: Let V be an open set in R^m, $\lambda \in S^p$, x_b be a λ-optimal solution of (P_b), and G_b be a $p \times m$ matrix such that (λ, G_b) is an optimal solution of

(D_b) associated with x_b for every $b \in V$. If the function $h : Y \to R^{p \times m}$ defined by

$$h(b) = G_b \quad \text{for every } b \in V \tag{7.109}$$

is differentiable, then the function $F : V \to R^p$ such that

$$F(b) = f(x_b) \quad \text{for every } b \in V \tag{7.110}$$

is also differentiable and moreover, the equality

$$F'(b) = -G_b - K(b) \tag{7.111}$$

holds for every $b \in V$, where $K(b)$ denotes the projection of $h'(b)b$ onto $\{\lambda\}^\perp$ (orthogonal compliment of the linear subspace spanned by the vector λ) for every $b \in V$.

An illustrative example is given in Balbás and Jiménez Guerra [5].

Example 7.3: Let $b \in (-\infty, 0)$ and consider a convex multiobjective problem

$$\begin{array}{ll} \text{minimize} & f(x) = (x_1^2, 2x_2^2)^T \\ \text{subject to} & g(x) = x_1 + x_2 \leqq b. \end{array}$$

Let $\lambda = (\frac{1}{2}, \frac{1}{2})$. Then λ-optimal solution of (P_b) is clearly $x_b = (\frac{2}{3}b, \frac{1}{3}b)$ and the Lagrangian λ-multiplier is $\mu_b = -\frac{2}{3}b$. Therefore the function F in the above theorem is

$$F(b) = f(\frac{2}{3}b, \frac{1}{3}b) = (\frac{4}{9}b^2, \frac{2}{9}b^2)^T$$

and

$$F'(b) = (\frac{8}{9}b, \frac{4}{9}b)^T.$$

On the other hand

$$\phi(\lambda, \mu) = -\frac{3}{4}\mu^2,$$

$$\psi(\lambda, \mu) = (-\frac{3}{4}\mu^2, -\frac{3}{4}\mu^2)^T,$$

$$G_b = (-\frac{7}{9}b, -\frac{5}{9}b)^T.$$

Hence

$$-G_b - K(b) = -\begin{pmatrix} -\frac{7}{9}b \\ -\frac{5}{9}b \end{pmatrix} - \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} -\frac{7}{9}b \\ -\frac{5}{9}b \end{pmatrix} = \begin{pmatrix} \frac{8}{9}b \\ \frac{4}{9}b \end{pmatrix} = F'(b)$$

and the result of Theorem 7.15 is illustrated. ■

7.7 SENSITIVITY ANALYSIS IN THE DISCRETE MCDM

In this section we assume that the set of feasible alternatives (decisions) is fixed as

$$A = \{a_1, a_2, \ldots, a_l\} \text{ (finite set)} \tag{7.112}$$

and consider a discrete multiple criteria decision making problem with an evaluation function $\Psi(\cdot, u)$ which depends on a parameter u which reflects uncertainty. We also assume that u satisfies the condition $u \in S$ which is a convex set in R^m.

In this case our attention is restricted mainly to the following analysis:

1. To obtain the region of u for which each alternative is considered to be best.
2. How does the best alternative change when the parameter value changes.

Ríos-Insua published a monograph dealing with this kind of analysis in detail ([30], see also Ríos-Insua and French [31]). His analysis was also extended by Proll et al. [28]. We will briefly introduce their results in the following.

First the non-dominated alternatives constitute the solution set to the problem at hand. Alternative a_j dominates a_k if

$$\begin{array}{ll} \Psi(a_j, u) \geqq \Psi(a_k, u) & \text{for all } u \in S \text{ and} \\ \Psi(a_j, u) > \Psi(a_k, u) & \text{for some } u \in S. \end{array} \tag{7.113}$$

The set of all non-dominated alternatives is denoted by A_1.

Definition 7.17: a_j is *potentially optimal* if

$$\Psi(a_j, u) \geqq \Psi(a_k, u) \text{ for all } a_k \in A \tag{7.114}$$

for some $u \in S$.

To find out whether $a_j \in A_1$ is potentially optimal, the following problem may be solved:

$$\begin{array}{ll} \text{minimize} & \max\{\Psi(a_k, u) - \Psi(a_j, u) : a_k \in A_1, k \neq j\} \\ \text{subject to} & u \in S. \end{array} \tag{7.115}$$

a_j is potentially optimal if and only if the optimal value of the above problem is $\leqq 0$.

Definition 7.18: The *optimality subset* S_j associated with a_j is

$$S_j = \{u \in S : \Psi(a_j, u) \geqq \Psi(a_k, u) \text{ for all } a_k \in A\}. \tag{7.116}$$

An alternative a_j is potentially optimal if and only if $S_j \neq \emptyset$. We assume that we now have an estimate $\hat{u}$ of u and the current best alternative is a_*, i.e.

$$\Psi(a_*, u) \geqq \Psi(a_k, u) \text{ for all } a_k \in A. \tag{7.117}$$

Definition 7.19: a_j is *adjacent potentially optimal* to a_* if

$$S_j \cap S_* \neq \emptyset. \tag{7.118}$$

To find whether a_j is adjacent potentially optimal to a_*, we may solve the following problem:

$$\begin{array}{ll} \text{minimize} & (\Psi(a_j,u) - \Psi(a_*,u))^2 \\ \text{subject to} & u \in S,\ \Psi(a_k,u) - \Psi(a_*,u) \leqq 0, \text{ for all } k \neq j. \end{array} \tag{7.119}$$

Then a_j is adjacent potentially optimal to a_* if and only if the optimal value of the above problem is 0. The solution of this problem gives us u for which both a_j and a_* are optimal.

Let A_* be the set of adjacent potentially optimal alternatives to a_*. Proll et al. stated that filtering of the initial set down to A_* is justified by the observation that the initial judgemental inputs $\hat{u}$ are not arbitrary but rather represent the values with which the DM is most comfortable. However, the DM may reflect upon these values and modify them as his understanding of the problem is enhanced by the decision analysis. To focus the DM's attention, it is not enough to list the immediate contenders of a_* for optimality. It is desirable to have a quantitative measure of how close each contender is to a_*. Equivalently, one would like to know how sensitive a_* is to changes in the judgemental data. The obvious approach is to find the distance between $\hat{u}$ and each of the common boundaries of S_* and its adjacent optimality subsets. This can be done by solving, for each $a_j \in A_1$, a problem

$$\begin{array}{ll} \text{minimize} & d(u,\hat{u}) \\ \text{subject to} & u \in S,\ \Psi(a_j,u) - \Psi(a_*,u) = 0 \end{array} \tag{7.120}$$

where $d(\cdot,\cdot)$ is some continuous metric. The optimal value d_j of the above problem is the least change in u such that the DM is indifferent between a_j and a_*.

Ríos Insua also proposed an index $r \in [0,1]$ in order to give an indication of the sensitivity of the DM's decisions to the judgemental data. This index is derived as follows. Let $\rho = \min\{d_j : a_j \in A_*\}$. Then ρ is the largest radius of the ball centerd at $\hat{u}$ contained in S_*. A relative measure of insensitivity can be defined as the ratio $r = \frac{\rho}{\delta}$, where δ is the radius of the smallest ball centered at $\hat{u}$ containing S, i.e. $\delta = \max_{u \in S} d(u,\hat{u})$.

7.8 CONCLUSIONS

We have explained several interesting results in stability and sensitivity analysis in MCDM or vector optimization. Selection of the topics is based on the author's personal preference. Mainly because of page limitation there remain several other topics which are not included in this chapter. For example, Jurkiewicz [18] and Steuer [36] dealt with sensitivity analysis in goal programming (see also Dauer and Liu [6]). Rarig and Haimes [29] regarded the parameter

as a random variable and sensitivity analysis as a kind of a stochastic problem. Fuller and Fedrizzi [9] dealt with stability in multiobjective possibilistic (i.e. fuzzy) linear programming. Guddat et al. [14] provided unified approach to both multiobjective and stochastic optimization based on parametric optimization. Results based on sensitivity analysis are often useful in hierarchical optimization (see, e.g. Shimizu et al. [37]). Research in this direction will be expected in the future.

Finally a special issue of sensitivity analysis in the Journal of Multi-Criteria Decision Analysis is to be published (guest editor Ríos Insua). It will provide several interesting new results in this field.

Acknowledgments

The author wishes to thank Professors Tomas Gal, Thomas Hanne and Theodor Stewart for encouraging him to write this chapter. He would also like to thank Professor Hirotaka Nakayama and Dr. Hun Kuk for their collaboration in his research.

References

[1] Antunes, C.H., J.N. Clímaco: Sensitivity analysis in MCDM using the weight space, Europ. J. Oper. Res. 12(1992)187-196

[2] Aubin, J.-P.: "Contingent derivatives of set-valued maps and existence of solutions to nonlinear inclusions and differential inclusions". In: L. Nachbin (ed): *Advances in Mathematics, Supplementary Studies*, Academic Press, New York, 1981, pp. 160-232

[3] Aubin, J.-P., I. Ekeland: *Applied Nonlinear Analysis*, John Wiley, New York, 1984

[4] Aubin, J.-P., H. Frankowska: *Set-Valued Analysis*, Birkhäuser, Boston, 1990

[5] Balbás, A., P. Jiménez Guerra: Sensitivity analysis for convex multiobjective programming in abstract spaces, J. Math. Anal. Appl. 202(1996)645-658

[6] Dauer, J., Y.-H. Liu: "Multi-criteria and goal programming". In: T. Gal, H.J. Greenberg (eds): *Advances in Sensitivity Analysis and Parametric Programming*, Kluwer Academic, Amsterdam 1997, pp. 11-1 - 11-31

[7] Deshpande, D.V., S. Zionts: Sensitivity analysis in multiple objective linear programming: Changes in the objective function matrix, Working Paper no. 399 (1979), State Univ. New York

[8] Fiacco, A.V.: *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*, Academic Press, New York, 1983

[9] Fuller, R., M. Fedrizzi: Stability in multiobjective possibilistic linear programs, Europ. J. Oper. Res. 74(1994)179-187

[10] Gal, T.: A general method for determining the set of all efficient solutions to a linear vectormaximum problem, Europ. J. Oper. Res. 1(1977)307-322

[11] Gal, T.: *Postoptimal Analyses, Parametric Programming, and Related Topics*, 2nd ed., Walter de Gruyter, Berlin, 1994

[12] Gal, T., H. Leberling: Relaxation analysis in linear vector-valued maximization, Euro. J. Oper. Res. 8(1981)274-284

[13] Gal, T., K. Wolf: Stability in vector maximization – A survey, Europ. J. Oper. Res. 25(1986)169-182

[14] Guddat, J., F. Guerra Vazquez, K. Tammer, K. Wendler: *Multiobjective and Stochastic Optimization Based on Parametric Optimization*, Akademie Verlag, Berlin 1985

[15] Hannan, E.L.: Using duality theory for identification of primal efficient points and for sensitivity analysis in multiple objective linear programming, J. Oper. Res. Soc. 29(1978)643-649

[16] Hansen, P., M. Labbé, R.E. Wendell: Sensitivity analysis in multiple objective linear programming, Europ. J. Oper. Res. 38(1989)63-69

[17] Hogan, W.W.: Point-to-set maps in mathematical programming, SIAM Rev. 15(1973)591-603

[18] Jurkiewicz, E.: Stability of compromise solution in multicriteria decision-making problems, J. Optim. Theory Appl. 40(1983)77-83

[19] Klose, J.: Sensitivity analysis using the tangent derivative, Numer. Funct. Anal. Optim. 13(1992)143-153

[20] Kornbluth, S.H.: Duality, indifference and sensitivity analysis in multiple objective linear programming, Oper. Res. Quart. 25(1974)599-614

[21] Kuk, K., T. Tanino, M. Tanaka: Sensitivity analysis in vector optimization, J. Optim. Theory Appl. 89(1996)713-730

[22] Kuk, K., T. Tanino, M. Tanaka: Sensitivity analysis in parametrized convex vector optimization, J. Math. Anal. Appl. 202(1996)511-522

[23] Luc, D.T.: *Theory of Vector Optimization*, Springer-Verlag, Berlin, 1989

[24] Maeda, T.: *Multiobjective Decision Making Theory and Economic Analysis* (in Japanese), Makino-Shoten, Tokyo (1996)

[25] Naccache, P.H.: Stability in multicriteria optimization, J. Math. Anal. Optim. 68(1979)441-453

[26] Penot, J.-P.: Differentiability of relations and differentiable stability of perturbed optimization problems, SIAM J. Control Optim. 22(1984)529-551

[27] Penot, J.-P., A. Sterna-Karwat: Parametrized multicriteria optimization: Continuity and closedness of optimal multifunctions, J. Math. Anal. Appl. 120(1986)150-168

[28] Proll, L.G., D. Ríos Insua, A. Salhi: Mathematical programming and the sensitivity in multi-criteria decisions, Ann. Oper. Res. 43(1993)109-122

[29] Rarig, H.M., Y.Y. Haimes: Risk/dispersion index method, IEEE Trans. Syst. Man Cybern. 13(1983)317-328

[30] Ríos Insua, D.: *Sensitivity Analysis in Multiobjective Decision Making*, Springer-Verlag, Berlin, 1990

[31] Ríos Insua, D., S. French: A framework for sensitivity analysis in discrete multi-objective decision-making, Europ. J. Oper. Res. 54(1991)176-190

[32] Rockafellar, R.T., R.J.-B. Wets: *Variational Analysis*, Springer-Verlag, Berlin, 1998

[33] Sawaragi, Y., H. Nakayama, T. Tanino: *Theory of Multiobjective Optimization*, Academic Press, New York, 1985

[34] Shi, D.S.: Contingent derivative of the perturbation map in multiobjective optimization, J. Optim. Theory Appl. 70(1991)385-396

[35] Shi, D.S.: Sensitivity analysis in convex vector optimization, J. Optim. Theory Appl. 77(1993)145-159

[36] Steuer, R.E.: Goal programming sensitivity analysis using interval penalty weights, Math. Prog. 17(1979)16-31

[37] Shimizu, K., Y. Ishizuka, J.F. Bard: *Nondifferentiable and Two-Level Mathematical Programming*, Kluwer Academic, Amsterdam, 1997

[38] Tanino, T.: Sensitivity analysis in multiobjective optimization, J. Optim. Theory Appl. 56(1988)479-499

[39] Tanino, T.: Stability and sensitivity analysis in convex vector optimization, SIAM J. Control Optim. 26(1988)521-536

[40] Tanino, T.: Stability and sensitivity analysis in multiobjective nonlinear programming, Ann. Oper. Res. 27(1990)97-114

[41] Tanino, T., Y. Sawaragi: Stability of nondominated solutions in multicriteria decision-making, J. Optim. Theory Appl. 30(1980)229-253

[42] Todorov, M.I.: Kuratowski convergence of the efficient sets in the parametric linear vector semi-infinite optimization, Europ. J. Oper. Res. 94(1996)610-617

[43] Yu, P.L.: Cone convexity, cone extreme points, and nondominated solutions in decision problems with multiple objectives, J. Optim. Theory Appl. 14(1974)319-377

8 GOAL PROGRAMMING

Sang M. Lee

David L. Olson

CONTENTS

Abstract: A review of goal programming formulations and of current applications is presented. Some issues in the use of goal programming are discussed along with current research streams, including interactive goal programming, the use of goal programming to incorporate decision maker preference, and the use of goal programming as a tool to aid in multiple objective decision making.

T. Gal et al. (eds.), *Multicriteria Decision Making*

8.1 GOAL PROGRAMMING DEFINITIONS

Goal programming is one of the first management science approaches to consider decision problems with multiple objectives. Charnes, Cooper and Ferguson [28] introduced the concept in 1955 when analyzing a salary schedule intended to increase pay sufficiently to attract those with special skills, subject to disrupting the current pay schedule as little as possible. The infeasibility associated with meeting the targets led to the consideration of deviational variables, which could be minimized while still taking on some value reflecting the amount of underachievement with respect to each target.

Goal programming has thus been around for a considerable length of time. The ability to consider a variety of objectives is attractive in treating managerial decision problems. Managers have long been involved in the process of setting targets, especially those requiring strategic considerations. Managers also are adept at making plans that are infeasible in the sense that all objectives cannot be attained simultaneously. The process of goal programming involves decision makers identifying goals, setting target levels for these goals, and identifying the importance of these targeted goals.

In this paper, we first review goal programming formulations that have widely been used. Some of the many recent applications of goal programming are considered, categorizing them in broad functional areas. This is followed by a discussion of some of the issues involved in the use of goal programming. This leads to review of current research in the interactive use of goal programming and preference functions. We will then briefly identify a few other areas of active research, and conclude with some comments about the use of goal programming to aid multiple objective decision making.

There are two major approaches taken to goal programming: (1) minimizing a weighted function of goals, and (2) preemptive goal programming. Other mathematical forms have also been used, including minimizing the maximum deviation and fractional goal programming. Other nonlinear goal programming has also been used, but the formulations are not distinctly different than single objective forms other than the addition of deviational variables.

8.2 MINSUM Goal Programming

The minimum sum method of goal programming is essentially an application of a linear preference function to linear programming models. This is the approach presented by Charnes and Cooper [26] in the first textbook presentation of goal programming. The process involves identifying objectives, setting a target for each objective, and weighting each of the targets. When objectives are identified, the function F_i measuring attainment for each objective i is set equal to the target T_i. Because overachievement or underachievement of the target are possible, deviational variables d_i^- (for underachievement) and d_i^+ (for overachievement) are introduced:

$$F_i + d_i^- - d_i^+ = T_i . \tag{8.1}$$

For all objectives where less is better, d_i^+ is minimized, while d_i^- is allowed to take on any positive value. For objectives where more is better, the reverse is true (d_i^- is minimized while d_i^+ is allowed to take on any value). The objective function is then:

$$\text{Minimize} \sum_{i=1}^{K} \left(w_{in} d_i^- + w_{ip} d_i^+ \right), \tag{8.2}$$

where K is the total number of objectives, and w_i is the weight given to minimizing the selected deviational variable from target i. Both deviational variables from a particular constraint can be minimized, possibly with different weights.

This form of goal programming is commonly referred to as MINSUM goal programming. Those researchers who prefer the preference function approach to multiple criteria analysis typically prefer this form. It guarantees nondominated solutions as long as each weight $w_i > 0$. MINSUM goal programming is a direct application of the idea of preference functions to the field of mathematical programming.

8.3 Least Absolute Value Regression

Another application of the MINSUM form of goal programming is least absolute value regression. For data consisting of n observations c_{iY}, c_{ix1} and c_{ix2} of the dependent variable Y and independent variables x_1 and x_2 the formulation is:

$$\text{Minimize } d_1^- + d_1^+ + d_2^- + d_2^+ + \ldots + d_n^- + d_n^+ \tag{8.3}$$

$$\text{Subject to} \quad c_{1x1}\beta_1 + c_{1x2}\beta_2 + d_1^- + d_1^+ = c_{1Y}$$

$$c_{2x1}\beta_1 + c_{2x2}\beta_2 + d_2^- + d_2^+ = c_{2Y}$$

$$\ldots$$

$$c_{nx1}\beta_1 + c_{nx2}\beta_2 + d_n^- + d_n^+ = c_{nY}$$

$$\beta_1, \beta_2, d_i^-, d_i^+ \geq 0 \quad \text{for all } i,j.$$

The solution vector β_j gives the coefficients in the regression model:

(8.4) $$Y = \beta_1 x_1 + \beta_2 x_2 + \varepsilon.$$

Note that this formulation is exactly the same as MINSUM goal programming, except that in MINSUM goal programming weights are usually applied to the deviational variables in the objective function, reflecting the decision maker's preference function. Early works in least absolute value analysis include Freed and Glover [56, 57] and Dielman and Pfaffenberger [46]. Charnes, Cooper and Sueyoshi [30] investigated some technical aspects related to least absolute value regression. Least absolute value regressions are more robust than ordinary least squares regressions, in that they are less affected by outlier observations. Using goal programming for least absolute regression models also allows including additional constraints (as done by Charnes, Cooper and Ferguson, [28]).

Recent publications using goal programming as a regression-related tool include Lam, et al. [105], Lam and Choo [103] (classification); and Love and Lam [123] (combining forecasts). Lam and Moy [106] developed a goal programming model to simultaneously determine cut-off values for different classification functions. Sueyoshi [195] used the median as the basis of a regression measuring salary discrimination. Caples, Hanna and Premeaux [18] analyzed the use of goal programming as a regression tool in real estate appraisal. Golany and Yu [70] used a discriminant function obtained from goal programming to estimate the efficient frontier in data envelopment analysis.

8.4 MINMAX Goal Programming

MINMAX goal programming is very similar to MINSUM goal programming, except that the objective function is to minimize the maximum deviational variable value. This can be easily modeled by creating a new variable *Max*, which is constrained to be greater than or equal to each deviational variables to be minimized. The objective function is simply to minimize the new variable *Max*. The formulation:

(8.5) Minimize *Max*

$$\text{subject to: } \sum_{j=1}^{n} a_{ij} x_j + d_i^- - d_i^+ = b_i \quad \text{for } i = 1 \text{ to } m$$

$$Max \geq d_i^- \quad \text{for } i = \text{negative deviations to be minimized}$$

$$Max \geq d_i^+ \quad \text{for } i = \text{positive deviations to be minimized}$$

$$x_j, d_i^-, d_i^+ \geq 0 \text{ for all } i, j.$$

Relative to human decision making, the MINMAX idea is seen primarily in a game theoretic context. Yang [218] demonstrated that a bilateral monopoly model was a special case of a generalized goal programming model with conflicting objectives between two monopolists. Ogryczak [150] presented algorithms for lexicographic minmax solution in the context of locating public facilities to minimize the travel distances among service recipients. While not presented as goal programming per se, the lexicographic solution method would adapt to preemptive goal programming. Ogryczak found that the standard minmax approach could violate desired Pareto optimality and equity characteristics, while the lexicographic method yielded solutions satisfying these requirements. Malczewski and Ogryczak [125] compared utility function based methods with goal programming approaches in multiple criteria location problems, noting advantages and disadvantages for each. A framework for an interactive decision support system was suggested.

8.5 Preemptive Goal Programming

The preemptive version of goal programming was presented by Lee [110] and Ignizio [80, 82]. Preemptive goal programming involves a slightly different modeling process.

1. **Identify objectives**
2. **Set targets for objectives (multiple targets per objective allowed)**
3. **Prioritize these objective-target pairs**
   At a particular priority level, more than one target can be considered. If they are not measurable in common terms (commensurate), tradeoff weights need to be identified for use within a priority level.
4. **Solve a sequence of linear programming models by priority level**

The formulation is:

$$(8.6)\quad \text{Minimize } P_1 \sum_i^m \left(w_{1in} d_i^- + w_{1ip} d_i^+\right); \ldots P_k \sum_i^m \left(w_{2in} d_i^- + w_{2ip} d_i^+\right)$$

$$\text{subject to: } \sum_j^n a_{ij} x_j + d_i^- - d_i^+ = b_i \quad \text{for i} = 1 \text{ to m}$$

$$x_j, d_i^-, d_i^+ \geq 0 \text{ for all } i, j$$

$$P_1 >>> P_2 >>> \ldots >>> P_k,$$

where k is the number of preemptive levels, m is the number of constraints (including goal constraints), and n is the number of decision variables x_j. This model is solved in effect through a series of linear programs. The first LP is to minimize the weighted set of deviational variables at P_1. If all deviational variables

to be minimized at this priority level are satisfied, those deviational variables are eliminated from the model and the LP minimizing the weighted set of deviational variables at P_2 is considered. Attainment of first priority goals is maintained because deviational variables allowing violation of priority one targets are not allowed to take on values. This process continues until target deviational variables at the specific priority level are not minimized at zero value. If there are not multiple solutions minimizing nonattainment, the solution to the model is identified. If there are multiple solutions minimizing the weighted function of target deviational variables, these deviational variables are constrained to be no greater than the current level, and the next priority is considered until there is a unique solution. If all goals in a model are satisfied, the solution is likely to be dominated. It is trivial to avoid this by ensuring that goals that are not capable of complete satisfaction are included in the model.

While this formulation allows weights, weights are usually not used in preemptive goal programming. As Martel and Aouni [128] pointed out, weight aggregation can be difficult for human decision makers. Barnett, et al. [6] used multidimensional scaling to identify a cardinal utility function of weights, arguing that other methods were difficult to implement in an accurate manner.

8.6 Fractional Goal Programming

Fractional goal programming involves targets for ratios. Charnes and Cooper [27] introduced this form of goal programming, and Kornbluth and Steuer [100] extended the topic. A common application is where a manager would like to adopt those policies that best meet target performance ratios such as return on investment, or some other of the many financial ratios used to evaluate company performance by stockholders. Goedhart and Spronk [66] gave one financial application of fractional goal programming. A formulation for a fractional goal programming model is:

$$(8.7) \qquad \text{Minimize} \sum_{i=1}^{m} \left(w_{1in} d_i^- + w_{1ip} d_i^+ \right)$$

$$\text{subject to: } \frac{f_{1i}(x_j)}{f_{2i}(x_j)} + d_i^- - d_i^+ = r_i \quad \text{for } i = 1 \text{ to } m$$

$$x_i, d_i^-, d_i^+ \geq 0 \quad \text{for all } i, j.$$

Sometimes it is possible to multiply through by the function of x_j in the denominator prior to adding the deviational variables, in which case this model becomes linear. However, interpretation of the deviations is complicated, and assigning weights, always a difficult step, is made much more indirect. Therefore, there are good reasons for treating this formulation as a nonlinear programming problem.

Goedhart and Spronk [67] presented an interactive solution approach for fractional goal programming models as part of an interactive, heuristic planning procedure for decentralized organizational planning and control. Despotis [43] used fractional goal programming as a tool to identify a minmax solution with the purpose of assessing utility. Ohta and Yamaguchi [151] addressed fuzzy fractional goal programming as a way of more accurately reflecting decision making considering natural goals of performance ratios, and dealing with vague data.

Ellis [49] and Ellis and Bowman [50] applied fractional functions to reflect target ratios of pollution emission reduction in an air pollution control model. These fractional functions were included to consider equity considerations across regions.

8.7 Nonlinear Goal Programming

Saber and Ravindran [176] gave a thorough review of nonlinear goal programming, including solution methodologies as well as applications. There have been a number of goal programming models presented that involve nonlinearities. These nonlinearities can arise due to nonlinear relationships between variables such as are found in engineering design [51, 101, 154], or sometimes due to economic formulations [143] or probability calculation [207]. Another source of nonlinearity comes from stochastic models, where technological coefficients can be probabilistic [47, 119, 152, 216], resulting in chance constrained goal programming models.

Solution methodologies for the most part draw upon the general set of tools available for nonlinear programming, just as in the case of linear programming. The exceptions are preemptive goal programming simplex [85, 109] and the Arthur and Ravindran algorithm [3], which utilize the specific features of preemptive goal programming. In nonlinear goal programming, four basic strategies were identified: simplex based, direct search, gradient based, and interactive.

Of the simplex based approaches, Ignizio [80] modified a method of approximation programming for goal programming. Taylor series expansion is used to linearize nonlinear functions in the vicinity of a given feasible point, and the resulting LP model solved. Separable programming can be used to solve models with nonlinearity due to chance constraints. Quadratic programming can also be modified for goal programming models.

Direct search has been applied by Nanda, et al. [143] to a nonlinear energy goal programming problem. Ignizio [80] used the Hooke and Jeeves approach for nonlinear goal programming, followed by Hwang and Masud [79]. This method was used for solution of a nonlinear quality control problem [164].

Gradient based methods using feasible directions with optimal step length determination have been adopted for nonlinear goal programming [120]. That algorithm was applied to chance constrained goal programming models. The gradient approach requires differentiable nonlinear functions, but can be significantly faster than other nonlinear methods (although for some problems it can be significantly slower). Clayton, et al. [36] developed a systematic search

procedure to optimize a multiresponse simulation model structured as a preemptive goal programming model. Modified pattern search and gradient search techniques were combined in this approach, which was used by Taylor, et al. [207, 208] to solve nonlinear integer goal programming models. Roljic and Dujsic [169] have developed a gradient nonlinear goal programming algorithm using the feasible directions method with optimal step length for nonlinear goal programming models arising from Cobb-Douglas production functions.

Interactive approaches operate through decision maker revision of preference information. Weistroffer [217] developed a method converting a nonlinear multiple objective problem into a sequence of unconstrained single objective subproblems to minimize squared deviation from targets and hard constraints. Masud and Hwang [129] proposed a nonlinear interactive goal programming method where no priority levels or weights were used initially. Upper and lower bound values for each objective were generated through a learning process on the part of the decision maker.

Other applications of nonlinear goal programming have been presented by Lewis and Taha [122] who addressed response functions. Chunhachinda, et al. [34] dealt with Markowitz portfolio theory. They studied fourteen international stock markets over the period January 1988 through December 1993. Returns were found to be stable, but not normally distributed. Polynomial goal programming incorporating investor preferences for skewness was used to determine an optimal portfolio. These researchers noted that investors trade expected return for skewness.

8.8 Application of Goal Programming

Goal programming is probably the most widely used multiple objective mathematical programming technique. There have been many applications, some reviewed and classified by Zanakis and Gupta [224], Romero [170], Schniederjans [179, 180], and Tamiz, et al. [205]. A number of agricultural goal programming applications were included in Romero and Rehman [174].

While we do not intend a complete and thorough review of goal programming applications (refer to the references in the prior paragraphs for such reviews), we would like to point out the variety of problem environments in which goal programming has been applied across the spectrum of mathematical programming problems where multiple objectives are present. The applications we consider are primarily very recently published articles. Schniederjans [181] noted almost one thousand goal programming articles through 1994, arguing that the product life cycle applied and that he noted a decline in such articles in recent times. However, a quick review of journals since that time does not indicate the slowdown in interest in goal programming expected by Schniederjans.

8.8.1 Engineering

In the engineering field, goal programming models have appeared in many applications. Civil engineering offers many opportunities, because of many public policy decisions involving complex sets of objectives in issues such as water resources management and other public works administration problems. Gu and Tang [190] presented a MINSUM goal programming network model to schedule bus trips in heavily congested cities. Goal programming models have also been applied to problems in other engineering disciplines. Shiau, et al. [187] applied goal programming in a mechanical engineering application. Chang, Chen and Chen [21] used fuzzy MINSUM goal programming in a cost estimation model. Ehie and Benjamin developed a model to plan Zambian copper mining [48].

A common decision problem involves selection. Ravirala and Grivas [165] applied MINSUM goal programming to select pavement and bridge projects in an annual operation involving goals including balance of maintenance and rehabilitation projects, geographic dispersion of funding, safety improvement, and minimization of cost and delay. Qiu [159] used a genetic algorithm to solve a zero-one project ranking problem. Taplin, et al. [206] also applied goal programming to road project selection, and Al-Faraj et al. to selection of traffic centers in Saudi Arabia [1]. Mukherjee and Bera [140] used MINSUM goal programming for selection of alternative approaches in the coal industry, while Chen presented a 0-1 model for scheduling maintenance projects [32]. The line between engineering and business disappears in selection problems. The same basic problem in a non-engineering context was tackled with preemptive forms of goal programming in models for selection of a brewery site [75], management information system projects [177, 178, 186], and selection of a house [184].

Computer science is a field with contacts in both the engineering and management environments. Goal programming has been applied to the computer field in a number of application areas (hard disk production [219, 220], robotic control [202], and technology selection [189]).

8.8.2 Operations Management

Goal programming is widely used in operations management. It has been applied to such diverse decisions as selection of flexible manufacturing products [117, 142], line balancing [68], reliability [60], warranty analysis [136], and quality control [185]. In a production planning application, goal programming was used to resolve infeasibility to repetitively develop a production plan by minimizing deviations below target product levels [89]. Project scheduling applications are widespread [44, 62, 157] as are worker scheduling applications [16, 61]. Yura [222] gave a model to seek satisfaction of worker preferences under due-date constraints. Lee and Luebbe presented a zero-one approach for warehouse location [118]. Nudtasomboon and Randhawa [149] modeled a resource-constrained project scheduling problem with a zero-one preemptive goal programming model. Lashine, et al. [107] modeled a machine shop scheduling model with mixed integer, nonlinear goal programming. Min [133]

presented a chance-constrained goal programming model for selection of transportation mode in international logistics. Goal programming models have been constructed to aid management of perishable inventory [93, 94], the master production scheduling problem [212, 213] and for just-in-time analysis [58, 98, 116, 112, 215].

8.8.3 Business

The applications of goal programming to business related decisions overlaps that of engineering to a large degree. There are shared interests in computer science related disciplines, as well as operations management. But there have been applications specific to business decision making as well. Lee and Shim [121] addressed zero-based budgeting with preemptive goal programming. Schniederjans and Garvin [182] applied a preemptive goal programming model to activity-based accounting. In the marketing field, Golany, et al. [69] dealt with a problem of selecting marketing panels considering the conflicting goals of cost and accuracy, while Kwak, et al. [102] presented a goal programming model to aid assignment of marketing distribution patterns. In the organizational management setting, there have been articles on force planning [59, 97], employee scheduling [47, 96, 162], and many other managerial decision problems. MINSUM goal programming was found to yield telecommunications pricing more reflective of fairness and efficiency, resulting in lower prices to consumers [14]. A preemptive model was presented to structure municipal bonds with the intent of providing an interactive system to minimize risk while assuring adequate return [158].

The field of finance has seen an especially rich number of goal programming applications. Portfolio analysis has been the subject of articles, including a MINSUM polynomial goal programming model [34]. Cooper, et al. [39] used dual variables from a MINSUM goal programming model measuring deviations from risk and return targets. Michnik and Trzaskalik [132] presented a model for bank portfolio management. Fund management involves a variety of goals, and goal programming models have been developed to support fund allocation [10, 64, 86, 95]. Global financial planning has been addressed [114, [115, 113]. Al-Saffar and Osman [2] analyzed portfolio investments in the Egyptian insurance industry. Schniederjans and Hoffman [183] presented a zero-one goal programming model for acquisition of multinational firms. A goal programming model to budget water resource investment was published in a joint engineering/business application [200]. Least absolute value regression was applied to a divestiture problem in Japan [196] and to the problem of appraising real estate [18].

The application of goal programming to portfolio management has been especially active [11, 111]. Feinstein and Thapa [53] developed a MINSUM goal programming model allowing interpretation of dual prices to aid in analysis of this class of problem.

8.8.4 Agriculture

There has been a long tradition of goal programming in agricultural studies [174]. Applications range from animal husbandry [84], forestry [210, 211], and fisheries [141]. Fiske, et al. [54] developed a zero-one goal programming model to simultaneously consider profitability, risk, and environmental factors in forage-resource production systems for cattle in West Virginia. Van Berlo and Van Berlo [209] developed a goal programming model to ensure consistency in a model for the vegetable processing industry considering the sectors of market, industry, and agriculture. Broad [13] applied goal programming to explore efficient means of production of milk concentrate. Damij [41] developed an interactive goal programming decision support system (DSS) to generate an animal diet with specified nutritional balance. Johnson, et al. [87] developed an internal parasite control model to fight animal disease. Brown, et al. [15] used a goal programming model to estimate the number of insect pests from crops. Their model combined goal programming with simulation to interactively evaluate alternative policies. Berbel, et al. [7] developed an interactive decision support system for a large horticultural company in Spain for annual crop planning, focusing on the tradeoffs between gross margin and marketing goals. The system combined goal programming with another form of multiobjective programming. Zamora and Berbel [223] used preemptive goal programming for wildlife management in Spain. Rossing, et al. [175] applied interactive goal programming to flower bulb production planning. Sumpsi, et al. [197] presented a formulation to convert preemptive goal structure into a MINSUM formulation in an agricultural setting.

In the field of forestry, Bernetti [9] reviewed goal programming and linear programming models for forest planning, noting the flexibility of goal programming models. Kangas and Pukkala [92] formulated a goal programming model for forest management planning in Finland. Suter and Calloway [201] developed a DSS incorporating goal programming for rough mill planning in the lumber industry. Kolenka [99] formulated a goal programming model for forest management in Slovakia. Buongiorno, et al. [17] developed goal programming models for improvement of forests in the Jura Mountains.

Goal programming models for water use planning in agriculture have been applied. McGregor and Dent [131] used preemptive goal programming (as well as another multiple objective programming model) to resolve water conflicts on the Rakaia River in New Zealand. Peralta, et al. [153] used a goal programming-based DSS for identifying water supply strategies while achieving desired groundwater quality goals in Arkansas.

There has been heavy use of goal programming for both macroeconomic and microeconomic analysis. At the macro level, goal programming has been applied to land-use planning in Egypt [52] and Canada [221]. He, et al. presented three goal programming models for 11 varieties of vegetables in China [74] that was estimated to have the potential of saving millions of Chinese dollars. Holden [76] analyzed fertilizer subsidy policies using a goal programming model in Zambia. Njiti and Sharpe [146] developed a goal programming model to resolve competition and conflicts among land uses in the Cameroon. Zekri and Romero

[225] developed two goal programming models to evaluate crop policies in Spain considering water and energy use. Sutardi, Bector and Goulter [198] developed an integer goal programming model to consider economic efficiency along with agricultural development policies in Indonesia.

Microeconomic analysis support from goal programming models of agricultural operations have also been common. Cornett and Williams [40] used goal programming for multiple land use planning considering outdoor recreation, timber, livestock production, and deer management in southern California. DeKoeijer, et al. [42] applied goal programming to gain insight into the tradeoff between income and pollution in Dutch agriculture. Ciuchi and Pennacchi [35] used a weighted goal programming model for evaluating alternative farm management policies in Italy. Coetzee, et al. [37] considered net farmer income, minimum lending, and minimum area planted in South Africa, finding that both the goal programming and the other multiple objective programming models yielded plans less risky than the linear programming generated plan. Nino de Zepeda, et al. [145] used goal programming to compare soil erosion policies in Chile.

8.8.5 Public Policy

Perhaps one should expect the greatest number of goal programming applications to be in the field of policy decision making. It is in this field that the most attention is usually given to multiple objectives, with the need to match diverse targets as much as possible.

Waste management is a critical societal issue involving difficult tradeoffs. Many models have been presented to aid waste management throughout the world [19, 24, 77, 138, 160, 194]. A number of fuzzy goal programming models have been proposed to deal with the sometimes uncertain relationships between actions and response [20, 108]. More advanced nonlinear [22] and stochastic [25] goal programming models have recently been proposed for such issues as water planning and other waste management related topics. Chang and Wang [23] applied fuzzy goal programming to reflect uncertainties in both priority and scale in economic and environmental goals.

Energy management continues to be an important public policy question [12, 90, 124, 161]. Kalu [91] developed a model of individual oil firms considering the supply side of energy, using the model to examine the effects of crude oil price changes. Goal programming models have also been applied to the problem of air pollution (see [38] for a comprehensive review, including goal programming) and radiation (see [33], where a genetic algorithm was implemented).

Goal programming models have been applied to a variety of other interesting public policy questions. These include mental health services planning [191], a DEA study of the European Union [5], wildlife management [8], ambulance allocation [227], and school districting [148]. Yin, et al. [221] applied a multisector goal programming model for land conversion impact assessment in British Columbia.

8.9 Issues in the Use of Goal Programming

Romero [171] addressed some of the issues involved across various forms of goal programming models. There have been three primary criticisms of preemptive goal programming. The first is that it might generate dominated solutions. The second is that when you move between priority levels, there is an infinite tradeoff in value implied. Third, some question the ability of decision makers to set target levels.

8.9.1 Avoiding Dominated Solutions in Goal Programming

Zeleny [226] recommended using MINSUM goal programming rather than preemptive goal programming to avoid the possibility of recommending dominated solutions. There is a feeling on the part of those who feel strongly that preemptive goal programming is not an appropriate modeling method that avoiding nondominated solutions is an extremely urgent characteristic. In response to these arguments, Hannan [71] provided formulations to ensure nondominated solutions when either MINSUM or preemptive versions of goal programming.

Ignizio [81] suggested the obvious approach of setting more ambitious targets, although this would require an interactive process. Hannan [73] proposed sensitivity analysis to identify dominated solutions, and then correction by adjusting the model.

8.9.2 Tradeoffs Across Preemptive Priority Levels

A second major criticism of preemptive goal programming is the infinite tradeoff between preemptive priority levels. This violates the principles of normative preference functions, in that it bars a small sacrifice in the more important objective for a possibly major gain in the objective at the lower priority (Stewart [193]). We agree that this is a weak point in the method, and that goal programming sensitivity analysis should focus on this tradeoff. This ties in with a related weakness of goal programming in general, the setting of target levels. However, we feel that interactive methods make setting targets no more problematic than setting weights across incommensurate objectives.

Analysis of the sensitivity of goal programming models has been approached in a variety of ways. Ignizio [80] presented a parametric approach to check the sensitivity of models to changes in parameters. Markowski and Ignizio [126, 127] analyzed duality in the context of preemptive goal programming models. Goedhart [65] used shadow prices on goal targets and resource budgets. In MINSUM goal programming, weights on different objectives as objective function coefficients result in composite objectives that are quite abstract. Preckel, et al.

[156] suggested a procedure for analysis of objective function coefficients in this context. McCarl, et al. [130] analyzed goal programming shadow prices in depth.

8.9.3 Setting Targets a priori

Preemptive goal programming has also been criticized for requiring setting targets *a priori* (although it was never a requirement to our knowledge that goal programming could not be used interactively). Min and Storbeck [135] argued that the philosophy of goal programming is more descriptive than normative. It is not based on the same philosophy as those methods focusing on preference functions. The concept of preference requires definition of the objectives to be considered prior to the analysis. And yet the most promising use of computer models to aid decision models is to enhance decision maker learning. Learning implies that the measures of importance distinguishing between alternatives could be many, and those that are important may well depend on the context of the problem. Therefore, it seems hollow to worry about dominance when not all objectives may yet be included in the model. Romero [172] debated the merits of this argument with Min [134].

8.10 Interactive Goal Programming

Interactive goal programming is a natural implementation of aiding decision makers identify solutions matching their preferences [188]. Adjusting targets for objectives in light of initial model results is an example of applying a mathematical programming model as a decision support system. Piech and Rehman [155] compared MINSUM goal programming with two other multicriteria methods, and concluded that while goal programming was easier to solve, it was inferior to the other methods with respect to the information provided to the decision maker because it gave only one solution. They did not consider the possibility of interactive use of goal programming. Piech and Rehman also noted that selecting weights for deviational variables was a problem. They overcame the difficulty of setting targets by using the individual optimal attainment levels for each objective.

The preemptive form of goal programming also can be used interactively, as demonstrated by Franz and Lee [55]. In this approach, there is less of a tie to the preference function view of decision making. Mote, et al. [139] demonstrated the use of multiple targets per objective in preemptive goal programming as a means to reflect the decision maker's preference. This allows tradeoffs among objectives to be expressed through the ordinal ranking of different target levels across objectives. Furthermore, interactive use of preemptive goal programming allows the decision maker to see the impact of various sets of priority rankings. The preemptive goal programming approach is viewed by some (including the authors of this paper) as being more appropriate for learning tools, in the spirit of decision support. This is best accomplished in an interactive setting.

Spronk and Matarazzo [192] developed a modified interactive goal programming model custom designed for financial planning. The model considered various means of reflecting the probabilistic aspects of an agricultural environment. Expected cash flow was corrected to reflect unexpected variations in factors influencing it.

Reeves and Hedin [166] gave the following process for interactive implementation of goal programming. Nondominated solutions were guaranteed by using MINSUM forms of goal programming. Alternative solutions were obtained by adding constraints to assure that targets for each specific objective was attained for that subproblem. If none of the solutions obtained were satisfactory to the decision maker, the process involved setting new target values, and regenerating solutions:

1. **Specify initial goal levels**
2. **Generate ideal solutions**
   - primary initial solution
   - alternatives by constraining each objective in turn
3. **If solution satisfactory, stop**
4. **Revise goal levels**
5. **Regenerate primary, alternative solutions**

Tamiz and Jones [204] gave a similar, more general interactive goal programming process:

1. **Generate initial feasible solution**
2. **Decision maker review of initial solution**
3. **If solution satisfactory, stop**
4. **Obtain additional decision maker preference information**
5. **Reformulate goal programming model**
6. **Resolve model and return to decision maker review**

Tamiz and Jones went on to discuss design issues, issues involving communication with the decision maker, and issues in parameter alteration. Goedhart and Spronk [67] extended the idea to match decentralized organizations. Their interactive planning process was:

1. **Central management set resources and goals**

   Divisional management optimizes for their division

2. **Central management incorporates divisional proposals**

   - selects intermediate solutions

   - imposes targets and budgets

   Divisional management selects projects

   - within targets and budgets

3. **Central management selects compromise solution**

   - sets final budgets and targets

   Divisional management makes final activity selection

This process was intended as an interactive framework where central management and divisional management would iterate within each of the three steps. Goedhart and Spronk called for more research on the implementation of this process.

8.11 Goal Programming Modeling of Decision Maker Preference

One of the most interesting approaches to modeling decision maker preference within goal programming models was one of the first, goal interval programming [29]. A step function of preference for different objectives was modeled. The model was not particularly tractible, however.

The MINSUM form of goal programming requires preference weights be identified, reflecting both objective importance, as well as measurement scale. This

is the approach preferred by those who follow the normative view of decision analysis, where a preference function is assumed to exist, and once it is measured, the MINSUM goal programming model will yield a nondominated solution with respect to those objectives included in the model. Romero [173] argued that it makes no sense to include multiple targets for any one objective, stating that those targets that did not affect the final solution were redundant. However, low weights may not turn out to influence solutions either. Minimization of cost may be an objective of a manager. That manager may give a much heavier weight to avoiding bad publicity, seeking to maintain the firm's image, which in MINSUM goal programming may jump to an extreme point of the model that involves high cost. That does not mean that cost is not a factor of consideration in the manager's preference function. On the other hand, there are contexts in which even small weights may make a difference in the final solution. Similarly, there are contexts where less important priority targets may make a difference in preemptive goal programming. Further note that in a normative context, multiple targets for the same objective can be used to reflect relative emphasis of strongly concave preference models.

The identification of weights in MINSUM goal programming has been a difficult task. Martel and Aouni [128] applied a form of fuzzy goal programming to incorporate PROMETHEE preference functions into a mathematical programming model. This was done in an attempt to answer the criticism that *a priori* aspiration levels are difficult to set, and that weight aggregation was difficult because of incommensurability. The PROMETHEE preference functions were presented as a means to flexibly assess decision maker preference (six forms of utility functions were available). However, in their simple example involving two decision variables and four goal constraints and thus 8 deviational variables (with 11 feasible corner points), they ended up with a model involving 27 constraints, 11 continuous variables, 19 0-1 variables, and 2 integer variables that took 270 iterations to solve. Tamiz and Jones [203] improved this model to the point that they had 15 constraints, 12 continuous variables, 6 0-1 variables, 2 integer variables, that took only 56 iterations to solve. The idea of incorporating utility functions to represent decision maker preference appears to require substantial additional work. Jones and Tamiz [88] and Diaby and Martel [45] have presented additional work in this direction. Lam and Choo [104] presented other work applying goal programming to preference decomposition.

8.12 Other Uses of Goal Programming as a Tool

In addition to its use as a tool to support decision making, goal programming has been used in a variety of other ways. Active research into the use of goal programming as a mathematical tool continues in a variety of fields, including least absolute value regression, data envelopment analysis, and fuzzy set analysis.

The use of goal programming in data envelopment analysis has also been quite heavy in recent years. Athanassopoulos [4] developed an interface between goal programming and data envelopment analysis to integrate target setting and resource allocation in planning. When implemented in Greek local governmental planning, the method was shown to incorporate decision maker preferences while selecting efficient technologies. Athanassopoulos [5] proposed a goal programming formulation for assessing production functions within the European Union. Charnes, Gallegos and Li [31] applied a multiplicative data envelopment analysis model to the Latin American airline industry. The method included a goal programming model to estimate the coefficients to use in the DEA model. Huang and Li [78] also used goal programming to derive linear deterministic equivalents for a chance constrained formulation of data envelopment analysis. Retzlaff-Roberts and Morey [167] used goal programming in a stochastic two-stage DEA model because of alternate optimal solutions identifying significantly inefficient units. Giokas [63] compared four methods of analysis to measure the efficiency of hospitals. The method using goal programming in conjunction with data envelopment analysis was found to be more reliable than the other methods for the data analyzed.

Goal programming has been widely used in fuzzy set analysis. Early formulations were presented by Narasimhan [144] and Hannan [72]. Sutardi, Bector and Goulter [199] used fuzzy integer goal programming as a framework for sequential budgetary decision making in water investment planning. This model determined optimal return for each combination of funding levels while reflecting uncertain socio-technical and political factors. Mohammed [137] used deviational variables to transform a linear fuzzy programming model into a crisp formulation. Wang and Fu [214] presented a preemptive fuzzy goal programming model to reflect decision maker preference in an uncertain environment. Inuiguchi, et al. [83] applied possibilistic programming as a solution approach for a production planning problem containing ambiguous data and vague decision maker aspirations. As noted in the public policy application section, there have been many fuzzy goal programming applications applied to waste management and water quality management [108], forest management [147], and for other applications [163].

8.13 Conclusions

Goal programming has proven to be very useful in many contexts. MINSUM goal programming provides a simple additive model of preference, complete with the requirement to identify weights describing tradeoffs among objectives. Preemptive goal programming provides a workable means for decision makers to learn about the implications of targets, given that interactive approaches are used. Goal programming has proven valuable as a means of implementing least absolute value regression, in data envelopment analysis, in stochastic modeling, and in many other applications. It continues to be applied to a wide variety of decision problems involving multiple, conflicting goals.

Many studies have compared goal programming with other methods. We do not attempt to enumerate these studies, many of which focus on other multiple objective programming methods found in this volume. We would encourage everyone to look at these alternative methods. The following chapter, presenting reference point approaches, combines the goal programming idea of target levels with approaches to assure nondominated solutions (although as noted earlier in this chapter, we contend that this can be accomplished within goal programming as well). Ringuest [168] gave his view of the implications of goal programming to decision maker behavior, and presented a general goal programming model.

Preemptive goal programming involves some controversy. It appears to us to be an excellent tool for decision maker learning. Further research into the interactive use of preemptive goal programming should enhance this value. On the other hand, if one believes strongly in normative preference functions, he or she should stick to MINSUM goal programming. Further research into preference function identification also holds promise.

REFERENCES

[1] Al-Faraj, T. N., A. S. Alidi, A. A. Al-Ibrahim: A planning model for determining the optimal location and size of traffic centers. European Journal of Operational Research 66:3 (1993) 272-278.

[2] Al-Saffar, A. A., M. A. A. Osman: A goal programming model for investment allocation in the Egyptian general insurance market. Middle East Business and Economic Review 8:2 (1996) 62-74.

[3] Arthur, J., A. Ravindran: PAGP: An efficient algorithm for linear goal programming problems. ACM Transactions of Mathematical Software 6 (1980) 378-386.

[4] Athanassopoulos, A. D.: Goal programming & data envelopment analysis (GoDEA) for target-based multi-level planning: Allocating central grants to the Greek local authorities. European Journal of Operational Research 87 (1995) 535-550.

[5] Athanassopoulos, A. D.: Assessing the comparative spatial disadvantage (CSD) of regions in the European Union using non-radial data envelopment analysis methods. European Journal of Operational Research 94 (1996) 439-452.

[6] Barnett, D., B. Blake, B. A. McCarl: Goal programming via multidimensional scaling applied to Senegalese subsistence farms. American Journal of Agricultural Economics 64:4 (1982) 720-727.

[7] Berbel, J., J. Gallego, H. Sagues: Marketing goals vs. business profitability: An interactive multiple criteria decision making approach. Agribusiness 7:6 (1991) 537-549.

[8] Berbel, J., R. Zamora: An application of MOP and GP to wildlife management. Journal of Environmental Management 44:1 (1995) 29-38.

[9] Bernetti, I.: Multiobjective modelling for forest management for timber production. Studi di Economia e Diritto 39:3 (1990) 523-552.
[10] Bessler, W., G. G. Booth: An interest rate risk management model for commercial banks. European Journal of Operational Research 74:2 (1994) 243-256.
[11] Booth, C. G., W. Bessler: Goal programming models for managing interest rate risk. *Omega* 17 (1989) 81-89.
[12] Bose, R. K., G. Anandalingam: Sustainable urban energy-environment management with multiple objectives. Energy 21:4 (1996) 305-318.
[13] Broad, L. R.: Spray drier optimisation using goal programming. Irish Journal of Food Science and Technology 15:1-2 (1991) 1-2, 17-25.
[14] Brown, K., R. Norgaard: Modelling the telecommunication pricing decision. Decision Sciences 23:3 (1992) 673-686.
[15] Brown, L. G., R. W. McClendon, K. S. Akbay: Goal programming for estimating initial influx of insect pests. Agricultural Systems 34:4 (1990) 337-348.
[16] Brusco, M.J., T. R. Johns: Improving the dispersion of surplus labor in personnel scheduling solutions. Computers & Industrial Engineering 28:4 (1995) 745-754.
[17] Buongiorno, J., J. L. Peyron, F. Houllier, M. Bruciamacchie: Growth and management of mixed-species, uneven-aged forests in the French Jura: Implications for economic returns and tree diversity. Forest Science 41:3 (1995) 397-429.
[18] Caples, S. C., M. E. Hanna, S. R. Premeaux: Least squares versus least absolute value in real estate appraisals. The Appraisal Journal 65:1 (1997) 18-24.
[19] Chang, C. T., J. R. Hwang: A multiobjective programming approach to waste minimization in the utility systems of chemical processes. Chemical Engineering Science 51:16 (1996) 3951-3965.
[20] Chang, N. B., Y. L. Chen: Optimal balance between recycling and waste to energy programs by grey fuzzy nonlinear programming. Journal of Environmental Science & Health A32:3 (1997) 633-647.
[21] Chang, N. B., Y. L. Chen, H. W. Chen: A fuzzy regression analysis for the construction cost estimation of wastewater treatment plants. 1. Theoretical development. Journal of Environmental Science and Health A32:4 (1997) 885-899.
[22] Chang, N. B., S. F. Wang: Managerial fuzzy optimal planning for solid-waste management systems. Journal of Environmental Engineering – ASCE 122:7 (1996) 649-658.
[23] Chang, N. B., S. F. Wang: A fuzzy goal programming approach for the optimal planning of metropolitan solid waste management systems. European Journal of Operational Research 99:2 (1997) 303-321.
[24] Chang, N. B., S. F. Wang: Integrated analysis of recycling and incineration programs by goal programming techniques. Waste Management & Research 15:2 (1997) 121-136.

[25] Changchit, C., M. P. Terrell: A multiobjective reservoir operation model with stochastic inflows. Computers & Industrial Engineering 24:2 (1993) 303-313.

[26] Charnes, A., W. W. Cooper: Management Models and Industrial Applications of Linear Programming. Wiley, New York, 1961.

[27] Charnes, A., W. W. Cooper: Programming with linear fractional functionals. Naval Research Logistics Quarterly 9 (1962) 181-186.

[28] Charnes, A., W. W. Cooper, R. O. Ferguson: Optimal estimation of executive compensation by linear programming. Management Science 1:2 (1955) 138-151.

[29] Charnes, A., W. W. Cooper, K. R. Karwan, W. A. Wallace: A goal interval programming model for resource allocation in a marine environmental protection program. Journal of Environmental Economy and Management 3:4 (1976) 347-362.

[30] Charnes, A., W. W. Cooper, T. Sueyoshi: Least squares/ridge regression and goal programming/constraint regression analysis. European Journal of Operational Research 27 (1986) 146-157.

[31] Charnes, A., A. Gallegos, H. Li: Robustly efficient parametric frontiers via multiplicative DEA for domestic and international operations of the Latin American airline industry. European Journal of Operational Research 88 (1996) 525-536.

[32] Chen, V.: A 0-1 goal programming model for scheduling multiple maintenance projects at a copper mine. European Journal of Operational Research 76:1 (1994) 176-191.

[33] Chen, Y., M. Narita, M. Tsuji, S. Sa: A genetic algorithm approach to optimization for the radiological worker allocation problem. Health Physics 70:2 (1996) 180-186.

[34] Chunhachinda, P., K. Dandapani, S. Hamid, A. J. Prakash: Portfolio selection and skewness: Evidence from international stock markets. Journal of Banking and Finance 21 (1997) 143-167.

[35] Ciuchi, P., F. Pennacchi: A farm planning problem with multiple objectives. Rivista di Economia Agraria 45:2 (1990) 295-332.

[36] Clayton, E., W. Weber, B. Taylor, III: A goal programming approach to the optimization of multiresponse simulation models. IIE Transactions 14 (1982) 1149-1158.

[37] Coetzee, K., J. Heckroodt, J. J. Janse van Rensburg, M. F. Viljoen: Multi-criteria programming models: Suitable planning strategies for entrepreneurs in a changing agricultural environment. Agrekon 32:4 (1993) 217-222.

[38] Cooper, W. W., H. Hemphill, Z. Huang, S. Li, V. Lelas, D. W. Sullivan: Survey of mathematical programming models in air pollution management. European Journal of Operational Research 96:1 (1997) 1-34.

[39] Cooper, W. W., V. Lelas, T. Sueyoshi: Goal programming models and their duality relations for use in evaluating security portfolio and regression relations. European Journal of Operational Research 98:2 (1997) 431-443.

[40] Cornett, D., W. A. Williams: Goal programming for multiple land use planning at Mineral King, California. Journal of Soil and Water Conservation 46:5 (1991) 373-376.

[41] Damij, T.: A nutrition optimization and planning system. Zbornik Biotechniske Fakultete Universe v. Ljubljani #63 (1994) 257-268.

[42] deKoeijer, T.J., J. A. Renkema, J. J. M. van Mensvoort: Environmental economic analysis of mixed crop-livestock farming. Agricultural Systems 48:4 (1995) 515-530.

[43] Despotis, D. K.: Fractional minmax goal programming: A unified approach to priority estimation and preference analysis in MCDM. Journal of the Operational Research Society 47 (1996) 989-999.

[44] Dey, P. K., M. T. Tabucanon, S. O. Ogunlana: Hierarchical approach to project planning: The case of a petroleum pipeline construction. Applied Mathematical Modelling 20:9 (1996) 683-698.

[45] Diaby, M., J.-M. Martel: Preference structure modelling for multi-objective decision making: A goal-programming approach. Journal of Multi-Criteria Decision Analysis 6 (1997) 150-154.

[46] Dielman, T., R. Pfaffenberger: LAV (least absolute value) estimation in the regression model: A review. TIMS Studies in the Management Sciences 19 (1982) 31-52.

[47] Easton, F. F., D. F. Rossin: A stochastic goal program for employee scheduling. Decision Sciences 27:3 (1996) 541-563.

[48] Ehie, I.C., C. O. Benjamin: An integrated multiobjective planning model: A case study of the Zambian copper mining industry. European Journal of Operational Research 68:2 (1993) 160-172.

[49] Ellis, J. H.: Multiobjective mathematical programming models for acid rain control. European Journal of Operational Research 35 (1988) 365-377.

[50] Ellis, J. H., M. L. Bowman: Critical loads and development of acid rain control options. ASCE Journal of Environmental Engineering 120:2 (1994) 273-290.

[51] El-Sayed, M., B. Ridgely, E. Sandgren: Nonlinear structural optimization using goal programming. Computers Structures 32 (1989) 69-73.

[52] Fahmy, D., H. El Shishiny: A goal programming model for desert farm planning. Advances in Desert and Arid Land Technology Development 5 (1991) 69-86.

[53] Feinstein, C. D., M. N. Thapa: Notes: A reformulation of a mean-absolute deviation portfolio optimization model. Management Science 37 (1993) 1552-1553.

[54] Fiske, W. A., G. E. D'Souza, J. J. Fletcher, T. T. Phipps, W. B. Bryan, E. C. Prigg: An economic and environmental assessment of alternative forage-resource production systems: A goal programming approach. Agricultural Systems 45:3 (1994) 259-270.

[55] Franz, L., S. M. Lee: "A goal programming based interactive support system". In J. N. Morse (ed) :Organizations: Multiple Agents with Multiple Criteria, Springer Lecture Notes on Economic Mathematical Systems Vol. 190, Springer-Verlag, Berlin 1981, pp. 110-115.

[56] Freed, N., F. Glover: Simple but powerful goal programming models for discriminant problems. European Journal of Operational Research 7 (1981) 44-60.

[57] Freed, N., F. Glover: Evaluating alternative linear programming models to solve the two-group discriminant problem. Decision Sciences 17 (1986) 151-162.

[58] Fukukawa, T., S.-C. Hong: The determination of the optimal number of kanbans in a just-in-time production system. Computers & Industrial Engineering 24:4 (1993) 551-559.

[59] Gagnon, R. J., C. Sheu: A strategic MIGP model for acquiring advanced technologies. Computers & Industrial Engineering 32:1 (1997) 145-168.

[60] Gen, M., K. Ida, Y. Tsujimura, C. E. Kim: Large-scale 0-1 fuzzy goal programming and its application to reliability optimization problems. Computers & Industrial Engineering 24:4 (1993) 539-549.

[61] Ghandforoush, P.: Optimal allocation of time in a hospital pharmacy using goal programming. Journal of the Operational Research Society 70:2 (1993) 191-198.

[62] Geiger, J. J., N. Pendegraft, L. M. Geiger: A pc-based project management tool. Journal of Systems Management 47:3 (1996) 52-57.

[63] Giokas, D.: The use of goal programming and data envelopment analysis for estimating efficient marginal costs of outputs. Journal of the Operational Research Society 48 (1997) 319-323.

[64] Giokas, D., M. Vassiloglou: A goal programming model for bank assets and liabilities management. European Journal of Operational Research 50:1 (1991) 48-60.

[65] Goedhart, M. H.: Financial Planning in Divisionalised Firms: Models and Methods. Thesis Publishers, Amsterdam, 1994.

[66] Goedhart, M. H., J. Spronk: Financial planning with fractional goals. European Journal of Operational Research 82 (1995a) 111-124.

[67] Goedhart, M. H., J. Spronk: An interactive heuristic for financial planning in decentralized organizations. European Journal of Operational Research 86 (1995b) 162-175.

[68] Gokcen, H., E. Erel: A goal programming approach to mixed-model assembly line balancing problem. International Journal of Production Economics 48:2 (1997) 177-185.

[69] Golany, B., F. Y. Phillips, J. J. Rousseau: Optimal design of syndicated panels: A mathematical programming approach. European Journal of Operational Research 87:1 (1995) 148-165.

[70] Golany, B., G. Yu: A goal programming-discriminant function approach to the estimation of an empirical production function based on DEA results. Journal of Productivity Analysis 6:2 (1995) 171-186.

[71] Hannan, E. L.: Nondominance in goal programming. INFOR 18 (1980) 300-309.

[72] Hannan, E. L.: On Fuzzy Goal Programming. Decision Sciences 12:3 (1981) 522-531.

[73] Hannan, E. L.: An assessment of some criticisms of goal programming. Computers and Operational Research 12 (1985) 525-541.

[74] He, X. P., X. D. Pei, X.D. Guan: Trends in vegetable supply in the peak season in Beijing. Acta Agriculturae Univesitatis Pekinensis 16:3 (1990) 301-306.

[75] Hoffman, J. J., M. J. Schniederjans: A two-stage model for structuring global facility site selection decisions: The case of the brewing industry. International Journal of Operations & Production Management 14:4 (1994) 79-96.

[76] Holden, S. T.: Peasant household modelling: Farming systems evolution and sustainability in northern Zambia. Agricultural Economics 9:3 (1993) 241-267.

[77] Hoshino, T., K. Yura, K. Hitomi: Optimization analysis for recycle-oriented manufacturing systems. International Journal of Production Research 33:8 (1995) 2069-2078.

[78] Huang, Z., S. X. Li: Dominance stochastic models in data envelopment analysis. European Journal of Operational Research 95 (1996) 390-403.

[79] Hwang, C., A. Masud: Multiple Objective Decision Making Methods and Applications. Springer, Berlin, 1979.

[80] Ignizio, J. P.: Goal Programming and Extensions. Lexington Books, Lexington, MA, 1976.

[81] Ignizio, J. P.: The determination of a subset of efficient solutions via goal programming. Computers and Operations Research 8 (1981) 9-16.

[82] Ignizio, J. P.: Linear Programming in Single and Multiple Objective Systems. Prentice-Hall, Englewood Cliffs, NJ, 1982.

[83] Inuiguchi, M., M. Sakawa, Y. Kume: The usefulness of possibilistic programming in production planning problems. International Journal of Production Economics 33:1-3 (1994) 45-52.

[84] Irwin, S. C., J. L. Scott, R. G. Jackson: A semen allocation system for livestock improvement: A case study in goal programming. Computers and Electronics in Agriculture 16:1 (1996) 87-101.

[85] Jaaskelainen, V.: Accounting and Mathematical Programming. The Helsinki School of Economics, Helsinki, 1969.

[86] Jedicke, V., W. L. Wilbur, A. K. Rifai: Goal programming: An effective tool in commercial bank management. Mid-American Journal of Business 14:10 (1994) 31-40.

[87] Johnson, P. J., P. A. Oltenacu, H. M. Kaiser, R. W. Blake: Modeling parasite control programs for developing nations using goal programming. Journal of Production in Agriculture 4:1 (1991) 33-38.

[88] Jones, D. F., M. Tamiz: Expanding the flexibility of goal programming via preference modelling techniques. Omega 23:1 (1995) 41-48.

[89] Kalpic, D., V. Mornar, M. Baranovic: Case study based on a multi-period multi-criteria production planning model. European Journal of Operational Research 87:3 (1995) 658-669.

[90] Kalu, T. C. U.: Determining the impact of Nigeria's economic crisis on the multinational oil companies: A goal programming approach. Journal of the Operational Research Society 45:2 (1994) 165-177.

[91] Kalu, T. C. U.: Petroleum industry sensitivity and world oil market prices: The Nigerian example. Journal of Energy and Development 21:1 (1995) 45-62.

[92] Kangas, J., T. Pukkala: A decision theoretic approach applied to goal programming of forest management. Silva Fennica 26:3 (1992) 169-176.

[93] Kendall, K., S. M. Lee: A goal programming approach to blood rotation. Management Science 26:11 (1980) 1145-1157.

[94] Kendall, K., S. M. Lee: Improving perishable product inventory management using goal programming. Journal of Operations Management 1:2 (1980) 77-84.

[95] Khorramshahgol, R., A. A. Okuruwa: Goal programming approach to investment decisions: A case study of fund allocation among different shopping malls. European Journal of Operational Research73:1 (1994) 17-22.

[96] Khouja, M., R. Conrad: Balancing the assignment of customer groups among employees: Zero-one goal programming and heuristic approaches. International Journal of Operations & Production Management 15:3 (1995) 76-87.

[97] Kim, S. H., B. S. Ahn, S. H. Choi: An efficient force planning system using multi-objective lineear goal programming. Computers & Operations Research 24:6 (1997) 569-580.

[98] Kim, G. C., M. J. Schniederjans: A multiple objective model for a just-in-time manufacturing system environment. Journal of Operations & Production Management 13:12 (1993) 47-61.

[99] Kolenka, I.: The choice of optimal strategy of forest management from the viewpoint of economic prosperity. Acta Facultatis Forestalis Zvolen #35 (1993) 369-377.

[100] Kornbluth, J. S. H., R. E. Steuer: Goal programming with linear fractional functionals. European Journal of Operational Research 8 (1981) 58-65.

[101] Kumar, P., N. Singh, N. K. Tewari: A nonlinear goal programming model for multistage, multiobjective decision problems with application to grouping and loading problem in a flexible manufacturing system. European Journal of Operational Research 53:2 (1991) 166-171.

[102] Kwak, N. K., M. J. Schniederjans, K. S. Warkentin: An application of linear goal programming to the marketing distribution decision. European Journal of Operational Research 52:3 (1991) 334-344.

[103] Lam, K. F., E. U. Choo: A linear goal programming model for classification with non-monotone attributes. Computers and Operations Research 20:4 (1993) 403-408.

[104] Lam, K. F., E. U. Choo: Goal programming in preference decomposition. Journal of the Operational Research Society 46 (1995) 205-213.

[105] Lam, K. F., E. U. Choo, W. C. Wedley: Linear goal programming in estimation of classification probability. European Journal of Operational Research 67:1 (1993) 101-110.

[106] Lam, K. F., J. W. Moy: Improved linear programming formulations for the multi-group discriminant problem. Journal of the Operational Research Society 47 (1996) 1526-1529.

[107] Lashine, S., B. Foote, A. Ravindran: A nonlinear mixed integer goal programming model for the 2-machine closed flow-shop. European Journal of Operational Research 55:1 (1991) 57-70.

[108] Lee, C. S., C. G. Wen: Fuzzy goal programming approach for water quality management in a river basin. Fuzzy Sets and Systems 89:2 (1997) 181-192.

[109] Lee, S. M.: Decision analysis through goal programming. Decision Sciences 2:2 (1971) 172-180.

[110] Lee, S. M.: Goal Programming for Decision Analysis. Auerbach, Philadelphia, 1972.

[111] Lee, S. M., D. L. Chesser: Goal programming for portfolio selection. Journal of Portfolio Management 6 (1980) 22-26.

[112] Lee, S. M., S. Chung, A. Everett: Goal programming methods for implementation of just-in-time production. Production Planning and Control 3:2 (1992) 175-182.

[113] Lee, S. M., H. Eom: A large scale goal programming-based decision support for global financing strategy formulation. Information & Management 12:1 (1987) 33-44.

[114] Lee, S. M., H. Eom: A multicriteria approach to formulating international project-financing strategies. Journal of the Operational Research Society 40:6 (1989) 519-528.

[115] Lee, S. M., H. B. Eom, C. A. Snyder, F. N. Ford: A multiple criteria decision support system for global finance and planning. Journal of Management Information Systems 4:4 (1987-1988) 94-113.

[116] Lee, S. M., A. Everett, J. Melkonian: Optimizing lot size and setup time in a JIT environment: A multiple objective application. Production Planning and Control 5:3 (1994) 308-319.

[117] Lee, S. M., H. Jung: A multi-objective production planning model in a flexible manufacturing environment. International Journal of Production Research 27:11 (1989) 1981-1992.

[118] Lee, S. M., R. Luebbe: The multi-criteria warehouse location problem revisited. International Journal of Physical Distribution & Materials Management 17:3 (1987) 56-59.

[119] Lee, S. M., D. L. Olson: Chance constrained aggregate blending. Journal of Construction Engineering and Management 109:1 (1983) 39-47.

[120] Lee, S. M., D. L. Olson: A gradient algorithm for chance constrained nonlinear goal programming. European Journal of Operational Research 22 (1985) 359-369.

[121] Lee, S. M., J.-P. Shim: Planning for multiple objectives with zero-base budgeting. Long Range Planning 17:5 (1984) 103-110.

[122] Lewis, R. P., Jr., H. A. Taha: An investigation of the use of goal programming to fit response surfaces. European Journal of Operational Research 86 (1995) 537-548.

[123] Love, C. E., K. F. Lam: Classifying and controlling errors in forecasting using multiple criteria goal programming. Computers and Operations Research 21:9 (1994) 979-989.

[124] Lyu, J., A. Gunasekaran, C. Y. Chen, C. Kao: A goal programming model for the coal blending problem. Computers & Industrial Engineering 6:2 (1995) 861-868.

[125] Malczewski, J., W. Ogryczak: The multiple criteria location problem: 2. Preference-based techniques and interactive decision support. Environment and Planning A-28:1 (1996) 69-98.

[126] Markowski, C. A., J. P. Ignizio: Duality and transformations in multiphase and sequential linear goal programming. Computers and Operations Research 10:4 (1983a) 321-333.

[127] Markowski, C. A., J. P. Ignizio: Theory and properties of the lexicographic linear goal programming dual. Large Scale Systems 5 (1983b) 115-121.

[128] Martel, J. M., B. Aouni: Incorporating the decision-maker's preferences in the goal-programming model. Journal of the Operational Research Society 41 (1990) 1121-1132.

[129] Masud, A. S., C. L. Hwang: Interactive sequential goal programming. Journal of the Operational Research Society 32 (1981) 391-400.

[130] McCarl, B. A., M. E. Rister, J. R. Stokes, H. A. Ziari: Improving shadow price information: Obtaining relevant shadow prices from multiple objective programming problems. American Journal of Agricultural Economics 78 (1996) 699-705.

[131] McGregor, M. J., J. B. Dent: An application of lexicographic goal programming to resolve the allocation of water from the Rakaia River. Agricultural Systems 41:3 (1993) 349-367.

[132] Michnik, J., T. Trzaskalik: Linear goal programming model for portfolio management in Polish commercial banks. Central European Journal for Operations Research and Economics 3:3 (1995-1995) 243-248.

[133] Min, H.: International intermodal choices via chance-constrained goal programming. Transportation Research Part A: Policy and Practice 25:6 (1991) 351-362.

[134] Min, H.: On misconceptions in goal programming – Response. Journal of the Operational Research Society 42:10 (1991) 928-929.

[135] Min, H., J. Storbeck: On the origin and persistence of misconceptions in goal programming. Journal of the Operational Research Society 42:4 (1991) 301-312.

[136] Mitra, A., J. G. Patankar: An integrated multicriteria model for warranty cost estimation and production. IEEE Transactions on Engineering Management 40:3 (1993) 300-311.

[137] Mohammed, R. H.: The relationship between goal programming and fuzzy programming. Fuzzy Sets and Systems 89 (1997) 215-222.

[138] Moore, L. J., J. C. Davis, A. E. Harvey, J. J. Imholz: Multicriteria models for analysis of the environmental restoration and waste management program. Computers & Industrial Engineering 25:1-4 (1993) 21-24.

[139] Mote, J., D. L. Olson, M. A. Venkataramanan: A comparative multiobjective programming study. Mathematical and Computer Modelling 10:10 (1988) 719-729.

[140] Mukherjee, K., A. Bera: Application of goal programming in project selection decision – A case study from the Indian coal mining industry. European Journal of Operational Research 82:1 (1995) 18-25.

[141] Muthukude, P., J. L. Novak, C. and Jolly: A goal programming evaluation of the fisheries development plan for Sri Lanka's coastal fishing fleet 1988-1991. American Journal of Agricultural Economics 72:5 (1990) 1367.

[142] Myint, S., M. T. Tabucanon: A multiple-criteria approach to machine selection for flexible manufacturing systems. International Journal of Production Economics 33:1-3 (1994) 121-131.

[143] Nanda, J., D. Kothari, K. Lingamurthy: Economic-emission load dispatch through goal programming techniques. IEEE Transactions on Energy Conversion 3 (1988) 26-32.

[144] Narasimhan, R.: Goal programming in a fuzzy environment. Decision Sciences 11 (1980) 325-336.

[145] Nino de Zepeda, A., M. Maino, F. di Silvestri, J. Berdegue, F. de Silvestre: The trade-off between sustainability and productivity: A multicriteria decision making approach. Investigacion Agraria Economia 9:1 (1994) 143-155.

[146] Njiti, C. F., D. M. Sharpe: A goal-programming approach to the management of competition and conflict among land uses in the tropics: The Cameroon example. Ambio 23:2 (1994) 112-119.

[147] Nogami, K.: Forest management planning and fuzzy goal programming. Journal of the Japanese Forestry Society 73:1 (1991) 34-39.

[148] November, S. M., R. G. Cromley, E. K. Cromley: Multi-objective analysis of school district regionalization alternatives in Connecticut. Professional Geographer 48:1 (1996) 1-14.

[149] Nudtasomboon, N., S. U. Randhaawa: Resource-constrained project scheduling with renewable and non-renewable resources and time-resource tradeoffs. Computers & Industrial Engineering 32:1 (1997) 227-242.

[150] Ogryczak, W.: On the lexicographic minimax approach to location problems. European Journal of Operational Research 100:3 (1997) 566-585.

[151] Ohta, H., T. Yamaguchi: Linear fractional goal programming in consideration of fuzzy solution. European Journal of Operational Research 92 (1996) 157-165.

[152] Olson, D. L., E. White, E., C.-Y. Wei: Prediction interval constrained programming. Journal of Quality Technology 20:2 (1988) 90-97.

[153] Peralta, R. C., J. Solaimanian, G. R. Musharrafich: Optimal dispersed groundwater contaminant management: MODCON method. Journal of Water Resources Planning and Management – ASCE 121:6 (1995) 490-498.

[154] Philipson, R., A. Ravindran: Application of goal programming to machinability data optimization. Transactions of the ASME 100 (1978) 286-291.

[155] Piech, B., T. Rehman: Application of multiple criteria decision making methods to farm planning: A case study. Agricultural Systems 41 (1993) 305-319.

[156] Preckel, P., A. Featherstone, T. Baker: Interpreting dual variables for optimization with nonmonetary objectives. American Journal of Agricultural Economics 69 (1987) 849-851.

[157] Premachandra, I. M.: A goal-programming model for activity crashing in project networks. Journal of Operations & Production Management 13:6 (1993) 79-85.

[158] Puelz, A. V., S. M. Lee: A multiple-objective programming technique for structuring tax-exempt serial revenue debt issues. Management Science 38:8 (1992) 1186-1200.

[159] Qiu, M.: Prioritising and scheduling road projects by genetic algorithm. Mathematics and Computers in Simulation 43:3-6, 1997, 569-574.

[160] Ramanathan, R., L. S. Ganesh: A multiobjective evaluation of energy alternatives for water pumping in urban households. Energy Sources 17:2 (1995) 223-239.

[161] Ramanathan, R., L. S. Ganesh: Energy resource allocation incorporating qualitative and quantitative criteria: An integrated model using goal programming and AHP. Socio-Economic Planning Sciences 29:3 (1995) 197-218.

[162] Randhawa, S. U., D. Sitompul: A heuristic-based computerized nurse scheduling system. Computers & Operations Research 20:8 (1993) 837-844.

[163] Rao, S. S., K. Sundararaju, B. G. Prakash, C. Balakrishna: Fuzzy goal programming approach for structural optimization. AIAA Journal 30:5 (1992) 1425-1432.

[164] Ravindran, A., W. Shin, J. Arthur, H. Moskowitz: Nonlinear integer goal programming models for acceptance sampling. Computers and Operations Research 13 (1986) 611-622.

[165] Ravirala, V., D. A. Grivas: Goal-programming methodology for integrating pavement and bridge programs. Journal of Transportation Engineering 121:4 (1995) 345-351.

[166] Reeves, G. R., S. R. Hedin: A generalized interactive goal programming procedure. Computers and Operations Research 20:7 (1993) 747-753.

[167] Retzlaff-Roberts, D. L., R. C. Morey: A goal-programming method of stochastic allocative data envelopment analysis. European Journal of Operational Research 71 (1993) 379-397.

[168] Ringuest, J. L.: Multiobjective optimization: Behavioral and computational considerations. Kluwer Academic, Norwell, MA, 1992.

[169] Roljic, L., M. Dujsic: Nonlinear goal programming model application in optimal economy development policy choosing. Economic Analysis and Workers' Management 25:2 (1991) 155-170.

[170] Romero, C.: A survey of generalized goal programming (1970-1982). European Journal of Operational Research 25 (1986) 183-191.

[171] Romero, C.: Handbook of Critical Issues Within Goal Programming. Pergamon Press, Oxford, 1991.

[172] Romero, C.: On misconceptions in goal programming. Journal of the Operational Research Society 42:10 (1991) 927-928.

[173] Romero, C.: Carry on with redundancy in lexicographic goal programming. European Journal of Operational Research 78 (1994) 441-442.
[174] Romero, C., T. Rehman: Multiple Criteria Analysis for Agricultural Decisions. Elsevier, Amsterdam, 1989.
[175] Rossing, W. A. H., J. W. E. Jansma, F. J. de Ruijter, J. Schans: Operationalizing sustainability: Exploring options for environmentally friendly flower bulb production systems. European Journal of Plant Pathology 103:3 (1997) 217-234.
[176] Saber, H. M., A. Ravindran: Nonlinear goal programming theory and practice: A survey. Computers & Operations Research 20:3 (1993) 275-271.
[177] Santhanam, R., J. Kyparisis: A multiple criteria decision model for information system project selection. Computers & Operations Research 22:8 (1995) 807-818.
[178] Santhanam, R., M. J. Schniederjans: A model formulation system for information system project selection. Computers & Operations Research 20:7 (1993) 755-767.
[179] Schniederjans, M. J.: Linear Goal Programming. Petrocelli Books, Princeton, NJ, 1984.
[180] Schniederjans, M. J.: Goal Programming: Methodology and Applications. Kluwer Academic Publishers, Boston, 1995.
[181] Schniederjans, M. J.: The life cycle of goal programming research as recorded in journal articles. Operations Research 43:4 (1995) 551-558.
[182] Schniederjans, M. J., T. Garvin: Using the analytic hierarchy process and multi-objective programming for the selection of cost drivers in activity-based costing. European Journal of Operational Research 100:1 (1997) 72-80.
[183] Schniederjans, M. J., J. J. Hoffman: Multinational acquisition analysis – A zero-one goal programming model. European Journal of Operational Research 62:2 (1992) 175-185.
[184] Schniederjans, M. J., J. J. Hoffman, G. S. Sirmans: Using goal programming and the analytic hierarchy process in house selection. Journal of Real Estate Finance and Economics 11:2 (1995) 167-176.
[185] Schniederjans, M. J., C. M. Karuppan: Designing a quality control system in a service organization: A goal programming case study. European Journal of Operational Research 81:2 (1995) 249-258.
[186] Schniederjans, M. J., R. Santhanam: A multi-objective constrained resource information system project selection method. European Journal of Operational Research 70:2 (1993) 244-253.
[187] Shiau, T. N., C. P. Kuo, J. R. Hwang: Multi-objective optimization of a flexible rotor in magnetic bearings with critical speeds and control current constraints. Journal of Engineering for Gas Turbines and Power - Transactions of the ASME 119:1, 186-195.
[188] Shin, W., A. Ravindran: An interactive MOMP method: A survey. Computers & Operations Research 18 (1991) 97-114.
[189] Siha, S.: A decision model for selecting mutually exclusive alternative technologies. Computers & Operations Research 24:3 (1993) 459-475.

[190] Sinclair, M., D. L. van Oudheusden: A103(1997) 18-27 in heavily congested cities. European Journal of Operational Research 103:1 (1997) 18-27.

[191] Specht, P. H.: Multicriteria planning model for mental health services delivery. International Journal of Operations & Production Management 13:9 (1993) 62-71.

[192] Spronk, J., B. Matarazzo: Interactive multifunctional planning for agriculture: State of the art. Rivista di Economia Agraria 47:4 (1992) 531-543.

[193] Stewart, T. J.: A critical survey on the status of multiple criteria decision making: Theory and practice. Omega 20:5 (1992) 569-586.

[194] Sudhir, V., V. R. Muraleedharan, G. Srinivasan: Integrated solid waste management in urban India: A critical operational research framework. Socio-Economic Planning Sciences 30:3 (1996) 163-181.

[195] Sueyoshi, T.: Constrained regression median for measuring possible salary discrimination. European Journal of Operational Research 77:2 (1994) 253-271.

[196] Sueyoshi, T.: Divestiture of Nippon Telegraph and Telephone. Management Science 42:9 (1996) 1326-1351.

[197] Sumpsi, J. M., F. Amador, C. Romero: On farmers' objectives: A multi-criteria approach. European Journal of Operational Research 96:1 (1997) 64-71.

[198] Sutardi, C. R. Bector, I. Goulter: Model for optimal investment for irrigation projects in Bengkulu Province, Indonesia. Water International 16:2 (1991) 64-74.

[199] Sutardi, C. R. Bector, I. Goulter: Multiobjective water resources investment planning under budgetary uncertainty and fuzzy environment. European Journal of Operational Research 82 (1995) 556-591.

[200] Sutardi, C. R. Bector, I. Goulter, T. C. E. Cheng: Multiobjective water resources investment planning under budgetary and socio-technical uncertainties. IEEE Transactions on Engineering Management 41:1 (1994) 50-68.

[201] Suter, W. C., Jr., J. A. Calloway: Rough mill policies and practices examined by a multiple-criteria goal program called ROMGOP. Forest Products Journal 44:10 (1994) 19-28.

[202] Suzuki, S.: Multigoal programming for structure control design synthesis with fuzzy goals. JSME International Journal C38:3 (1995) 450-456.

[203] Tamiz, M., D. F. Jones: Algorithmic improvements to the method of Martel and Aouni. Journal of the Operational Research Society 46 (1995) 254-257.

[204] Tamiz, M., D. F. Jones: Interactive framework for investigation of goal programming models: Theory and practice. Journal of Multi-Criteria Decision Analysis 6 (1997) 52-60.

[205] Tamiz, M., D. F. Jones, E. Eldarzi: A review of goal programming and its applications. Annals of Operations Research 58 (1995) 39-53.

[206] Taplin, J. H. E, M. Qiu, Z. Zhang: Allocation of public road funds by goal programming on multiple criteria. Australian Journal of Public Administration 54:1 (1995) 58-64.

[207] Taylor, B., L. Moore, E. Clayton: R&D project selection and manpower allocation with integer nonlinear goal programming. Management Science 28 (1982) 1149-1158.

[208] Taylor, B., L. Moore, E. Clayton, K. R. Davis, T. Rakes: An integer goal programming models for the deployment of state highway patrol units. Management Science 31 (1985) 1335-1347.

[209] Van Berlo, J. M., J. M. Van Berlo: A decision support tool for the vegetable processing industry: An integrative approach of market, industry and agriculture. Agricultural Systems 43:1 (1993) 91-109.

[210] Vankooten, G. C.: How to preserve wilderness areas without an endangered species act – Examining tradeoffs on Vancouver Island with goal programming. American Journal of Agricultural Economics 76:5 (1994) 1262.

[211] Vankooten, G. C.: Modeling public forest land use tradeoffs on Vancouver Island. Journal of Forest Economics 1:2 (1995) 191-217.

[212] Venkataraman, R., J. Nathan: Master production scheduling for a process industry environment: A case study. International Journal of Operations & Production Management 14:10 (1994) 44-54.

[213] Venkataraman, R., S. B. Smith: Disaggregation to a rolling horizon master production schedule with minimum batch-size production restrictions. International Journal of Production Research 34:6 (1996) 1517-1537.

[214] Wang, H.-F., C.-C. Fu: A generalization of fuzzy goal programming with preemptive structure. Computers and Operations Research 24:9 (1997) 819-828.

[215] Watanabe, N., S. Hiraki: An approximate solution to a JIT-based ordering system. Computers & Industrial Engineering 31:3-4 (1996) 565-569.

[216] Wei, J.-C., D. L. Olson, E. M. White: Simultaneous optimization in process quality control via prediction-interval constrained programming. Journal of the Operational Research Society 41:12 (1990) 1161-1167.

[217] Weistroffer, H.: An interactive goal programming method for nonlinear multiple-criteria decision-making problems. Computers & Operations Research 10 (1983) 311-320.

[218] Yang, C. W.: A note on the bounded solution of the bilateral monopoly mode. Eastern Economic Journal 17:3 (1991) 367-372.

[219] Yang, Y. P., Y. A. Chen: Multiobjective optimization of hard disk suspension assemblies: 2. Integrated structure and control design. Computers & Structures 59:4 (1996) 771-782.

[220] Yang, Y. P., C. C. Kuo: Passive and active design of hard disk suspension assemblies using multiobjective optimization techniques. Computer Methods in Applied Mechanics and Engineering 145:1-2 (1997) 147-166.

[221] Yin, Y. Y., J. T. Pierce, E. Love: Designing a multisector model for land conversion study. Journal of Environmental Management 44:3 (1995) 249-266.

[222] Yura, K.: Production scheduling to satisfy worker's preferences for days off and overtime under due-date constraints. International Journal of Production Economics 33:1-3 (1994) 265-270.

[223] Zamora, R., J. Berbel: Multiobjective and goal programming applied to wildlife management: A case study on a Mediterranean ecosystem. Investigacion Agraria Economia 10:2 (1995) 253-263.

[224] Zanakis, S. H., S. K. Gupta: A categorized bibliographic survey of goal programming. Omega 13 (1985) 211-222.

[225] Zekri, S., C. Romero: Influence of decision making preferences and economic incentives upon the reduction of river pollution. Investigacion Agraria Economia 6:2 (1991) 223-239.

[226] Zeleny, M.: The pros and cons of goal programming. Computers and Operational Research 8 (1981) 357-359.

[227] Zhu, Z. W., M. A. Mcknew: A goal programming workload balancing optimization model for ambulance allocation – An application to Shanghai, PR China. Socio-Economic Planning Sciences 27:2 (1993) 137-148.

9 REFERENCE POINT APPROACHES

Andrzej P. Wierzbicki

Contents

Abstract: This chapter presents a summary of reference point methodology in vector optimization and decision support. The methodology has been developed at the International Institute for Applied Systems Analysis (IIASA) since 1980 and found numerous applications, both in IIASA and elsewhere. The chapter presents methodological foundations, basic concepts and notation, reference points and achievement functions, neutral and weighted compromise solutions, issues of modeling for multi-objective analysis, some basic applications of reference point methods and a discussion of a decision process type supported by reference point methodology.

T. Gal et al. (eds.), *Multicriteria Decision Making*

9.1 GENERAL ASSUMPTIONS AND FEATURES

Reference point approaches were developed starting with research done at the International Institute for Applied Systems Analysis (IIASA) in Laxenburg n. Vienna, Austria, since 1980 – see Wierzbicki [45], Kallio *et al.* [20] – specifically as a tool of environmental model analysis, although these approaches have found numerous other applications since that time. Soon, similar or equivalent approaches were developed, *e.g.* the *weighted Chebyshev*[1] procedure by Steuer and Cho [42] or the *satisficing trade-off* method by Nakayama and Sawaragi [31]. Later, Korhonen and Laakso [22] drew attention to the fact that reference point methods can be considered as *generalized goal programming.* This generalization tries to preserve main advantages of goal programming (see also Chapter 8 in this volume) and to overcome its basic disadvantages.

The main advantages of goal programming are related to the psychologically appealing idea that we should set a goal in objective space and try to come close to it. Coming close to a goal suggests minimizing a distance measure between an attainable objective vector (decision outcome) and the goal vector.

The basic disadvantage relates to the fact that this idea is mathematically inconsistent with the concept of *vector-optimality* or *efficiency* . One of the basic requirements – a generally sufficient condition for efficiency – for a function to produce a vector-optimal outcome (when minimized or maximized) is an appropriate monotonicity of this function. But any distance measure is obviously not monotone when its argument crosses zero. Therefore, distance minimization cannot, without additional assumptions, provide vector-optimal or efficient solutions.

Consider, for example, the simplest case when the goal vector is in itself an attainable decision outcome but not an efficient objective vector; then distance minimization leads to the obvious solution with objectives equal to the goals. Even for convex outcome sets, either special tricks or rather restrictive assumptions are needed in goal programming to provide for efficiency of obtained decision outcomes. If, however, the set of attainable objectives is not convex – for example, discrete, as in Fig. 9.1a – then distance minimization cannot result, generally, in efficient outcomes. Both components of decision outcomes or objectives y_1 and y_2 in this Figure are to be maximized and the efficient outcomes, denoted by circles, are to the "North-East" of the attainable outcome set; there are many intuitively reasonable vectors of goals, such as $\bar{\mathbf{y}}^1$, which would produce inefficient outcomes, such as $\mathbf{y}^1$, if a norm as a measure of the distance is minimized.

Nevertheless, setting a goal and trying to come close to it is psychologically appealing; the problem is "only" how to provide for efficiency of resulting outcomes. There are two ways to do it: either to limit the goals or to change the sense of coming close to the goals.

[1] In the original paper, the authors used the word Tchebycheff, not Chebyshev; the former is a German transliteration of this Russian name.

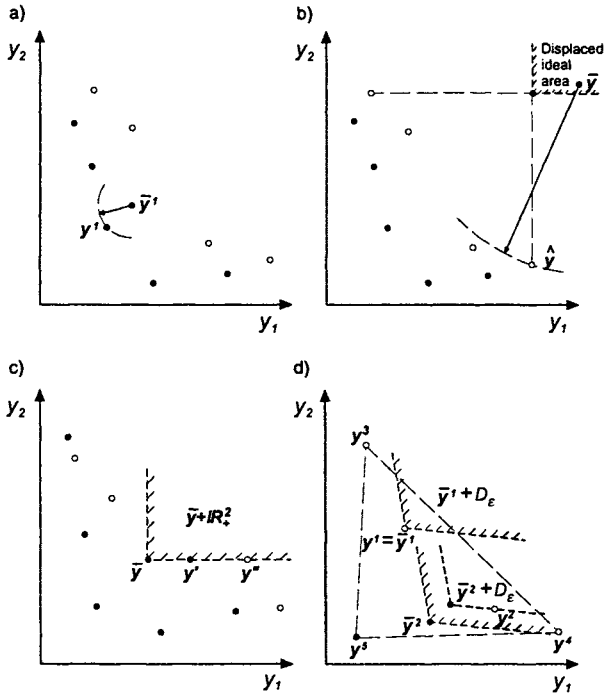


Figure 9.1: Examples of selections of discrete outcomes by using various approaches: a) goal programming or norm minimization; b) displaced ideal; c) max-min approach d) reference point approach

Trying to limit the set of goals is the essence of the *displaced ideal* method of Zeleny [54]: if we select goals that are sufficiently distant from the set of attainable outcomes, then we can prove that norm minimization will result only in efficient outcomes, no matter what norm we use or what properties the sets of attainable outcomes have. This is illustrated by Fig. 9.1b, where the goal $\bar{y}$ is in the displaced ideal area and the outcomes resulting from norm minimization are efficient. However, such limitation means precisely loosing the intuitive appeal of the goal programming approach: if we can set only unrealistic goals, the approach looses its basic advantages.

Trying to change the sense of coming close to the goal changes the nature of the goal. *Reference points are goals interpreted consistently with basic concepts of vector optimality;* the sense of "coming close" to it is rather special and certainly does not mean distance minimization. If we accept the logic of various concepts of vector optimality, as discussed in earlier Chapters, then "coming close" to a given reference point should mean:

- decision outcomes in some sense uniformly close to the given reference point, if the latter is not attainable – while the precise sense of uniform

closeness might be modified by demanding that the resulting decisions and their outcomes remain efficient i.e. vector-optimal;

- decision outcomes precisely equal to the given reference point, if the latter is efficient, vector-optimal – which, somewhat simplifying, means attainable without any surplus;

- decision outcomes in some sense uniformly better than the given reference point, if the latter is attainable with some surplus – thus inefficient, not vector-optimal (where the sense of uniform improvement can be again variously interpreted).

The first two cases coincide (almost) with goal programming; the third case is, however, essentially different: it means not "coming close" in any traditional sense, but "coming close or better".

This change of the sense of "coming close" is in fact deeply related to the discussion how people make decisions in reality and how computers should support decisions. In turn, this is related to the concept of *satisficing* decisions of Simon [40] which was used as a description how people make actual decisions (particularly in large organizations) and the concept of *quasi-satisficing* decisions of Wierzbicki [46] which describes how a computerized decision support system should help a human decision maker.

According to Simon, real decision makers do not optimize their utility when making decisions, for many reasons. Simon postulated that actual decision makers, through learning, adaptively develop *aspiration levels* for various important outcomes of their decisions. Then they seek decisions that would result either:

- in outcomes as close as possible to the aspiration levels, if the latter are not attainable (which corresponds to an optimization of decisions, but in the sense of the distance from aspiration levels);

- in outcomes equal to aspiration levels, if the latter are attainable (which corresponds to stopping improvements in this case).

We see that satisficing decision making can be in fact mathematically represented by goal programming. In the case of attainable aspiration levels, the decision maker might learn to increase them, but usually not for current, only for future decisions. One can ask why; the most probable answer is that decision making processes are difficult and this assumption reflects some inherent human laziness. Many further studies have shown that such a satisficing behavior of a decision maker, though might seem peculiar, is very often observed in practice. In particular, the use of various reference levels by decision makers – such as aspiration levels, but including also *reservation levels,* very important e.g. in the theory of negotiations – has been repeatedly confirmed in practice.

Independently, however, from the issue whether a real, human decision maker would (or could, or should) optimize in all cases, we can require that a good

computer program supporting decisions through model analysis should behave like a hypothetical, perfectly rational decision maker – with one important exception: the program should not "outguess" its user, the real decision maker, by trying to construct a model of her/his preferences or utility function, but should instead accept simple instructions which characterize such preferences.

Thus, the methodology of reference point approaches assumes that the instructions from a user to the computerized decision support system (DSS) have the convenient form of reference points, including aspiration levels and, possibly, reservation levels – and that the user is not asked how she/he determines the reference points. An essential departure from Simon asumptions and from goal programming techniques, however, is as follows: the methodology of reference point approaches assumes that *the computerized DSS tries to improve a given reference point, if this point is attainable.*

Therefore, the behavior of the DSS – not that of its user – is in a sense similar to perfect rationality. It does not minimize a norm, but optimizes a special function, called *achievement function* which is a kind of a proxy utility or value function (of the DSS) such that the decisions proposed by the DSS satisfy the three cases of "coming close or better" described above. Because of the difference – in the last case of "coming better" – to the satisficing behavior, we call such behavior quasi-satisficing. It can be compared to the behavior of a perfect staff which supports a manager or boss, who gives instructions to this staff in the form of reference levels. The staff works out detailed decisions which are guided by the given reference point.

However, being perfect, the staff does not correct attainability estimates (a real, human staff might behave otherwise) and does not report to the boss that the reference point is attainable when it really is not. Instead, the staff proposes decisions that result in outcomes as close as possible to the desired reference point and reports these decisions together with their not quite satisfactory outcomes to the boss. If the reference point is attainable without any surplus, the perfect staff just works out the decisions how to reach this point and does not argue with the boss that a different point and different decisions might be better (if not specifically asked about such opinion). If the reference point is attainable with surplus, the perfect staff does not stop working and start gossiping over drinks – as Simon's model of satisficing behavior would suggest – but works out decisions that would result in a uniform improvement of outcomes as compared to reference levels, and proposes such decisions together with improved outcomes to the boss. Obviously, only a computer program could behave all times in this perfect, quasi-satisficing manner.

On the other hand, goal programming corresponds precisely to satisficing behavior: if the aspiration levels are attainable, then there exist attainable outcomes precisely equal to them, thus the corresponding distance is zero; since we cannot get distance less than zero, the optimization is stopped (the staff prepares drinks for relaxation).

Thus, reference point optimization is a generalization of the goal programming approach to such cases when we can and want to improve (minimize or

maximize) certain outcomes beyond their reference points. For this purpose, a special class of *order-consistent achievement functions,* similar but not equivalent to distance functions, was developed, investigated in detail and applied in many examples and DSS's.

We shall briefly indicate here some of general properties of such achievement functions. Vector optimization corresponds to a partial order of the objective space, which might be defined with the help of a *positive cone D;* if we e.g. want to maximize two objectives, the positive cone is just $\boldsymbol{R}^2_+$, the positive ortant of the plane. The mathematical definition of vector optimality implies the idea of choosing an achievement function whose level-sets represent or closely approximate the positive cone, possibly with vertex shifted to the reference point.

Actually, the idea of using an achievement function with level sets precisely representing the positive cone is rather old and corresponds to the max-min approach[2], see e.g. Polak [34]. However, if the level sets of an achievement function precisely represent the shifted positive cone, the decisions and their outcomes obtained by a maximization of this function are only *weakly efficient,* i.e. the decision outcomes cannot be improved jointly but can be improved componentwise. This is illustrated in Fig. 9.1c: the decision outcomes $\mathbf{y}'$ and $\mathbf{y}''$ differ only in the coordinate y_1, hence might be both on the boundary of the cone $\bar{\mathbf{y}} + \boldsymbol{R}^2_+$; therefore, a max-min approach might produce as well $\mathbf{y}'$ as $\mathbf{y}''$, while $\mathbf{y}'$ is clearly worse than $\mathbf{y}''$ (if we maximize both objectives y_1 and y_2). Such situations often occur in practical applications, particularly with linear or discrete-linear models; therefore, the max-min approach should be used with extreme care, if at all.

For this reason, typical achievement functions used in reference point approaches do not precisely represent, but only approximate the shifted positive cone $\bar{\mathbf{y}} + D$. A specific way of this approximation was developed to obtain an important theoretical property that each properly efficient decision outcome with a given prior bound on trade-off coefficients between objectives can be obtained when maximizing an achievement function with suitably chosen reference point. This property can be guaranteed by selecting a cone D_ε "slightly broader" than the cone D and choosing an achievement function which level sets precisely represent not the cone D, but the slightly broader cone D_ε. Such theoretical property has two important practical consequences.

The first consequence concerns the concept of *proper efficiency with a prior bound* on trade-off coefficients. This is, in fact, the most practical concept of efficiency or vector-optimality[3]: decision makers do not usually care if an objective might be worsened by a small percentage of its value, if other objectives could be considerably improved instead. The second consequence concerns the

[2]Contemporary, the max-min approach is used as a tool for multi-objective optimization e.g. in the OPTIX toolbox of MATLAB – however, without warning the user that it might result in weakly efficient outcomes.

[3]Although this concept might be the most difficult to express theoretically.

possibility of obtaining *any* of such properly efficient objective outcomes. As opposed, for example, to a weighted linear aggregation of objectives, achievement functions in reference point methods can produce any desired properly efficient outcome also in nonconvex, in particular in discrete cases. This is illustrated in Fig. 9.1d: the properly efficient outcomes $\mathbf{y}^1$ and $\mathbf{y}^2$ cannot be obtained by the maximization of a linear combination of their components y_1 and y_2 with linear level sets (because $\mathbf{y}^1$ and $\mathbf{y}^2$ are contained in the convex cover of $\mathbf{y}^3$, $\mathbf{y}^4$ and $\mathbf{y}^5$), but they can be reached by maximizing an achievement function with level sets either $\bar{\mathbf{y}}^1 + D_\varepsilon$ or $\bar{\mathbf{y}}^2 + D_\varepsilon$. Observe that we can either choose $\bar{\mathbf{y}}^1 = \mathbf{y}^1$ or, more broadly, $\bar{\mathbf{y}}^2 \neq \mathbf{y}^2$; in the latter case, the maximal value of the achievement function indicates whether $\mathbf{y}^2$ is "more attainable" or "less attainable" than $\bar{\mathbf{y}}^2$.

9.2 BASIC CONCEPTS AND NOTATION

In order to discuss the above general ideas and properties in more mathematical detail we need some notation and concepts.

We distinguish here two parts of a model of a decision situation. One part, called here a *preferential model*, represents the preferences of the decision maker or DSS user (most often, the real users of decision support systems are not the final decision makers, but their advisors – analysts, modelers, designers etc.). In reference point methodology, the attention is not concentrated on the precise form of a preferential model; on the contrary, it is assumed that the preferential model might change during the decision process and the decision support tools should be flexible enough to accommodate such changes. Therefore, we typically assume that the preferential model is very general, similar to the partial order of Pareto type (which corresponds just to the desire to maximize all decision outcomes) and that the specifics of this model (say, the selection of decision outcomes to be maximized) might also change during the decision process.

The second part of a model of decision situation is called here a *substantive model* which represents the available knowledge about possible decisions and their possible outcomes. Therefore, we assume here that the general form of a substantive model is:

$$\mathbf{y} = \mathbf{f}(\mathbf{x}, \mathbf{z}), \quad \mathbf{x} \in X_0, \quad \mathbf{z} \in Z_0, \tag{9.1}$$

where $\mathbf{x} \in \boldsymbol{R}^n$ denotes a vector of *decision variables*, $\mathbf{z}$ is a *parameter vector* fixed by the modeler, X_0 is a *set of admissible decisions* which is usually defined by a set of additional inequalities or equations called *constraints*, $\mathbf{y} \in \boldsymbol{R}^m$ is a vector of *model outputs* or *decision outcomes* which includes also various intermediary variables that are useful when formulating the model, even when determining the constraints – thus, the set X_0 is often defined implicitly. The function $\mathbf{f} : \boldsymbol{R}^n \times Z_0 \to \boldsymbol{R}^m$ that determines model outputs is usually defined also implicitly, often by a quite complicated model structure. In actual applications, substantive models might express dynamic system behavior, uncertainty of results of decisions (while the outcomes $\mathbf{y}$ might be understood e.g. as math-

ematical expectations of such results, see e.g. Ermolev et al. [6]), etc. Here we assume the abstract, simple form of substantive models; we shall even suppress the dependence of this model on parameters $\mathbf{z}$ by writing $\mathbf{y} = \mathbf{f}(\mathbf{x})$.

In such a case, $Y_0 = \mathbf{f}(X_0)$ is called the set of *attainable outcomes*. It should be stressed that this set is not given explicitly (even in the simple case when $\mathbf{f}$ is given explicitly) and we can only compute its elements by assuming some $\mathbf{x} \in X_0$ and then determining the corresponding $\mathbf{y} = \mathbf{f}(\mathbf{x})$ by simulating the model.

The modeler, when analyzing the substantive model, might specify several model outputs as especially interesting – we call them *objectives* or *criteria* and shall denote by $q_i = y_j$, forming an *objective vector* $\mathbf{q} \in R^k$ – a vector in the *objective space*. While this vector and space might change during the decision process, we shall denote the relation between decisions and objectives by $\mathbf{q} = \mathbf{F}(\mathbf{x}, \mathbf{z})$ or shorten it to $\mathbf{q} = \mathbf{F}(\mathbf{x})$. $Q_0 = \mathbf{F}(X_0)$ is called the *set of attainable objectives*.

Since we can change minimization to maximization by changing the sign of an objective, we can as well assume that all objectives are, say, maximized. Recall that a *Pareto-optimal* decision and its outcomes are such that there are no other admissible decisions which would improve any objective without deteriorating other objectives. A closely related, but slightly broader and weaker concept is that of *weakly Pareto-optimal* decision and outcomes: these are such that there are no other admissible decisions which would result in a joint improvement of all objectives. This concept is actually too weak for applications, as already indicated.

In fact, even the concept of Pareto-optimality is sometimes too weak for applications, in cases where we could improve significantly one objective component at the cost of an infinitesimally small deterioration of another objective. The (limits of) ratios of improvements and deteriorations of objectives, determined at a Pareto-optimal decision, are called *trade-off coefficients*; we define *properly Pareto-optimal* decisions and outcomes as such that the corresponding trade-off coefficients are bounded. Even this concept is too weak for applications, since the mathematical sense of "bounded" means "anything smaller than infinity". Truly important for applications are rather decisions and outcomes which are *properly Pareto-optimal with a prior bound*, i.e. such that a finite bound on trade-off coefficients is *a priori* given and satisfied.

In each of these specific cases of Pareto-optimality (weak, proper, etc.), the sets of Pareto-optimal decisions and outcomes contain typically many elements, not just a singleton decision and its outcome. Thus, Pareto-optimality does not tell us, which decision to choose, it tells us only which decisions to avoid. This non-uniqueness of Pareto-optimal decisions has been considered a drawback in the classical decision analysis; thus, on top of a substantive model, a preferential model was usually assumed in the form of at least weak order which could be specified by a given utility or value function whose maximum defined – hopefully, uniquely – "the optimal" decision and outcome. However, in *interactive decision support*, when we assume that the preferences of the user of the DSS

(or the modeler, the analysts etc.) can change during the decision process, the non-uniqueness of Pareto-optimal decisions is an advantage, not a drawback. We need only an additional way of controlling the selection of Pareto-optimal decisions by parameters specified by the user.

We recall that Pareto-optimality can be generalized by using a partial order implied by a positive cone, while the positive cone indicates what we understand by an improvement in the space of objectives. In the case of Pareto-optimality (if all objectives are maximized), the positive cone is the positive "part" of the objective space:

$$D = \boldsymbol{R}^k_+ = \{\mathbf{q} \in \boldsymbol{R}^k : \; q_i \geq 0 \; \forall i = 1, \dots k\}. \tag{9.2}$$

A *strictly positive cone* (assuming an improvement of at least one objective component, which is needed for the definition of Pareto-optimality) can be written as:

$$\tilde{D} = \boldsymbol{R}^k_+ \setminus \{0\} = \{\mathbf{q} \in \boldsymbol{R}^k : \; q_i \geq 0 \forall i = 1, \dots k; \; \exists i = 1, \dots k : q_i > 0\}. \tag{9.3}$$

A *strongly positive cone* (assuming an improvement of all objective components, as needed in the definition of weak Pareto-optimality) is defined simply as the interior of the positive cone, $IntD = Int\boldsymbol{R}^k_+$.

In the case when some objectives (from 1 to k_1) are maximized, some (from $k_1 + 1$ to k_2) are minimized and some (from $k_2 + 1$ to k) are *stabilized* (i.e. kept close to a given reference level), the positive cone can be defined as:

$$\begin{aligned} D \;=\; & \{\mathbf{q} \in \boldsymbol{R}^k : \; q_i \geq 0, \; i = 1, \dots k_1, \; q_i \leq 0, \; i = k_1 + 1, \dots k_2, \\ & q_i = 0, \; i = k_2 + 1, \dots k\}. \end{aligned} \tag{9.4}$$

Note that the cone describes only changes in objective values, hence $q_i = 0$ means that the objective component is kept equal to its reference level. If we define similarly the strictly positive cone as $\tilde{D} = D \setminus \{0\}$ and the strongly positive cone as $IntD$, we can give a more general definition of Pareto-optimality, called *efficiency with respect to the cone* D; the *set of efficient objectives or outcomes* is defined as:

$$\hat{Q}_0 = \{\hat{\mathbf{q}} \in Q_0 : \; (\hat{\mathbf{q}} + \tilde{D}) \cap Q_0 = \emptyset\} \tag{9.5}$$

and the set of efficient decisions is defined equivalently, while taking into account that $\hat{\mathbf{q}} = \mathbf{F}(\hat{\mathbf{x}})$, as:

$$\hat{X}_0 = \{\hat{\mathbf{x}} \in X_0 : \; (\mathbf{F}(\hat{\mathbf{x}}) + \tilde{D}) \cap \mathbf{F}(X_0) = \emptyset\}. \tag{9.6}$$

If $D = \boldsymbol{R}^k_+$ and $\tilde{D} = \boldsymbol{R}^k_+ \setminus \{0\}$, then the above definition of efficiency coincides with the descriptive definition of Pareto-optimality given earlier. The generalization of weak Pareto optimality to weak efficiency is obtained by simply replacing the strictly positive cone $\tilde{D}$ with the strongly positive cone $IntD$:

$$\hat{Q}^w_0 = \{\hat{\mathbf{q}} \in Q_0 : \; (\hat{\mathbf{q}} + IntD) \cap Q_0 = \emptyset\} \tag{9.7}$$

and:

$$\hat{X}_0^w = \{\hat{\mathbf{x}} \in X_0 : (\mathbf{F}(\hat{\mathbf{x}}) + IntD) \cap \mathbf{F}(X_0) = \emptyset\}. \tag{9.8}$$

Note that if $k > k_2$ (there are stabilized objectives), then the cone (9.4) has empty interior, hence $\hat{Q}_0^w = Q_0$ and the concept of weak efficiency is quite useless in such a case.

In order to define trade-off coefficients, we shall assume here, for simplicity, that all objectives are dimension-free and can be directly compared (we shall relax this assumption later). At an efficient point $\hat{\mathbf{x}} \in \hat{X}_0$ with $\hat{\mathbf{q}} = F(\hat{\mathbf{x}}) \in \hat{Q}_0$, if the efficient frontier is smooth at this point, the *local trade-off coefficient* $t_{ij}(\hat{\mathbf{q}})$ between maximized objectives q_i, q_j is defined as:

$$t_{ij}(\hat{\mathbf{q}}) = \lim_{l\to\infty} \sup_{\mathbf{q}^{(l)} \in \hat{Q}_0} \frac{q_i^{(l)} - \hat{q}_i}{\hat{q}_j - q_j^{(l)}};\ \lim_{l\to\infty} \mathbf{q}^{(l)} = \hat{\mathbf{q}}, \tag{9.9}$$

where the supremum is taken over all sequences $\{\mathbf{q}^{(l)}\}_{l=1}^{\infty} \subset \hat{Q}_0$ converging to $\hat{\mathbf{q}}$. If an objective i or j is minimized, the sign of the appropriate increment in the above equation must be changed, etc. For non-convex sets Q_0, it is useful to define also *global trade-off coefficients* which might be greater (but not in the convex case) than the local ones:

$$\begin{aligned} t_{ij}(\hat{\mathbf{q}}) &= \sup_{\mathbf{q} \in Q^{(j)}(\hat{\mathbf{q}})} \frac{q_i - \hat{q}_i}{\hat{q}_j - q_j}, \\ Q^{(j)}(\hat{\mathbf{q}}) &= \{\mathbf{q} \in Q_0 : q_j < \hat{q}_j,\ q_i \geq \hat{q}_i,\} \end{aligned} \tag{9.10}$$

with the signs of inequalities in the definition of $Q^{(j)}(\hat{\mathbf{q}})$ appropriately changed for minimized objectives.

The computation of trade-off coefficients according to their definitions is a difficult problem, see e.g. Kaliszewski [19]. It turns out that we can obtain bounds on trade-off coefficients if we express the concept of proper efficiency in terms of modified positive cones. There are various approaches to such representation – see e.g. Henig [16], Sawaragi et al. [38]. It can be shown – see Wierzbicki [48, 50], Kaliszewski [19] – that properly efficient outcomes and decisions with a prior bound M on trade-off coefficients can be defined as weakly efficient outcomes and decisions with respect to a "slightly broader" positive cone. For this purpose, we define first $\varepsilon = 1/(M-1)$ (note that there is no sense in considering $M \leq 1$) and define an ε-neighborhood $IntD_\varepsilon$ of the positive cone D:

$$IntD_\varepsilon = \{\mathbf{q} \in \boldsymbol{R}^k : dist(\mathbf{q}, D) < \varepsilon \parallel \mathbf{q} \parallel\}, \tag{9.11}$$

where we could choose any norm in $\boldsymbol{R}^k$ and a (Haussdorff) concept of distance between the point $\mathbf{q}$ and the set D in order to obtain an open[4] set $IntD_\varepsilon$.

[4]The concept of distance can correspond even to another norm in $\boldsymbol{R}^k$ than on the right-side, since all norms in $\boldsymbol{R}^k$ are topological equivalent.

However, in order to obtain the needed bound on trade-off coefficients, it is useful to choose rather specific norms: l_1 on the right-hand side and mixed l_1 and l_∞ for the distance on the left-hand side. Without specifying the details here[5] we note that, if $D = \boldsymbol{R}^k_+$, in case of Pareto-optimality, the cone D_ε of the form (9.11) can be also written as:

$$\begin{aligned} D_\varepsilon &= \{\mathbf{q} \in \boldsymbol{R}^k : \mathbf{q} = \sum_{j=1}^{k} \lambda_j \mathbf{q}_\varepsilon^{(j)},\ \lambda_j \geq 0\}, \\ \mathbf{q}_\varepsilon^{(j)} &= (-\varepsilon, -\varepsilon, \ldots 1 + (k-1)\varepsilon_{(j)}, \ldots, -\varepsilon, -\varepsilon)^T; \\ D_\varepsilon &= \{\mathbf{q} \in \boldsymbol{R}^k : -q_j \leq \varepsilon \sum_{i=1}^{k} q_i,\ j = 1, \ldots k\} \\ &= \{\mathbf{q} \in \boldsymbol{R}^k : \min_{1 \leq i \leq k} q_i + \varepsilon \sum_{i=1}^{k} q_i \geq 0\}. \end{aligned} \tag{9.12}$$

The last representation is particularly important: D_ε can be represented as a zero-level set of the function $\min_{1 \leq i \leq k} q_i + \varepsilon \sum_{i=1}^{k} q_i$.

Note that $IntD_\varepsilon \neq \emptyset$ even if $IntD = \emptyset$ (as stressed before, the later holds for cones D of the form (9.4) including some stabilized objectives). Moreover, if we define weakly efficient solutions with respect to the "broader" cone $IntD_\varepsilon$, we can prove that they are equivalent to properly efficient solutions with (global, not only local) trade-off coefficients bounded a priori by $M = 1 + 1/\varepsilon$; we shall call such outcomes and decisions *ε-properly efficient*. Thus, the sets of ε-properly efficient outcomes and ε-properly efficient decisions can be defined as:

$$\hat{Q}_0^{p\varepsilon} = \{\hat{\mathbf{q}} \in Q_0 : (\hat{\mathbf{q}} + IntD_\varepsilon) \cap Q_0 = \emptyset\} \tag{9.13}$$

and:

$$\hat{X}_0^{p\varepsilon} = \{\hat{\mathbf{x}} \in X_0 : (\mathbf{F}(\hat{\mathbf{x}}) + IntD_\varepsilon) \cap \mathbf{F}(X_0) = \emptyset\}. \tag{9.14}$$

The traditional proper efficiency – with only an existential bound on trade-off coefficients – can be then defined by:

$$\hat{Q}_0^{p} = \bigcup_{\varepsilon > 0} \hat{Q}_0^{p\varepsilon},\ \hat{X}_0^{p} = \bigcup_{\varepsilon > 0} \hat{X}_0^{p\varepsilon}. \tag{9.15}$$

In any case (convex or not) the definitions of various types of efficiency imply that:

$$\hat{Q}_0^{p\varepsilon} \subseteq \hat{Q}_0^{p} \subseteq \hat{Q}_0 \subseteq \hat{Q}_0^{w};\ \hat{X}_0^{p\varepsilon} \subseteq \hat{X}_0^{p} \subseteq \hat{X}_0 \subseteq \hat{X}_0^{w}. \tag{9.16}$$

After specifying any variables in a model as objectives q_i, we should first know – at least approximately – the ranges in which these variables might

[5] See Wierzbicki [48, 50] for more detailed derivation.

vary. This is also important because we shall often *aggregate objectives* – that is, combine them into one function (not necessarily by summation) – and many objectives might have various units of measurement and must be re-scaled to dimension-free units before aggregation. Thus, any system supporting vector optimization must include a function of estimating such ranges.

The usual way of such estimation is to compute the *ideal* or *utopia point* by optimizing separately each objective and to estimate its counterpart – the *nadir point* (a lower bound on objectives that are maximized, upper on those minimized). While the utopia point components do not usually change, if we change the number of objectives selected (we might need to compute utopia components for new objectives, but they do not influence old objectives), the nadir point components do change.

This is because of the difference in definitions of the utopia and nadir points. The utopia point consists of best values of objectives in both the sets Q_0 and $\hat{Q}_0$ – but it is simpler to compute the best values in the larger set Q_0. The nadir point consist of worst values of objectives, but only in the smaller set of efficient outcomes $\hat{Q}_0$ – there might be worse values, than at nadir point, in the set of non-efficient points $Q_0 \setminus \hat{Q}_0$. Although the computations of the nadir point might be quite difficult (see e.g. Korhonen et al. [23]), the information contained in this point is quite important; therefore, we need at least an approximation of the nadir point.

A simple way (though certainly not the best) of such approximation of nadir components is to take the worst values of objective components that occur while computing the best values of other components during the calculations of the utopia point:

$$
\begin{aligned}
q_{i,uto} &= \max_{\mathbf{q}\in Q_0} q_i,\ \hat{\mathbf{q}}^{(i)} = \operatorname*{argmax}_{\mathbf{q}\in Q_0} q_i, \\
i &= 1,\ldots,k_1 \text{ (for maximized objectives)}, \\
q_{i,uto} &= \min_{\mathbf{q}\in Q_0} q_i,\ \hat{\mathbf{q}}^{(i)} = \operatorname*{argmin}_{\mathbf{q}\in Q_0} q_i, \\
i &= k_1+1,\ldots,k_2 \text{ (for minimized objectives)}, \\
q^{(1)}_{i,nad} &= \min_{1\le j\le k} \hat{q}_i^{(j)},\ i = 1,\ldots,k_1 \text{ (max)}, \\
q^{(1)}_{i,nad} &= \max_{1\le j\le k} \hat{q}_i^{(j)},\ i = k_1+1,\ldots,k_2 \text{ (min)}.
\end{aligned}
\tag{9.17}
$$

Such worst values $q^{(1)}_{i,nad}$ might be still better than the actual nadir components (they are equal to nadir components only in some special cases, including the case $k = 2$). Thus, in order to estimate the nadir approximately, it is sufficient to increase the range $q_{i,uto} - q^{(1)}_{i,nad}$ somewhat arbitrarily. There exist various ways of further improvement of estimates of nadir components, see e.g. Lewandowski at al. [25]. In any case, we can assume that there are defined (either arbitrarily or by computing the utopia point and estimating the nadir point) some estimates of ranges of each objective values:

$$q_{i,lo} \leq q_i \leq q_{i,up}, \quad i = 1, \ldots, k, \tag{9.18}$$

where $q_{i,up}$ for maximized objectives ($q_{i,lo}$ for minimized ones) is at least as high (low) as the corresponding utopia point component and the range $q_{i,up} - q_{i,lo}$ is approximately as large as the range utopia-nadir. First after specifying such ranges, we can reduce objectives to dimension-free (e.g. percentage) scales and then speak about relative importance of criteria, their weights, interpret the trade-off coefficients, etc.

9.3 REFERENCE POINTS AND ACHIEVEMENT FUNCTIONS

We assume here that for each objective – which can be maximized, minimized or stabilized – *reference levels* in the form of either *aspiration levels* $\bar{q}_i$ (which would be good to achieve) or, additionally, *reservation levels* $\bar{\bar{q}}_i$ (which should be achieved if it is at all possible) are specified by the modeler. These reference levels will be used as main interaction parameters by which the user of a DSS controls the selection of decisions and their outcomes. The values of these reference levels are subject to reasonability constraints only, given lower and upper bounds $q_{i,lo}$, $q_{i,up}$ for each objective:

$$\begin{array}{lllllll} q_{i,lo} & < & \bar{\bar{q}}_i & < & \bar{q}_i & < & q_{i,up}, \quad i = 1, \ldots, k_1 \quad \text{(max)}, \\ q_{i,lo} & < & \bar{q}_i & < & \bar{\bar{q}}_i & < & q_{i,up}, \quad i = k_1 + 1, \ldots, k_2 \quad \text{(min)}. \end{array} \tag{9.19}$$

For stabilized outcomes we can use two pairs of reservation and aspiration levels: one "lower" pair $\bar{\bar{q}}_{i,lo} < \bar{q}_{i,lo}$ as for maximized outcomes and one "upper" pair $\bar{q}_{i,up} < \bar{\bar{q}}_{i,up}$ as for minimized ones.

A way of aggregating the objectives into an order-consistent achievement function[6] consists in specifying *partial achievement functions* $\sigma_i(q_i, \bar{q}_i)$ or $\sigma_i(q_i, \bar{q}_i, \bar{\bar{q}}_i)$ which should:

a) *be strictly monotone* consistently with the specified partial order – increasing for maximized objectives, decreasing for minimized ones, increasing below (lower) aspiration level and decreasing above (upper) aspiration level for stabilized ones;

b) *assume value 0* if $q_i = \bar{q}_i$, $\forall i = 1, \ldots, k$, and aspiration levels are used alone – or *assume value 0* if $q_i = \bar{\bar{q}}_i$, $\forall i = 1, \ldots, k$, *and assume value 1* if $q_i = \bar{q}_i$, $\forall i = 1, \ldots, k$, if both aspiration and reservation levels are used.

This seeming inconsistency results from the fact that the number 0 is more important than the number 1: if the aspiration levels are used alone, we just check with the help of the sign of an achievement function, whether they could be reached. In such a case, it is useful to define partial achievement functions with a slope that is larger if the aspiration levels are closer to their extreme levels:

[6]For a more detailed theory of such functions see e.g. Wierzbicki, [48].

$$
\begin{aligned}
\sigma_i(q_i, \bar{q}_i) &= (q_i - \bar{q}_i)/(q_{i,up} - \bar{q}_i) \quad (\max), \\
\sigma_i(q_i, \bar{q}_i) &= (\bar{q}_i - q_i)/(\bar{q}_i - q_{i,lo}) \quad (\min), \\
\sigma_i(q_i, \bar{q}_i) &= \left\{ \begin{array}{l} (\bar{q}_i - q_i)/(q_{i,up} - \bar{q}_i), \text{ if } q_i > \bar{q}_i \\ (q_i - \bar{q}_i)/(\bar{q}_i - q_{i,lo}), \text{ if } q_i \leq \bar{q}_i \end{array} \right\} \text{(stab)}.
\end{aligned}
\tag{9.20}
$$

where $\bar{q}_{i,lo} = \bar{q}_{i,up} = \bar{q}_i$ was assumed for stabilized objectives. An alternative way is to use piece-wise linear functions, e.g. to change the slope of the partial achievement function depending on whether the current point is above or below the aspiration point:

$$
\begin{aligned}
\underset{(\max)}{\sigma_i(q_i, \bar{q}_i)} &= \left\{ \begin{array}{llll} (q_i - \bar{q}_i)/(q_{i,up} - \bar{q}_i), & \bar{q}_i \leq q_i \leq q_{i,up} \\ \beta(q_i - \bar{q}_i)/(\bar{q}_i - q_{i,lo}), & q_{i,lo} < q_i \leq \bar{q}_i \end{array} \right\}, \\
\underset{(\min)}{\sigma_i(q_i, \bar{q}_i)} &= \left\{ \begin{array}{llll} (\bar{q}_i - q_i)/(\bar{q}_i - q_{i,lo}), & q_{i,lo} \leq q_i \leq \bar{q}_i \\ \beta(\bar{q}_i - q_i)/(q_{i,lo} - \bar{q}_i), & \bar{q}_i < q_i \leq q_{i,up} \end{array} \right\}, \\
\underset{(\text{stab})}{\sigma_i(q_i, \bar{q}_i)} &= \left\{ \begin{array}{llll} (q_{i,up} - q_i)/(q_{i,up} - \bar{q}_i), & \bar{q}_i \leq q_i \leq q_{i,up} \\ (q_i - q_{i,lo})/(\bar{q}_i - q_{i,lo}), & q_{i,lo} \leq q_i < \bar{q}_i \end{array} \right\},
\end{aligned}
\tag{9.21}
$$

where the coefficient $\beta > 0$ is selected in such a way that the functions are not only monotone, but also concave (thus can be expressed as minima of their component linear functions, which is useful for their applications together with linear models).

If both aspiration and reservation levels are used, it is more useful to define the partial achievement functions as piece-wise linear functions e.g. of the form:

$$
\begin{aligned}
\underset{(\max)}{\sigma_i(q_i, \bar{q}_i, \bar{\bar{q}}_i)} &= \left\{ \begin{array}{llll} 1 + \alpha(q_i - \bar{q}_i)/(q_{i,up} - \bar{q}_i), & \bar{q}_i \leq q_i \leq q_{i,up} \\ (q_i - \bar{\bar{q}}_i)/(\bar{q}_i - \bar{\bar{q}}_i), & \bar{\bar{q}}_i < q_i < \bar{q}_i \\ \beta(q_i - \bar{\bar{q}}_i)/(\bar{\bar{q}}_i - q_{i,lo}), & q_{i,lo} \leq q_i \leq \bar{\bar{q}}_i \end{array} \right\}, \\
\underset{(\min)}{\sigma_i(q_i, \bar{q}_i, \bar{\bar{q}}_i)} &= \left\{ \begin{array}{llll} 1 + \alpha(\bar{q}_i - q_i)/(\bar{q}_i - q_{i,lo}), & q_{i,lo} \leq q_i \leq \bar{q}_i \\ (\bar{\bar{q}}_i - q_i)/(\bar{\bar{q}}_i - \bar{q}_i), & \bar{q}_i < q_i < \bar{\bar{q}}_i \\ \beta(\bar{\bar{q}}_i - q_i)/(q_{i,lo} - \bar{\bar{q}}_i), & \bar{\bar{q}}_i \leq q_i \leq q_{i,up} \end{array} \right\}, \\
\underset{(\text{stab})}{\sigma_i(q_i, \bar{q}_i, \bar{\bar{q}}_i)} &= \left\{ \begin{array}{llll} \beta(\bar{\bar{q}}_{i,up} - q_i)/(q_{i,up} - \bar{\bar{q}}_{i,up}), & \bar{\bar{q}}_{i,up} \leq q_i \leq q_{i,up} \\ (\bar{\bar{q}}_{i,up} - q_i)/(\bar{\bar{q}}_{i,up} - \bar{q}_i), & \bar{q}_i < q_i < \bar{\bar{q}}_{i,up} \\ (q_i - \bar{\bar{q}}_{i,lo})/(\bar{q}_i - \bar{\bar{q}}_{i,lo}), & \bar{\bar{q}}_{i,lo} < q_i < \bar{q}_i \\ \beta(q_i - \bar{\bar{q}}_{i,lo})/(\bar{\bar{q}}_{i,lo} - q_{i,lo}), & q_{i,lo} \leq q_i \leq \bar{\bar{q}}_{i,lo} \end{array} \right\}.
\end{aligned}
\tag{9.22}
$$

The coefficients α, β should be positive and chosen in such a way that partial achievement functions are not only monotone, but also concave. Other forms of piece-wise linear partial achievement functions satisfying these conditions are

also possible – e.g. an achievement function for stabilized objectives might be defined as greater than 1 inside the interval $[\bar{q}_{i,lo}; \bar{q}_{i,up}]$ if $\bar{q}_{i,lo} < \bar{q}_{i,up}$, see Fig. 9.2.

If the values of $\sigma_i(q_i, \bar{q}_i, \bar{\bar{q}}_i)$ would be restricted to the interval [0;1], then they could be interpreted as *fuzzy membership functions* $\mu_i(q_i, \bar{q}_i, \bar{\bar{q}}_i)$ (see e.g. Zadeh [53], Seo et al. [39], Zimmermann et al. [55]) which express the degree of satisfaction of the modeler with the value of the objective q_i. More complicated forms of such fuzzy membership functions can be also used, see e.g. Vincke [44], Fodor and Roubens [7], Granat et al. [12]; for illustrative simplicity, we shall not consider these more complicated forms here.

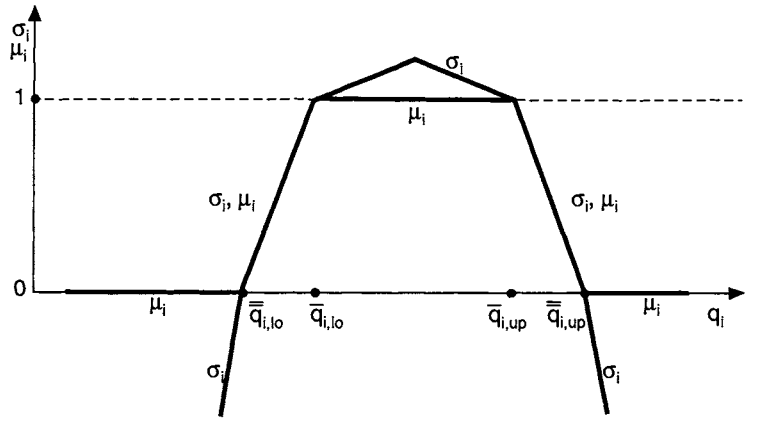


Figure 9.2: The difference between a partial achievement function $\sigma_i(q_i, \bar{q}_i, \bar{\bar{q}}_i)$ and a corresponding fuzzy membership function $\mu_i(q_i, \bar{q}_i, \bar{\bar{q}}_i)$ in the case of a stabilized objective.

A partial achievement function can be looked upon as simply a nonlinear transformation of the objective range satisfying some monotonicity requirements. The essential issue is how to aggregate these functions as to obtain a scalarizing achievement function with good properties for vector optimization or multi-objective model analysis. There are several ways of such aggregation. One way is to use fuzzy logic and select an appropriate representation of the "fuzzy and" operator[7]. The simplest operator of this type is the minimum operator:

$$\mu(\mathbf{q}, \bar{\mathbf{q}}, \bar{\bar{\mathbf{q}}}) = \bigwedge_{1 \le i \le k} \mu_i(q_i, \bar{q}_i, \bar{\bar{q}}_i) = \min_{1 \le i \le k} \mu_i(q_i, \bar{q}_i, \bar{\bar{q}}_i,) \tag{9.23}$$

[7]In selecting "fuzzy and" operator for aggregation, we actually assume that all objectives are similarly important and non-compensative. This assumption is fully justified in multi-objective model analysis (we do not ask the modeler for reasons why she/he has selected a given set of objectives), but it might be not necessarily satisfied in other cases of aggregation of attributes.

which, however, would result only in weakly Pareto-optimal or weakly efficient outcomes when used for multi-objective analysis. To secure obtaining ε-properly efficient outcomes, we have to augment this operator by some linear part (compare the last expression for the cone D_ε in (9.12)). The corresponding overall membership function would then have the form:

$$(9.24) \qquad \mu(\mathbf{q}, \bar{\mathbf{q}}, \bar{\bar{\mathbf{q}}}) = (\min_{1 \le i \le k} \mu_i(q_i, \bar{q}_i, \bar{\bar{q}}_i) + \varepsilon \sum_{i=1}^{k} \mu_i(q_i, \bar{q}_i, \bar{\bar{q}}_i))/(1 + k\varepsilon).$$

An interpretation in terms of membership functions can be in fact used in a graphic interaction with the modeler; however, membership functions $\mu_i(q_i, \bar{q}_i, \bar{\bar{q}}_i)$ and $\mu(\mathbf{q}, \bar{\mathbf{q}}, \bar{\bar{\mathbf{q}}})$ are not strictly monotone if they are equal to 0 or 1. Therefore, inside a vector optimization system, a slightly different overall achievement function must be used, with values not restricted to the interval [0;1]:

$$(9.25) \qquad \sigma(\mathbf{q}, \bar{\mathbf{q}}, \bar{\bar{\mathbf{q}}}) = (\min_{1 \le i \le k} \sigma_i(q_i, \bar{q}_i, \bar{\bar{q}}_i) + \varepsilon \sum_{i=1}^{k} \sigma_i(q_i, \bar{q}_i, \bar{\bar{q}}_i))/(1 + k\varepsilon).$$

In both above equations, $\varepsilon > 0$ is the same coefficient as the one used when defining the proper ε-efficiency, with a prior bound $M = 1 + 1/\varepsilon$ on corresponding trade-off coefficients. Actually, this bound limits here trade-off coefficients not between various objectives q_i and q_j, but between their transformed values $\sigma_i(q_i, \bar{q}_i, \bar{\bar{q}}_i)$ and $\sigma_j(q_j, \bar{q}_j, \bar{\bar{q}}_j)$; in order to obtain bounds on original trade-off coefficients between q_i and q_j, it is necessary to take into account the current slopes of partial achievement functions. However, if these slopes have prior bounds, the original trade-off coefficients will also have prior bounds.

The above derivation of an order-consistent achievement function from a "fuzzy and" operator is not the only one possible. In fact, simpler versions of order-consistent achievement functions were used originally. Some of such versions can be looked upon as a simplification of function (9.25). For example, suppose only aspiration levels $\bar{q}_i$ are used, all objectives are maximized and dimension-free and the partial achievement functions have a simple form $\sigma_i(q_i, \bar{q}_i) = q_i - \bar{q}_i$. Then the order-consistent achievement function takes on the form:

$$(9.26) \qquad \sigma(\mathbf{q}, \bar{\mathbf{q}}) = \min_{1 \le i \le k} (q_i - \bar{q}_i) + \varepsilon \sum_{i=1}^{k} (q_i - \bar{q}_i),$$

where we do not have to subdivide by $1 + k\varepsilon$ because only the value 0, not 1, of this function is significant. This function can be seen as a prototype order-consistent achievement scalarizing function. It is monotone with respect to the cone $IntD_\varepsilon$ and its zero-level set represents this cone – compare (9.12):

$$(9.27) \qquad \bar{\mathbf{q}} + IntD_\varepsilon = \{\mathbf{q} \in \boldsymbol{R}^k : \sigma(\mathbf{q}, \bar{\mathbf{q}}) > 0\}.$$

Since function (9.25) is also strictly monotone with respect to the cone $IntD_\varepsilon$, we have:

• **Sufficient condition for ε-proper efficiency.** For any $\bar{\mathbf{q}}, \bar{\bar{\mathbf{q}}}$ (with components strictly contained in the ranges $[q_{i,lo}; q_{i,up}]$) a maximal point of $\sigma(\mathbf{q}, \bar{\mathbf{q}}, \bar{\bar{\mathbf{q}}})$ with respect to $\mathbf{q} \in Q_0 = \mathbf{F}(X_0)$ is a properly efficient objective vector with a prior bound on trade-off coefficients and, equivalently, a maximal point of $\sigma(\mathbf{F}(\mathbf{x}), \bar{\mathbf{q}}, \bar{\bar{\mathbf{q}}})$ with respect to $\mathbf{x} \in X_0$ is a properly efficient decision with a prior bound.

In order to derive a corresponding necessary condition, consider $\sigma(\mathbf{q}, \bar{\mathbf{q}}, \bar{\bar{\mathbf{q}}})$ as a function not of $\mathbf{q} = (q_1, \ldots q_i, \ldots q_k)^T$ but of their transformed values $\mathbf{y} = (y_1, \ldots y_i, \ldots y_k)^T, y_i = \sigma_i(q_i, \bar{q}_i, \bar{\bar{q}}_i)$. In the transformed space, the reservation point $\bar{\bar{\mathbf{y}}} = 0$, since $\bar{\bar{y}}_i = \sigma_i(\bar{\bar{q}}_i, \bar{q}_i, \bar{\bar{q}}_i) = 0$. Denote by $\rho(\mathbf{y}) = \sigma(\mathbf{q}, \bar{\mathbf{q}}, \bar{\bar{\mathbf{q}}})$ the achievement scalarizing function in the transformed space. Then, according to (9.12), we can write the cone D_ε (actually, with its vertex shifted to $\bar{\bar{\mathbf{y}}}$, which is conveniently equal to 0 in this case) in the following form:

$$\bar{\bar{\mathbf{y}}} + D_\varepsilon = \{\mathbf{y} \in \boldsymbol{R}^k : \ \rho(\mathbf{y}) \geq 0\} \tag{9.28}$$

and, when taking into account the monotonicity of $\rho(\mathbf{y})$, we obtain similarly as in (9.27):

$$\bar{\bar{\mathbf{y}}} + IntD_\varepsilon = \{\mathbf{y} \in \boldsymbol{R}^k : \ \rho(\mathbf{y}) > 0\}. \tag{9.29}$$

Now, suppose $\hat{\mathbf{q}} = \mathbf{F}(\hat{\mathbf{x}})$ is a properly efficient outcome of an admissible decision $\hat{\mathbf{x}} \in Q_0$ with such bounds on trade-off coefficients that they are less than $M = 1 + 1/\varepsilon$ in the transformed space of $y_i = \sigma_i(q_i, \bar{q}_i, \bar{\bar{q}}_i)$. Let us shift the reservation point to this properly efficient point, $\bar{\bar{\mathbf{q}}} = \hat{\mathbf{q}}$. According to the definition of the ε-proper efficiency, the cone $\bar{\bar{\mathbf{y}}} + IntD_\varepsilon$ cannot intersect the (transformed by $y_i = \sigma_i(q_i, \bar{q}_i, \bar{\bar{q}}_i)$) set Q_0. However, relation (9.29) indicates that, in such a case, $\hat{\mathbf{y}} = \bar{\bar{\mathbf{y}}} = 0$ corresponding to $\hat{\mathbf{q}} = \bar{\bar{\mathbf{q}}}$ will be a maximal point of $\rho(\mathbf{y})$ in the transformed set Q_0, or, equivalently, $\hat{\mathbf{q}}$ *will be a maximal point of* $\sigma(\mathbf{q}, \bar{\mathbf{q}}, \bar{\bar{\mathbf{q}}})$ *with respect to* $\mathbf{q} \in Q_0$.

Such a way of deriving the necessary conditions of efficiency is actually an adaptation of the concept of separation of sets to the case of nonlinear separating functions which represent conical sets: the function $\rho(\mathbf{y})$ separates (by a cone) the sets $\bar{\bar{\mathbf{y}}} + D_\varepsilon$ and transformed Q_0, if $\bar{\bar{\mathbf{q}}} = \hat{\mathbf{q}}$, see Wierzbicki [48, 50]. We conclude that we have:

• **Necessary condition for ε-proper efficiency.** For any properly efficient $\hat{\mathbf{q}} = \mathbf{F}(\mathbf{x})$ with appropriate prior bounds on trade-off coefficients, there exist $\bar{\bar{\mathbf{q}}}$ and/or $\bar{\mathbf{q}}$ such that $\hat{\mathbf{q}}$ maximizes $\sigma(\mathbf{q}, \bar{\mathbf{q}}, \bar{\bar{\mathbf{q}}})$ with respect to $\mathbf{q} \in Q_0 = \mathbf{F}(X_0)$.

Actually, we can prove even more – see Wierzbicki [48]: the user can influence the selection of $\hat{\mathbf{q}} = \mathbf{F}(\hat{\mathbf{x}})$ Lipschitz-continuously by changing $\bar{\bar{\mathbf{q}}}$ and/or $\bar{\mathbf{q}}$ (except in cases when the set of properly efficient objectives is disjoint). We say that this selection is *continuously controllable.*

Moreover, the scaling of the partial achievement functions and the scalarizing achievement function is such that the user can draw easily:

• **Conclusions on the attainability of reservation and/or aspiration points.** If the maximal value of $\sigma(\mathbf{q}, \bar{\mathbf{q}}, \bar{\bar{\mathbf{q}}})$ with respect to $\mathbf{q} \in Q_0 = \mathbf{F}(X_0)$ is below 0, it indicates that the reservation point is not attainable, $\bar{\bar{\mathbf{q}}} \notin Q_0 =$

$\mathbf{F}(X_0)$, and also that there are no points $\mathbf{q} \in Q_0$ dominating $\bar{\bar{\mathbf{q}}}$, i.e. $\{\mathbf{q} \in R^k : \mathbf{q} \geq \bar{\bar{\mathbf{q}}}\} \cap Q_0 = \emptyset$. If this maximal value is 0, it indicates that the reservation point is attainable and properly efficient. If this maximal value is 1, the same can be said about the aspiration point $\bar{\mathbf{q}}$. Similar conclusions concerning the values between 0 and 1 and above 1 can be made. If we use aspiration levels alone, there is only one critical value 0 of the achievement function corresponding to the aspiration point $\bar{\mathbf{q}}$.

This property justifies the name "achievement function" since its values measure the achievement as compared to aspiration and reservation points. The name "order-consistent" achievement function is used to indicate that the function is strictly monotone with respect to the cone $IntD_\varepsilon$, hence it preserves the partial order implied by the cone, and its zero-level-set corresponds to the set $\bar{\bar{\mathbf{q}}} + D_\varepsilon$, hence it represents this order.

The achievement function $\sigma(\mathbf{q}, \bar{\mathbf{q}}, \bar{\bar{\mathbf{q}}})$ – and other similar functions – is nondifferentiable. Moreover, the maximum of this achievement function is in most cases attained at its "corner", i.e. at the point of nondifferentiability. In the case of linear models, the nondifferentiability of the achievement function $\sigma(\mathbf{q}, \bar{\mathbf{q}}, \bar{\bar{\mathbf{q}}})$ does not matter, since the function is concave and its maximization can be equivalently expressed as a linear programming problem by introducing dummy variables – see next section.

In the case of nonlinear models, however, optimization algorithms for smooth functions are more robust (work more reliably without the necessity of adjusting their specific parameters to obtain results) than algorithms for nonsmooth functions. Therefore, there are two approaches to the maximization of such achievement functions. One is to introduce additional constraints and dummy variables as for linear models. Another is a useful modification of the achievement function by its smooth approximation, which can be defined e.g. when using an l_p norm (with $p > 2$, because a circle or a ball rather badly approximates the piece-wise linear achievement function; it would be best approximated by very large p, but usually $p = 4 \ldots 8$ suffices, since larger p result in badly conditioned optimization problems). We quote here such an approximation only for the case of using aspiration point $\bar{\mathbf{q}}$ alone, assuming that partial achievement functions $\sigma_i(q_i, \bar{q}_i) \leq 1$ e.g. $\sigma_i(q_i, \bar{q}_i) = 1$ if $q_i = q_{i,up}$ for maximized objectives):

$$\sigma(\mathbf{q}, \bar{\mathbf{q}}) = 1 - \left(\frac{1}{k}\sum_{i=1}^{k}(1 - \sigma_i(q_i, \bar{q}_i))^p\right)^{1/p}, \tag{9.30}$$

although a similar formula can be given also when using both $\bar{\mathbf{q}}$ and $\bar{\bar{\mathbf{q}}}$, see J. Granat et al. [12].

We stress again that $\sigma(\mathbf{q}, \bar{\mathbf{q}})$, even in its above form, is not a norm or a distance function between $\mathbf{q}$ and $\bar{\mathbf{q}}$; it might be equivalent to such a distance function only if all objectives are stabilized. As discussed above, a norm would not preserve needed properties of monotonicity for maximized or minimized objectives.

Until now we discussed reference (reservation and/or aspiration) points as if they were simple collections of their components. However, for more compli-

cated models – e.g. with dynamic structure – it is often advantageous to use *reference profiles* or *reference trajectories* of the same outcome variable changing e.g. over time. Suppose (see Kallio et al. [20]) that a model describes ecological quality of forests in a region or country, expected demand for wood, forestry strategies and projected prices for some longer time – say, next fifty years because of the slow dynamic of forest growth. The user would then interpret all model variables and outcomes rather as their profiles over time or trajectories than as separate numbers in given years. Mathematically, we can represent such a profile as a vector in a – say, fifty-dimensional – space, hence the methodology presented above is fully applicable. From the user point of view, however, it is much easier to interpret model outcomes and their reference points as entire profiles. Psychological studies show that it is too difficult to evaluate jointly more than seven to nine outcomes. However, this applies to separate numbers, not to their profiles or trajectories. A mental evaluation of such profiles, particularly if they are graphically presented, is not as difficult as that of a large number of separate variables. There are also other issues related to applications of reference trajectory optimization to dynamic models. For example, the user might like to specify the growth ratios or the increments of selected model outcomes as an additional objective trajectory; the dynamic properties of the model might be exploited when preparing or executing optimization of model outcomes, etc. – see Makowski et al. [29] and Wierzbicki [49].

9.4 NEUTRAL AND WEIGHTED COMPROMISE SOLUTIONS

By a *neutral compromise solution*, we understand typically in multi-criteria analysis a decision with outcomes located somewhere in the middle of the efficient set; a more precise meaning of "somewhere in the middle" specifies the type of a compromise solution. This notion was investigated in detail first by Zeleny, see e.g. [54]. He has shown (for the case of maximizing all objectives) that, in order to be sure of efficiency of solutions minimizing the distance even if the set Q_0 is not convex, the reference point $\bar{\mathbf{q}}$ for a scalarizing function $s(\mathbf{q}, \bar{\mathbf{q}}) = \| \mathbf{q} - \bar{\mathbf{q}} \|$ should be taken at the utopia point, called also ideal point, $\bar{\mathbf{q}} = \mathbf{q}_{uto}$, or "above to the North-East" of this point, at a "displaced ideal" or simply upper bound $\bar{\mathbf{q}} = \mathbf{q}_{up} \in \mathbf{q}_{uto} + D$. Then, when minimizing a distance related to an l_p norm with $1 \leq p < \infty$, properly efficient (Pareto-optimal) compromise solutions are obtained. The Chebyshev (l_∞) norm results in only weakly efficient solutions, unless its minimization is supplemented by a lexicographic test. Dinkelbach and Isermann [5] have shown that an augmented Chebyshev norm – with a linear part added such as in achievement functions discussed earlier in this chapter – results in properly efficient solutions.

As we have stressed earlier, in order to use a norm we must be sure that all objective components are of the same dimension or dimension-free. Thus, we have anyway to rescale the increments of objectives – from $| q_i - q_{i,up} |$ to $| q_i - q_{i,up} | / | q_{i,up} - q_{i,lo} |$ (where $q_{i,lo}$ is a lower bound, e.g. an approximation of the nadir point components). After such rescaling, when fixing the reference point

at the upper bound point, we can define neutral compromise solutions (actually, neutral compromise solution outcomes) with equal weighting coefficients as:

$$
\begin{aligned}
(9.31)\quad \hat{\mathbf{q}}_{neu}^{(p)} &= \underset{\mathbf{q}\in Q_0}{\operatorname{argmin}} \left(\sum_{i=1}^{k} \frac{\mid q_{i,up} - q_i \mid^p}{\mid q_{i,up} - q_{i,lo} \mid^p} \right)^{1/p}, \quad 1 \le p < \infty, \\
\hat{\mathbf{q}}_{neu}^{(\infty)} &= \underset{\mathbf{q}\in Q_0}{\operatorname{argmin}} \left(\max_{1\le i\le k} \frac{\mid q_{i,up} - q_i \mid}{\mid q_{i,up} - q_{i,lo} \mid} \right), \\
\hat{\mathbf{q}}_{neu}^{(1,\infty)} &= \underset{\mathbf{q}\in Q_0}{\operatorname{argmin}} \left(\max_{1\le i\le k} \frac{\mid q_{i,up} - q_i \mid}{\mid q_{i,up} - q_{i,lo} \mid} + \varepsilon \sum_{i=1}^{k} \frac{\mid q_{i,up} - q_i \mid}{\mid q_{i,up} - q_{i,lo} \mid} \right),
\end{aligned}
$$

the last one with some small $\varepsilon > 0$. As noted above, $\hat{\mathbf{q}}_{neu}^{(\infty)}$ might be only weakly efficient – and not uniquely defined in such a case. However, we can select then an efficient solution by additional testing, e.g. an additional lexicographic optimization, see Ogryczak et al. [32].

The neutral solution $\hat{\mathbf{q}}_{neu}^{(1,\infty)}$ can be obtained also in a different way, since it can be shown that minimizing the distance induced by the augmented Chebyshev norm from the upper bound point is equivalent to maximizing the following order-consistent achievement function (which is just a re-scaled version of (9.26)):

$$
\begin{aligned}
(9.32)\quad \sigma(\mathbf{q}, \bar{\mathbf{q}}_{mid}) &= \min_{1\le i\le k} \frac{q_i - \bar{q}_{i,mid}}{q_{i,up} - q_{i,lo}} + \varepsilon \sum_{i=1}^{k} \frac{q_i - \bar{q}_{i,mid}}{q_{i,up} - q_{i,lo}}, \\
\bar{q}_{i,mid} &= \frac{q_{i,up} + q_{i,lo}}{2}
\end{aligned}
$$

with the reference point $\bar{\mathbf{q}}_{mid}$ located precisely in the middle of the ranges between upper bound and lower bound point.

However, such neutral solutions as defined above might serve only as a starting point for interaction with the user. More general is the concept of *weighted compromise solutions*. Suppose that the *weighting coefficients* $\alpha_i > 0$, $\forall i = 1, \ldots k$, $\sum_{i=1}^{k} \alpha_i = 1$, are given (by the decision maker or by a special identification method, e.g. *Analytical Hierarchy Process* as proposed by Saaty [37]). Once the weighting coefficients α are determined, the weighted compromise solutions $\hat{\mathbf{q}}_{\alpha}^{(p)}$ are defined by:

$$
\begin{aligned}
(9.33)\quad \hat{\mathbf{q}}_{\alpha}^{(p)} &= \underset{\mathbf{q}\in Q_0}{\operatorname{argmin}} \left(\sum_{i=1}^{k} \alpha_i^p \frac{\mid q_{i,up} - q_i \mid^p}{\mid q_{i,up} - q_{i,lo} \mid^p} \right)^{1/p}, \; 1 \le p < \infty, \\
\hat{\mathbf{q}}_{\alpha}^{(\infty)} &= \underset{\mathbf{q}\in Q_0}{\operatorname{argmin}} \left(\max_{1\le i\le k} \alpha_i \frac{\mid q_{i,up} - q_i \mid}{\mid q_{i,up} - q_{i,lo} \mid} \right), \\
\hat{\mathbf{q}}_{\alpha}^{(1,\infty)} &= \underset{\mathbf{q}\in Q_0}{\operatorname{argmin}} \left(\max_{1\le i\le k} \alpha_i \frac{\mid q_{i,up} - q_i \mid}{\mid q_{i,up} - q_{i,lo} \mid} + \varepsilon \sum_{i=1}^{k} \alpha_i \frac{\mid q_{i,up} - q_i \mid}{\mid q_{i,up} - q_{i,lo} \mid} \right).
\end{aligned}
$$

While the concept of weighted compromise results in sufficient conditions of Pareto-optimality – all weighted compromise solutions $\hat{\mathbf{q}}_{\alpha}^{(p)}$ for $1 \le p < \infty$ are

properly efficient (Pareto-optimal) – necessary conditions are more complicated. We cannot generally say that we can obtain any properly efficient outcome desired by the decision maker by changing weighting coefficients. Moreover, the character of the dependence of $\hat{\mathbf{q}}_\alpha^{(p)}$ on α is not easy to interpret. In some applications, this might be a procedural advantage; however, in the case of a decision maker who is an analyst, designer or a modeler, the lack of a clear interpretation of this dependence is disadvantageous.

There is, fortunately, one case in which the dependence of a weighted compromise solution on weighting coefficients is easy to interpret, namely the case of Chebyshev norms. We shall discuss here the augmented Chebyshev norm and $\hat{\mathbf{q}}_\alpha^{(1,\infty)}$. Suppose we choose a weighting coefficient vector α with $\alpha_i > 0$, $\sum_{i=1}^k \alpha_i = 1$, and a scalar coefficient $\eta \geq 1/\alpha_i$, $\forall i = 1, \ldots, k$, in order to assign to each α_i an aspiration level:

$$\bar{q}_i = q_{i,up} - \frac{q_{i,up} - q_{i,lo}}{\eta \alpha_i}. \tag{9.34}$$

Then we obtain $q_{i,lo} \leq \bar{q}_i < q_{i,up}$ $\forall i = 1, \ldots k$ because of the inequality satisfied by η – although the aspiration levels $\bar{q}_i$ might change with η, in which case equation (9.34) describes a line segment in $\boldsymbol{R}^k$ starting with $\mathbf{q}_{lo}$ and ending at $\mathbf{q}_{up}$ as $\eta \to \infty$. Conversely, for any aspiration point $\bar{\mathbf{q}}$ such that $q_{i,lo} \leq \bar{q}_i < q_{i,up}$ $\forall i = 1, \ldots k$ we can set:

$$\begin{aligned} \alpha_i &= \frac{q_{i,up} - q_{i,lo}}{q_{i,up} - \bar{q}_i} \Bigg/ \sum_{j=1}^k \frac{q_{j,up} - q_{j,lo}}{q_{j,up} - \bar{q}_j}, \\ \eta &= \sum_{j=1}^k \frac{q_{j,up} - q_{j,lo}}{q_{j,up} - \bar{q}_j}, \end{aligned} \tag{9.35}$$

which defines the inverse to the transformation (9.34). We can interpret this inverse transformation in the following way: the ratios $\omega_{ij} = \alpha_i/\alpha_j$ of importance of criteria are defined by selected aspiration levels as an inverse ratio of their relative distances from upper bound levels:

$$\omega_{ij} = \frac{\alpha_i}{\alpha_j} = \frac{q_{j,up} - \bar{q}_j}{q_{i,up} - \bar{q}_i}. \tag{9.36}$$

The transformation (9.35) has been used by Steuer and Choo [42] in a procedure using the augmented Chebyshev norm and $\hat{\mathbf{q}}_\alpha^{(1,\infty)}$, but controlled interactively by the decision maker who specified aspiration points (called *definition points* by Steuer and Choo) that were used to define the weighting coefficients. The outcomes of such a procedure are properly efficient; Steuer and Choo show that any properly efficient outcome can be obtained by this procedure for convex sets Q_0.

However, we can show more: under transformations (9.34), (9.35), the weighted compromise solution $\hat{\mathbf{q}}_\alpha^{(1,\infty)}$ can be equivalently obtained by the maximization of an order-consistent achievement scalarizing function with such aspiration levels used as a reference point. This is because we have:

$$\alpha_i \frac{q_{i,up} - q_i}{q_{i,up} - q_{i,lo}} = \left(\frac{q_{i,up} - q_i}{q_{i,up} - \bar{q}_i}\right) \Big/ \eta = \left(1 - \frac{q_i - \bar{q}_i}{q_{i,up} - \bar{q}_i}\right) \Big/ \eta, \tag{9.37}$$

therefore, since $q_{i,up} \geq q_i \geq q_{i,lo}$:

$$\max_{1 \leq i \leq k} \frac{| q_{i,up} - q_i |}{| q_{i,up} - q_{i,lo} |} + \varepsilon \sum_{i=1}^{k} \frac{| q_{i,up} - q_i |}{| q_{i,up} - q_{i,lo} |} =$$

$$(- \min_{1 \leq i \leq k} \frac{q_i - \bar{q}_i}{q_{i,up} - \bar{q}_i} - \varepsilon \sum_{i=1}^{k} \frac{q_i - \bar{q}_i}{q_{i,up} - \bar{q}_i} + 1 + \varepsilon k)/\eta. \tag{9.38}$$

Hence, minimizing the weighted distance induced by the augmented Chebyshev norm from the upper bound point is equivalent to maximizing the following order-consistent achievement function (again defined as in (9.26) with re-scaling):

$$\sigma(\mathbf{q}, \bar{\mathbf{q}}) = \min_{1 \leq i \leq k} \frac{q_i - \bar{q}_i}{q_{i,up} - \bar{q}_i} + \varepsilon \sum_{i=1}^{k} \frac{q_i - \bar{q}_i}{q_{i,up} - \bar{q}_i}, \tag{9.39}$$

with the aspiration point $\bar{\mathbf{q}}$ defined as the transformation (9.34) of the vector of weighting coefficients α with any (sufficiently large) parameter η.

Moreover, even if the set Q_0 is not convex (or even if it is a discrete set), we can take any properly efficient outcome $\hat{\mathbf{q}}$ with trade-off coefficients scaled down by the deviations from the upper levels and bounded by:

$$t_{ij}(\hat{\mathbf{q}}) \frac{q_{j,up} - \hat{q}_j}{q_{i,up} - \hat{q}_i} \leq (1 + 1/\varepsilon). \tag{9.40}$$

At any such point, we can define a reference point and weighting coefficients – by taking $\bar{\mathbf{q}} = \hat{\mathbf{q}}$ and applying the transformation (9.35) – in such a way that the maximal point of $\sigma(\mathbf{q}, \bar{\mathbf{q}})$, equal to the weighted compromise solution $\hat{\mathbf{q}}_\alpha^{(1,\infty)}$, coincides with $\hat{\mathbf{q}}$. This can be shown by an appropriate modification of the argument on separating the set Q_0 and the conical set $\hat{\mathbf{q}} + \operatorname{Int} D_\varepsilon$ by a level set of the function $\sigma(\mathbf{q}, \bar{\mathbf{q}})$, as discussed in the previous section, see also Wierzbicki [50].

9.5 MODELING FOR MULTIOBJECTIVE ANALYSIS

Reference point methods can be used for a wide variety of substantive model types. However, methods of optimization of an achievement function attached to a complicated model depend very much on the model type. Moreover, this concerns even model building: constructing a complicated model is an art and requires a good knowledge not only of the disciplinary field concerned, but also of the properties of models of the particular class.

There exist today special software tools for building analytical models, called generally modeling systems or algebraic modeling languages – such as GAMS, AIMMS, AMPL, see e.g. Brooke et al. [3], Bisschop et al. [2], Fourer et al. [8].

However, they usually represent the perspective of single-objective optimization and can be adapted to multiobjective model analysis only through additional tricks.

Linear models provide a good starting point in modeling. In the case of large-scale models, a practical way to develop a model is to prepare first a linear version and then augment it by necessary nonlinear parts.

In a textbook, the standard form of a multiobjective linear programming problem is usually presented as:

$$\text{"}\underset{\mathbf{x}\in X_0}{\text{maximize}}\text{"}(\mathbf{q} = \mathbf{C}\mathbf{x} \in \boldsymbol{R}^k), \tag{9.41}$$

$$X_0 = \{\mathbf{x} \in \boldsymbol{R}^n : \mathbf{A}\mathbf{x} = \mathbf{b} \in \boldsymbol{R}^m,\ \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}\}, \tag{9.42}$$

where "maximize" might either mean single-objective optimization if $\mathbf{q}$ is a scalar, or be understood in the Pareto sense, or in the sense of another partial order implied by a positive cone. Much research has been done on the specification of Pareto-optimal or efficient decisions and objectives for linear models, see e.g. Gal [9] or Steuer [43]. However, we must note that the standard form above uses the equality form of constraints $\mathbf{A}\mathbf{x} = \mathbf{b}$ in order to define X_0. Other forms of linear constraints can be converted theoretically to equality form by introducing dummy variables as additional components of the vector $\mathbf{x}$. In the practice of linear programming it is known, however, that the standard form is rather unfriendly to the modeler. Thus, specific formats of writing linear models have been proposed, such as MPS or LP-DIT format, see e.g. Makowski [27]. Without going into details of such formats, we shall note that they correspond to writing the set X_0 in the form:

$$X_0 = \{\mathbf{x} \in \boldsymbol{R}^n : \mathbf{b} \leq \mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{W}\mathbf{y} \leq \mathbf{b} + \mathbf{r} \in \boldsymbol{R}^m,\ \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}\}, \tag{9.43}$$

where the vector $\mathbf{x}$ denotes rather actual decisions than dummy variables, thus $\mathbf{x}, m, n$ denote different entities than those implied by the standard textbook form. The model output $\mathbf{y}$ is composed of various intermediary variables (hence it depends implicitly on itself, although usually through a lower-triangular matrix $\mathbf{W}$). Essential for the modeler is her/his freedom to choose any of outputs y_j, including actually decisions x_j, as an objective variable q_i and to use many objectives – not only one, which is typical for algebraic modeling languages.

Even more complicated formats of linear models are necessary if we allow for the repetition of some basic model blocks indexed by additional indices, as in the case of linear dynamic models:

$$\begin{aligned} X_0 = \{\mathbf{x} \in \boldsymbol{R}^n : &\mathbf{w}_{t+1} = \mathbf{A}_t\mathbf{w}_t + \mathbf{B}_t\mathbf{x}_t;\ \mathbf{b}_T \leq \mathbf{y}_t = \\ &= \mathbf{C}_t\mathbf{w}_t + \mathbf{D}_t\mathbf{x}_t \leq \mathbf{b}_t + \mathbf{r}_t \in \boldsymbol{R}^m,\ \mathbf{l}_t \leq \mathbf{x}_t \leq \mathbf{u}_t;\ t = 1, \ldots T\}, \end{aligned} \tag{9.44}$$

where $\mathbf{w}_t$ is called the *dynamic state* of the model (the initial condition $\mathbf{w}_1$ must be given), the index t has usually the interpretation of (discrete) time,

and $\mathbf{x} = (\mathbf{x}_1, \ldots \mathbf{x}_T)$ is a *decision trajectory* (called also control trajectory). Similarly, $\mathbf{w} = (\mathbf{w}_1, \ldots \mathbf{w}_{T+1})$ is a *state trajectory* while $\mathbf{y} = (\mathbf{y}_1, \ldots \mathbf{y}_T)$ is the *output trajectory*. Actually, the variable $\mathbf{w}$ should be considered a part of the vector $\mathbf{y}$ (it is an intermediary variable, always accessible to the modeler) but is denoted separately because of its special importance – e.g. when differentiating the model, we must account for the state variables in a special way, see e.g. Wierzbicki [47]. Other similarly complicated forms of linear models result e.g. from stochastic optimization.

A modeler who has developed or modified a complicated (say, dynamic) large scale linear model should first validate it by simple simulation – that is, assume some common sense decisions and check whether the outputs of the model make also sense to her/him. Because of multiplicity of constraints in large-scale models it might, however, happen that the common sense decisions are not admissible (in the model); thus, even simple simulation of large-scale linear models might be actually difficult.

An important help for the modeler can be *inverse simulation*, in which she/he assumes some desired model outcomes $\bar{\mathbf{y}}$ and checks – as in the classical goal programming – whether there exist admissible decisions which result in these outcomes. *Generalized inverse simulation* consists in specifying also some reference decision $\bar{\mathbf{x}}$ and in testing, whether this reference decision could result in the desired outcomes $\bar{\mathbf{y}}$. This can be written in the goal programming format of norm minimization, while it is useful to apply the augmented Chebyshev norm (with changed sign, because we keep to the convention that that achievement functions are usually maximized while norms are minimized):

$$\begin{aligned} \sigma(\mathbf{y}, \bar{\mathbf{y}}, \mathbf{x}, \bar{\mathbf{x}}) &= -(1-\rho)(\max_{1 \le i \le n} |x_i - \bar{x}_i| + \varepsilon \sum_{i=1}^{n} |x_i - \bar{x}_i|) \\ &- \rho(\max_{1 \le j \le m} |y_j - \bar{y}_j| + \varepsilon \sum_{j=1}^{m} |y_j - \bar{y}_j|). \end{aligned} \tag{9.45}$$

The coefficient $\rho \in [0; 1]$ indicates the weight given to achieving the desired output versus keeping close to reference decision. It is assumed for simplicity sake that all variables are already re-scaled to be dimension-free.

A multi-objective optimization system based on reference point methodology can clearly help in such inverse simulation. In such a case, we stabilize all outcomes and decisions of interest and use for them partial achievement functions of the form $\sigma_i(y_i, \bar{y}_i)$ (or even $\sigma_i(y_i, \bar{y}_i, \bar{\bar{y}}_i)$), similar to those defined in Section 9.3 in terms of objectives q_i. An overall achievement function has then the form:

$$\begin{aligned} \sigma(\mathbf{y}, \bar{\mathbf{y}}, \mathbf{x}, \bar{\mathbf{x}}) &= (1-\rho)(\min_{1 \le i \le n} \sigma_i(x_i, \bar{x}_i) + \varepsilon \sum_{i=1}^{n} \sigma_i(x_i, \bar{x}_i) \\ &+ \rho(\min_{1 \le j \le m} \sigma_j(y_j, \bar{y}_j) + \varepsilon \sum_{j=1}^{m} \sigma_j(y_j, \bar{y}_j)). \end{aligned} \tag{9.46}$$

It is more convenient for the modeler, if such functions are defined inside the decision support system which also has a special function *inverse simulation*, prompting her/him to define which (if not all) decisions and model outputs should be stabilized and at which reference levels.

Even more important for the modeler might be another interpretation of the above function, called *simulation with elastic constraints* or *softly constrained simulation*. Common sense decisions might appear inadmissible for the model, because it interprets all constraints as *hard* mathematical inequalities or equations. On the other hand, we have already stressed that it is a good modeling practice to distinguish between *hard constraints* that can never be violated and *soft constraints* which in fact represent some desired relations and are better represented as additional objectives with given aspiration levels. Thus, in order to check actual admissibility of some common-sense decision $\bar{\mathbf{x}}$, the modeler should answer first the question which constraints in her/his model are actually hard and which might be softened and included in the objective vector $\mathbf{q}$. Thereafter, simulation with elastic constraints might be performed by maximizing an overall achievement function similar as above.

If (9.46) is maximized with concave piece-wise linear partial achievement functions σ_i and for a linear model, then the underlying optimization problem can be converted to linear programming. In fact, if a partial achievement function – say, $\sigma_i(x_i, \bar{x}_i)$ – is piece-wise linear but concave, then it can be expressed as the minimum of a number of linear functions:

$$\sigma_i(x_i, \bar{x}_i) = \min_{l \in L_i} \sigma_{il}(x_i, \bar{x}_i), \tag{9.47}$$

where $\sigma_{il}(x_i, \bar{x}_i)$ are linear functions. Assume that a similar expression is valid for $\sigma_j(q_j, \bar{q}_j)$. The maximization of the function (9.46) can be then equivalently expressed as the maximization of the following function of additional variables z, z_i, w, w_j:

$$(1-\rho)(z + \varepsilon \sum_{i=1}^{n} z_i) + \rho(w + \varepsilon \sum_{j=1}^{m} w_j), \tag{9.48}$$

with additional constraints:

$$\begin{aligned} \sigma_{il}(x_i, \bar{x}_i) &\geq z_i, \ \forall l \in L_i, \\ z_i &\geq z, \ \forall i = 1, \ldots n, \\ \sigma_{jl}(q_j, \bar{q}_j) &\geq w_j, \ \forall l \in L_j, \\ w_j &\geq w, \ \forall j = 1, \ldots m. \end{aligned} \tag{9.49}$$

Similar conversion principles apply if we have a mixed integer linear programming model – that can even express piece-wise linear models which are not concave (or not convex in the case of function minimization). Thus, we can use inverse simulation or even softly constrained simulation for mixed integer programming models (although not all heuristic algorithms, related to some

specific forms of objective functions in mixed integer optimization, would work for such optimization problems).

Even less developed than user-friendly standards of defining linear models are such standards for nonlinear models. The classical textbook format for (multi-objective optimization of) such models is simple:

$$\text{"}\underset{\mathbf{x}\in X_0}{\text{maximize}}\text{"}(\mathbf{q} = \mathbf{f}(\mathbf{x}) \in \boldsymbol{R}^k), \tag{9.50}$$

$$X_0 = \{\mathbf{x} \in \boldsymbol{R}^n : \mathbf{g}(\mathbf{x}) \leq 0 \in \boldsymbol{R}^m\}, \tag{9.51}$$

where $f_i(\mathbf{x})$ models consecutive objective functions and $g_j(\mathbf{x})$ models consecutive constraints of the set of admissible decisions. However, such a format is seldom convenient for more complicated models, in which it is useful to consider various model outputs $\mathbf{y}$ and define both objectives and constraints in terms of such model outputs.

While there exist some standards for specific nonlinear optimization systems – such as in MINOS, GAMS, AIMMS, AMPL, see e.g. Brooke et al. [3], Bisschop et al. [2], Fourer et al. [8] – they are devised more for single-objective optimization purposes than for practical multi-objective modeling and analysis, while experience in modeling shows that a model should be analyzed multi-objectively even if it is later used for single-objective optimization only. A useful standard was developed in the multi-objective nonlinear optimization system DIDAS-N[8] (see e.g. Kręglewski at al. [24]). Briefly, it consists in defining subsequent nonlinear model output relations:

$$\begin{aligned} y_1 &= f_1(\mathbf{x}, \mathbf{z}), \\ \ldots &= \ldots \\ y_{j+1} &= f_{j+1}(\mathbf{x}, \mathbf{z}, y_1, \ldots y_j), \quad j = 1, \ldots m-1, \\ \ldots &= \ldots \\ y_m &= f_m(\mathbf{x}, \mathbf{z}, y_1, \ldots y_{m-1}), \end{aligned} \tag{9.52}$$

together with bounds for decision variables and outputs:

$$x_{i,lo} \leq x_i \leq x_{i,up}, \quad i = 1, \ldots, n; \quad y_{j,lo} \leq y_j \leq y_{j,up}, \quad j = 1, \ldots, m \tag{9.53}$$

(bounds for model parameters $\mathbf{z}$ are less essential). In this way, a directly computable (explicit, except for bounds) nonlinear model is defined. Implicit models can be defined by specifying $y_{j,lo} = y_{j,up}$ for some j, which is then

[8]Developed in the Institute of Control and Computation Engineering, Technical University of Warsaw, in cooperation with IIASA. Available as a public domain software from IIASA, see Appendix, in Poland from system authors.

taken into account and resolved during optimization. Any variable y_j (and x_i, if needed) can be specified as maximized, minimized or stabilized objective.

The model equations and bounds are specified using a computer spreadsheet format. The DIDAS-N system includes rather advanced automatic (algebraic) functions of model differentiation: it presents to the modeler all required partial and full derivatives and prepares an economical way of computing numerically the derivatives of the overall achievement function in a smooth form similar to Eq. (9.30). A specific robust optimization solver, based on a shifted penalty function approach, was developed and included in the system.

However, DIDAS-N is a closed, nonmodular system written in PASCAL, difficult for working with larger models, particularly when including large-scale linear model parts. Therefore, a new system called DIDAS-N++ was developed, see Granat et al. [12]. This system was written in C++ with a modular structure, includes the possibility of selecting optimization solvers, and a choice and customization of a graphical user interface, together with a preferred option of specifying user preferences in terms of fuzzy membership functions controlled by aspiration and reservation levels (as discussed in a previous section; the membership functions $\mu_i(q_i, \bar{q}_i, \bar{\bar{q}}_i)$ are displayed and modified graphically while the achievement functions $\sigma_i(q_i, \bar{q}_i, \bar{\bar{q}}_i)$ are in fact used in computations).

The format of the nonlinear model definition in DIDAS-N++ is similar to that of DIDAS-N. However, the nonlinear part can be also linked with a linear part, which is indicated by the general format:

$$\begin{aligned} \mathbf{y}_1 &= \mathbf{A}_1\mathbf{x}_1 + \mathbf{A}_c\mathbf{x}_c, \\ \mathbf{y}_2 &= \mathbf{f}(\mathbf{x}_2, \mathbf{x}_c, \mathbf{z}, \mathbf{y}_1, \mathbf{y}_2), \end{aligned} \tag{9.54}$$

where $\mathbf{y}_1, \mathbf{y}_2, \mathbf{x}_1, \mathbf{x}_2$ denote the vectors of model outputs and decision variables specific for the linear and nonlinear parts, while $\mathbf{x}_c$ is the vector of decision variables common for both parts.

The model is first analyzed by an algebraic processing and compiling module to produce an executable file easily linked with other modules of the system – the organizing module, a graphic interface, a selected solver – and containing all information how to compute model outputs and their derivatives (see e.g. Griewank [14]), together with the possibility of modifying values of decision variables, bounds and parameters. Thus the compiling process might be long for complicated models, but the repetitive runs of the compiled model needed in its simulation and optimization are relatively short.

9.6 APPLICATIONS OF REFERENCE POINT METHODS

The reasoning presented in previous sections might seem rather abstract. Nonetheless, all development of reference point methods was very much applications-oriented, starting with the original work of Kallio et al. [20] on forestry models, including many other applications to energy, land use and environmental models at IIASA, applications of satisficing trade-off methods by Nakayama et al. [31] to engineering design, various applications of *Pareto Race* of Korhonen et

al. [22], and many others. Recent applications of a reference (aspiration-reservation) point method have been developed at IIASA using a modular tool MCMA (MultiCriteria Model Analysis)[9] by Granat and Makowski [11], in relation to regional management of water quality (Makowski, Somlyódy and Watkins [28]), land use planning (Antoine, Fischer, Makowski [1]), urban land-use planning (Matsuhashi, [30]), and other examples [18].

Here we present only two short examples: one application to engineering design and another to ship navigation support.

The first case concerns a classical problem in mechanical design – the design of a spur gear transmission unit, see e.g. Osyczka [33]. The mechanical outlay of this unit is shown in Fig. 9.3:. The design problem consists in choosing some mechanical dimensions (the width of the rim of toothed wheel, the diameters of the input and output shafts, the number of teeth of the pinion wheel, etc.) in order to obtain a *best design*. However, there is no single measure of the quality of design of such a gear transmission. Even when trying only to make the unit as compact as possible – which can be expressed by minimizing the volume of the unit while satisfying various constraints related to mechanical stresses and to an expected lifetime of efficient work of the gear unit – we should take into account other objectives, such as the distance between the axes or even the width of the rim of toothed wheel (which is, at the same time, a decision variable).

The specification of a mathematical model that expresses the available knowledge on designing such gear units is obviously a question of expert opinion. After all, the modeler is a specialist in her/his specific field and knows best how to choose substantive models for a given problem; that is also the reason why we present here mostly methods for supporting the modeler in model analysis, not supplementing her/him in final decisions. Therefore, in the example of gear unit design, we follow a specialist who has selected a specific model in this case (Osyczka [33]) and comment only on the methodology of preparing the model for analysis and analyzing it.

The equations of the corresponding model contain some tables of coefficients obtained by empirical, mechanical studies. While such original data are very valuable, an analytic approximation of them might be more useful for model analysis. Thus, these tables were approximated by exponential functions. The problem might be then specified in a classical textbook format such as (9.50, 9.51) by defining three objective functions $f_i(\mathbf{x})$ and 14 constraints $g_j(\mathbf{x})$, some nonlinear and some expressing simple bounds. We present it here in a form similar to the textbook format (although the model was in fact rewritten in the DIDAS-N++ format, because this system was used for further analysis).

1. The decision variables are: the width of the toothed wheel rim b (which is also an objective), the diameters d_1 and d_2 of the input and output shafts,

[9]The MCMA tool is available from the URL: `www.iiasa.ac.at/~marek/soft` free of charge for research and educational purposes.

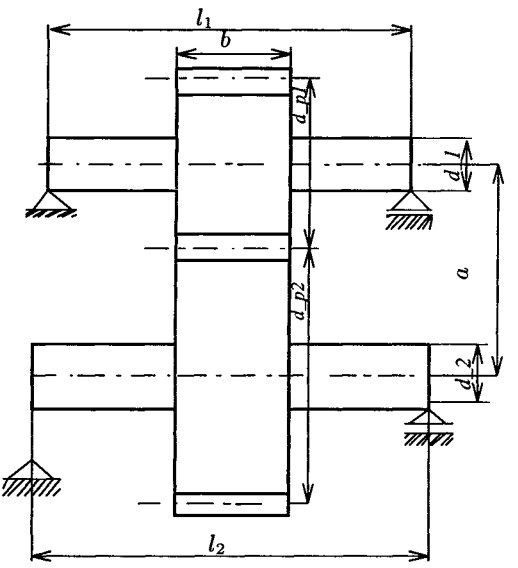


Figure 9.3: A diagram of the spur gear unit

the number of teeth of the pinion wheel $\tilde{z}_1$ and the pitch of gear teeth $\tilde{m}$ (the last two decision variables are in fact discrete).

2. The objectives are the volume of the gear unit $q_1 = f_1$ [mm^3], the distance between the axes $q_2 = f_2$ [mm], the width of the toothed wheel rim $q_3 = f_3$ [mm]:

$$
\begin{aligned}
q_1 &= ((\frac{\pi}{4}\tilde{m}^2(\tilde{z}_1^2 + \tilde{z}_2^2)b) + \frac{\pi}{2}d_1^3 + \frac{\pi}{2}d_2^3) * 10^{-5}, \\
q_2 &= \frac{(\tilde{z}_1 + \tilde{z}_2)}{2}\tilde{m}, \\
q_3 &= b.
\end{aligned} \tag{9.55}
$$

3. The constraints on the decisions concern various geometric relations and mechanical stresses.

- g_1 expresses the bending stress of the pinion,

$$g_1 = k_{g1} - P_{og} * w_1/(b * \tilde{m}), \tag{9.56}$$

where:

$$V = \pi * \tilde{m} * \tilde{z}_1 * \tilde{n}/60000, \quad K_d = (14.5 + V)/14.5,$$

$$P_{max} = 102 * N * 9.81/V, \quad P_{og} = P_{max} * K_p * K_b * K_d,$$

$$w_1 = 4.7607 * exp(-0.104531 * (\tilde{z}_1 + 1.28627)) + 1.67421.$$

- g_2 expresses the bending stress of the gear,

$$g_2 = k_{g2} - P_{og} * w_2/(b * \tilde{m}), \tag{9.57}$$

where:

$$w_2 = 4.7607 * exp(-0.104531 * (\tilde{z}_2 + 1.28627)) + 1.67421.$$

- g_3 expresses the surface pressure of smaller wheel,

$$g_3 = k_{o1} - P_{o1}/(b * \tilde{m} * \tilde{z}_1) * (1 + \tilde{z}_1/\tilde{z}_2) * y_1, \tag{9.58}$$

where:

$$P_{o1} = P_{max} * K_p * K_b * K_d * K_{\tilde{z}_1},$$
$$y_1 = 28.4869 * exp(-0.290085 * (\tilde{z}_1 - 1.78811)) + 3.31178.$$

- g_4 expresses the surface pressure of the greater wheel:

$$g_4 = k_{o2} - P_{o2}/(b * \tilde{m} * \tilde{z}_2) * (1 + \tilde{z}_2/\tilde{z}_1) * y_c, \tag{9.59}$$

where:

$$P_{o2} = P_{max} * K_p * K_b * K_d * K_{z2}.$$

- g_5, g_6 express the torsional stresses of input and output shafts,

$$g_5 = k_s - M_{s1}/W_{01}, \quad g_6 = k_s - M_{s2}/W_{02}, \tag{9.60}$$

where:

$$\begin{array}{rclrcl} M_{s1} & = & 9549296 * N/n; & W_{01} & = & (\pi * d_1^3)/16, \\ M_{s2} & = & M_{s1}/(z_1/z_2); & W_{02} & = & (\pi * d_2^3)/16. \end{array}$$

- g_7, g_8, g_9 express the deviations of the velocity ratio and the relation between $\tilde{m}$ and d_1:

$$\begin{array}{rcl} g_7 = i - \tilde{z}_1/\tilde{z}_2 + \Delta i,; \; g_8 & = & \tilde{z}_1/\tilde{z}_2 - i + \Delta i,; \\ g_9 & = & \tilde{m} * (\tilde{z}_1 - 2.4) - d_1. \end{array} \tag{9.61}$$

- Other constraints are:

$$\begin{array}{rcl} g_{10} = \tilde{m} * (\tilde{z}_2 - 2.4) - d_2, \; g_{11} = b/\tilde{m} - b\tilde{m}_{min}, \; g_{12} & = & b\tilde{m}_{max} - b/\tilde{m}, \\ g_{13} = a_{max} - (\tilde{z}_1 + \tilde{z}_2)/2 * \tilde{m}, \quad g_{14} & = & \tilde{z}_2 - \tilde{z}_1/i. \end{array} \tag{9.62}$$

4. In the above model, the following parameters were used:

$$N = 12.0, \quad \tilde{n} = 280.0, \quad i = 0.317, \quad \Delta i = 0.01, \quad \tilde{z}_1 = 20,$$

where N is the input power [kW], $\tilde{n}$ is the rotational input speed [rev/min], i is the velocity ratio, Δi is the allowable deviation of velocity ratio, $\tilde{z}_1$ is the number of teeth of the pinion.

Geometric data are:

$$b\tilde{m}_{min} = 5.0, \quad b\tilde{m}_{max} = 10.0, \quad a_{max} = 293.8,$$

where $b\tilde{m}_{min}$ is the minimum $b/\tilde{m}$ coefficient ($\tilde{m} = d_{pi}/z_i$, $i = 1, 2$, is the pitch of the gear teeth, while d_{pi} are the standard diameters of the gear wheels and b is the teeth width), $b\tilde{m}_{max}$ is the maximum $b/\tilde{m}$ coefficient, a_{max} is the maximum distance between the axes [mm].

Material data are:

$$k_{g1} = 105, \quad k_{g2} = 105, \quad k_{o1} = 62, \quad k_{o2} = 62, \quad k_s = 70,$$

where k_{g1} is the allowable bending stress for the pinion [MPa], k_{g2} is the allowable bending stress for the gear [MPa], k_{o1} is the allowable surface pressure for the pinion [MPa], k_{o2} is the allowable surface pressure for the gear [MPa], k_s is the allowable torsional stress of the shaft [MPa].

Other data are:

$$K_b = 1.12, \quad K_{z1} = 1.87, \quad K_{z2} = 1.3, \quad K_p = 1.25,$$

where K_b is the coefficient of the concentrated load, K_{z1} is the coefficient of the equivalent load for the pinion, K_{z2} is the coefficient of the equivalent load for the gear, K_p is an overload factor.

Calculated data are:

$$T = 8000, \quad y_c = 3.11,$$

where T is the time of efficient work of the gear, y_c is a coefficient for the assumed pressure angle.

The exponential approximations of empirical data tables are expressed by the functions w_1, w_2, y_1. We presented all these equations with a purpose: in order to stress that a computerized mathematical model might be very complicated. The model presented above is actually rather small – because it is static, not dynamic – as compared to other models used in applications. However, the model represents rather advanced knowledge in mechanical engineering and the selection of its various details relies on expert intuition: good modeling is an art. Moreover, even for such rather small model, the reader should imagine programming the model, supplying it with all necessary derivatives, selecting "by hand" such values of decision variables which would satisfy required constraints, all done without specialized software supporting model analysis.

When using such a specialized software, the modeler should use first a model generator, then model compiler; a good model compiler will automatically determine all needed derivatives. Even when such fast executable, compiled core model is available, the modeler might have trouble with simple model simulation. The form of the model is rather complicated (actually – not convex)

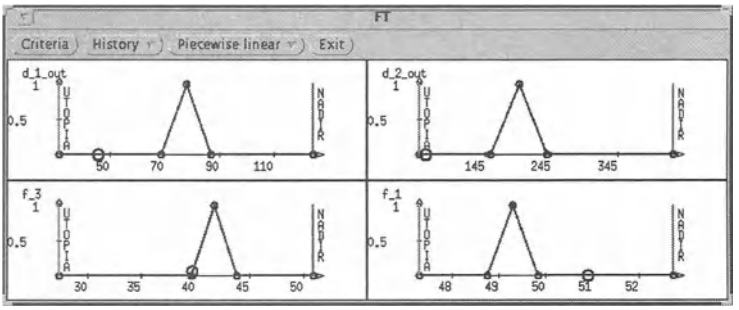


Figure 9.4: Interaction screen of DIDAS-N++ in the inverse simulation case, arbitrary aspiration levels

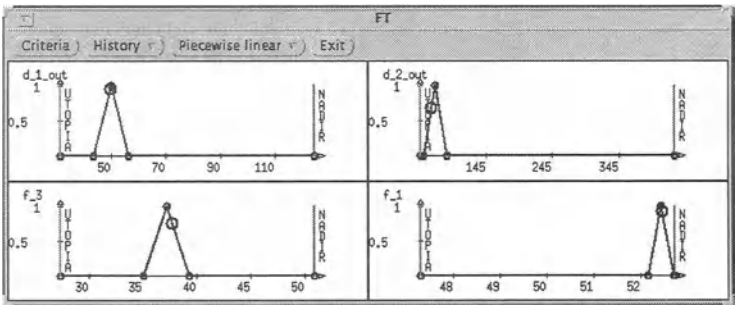


Figure 9.5: Interaction screen of DIDAS-N++ in the inverse simulation case, aspiration levels based on mechanical experience

and without a good experience in mechanical design it is difficult to select such values of decision variables which are acceptable.

This is illustrated in Fig. 9.4: which shows the results of an inverse simulation of the model with two model outcomes – objectives q_1 and q_3 denoted respectively by f_1 and f_3 – and two decision variables denoted by d_1 and d_2, all stabilized[10]. However, since the aspiration and reservation levels were arbitrarily selected, even the inverse simulation cannot give satisfactory results. The optimization of a corresponding achievement function indicates that such arbitrary reference levels cannot be realized in this model. The contours indicated in Fig. 9.4: represent the values of membership functions $\mu_i(q_i, \bar{q}_i, \bar{\bar{q}}_i)$ and the circles on these contours indicate the attained levels of objectives. Values 0 of these membership functions at circled points indicate that the requirements of the modeler cannot be satisfied.

[10] In Fig. 9.4: – Fig. 9.6: we use actual interaction screens of ISAAP-TOOL in DIDAS-N++.

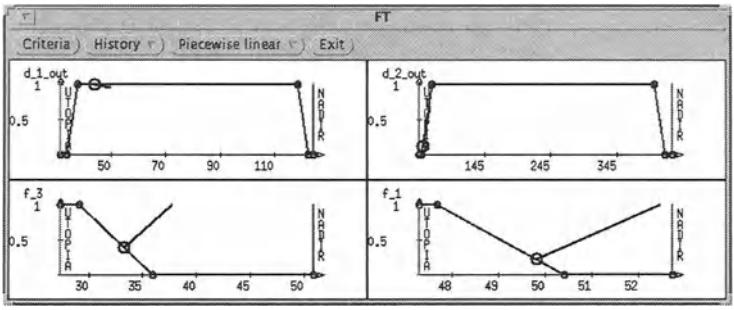


Figure 9.6: Interaction screen of DIDAS-N++ in the softly constrained simulation case, improvements of both objectives

In order to find results that are admissible for the model, other aspiration levels must be selected using the experience of a designer, see Fig. 9.5: where the aspirations were set according to data given by Osyczka [33]. Since the model was actually changed – by using the exponential approximation of data tables – from the one described by Osyczka, the results of the inverse simulation with membership values close to 1 indicate a positive validity test of the model. However, the inverse simulation results are not efficient in the sense of minimization of objectives (the results given by Osyczka might be efficient for his model, but the model was changed by including approximating functions).

Improvement of both (or even all three) objectives considered can be obtained by switching to softly constrained simulation, as shown in Fig. 9.6:, where the soft constraints on decision variables were relaxed in such a way as to obtain efficient results for the problem of minimizing both selected objectives. In Fig. 9.6:, the improvement of objective values is shown by line segments leading to circles that indicate the attained values. A serious model analysis would clearly not stop at the results of such an experiment – many other experiments, including post-optimal parametric analysis, might be necessary. However, the above example is presented only as an illustration of some basic functions of a system of computerized tools for multi-objective model analysis and decision support.

Another application example shows the usefulness of including dynamic formats of models. This case concerns ship navigation support (see Śmierzchalski et al. [41]): the problem is to control the course of a ship in such a way as to maximize the minimal distance from possible collision objects while minimizing the deviations from the initial course of the ship, see Fig. 9.7:.

This is a dynamic problem, with the equations of the model described initially by a set of differential equations for $t \in [0;T]$:

$$\begin{aligned} \dot{w}_1(t) &= v_1 \sin x(t), \\ \dot{w}_2(t) &= v_1 \cos x(t), \end{aligned}$$

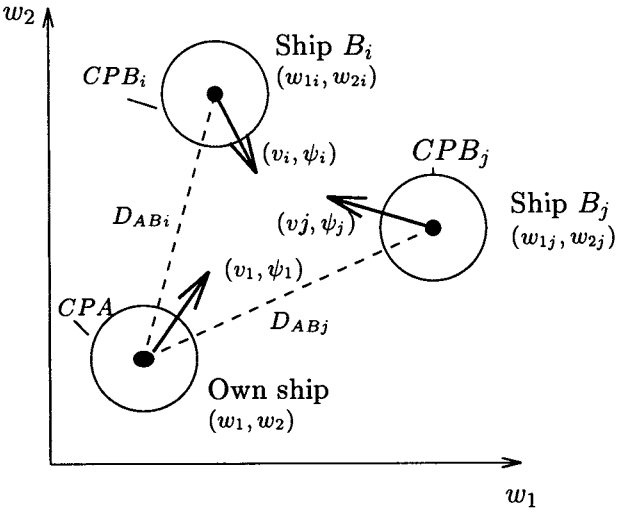


Figure 9.7: A diagram of ship collision control situation (CPA – safe zone for ship A)

$$\begin{aligned} \dot{w}_{1j}(t) &= v_j \sin\psi_j, \quad j = 2, \ldots \check{n}, \\ \dot{w}_{2j}(t) &= v_j \cos\psi_j, \quad j = 2, \ldots \check{n}, \end{aligned} \tag{9.63}$$

where $x(t)$ is the course of "our" ship, ψ_j – courses of other ships, with initial values of ship positions given as the vector $\mathbf{w}(0)$; between other model outcomes, the objectives can be modeled as:

$$\begin{aligned} q_1 &= \min_{t\in[0;T]} \min_{j=2,\ldots\check{n}} ((w_1(t) - w_{1j}(t))^2 + (w_2(t) - w_{2j}(t))^2), \\ q_2 &= \int_0^T (x(t) - \psi_1)^2 dt, \end{aligned} \tag{9.64}$$

where q_1 represents the (squared) minimal distance which should be maximized and q_2 represents the (squared) average deviation from initial course, which should be minimized.

To be used in a DIDAS-N system, this model was simply discretized in time. We do not describe the analysis of this model in more detail here (the results of such analysis are given e.g. in Śmierzchalski et al. [41]); this example was quoted only to show the practical sense of using dynamic models with multi-objective analysis and optimization.

9.7 A DECISION PROCESS IN REFERENCE POINT METHODS

We turn now back to a broader discussion and interpretation of the underlying methodological assumptions, theoretical results and the decision process considered in the reference point methodology.

We assume in this methodology that the decision maker – for example, a scientist analyzing environmental models, an analyst or an engineering designer – develops, modifies and uses substantive models which are specific for her/his profession and express essential aspects of the decision situation as perceived by her/him. In the decision process, the decision maker might have to specify at least partly her/his preferences and thus to define a preferential model. However, we assume that the decision maker preserves the right to change these preferences and thus the form of the preferential model is rather general, for example, restricted to specifying only which decision outcomes should be maximized or minimized.

Such a decision process might be subdivided into various phases. We might either include into it the early phases concerned with problem recognition and model building, or consider them as lying outside of the decision process. We include them for the sake of completeness and consider the following phases:

1. Problem recognition and formulation, data gathering and substantive model selection.

2. Formulation of a substantive model; initial analysis, including model validation.

3. Selection of a partial preferential model, detailed analysis of the substantive model, generation of scenarios or design options.

4. Final selection of a scenario or a design, implementation, feedback from practice.

Phase **1**, though extremely important, is not supported by reference point methodology. Many known methods of decision analysis and support can be applied for phase **4**; however, they require more detailed specification of preferential models. The reference point methodology concentrates on methods and techniques that might be used to support phases **2** and **3**.

While it is well known that vector optimization provides various techniques for supporting phase **3**, we stress that such techniques, in particular the reference point methodology, can be usefully extended to support also phase **2** – often very important and time-consuming for the modeler. The application of such techniques in phase **2** might be called multi-objective model analysis, which is understood here mainly as a tool of learning by the modeler of various possibilities and outcomes predicted by her/his models.

Such learning should enhance the intuitive capabilities of an analyst or decision maker as an expert in the field of his specialization. If we aim to support such learning by optimization and decision-analytical tools in the early stages of such a decision process, we cannot concentrate on modeling explicit preference or utility representation. We cannot even require that the decision maker should be consistent: the inconsistency of the decision maker is valuable in learning. We must rather concentrate on supporting various experiments performed with the help of the substantive model. During all such experiments, the final choice of decisions is not explicitly supported but even actually postponed. In fact, we suggest the use of "hard" optimization tools to support "soft" learning, deliberation and intuition formation.

This "hard" optimization concerns an achievement function – a proxy utility or value function of the computerized DSS[11] working in a quasi-satisficing manner – which, as already stressed, cannot be described just by a distance from the reference point. Reference point methods are a generalization of goal programming approach to such cases when we can and want to improve (minimize or maximize) certain outcomes beyond their reference points.

The main assumption of this approach is the use of multiple criteria optimization as a tool supporting not necessarily actual decision selection, but much rather facilitating learning about various possible outcomes of decisions as predicted by relevant models or helping in generating scenarios for possible development patterns in response to the accumulated expertise of the analyst. This approach is thus devised for a specific type of decision process which typically arises when using environmental or economic models for generating future development scenarios[12] or when using engineering models for computer-aided design.

The main conclusion of the reference point methodology is that, if we want to learn, we must postpone choice; if we postpone choice long enough, it might become self-evident. In this sense, optimization in the reference point methodology is used not necessarily in a sense of the goal of choice, but rather in the sense of a tool of learning. That does not mean that the decisions obtained by applying reference point methodology are arbitrary; if the decision maker learned enough, her/his value function has stabilized and she/he would like to have a support in the final stage of actual decision choice, such a support can be also provided by the reference point methodology, including interactive procedures of choosing best decisions with proven convergence (see Wierzbicki [52]).

References

[1] Antoine, J., G. Fischer and M. Makowski: Multiple Criteria Land Use Analysis. Applied Mathematics and Computation 83(1997)195–215

[2] Bisschop J., R. Entriken: *AIMMS, The Modeling System.* Paragon Decision Technology Haarlem, The Netherlands 1993

[3] Brooke A., D. Kendrick and A. Meeraus *GAMS, A User's Guide.* The Scientific Press Redwood City 1988

[4] Charnes, A., W.W. Cooper: Goal programming and multiple objective optimization. J. Oper. Res. Soc. 1(1977)39–54.

[11]The achievement function only approximates the value function of the user of DSS, which function can be changing during the interactive process of decision support.

[12]We should stress here that the concept of a scenario is usually understood in terms of scenario simulation and analysis – that is, specifying parameters and decisions as a scenario for a given model, then simulating the model for this scenario, finally analysing the outcomes of this scenario. However, here we use the concept in a broader sense, including scenario generation, or inverse scenario analysis – that is, specifying parameters and some desired model outcomes but not decisions and using vector optimization tools, in particular the reference point methodology described here, to generate the decisions which correspond to such inverse scenario.

[5] Dinkelbach, W., H. Isermann: "On decision making under multiple criteria and under incomplete information." In: J.L. Cochrane, M. Zeleny (eds.): *Multiple Criteria Decision Making.* University of South Carolina Press Columbia SC 1973, pp. 302-312

[6] Ermolev, Yu., R. J-B. Wets: *Numerical Techniques for Stochastic Optimization.* Springer Berlin etc. 1988

[7] Fodor, J., M. Roubens: *Fuzzy Preference Modeling and Multicriteria Decision Support.* Kluwer Academic Publishers Dordrecht 1994

[8] Fourer R., D.M. Gay and B.W. Kernighan: *AMPL, A Modeling Language for Mathematical Programming.* The Scientific Press San Francisco 1993

[9] Gal, T.: A general method of determining the set of all efficient solutions to a linear vectormaximum problem. European Journal of Operational Research 1 (1977) 307–322

[10] Geoffrion, A.M.: An Introduction to Structured Modeling. Operations Research 37(1987)30–51

[11] Granat J., M. Makowski: "ISAAP – Interactive Specification and Analysis of Aspiration-based Preferences." Interim Report IR-98-52, International Institute for Applied Systems Analysis, Laxenburg, Austria 1998

[12] Granat, J., A.P. Wierzbicki: Interactive Specification of DSS User Preferences in Terms of Fuzzy Sets. Archives of Control Sciences 5(1996)185–201

[13] Granat, J., T. Kręglewski, J. Paczyński and A. Stachurski: "IAC-DIDAS-N++ Modular Modeling and Optimization System. Part I: Theoretical Foundations, Part II: Users Guide." Internal Report, Institute of Automatic Control, Warsaw University of Technology, Warsaw, Poland 1994

[14] Griewank, A.: Computational Differentiation and Optimization. In: J.R. Birge, K.G. Murty (eds.): *Mathematical Programming: State of the Art 1994.* University of Michigan Press Ann Arbor 1994, pp. 102-131

[15] Haimes, Y.Y., W.A. Hall: Multiobjectives in Water Resources Systems Analysis: The Surrogate Worth Trade-Off Method. Water Resources Research 10(1974)615–623

[16] Henig, M.: Proper efficiency with respect to cones. JOTA 22(1982)618–630.

[17] Janssen, R.: *Multiobjective Decision Support for Environmental Management.* Kluwer Academic Publishers Dordrecht 1992

[18] Jaszkiewicz, A., R. Słowiński: "The LBS package – a microcomputer implementation of the Light Beam Search method for multi-objective nonlinear mathematical programming." Working Paper WP-94-07, International Institute for Applied Systems Analysis, Laxenburg, Austria 1985

[19] Kaliszewski, I.: *Quantitative Pareto Analysis by Cone Separation Techniques.* Kluwer Academic Publishers Dordrecht 1994

[20] Kallio, M., A. Lewandowski and W. Orchard-Hays: "An Implementation of the Reference Point Approach for Multi-Objective Optimization." Working Paper WP-80-35, International Institute for Applied Systems Analysis, Laxenburg, Austria 1980

[21] Kok, M., F.A. Lootsma: Pairwise comparison methods in multiple objective programming. European Journal of Operational Research 22(1985)44–55.

[22] Korhonen, P. and J. Laakso: A Visual Interactive Method for Solving the Multiple Criteria Problem. European Journal of Operational Research 24(1985)277–287.

[23] Korhonen, P., S. Salo and R.E. Steuer: A Heuristic for Estimating Nadir Criterion Values in Multiple Objective Linear Programming. Operations Research 45(1997)751–757

[24] Kręglewski, T., J. Paczyński, J. Granat and A.P. Wierzbicki: "IAC-DIDAS-N: A Dynamic Interactive Decision Support System for Multicriteria Analysis of Nonlinear Models with a Nonlinear Model Generator Supporting Model Analysis." Working Paper WP-88-112, International Institute for Applied Systems Analysis, Laxenburg, Austria 1988

[25] Lewandowski, A., A.P. Wierzbicki (eds.): *Aspiration Based Decision Support Systems.* Lecture Notes in Econ. and Math. Syst. No 331, Springer Berlin etc. 1989

[26] Lootsma, F.A.: Scale sensitivity in multiplicative AHP and SMART. Journal of Multi-Criteria Decision Analysis 2(1993)87–110.

[27] Makowski, M.: "LP-DIT, Data Interchange Tool for Linear Programming Problems, (version 1.20)." Working Paper WP-94-36, International Institute for Applied Systems Analysis, Laxenburg, Austria, Laxenburg 1994

[28] Makowski, M., L. Somlyódy and D. Watkins: Multiple Criteria Analysis for Water Quality Management in the Nitra Basin. Water Resources Bulletin 32(1996)937–951

[29] Makowski, M., J. Sosnowski: "HYBRID: Multicriteria Linear Programming System for Computers under DOS and UNIX." In: J. Wessels, A.P. Wierzbicki (eds.): *User-Oriented Methodology and Techniques for Decision Analysis and Support.* Lecture Notes in Econ. and Math. Syst. No 397, Springer Berlin etc. 1991, pp. 223-233

[30] Matsuhashi, K.: "Application of Multi-Criteria Analysis to Urban Land-Use Planning." Interim Report IR-97-091, International Institute for Applied Systems Analysis, Laxenburg, Austria 1997

[31] Nakayama, H. and Y. Sawaragi: "Satisficing trade-off method for multiobjective programming." In: M. Grauer, A.P. Wierzbicki (eds.): *Interactive Decision Analysis.* Lecture Notes in Econ. and Math. Syst. No 229, Springer Berlin etc. 1983, pp. 113-122

[32] Ogryczak, W., K. Studziński and K. Zorychta: "A generalized reference point approach to multiobjective transshipment problem with facility location." In: A. Lewandowski, A.P. Wierzbicki (eds.): *Aspiration Based Decision Support Systems.* Lecture Notes in Econ. and Math. Syst. No 331, Springer Berlin etc. 1989, pp. 230–250

[33] Osyczka, A.: C Version of Computer Aided Multicriterion Optimization System (CAMOS). Workshop on Advances in Methodology and Software in DSS, IIASA, Laxenburg 1994

[34] Polak, E.: "On the approximation of solutions to multiple criteria decision making problems." In: M. Zeleny (ed.): *Multiple Criteria Decision Making.* Springer New York 1976, pp. 271-282

[35] Rios, S.: *Decision Theory and Decision Analysis: Trends and Challenges.* Kluwer Academic Publishers Boston etc. 1994

[36] Roy, B., Ph. Vincke: Multicriteria Analysis: Survey and New Directions. European Journal of Operational Research 8(1981)207–218.

[37] Saaty, T.: *The Analytical Hierarchy Process.* McGraw-Hill New York 1980

[38] Sawaragi, Y., H. Nakayama and T. Tanino: *Theory of Multiobjective Optimization.* Academic Press New York 1985

[39] Seo, F., M. Sakawa: *Multiple Criteria Decision Analysis in Regional Planning.* Reidel Publishing Company Dordrecht 1988

[40] Simon, H.A.: *Models of Man.* Macmillan New York 1957

[41] Śmierzchalski, R., J. Lisowski: "Computer simulation of safe and optimal trajectory avoiding collision at sea." Joint Communication, WSM Gdynia and Hochschule Bremerhaven 1994

[42] Steuer, R.E., E.V. Choo: An interactive weighted Tchebycheff procedure for multiple objective programming. Mathematical Programming 26(1983)326–344

[43] Steuer, R.E.: *Multiple Criteria Optimization: Theory, Computation, and Application.* J.Wiley & Sons New York 1986

[44] Vincke, Ph.: *Multicriteria Decision Aid.* Wiley Chichester etc. 1989

[45] Wierzbicki, A.P.: "The use of reference objectives in multiobjective optimization." In: G. Fandel, T. Gal (eds.): *Multiple Criteria Decision Making; Theory and Applications.* Lecture Notes in Econ. and Math. Syst. No 177, Springer Berlin etc. 1980, pp. 468-486

[46] Wierzbicki, A.P.: A mathematical basis for satisficing decision making. Mathematical Modeling 3(1983)391–405.

[47] Wierzbicki, A.P.: *Models and Sensitivity of Control Systems.* Elsevier-WNT Amsterdam 1984

[48] Wierzbicki, A.P.: On the completeness and constructiveness of parametric characterizations to vector optimization problems. OR-Spektrum 8(1986)73–87.

[49] Wierzbicki, A.P.: Dynamic Aspects of Multi-Objective Optimization. In: A. Lewandowski, V. Volkovich (eds.): *Multiobjective Problems of Mathematical Programming.* Lecture Notes in Econ. and Math. Syst. No 351 Springer Berlin etc. 1988, pp. 154–174

[50] Wierzbicki, A.P.: Multiple Criteria Games: Theory and Applications. Working Paper WP-92-079, International Institute for Applied Systems Analysis, Laxenburg, Austria 1992

[51] Wierzbicki, A.P.: On the Role of Intuition in Decision Making and Some Ways of Multicriteria Aid of Intuition. Multiple Criteria Decision Making 6(1996)65–78.

[52] Wierzbicki, A.P.: "Convergence of Interactive Procedures of Multiobjective Optimization and Decision Support." In: M.H. Karwan, J.Spronk, J.Wallenius (eds): *Essays in Decision Making.* Springer Berlin etc. 1997, pp. 19–47

[53] Zadeh, L.A.: Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems 1(1978)3–28

[54] Zeleny, M.: A concept of compromise solutions and the method of the displaced ideal. Comput. Oper. Res. 1(1974)479–496

[55] Zimmermann, H.-J., H.-J. Sebastian: "Fuzzy Design – Integration of Fuzzy Theory with Knowledge-Based System Design." In: Fuzzy-IEEE 94 Proceedings, Orlando Florida, Pergamon Press 1994, pp. 352-357

10 CONCEPTS OF INTERACTIVE PROGRAMMING

Theodor J. Stewart

CONTENTS

Abstract: In this chapter we discuss some of the principles underlying what are often called "interactive" methods of MCDM (or "progressive articulation of preferences" in MCDM). These are methods in which the full preference structure of the decision maker is not structured and elicited *a priori*, but is evaluated progressively and locally in response to simple choices made by the decision maker. We differentiate methods in which the responses of the decision maker are expressed in terms of tradeoffs (directly, or indirectly by choices between pairs of outcomes), or in terms of aspiration levels (i.e. desired levels of performance for each criterion). In the final section, we report briefly on simulation studies which have been undertaken in order to assess convergence properties of the interactive methods.

T. Gal et al. (eds.), *Multicriteria Decision Making*

10.1 WHAT IS INTERACTIVE?

At first sight, it may appear strange to find a specific chapter devoted to "interactive" methods of MCDM, as in a very real sense all MCDM must involve interaction with the decision maker. Even the technical mathematical programming aspects of (for example) identifying all efficient solutions, needs to be preceeded by interaction with the decision maker (DM) to identify and to structure the criteria. Any form of preference modeling which may follow will involve a considerable degree of interaction with the decision maker.

Nevertheless, a number of MCDM approaches have come to be termed "interactive methods" (or sometimes "progressive articulation of preferences"), in that they involve the following characteristic steps (cf. also Gardiner and Steuer [4]):

(1) A feasible (and usually efficient) solution, or small number of solutions, is generated according to some specified procedure and presented to the decision maker.

(2) If the decision maker is satisfied with the solution (or one of these) generated, then the process stops. Otherwise, he/she is requested to provide some local preference information in the vicinity of the solution(s) presented, such as direct comparisons between (actual or hypothetical) solutions, tradeoffs, or desired directions of improvement.

(3) In the light of the local information provided, preference models are updated and/or parts of the decision space are eliminated, and the process returns to the first step.

There are certain advantages and disadvantages associated with the use of interactive methods for MCDM in this sense, and these have an influence on the types of MCDM problems for which they are or are not suited. The primary **advantage** of the interactive approach is that the value judgements which have to be made by the decision maker are set in a realistic context. This is in contrast to the situation with certain other MCDM methodologies in which global judgements (such as importance weights of criteria) have to be made, often involving the need to compare quite extreme combinations of outcomes not encountered in practice. The **disadvantage** is that process is relatively unstructured, with the result that it may be difficult to motivate clearly the solution obtained. The following observations may thus be made:

- The use of interactive methods is well suited to situations in which a single individual, or a small homogeneous group, needs quite quickly to find one or a few satisfactory alternatives to a decision problem, primarily when the criteria are well-represented in terms of quantitative attributes. The context may be that of taking a decision directly, or that of generating a shortlist of alternatives for more detailed evaluation later. The exploration of the efficient frontier which is facilitated by the interactive approach can assist the decision maker in learning quickly about the options and tradeoffs which are available.

- Interactive methods are not well suited to use in group decision making contexts in which there are substantial conflicts, in which many criteria are qualitative, or in which a clearly defensible justification for the solution obtained needs to be established (as is often the case in strategic public sector policy decisions). Of course, individual members of the group may well find that the interactive methods are useful in exploring the options for themselves prior to the group sessions.

Both observations above do clearly indicate that the most important benefit may not be in the generation of an optimal solution *per se*, but rather in the insights and understanding generated for the decision maker in exploring the decision space in a systematic and coherent manner.

Shin and Ravindran [25] provide a comprehensive review (including 116 references) of interactive methods published up to about 1990, by which time the basic approaches can be said to have been established. They also provide a classification of interactive methods into 10 categories, differentiated primarily in terms of:

- *Style of interaction with the decision maker*, which may be in the form of classifying criteria as satisfactory or otherwise, tradeoffs (which may be global or local, precise or imprecise), direct comparisons of solutions, or aspiration levels; and

- *Solution approach*, which they see as differentiated between "point assessment", in which a sequence of specific solutions is evaluated, and "interval reduction", in which the decision space is systematically pruned.

Gardiner and Steuer [4, 5] also review a number of the most prominent interactive procedures, most of which we shall refer to below. Their intention was to identify many of the common themes running through the different methodologies, and to use these to suggest "unified" procedures, linking various approaches and allowing users to switch between methodologies. This is an interesting concept worthy of further empirical research. One of the important research questions would have to concern the optimal balance between offering different means of exploring the decision space and the potential for confusing the user through too many different types of information.

The purpose of the present chapter is not to repeat the still quite current reviews of Shin and Ravindran, and of Gardiner and Steuer. The intention is rather to present an overview of the broad streams of thought and their rationale, and to extract general principles for the implementation of interactive methods. As indicated in the above-mentioned review papers, there are a number of different means of classifying interactive methods. For this paper, we have chosen to differentiate approaches primarily according to whether the preferences of the decision maker are modeled in terms of value functions or of aspiration levels. In the case of value function models, we differentiate further according to whether information from the decision maker is primarily in the form of tradeoffs or comparisons of alternatives. We have elsewhere (e.g.

Stewart [35]) argued that the use of value functions and of aspiration levels characterize two of the three fundamentally different approaches to (or "schools" of) MCDM. The third approach is that of outranking, but these methods are not generally expressed in the interactive framework defined above, and will thus not be dealt with further in this chapter (but see Vetschera [40] for at least one interactive outranking approach).

Let us now establish the notation to be used in this chapter. Let $\mathcal{A}$ represent the set of alternatives from which either a single element or a shortlist needs to be selected. Each action $a \in \mathcal{A}$ is associated with an m-dimensional vector of attributes $\boldsymbol{z}^a = (z_1^a, z_2^a, \ldots, z_m^a)$ (i.e. levels of achievement for each of m criteria). Unsuperscripted symbols $\boldsymbol{z}$ and z_i will be used to represent arbitrary attribute values, not necessarily associated with any specific alternative. Let $\mathcal{Z} = \{\boldsymbol{z} \in \mathbb{R}^m | \boldsymbol{z} = \boldsymbol{z}^a \text{ for some } a \in \mathcal{A}\}$, i.e. the set of attainable attribute vectors. Without loss of generality, we shall suppose that the attributes are defined in such a manner that increasing values are preferred. We suppose also that preferences between alternatives can be stated (by the decision maker) in terms of preferences between attribute vectors, and will use the notation $\boldsymbol{z}^a \succ \boldsymbol{z}^b$ to represent preference for alternative a over b, and $\boldsymbol{z}^a \succeq \boldsymbol{z}^b$ to represent preference or indifference, i.e. a is at least as good as b.

The set $\mathcal{A}$ is typically assumed either to be finite, consisting of N discrete choices, or to be a subset of $\mathbb{R}^n$, so that each element of $\mathcal{A}$ can be represented by a decision vector (i.e. vector of decision variables) $\boldsymbol{x} \in \mathbb{R}^n$. In the latter case, we shall for ease of notation often refer directly to "$\boldsymbol{x} \in \mathcal{A}$", where $\mathcal{A} \subset \mathbb{R}^n$ is now viewed as the set of decision vectors corresponding to the feasible set of alternatives. The attribute values can then be represented by real-valued criterion functions $z_i = f_i(\boldsymbol{x})$, or in vector form by $\boldsymbol{z} = \boldsymbol{f}(\boldsymbol{x})$.

10.2 VALUE FUNCTION METHODS WITH TRADEOFF INFORMATION

Many interactive methods start with the assumption of the existence of some form of value function,$V(\boldsymbol{z})$, which can represent preferences of the decision maker in the sense that $V(\boldsymbol{z}^a) > V(\boldsymbol{z}^b)$ if and only if $\boldsymbol{z}^a \succ \boldsymbol{z}^b$. No attempt is made to fully specify or estimate this value function, however, and generally relatively mild assumptions concerning the form of $V(\boldsymbol{z})$ are made, such as that of pseudoconcavity. Stronger assumptions such as additivity (implying the need for preferential independence of criteria) are often not required, so that interactive value-based methods are potentially usable in a "do-it-yourself" mode (i.e. not necessarily under the direct guidance of a skilled MCDM facilitator), without the need for careful tests for assumptions such as preferential independence. Furthermore, even the assumptions which are made are only applied in a local sense, i.e. to provide sufficient guidance to the DM concerning potentially better regions of the decision space which need to be explored. The underlying philosophy is that preferences tend to evolve and to develop as greater understanding of the problem is attained, with the result that the value function itself may change during the process.

Perhaps the earliest interactive procedure of this type is that described by Geoffrion, Dyer and Feinberg [6] for the case in which $\mathcal{A}$ is a compact, convex subset of $\mathbb{R}^n$. Their approach was based on maximizing $V(\boldsymbol{f}(\boldsymbol{x}))$ subject to $\boldsymbol{x} \in \mathcal{A}$ using the Franke-Wolfe algorithm, but using the decision maker, in effect, as a function and gradient evaluator. Each iteration of the process starts with a feasible point, say $\boldsymbol{x}^1 \in \mathcal{A}$, and involves in principle the following steps:

(1) Determine $\nabla V(\boldsymbol{f}(\boldsymbol{x}^1))$, i.e. the gradient vector of $V(\boldsymbol{f}(\boldsymbol{x}))$ at $\boldsymbol{x} = \boldsymbol{x}^1$;

(2) Obtain a direction of improvement $\boldsymbol{d} = \boldsymbol{y} - \boldsymbol{x}^1$, where $\boldsymbol{y}$ is the solution to:

$$\text{Maximize}_{\boldsymbol{y} \in \mathcal{A}} \ \nabla V(\boldsymbol{f}(\boldsymbol{x}^1)) \cdot \boldsymbol{y}.$$

(3) Solve the one-dimensional search problem:

$$\text{Maximize}_{0 \le t \le 1} \ V(\boldsymbol{f}(\boldsymbol{x}^1 + t\boldsymbol{d}))$$

and set $\boldsymbol{x}^2 = \boldsymbol{x}^1 + t\boldsymbol{d}$ as the starting point for the next iteration.

The problem is that the value function is assumed not to be known explicitly, so that neither the gradient (in the first step) nor the linear search (in the third step) can be evaluated explicitly. Note, however, that we can express the gradient in the form:

$$\nabla V(\boldsymbol{f}(\boldsymbol{x}^1)) = \sum_{i=1}^{m} \frac{\partial V(\boldsymbol{z})}{\partial z_i}\bigg|_{\boldsymbol{z}=\boldsymbol{f}(\boldsymbol{x}^1)} \nabla f_i(\boldsymbol{x}^1). \tag{10.1}$$

Geoffrion et al. pointed out that the linearized optimization problem in the second step is unaffected if all elements of $\nabla V(\boldsymbol{f}(\boldsymbol{x}^1))$ are divided through by a common positive term. Thus dividing through by $\partial V(\boldsymbol{z})/\partial z_1$ evaluated at $\boldsymbol{z} = \boldsymbol{f}(\boldsymbol{x}^1)$ (where choice of the attribute to be the reference z_1 is arbitrary), this optimization can be expressed in the form:

$$\text{Maximize}_{\boldsymbol{y} \in \mathcal{A}} \ \sum_{i=1}^{m} w_i \nabla f_i(\boldsymbol{x}^1) \cdot \boldsymbol{y} \tag{10.2}$$

where the weight w_i is simply the tradeoff between attributes 1 and i, i.e. the amount of attribute 1 (the reference) which the decision maker is prepared to sacrifice, in order to obtain a unit gain in attribute i, defined by:

$$w_i = \frac{\partial V(\boldsymbol{z})/\partial z_i}{\partial V(\boldsymbol{z})/\partial z_1}$$

evaluated at $\boldsymbol{z} = \boldsymbol{f}(\boldsymbol{x}^1)$. The required tradeoffs need to be provided by the decision maker. These may be assessed either directly (by comparing each attribute to a pre-selected reference attribute), or indirectly by means of a number of more natural "enquiries" as suggested by Rosinger [23]. Once the tradeoffs

have been established, the first two steps of the algorithm can immediately be implemented. The absence of an explicit value function is less of a problem for the final step, as the DM can be presented with a sequence of alternatives along the linear search direction (for example by using a Fibonacci search or secant methods), from which an approximate optimal position can be ascertained by direct evaluation.

The optimization steps are based on relatively imprecise information, and only a relatively small number of iterations can be executed in practice. Geoffrion et al. provide some practical evidence, nevertheless, that good solutions can be obtained. One difficulty in principle with the approach is, however, the fact that preference information from one iteration is discarded before the next iteration. This is very inefficient use of information, the assessment of which is both a difficult task for the DM and imprecise. While local tradeoffs may be expected to change as the DM moves through the decision space, it is also true that in most cases the changes would be gradual. With this in mind, Sakawa [24] suggested the use of "*proxy functions*" as local estimates to value function. He proposed a few families of functions (sums of exponential, power or log functions), from which an appropriate form could be selected. The parameters of the chosen family of functions could be estimated from the more recent tradeoffs provided by the DM, thus generating continually updated versions of the approximations. The proxy functions could be optimized directly instead of using the Frank-Wolfe algorithm, providing promise of much more rapid convergence to the DM's most preferred solution.

More effective use can be made of the tradeoff information if the assumption is made that the value function is pseudoconcave, which means that if for any two attribute vectors, say $\boldsymbol{z}^a$ and $\boldsymbol{z}^b$, we have that:

$$\nabla V(\boldsymbol{z}^a) \cdot (\boldsymbol{z}^b - \boldsymbol{z}^a) \leq 0 \tag{10.3}$$

then $V(\boldsymbol{z}^a) \geq V(\boldsymbol{z}^b)$, i.e. $\boldsymbol{z}^b \not\succ \boldsymbol{z}^a$. As in the Geoffrion-Dyer-Feinberg approach, we can divide through by $\partial V(\boldsymbol{z})/\partial z_1|_{\boldsymbol{z}=\boldsymbol{z}^a}$, so that the condition for concluding that $\boldsymbol{z}^a \succeq \boldsymbol{z}^b$ is equivalent to:

$$\sum_{i=1}^{m} w_i(z_i^b - z_i^a) \leq 0 \tag{10.4}$$

where the w_i are the tradeoffs assessed at the solution represented by the attribute vector $\boldsymbol{z}^a$. Note that this result applies equally well whether $\mathcal{A}$ is discrete or continuous.

The above property of pseudoconcave value functions can be used to enhance the Geoffrion-Dyer-Feinberg algorithm, by a process of generating "*tradeoff cuts*" (Musselman and Talavage [19]) to reduce the decision space following each iteration. Suppose that tradeoffs as defined above are assessed at the point $\boldsymbol{x}^1$, at which the attribute values are given by $\boldsymbol{f}(\boldsymbol{x}^1)$. By the assumed pseudoconcavity property, an alternative solution $\boldsymbol{x}$, say, can only be preferred to $\boldsymbol{x}^1$ if $\sum_{i=1}^{m} w_i(f_i(\boldsymbol{x}) - f_i(\boldsymbol{x}^1)) \geq 0$, where once again the weights are the tradeoffs

assessed by the decision maker at the point $\boldsymbol{x}^1$. If we define $\mathcal{A}^0 = \mathcal{A}$, and if at iteration k, tradeoff weights w_i^k are assessed at the point $\boldsymbol{x}^k$, then the search for $\boldsymbol{x}^{k+1}$ can be restricted to $\mathcal{A}^k = \mathcal{A}^{k-1} \cap \{\boldsymbol{x} \in \mathbb{R}^n | \sum_{i=1}^m w_i^k (f_i(\boldsymbol{x}) - f_i(\boldsymbol{x}^k)) \geq 0\}$. The search for the direction of improvement for the next iteration of the Geoffrion-Dyer-Feinberg approach is thus refined by limiting the optimization step in (10.2) to $\boldsymbol{y} \in \mathcal{A}^k$.

In introducing tradeoff cuts, however, Musselman and Talavage [19] replaced the optimization steps of the Geoffrion-Dyer-Feinberg algorithm by a search for the most central point of $\mathcal{A}^k$. The rationale behind this is that at each iteration of the algorithm it may be more advantageous for the decision maker to examine a solution which will lead to a large reduction in the decision space (through the next tradeoff cut), than to seek an immediately maximum rate of value improvement. This is an important insight, which has also been used in other interactive approaches, e.g. by Zionts and Wallenius [43] who introduced the term "*middlemost*" to denote such estimates. These basic themes are refined by Loganathan and Sherali [14], who argue for applying the middlemost concept to constraints generated by the tradeoff cuts, but not to the constraints defining the initial feasible region $\mathcal{A}$, so as to ensure that at each iteration the decision maker only needs to evaluate efficient solutions. They suggested that by obtaining the decision maker's tradeoffs at efficient solutions only, more meaningful marginal rates of substitution can be achieved. Numerical studies on the above tradeoff cut approaches, and some variations on them, have been reported by Shin and Ravindran [26].

10.3 VALUE FUNCTION METHODS USING DIRECT COMPARISONS

Instead of requiring the decision maker to provide tradeoffs between criteria, preference information can also be obtained from direct comparisons by the decision maker of two or more decision alternatives (represented in terms of their attribute vectors $\boldsymbol{z}$). Such information can be used either directly to draw inferences concerning the form of the underlying value function, or indirectly to eliminate parts of the decision space (as with the tradeoff cuts above).

Let $\boldsymbol{z}^*$ be the attribute vector which maximizes $V(\boldsymbol{z})$ over $\boldsymbol{z} \in \mathcal{Z}$. We continue to make the assumption that the value function is pseudoconcave, but now assume that the set $\mathcal{Z}$ is convex. Note that convexity of $\mathcal{Z}$ need not necessarily be implied by the convexity of $\mathcal{A}$, although in the case of multiple objective linear programming, for which most of the methods described in this section were originally developed, both $\mathcal{A}$ and $\mathcal{Z}$ are convex. Under these assumptions, it follows that there exist non-negative weights $\lambda_1, \lambda_2, \ldots, \lambda_m$ such that $\boldsymbol{z}^*$ maximizes $\sum_{i=1}^m \lambda_i z_i$ over $\mathcal{Z}$. This simple observation has been used as the basis for a number of interactive procedures, in which preference information obtained from the decision maker is used to increasingly restrict the set of allowable weight vectors, which implicitly then also restricts the range of solutions to be considered. As indicated above, these methods have generally been framed in the context of multiple objective linear programming, which has

the advantage of providing a natural point of termination of the process when all retained weight vectors generate the same optimal basis.

One such scheme is that introduced by Zionts and Wallenius [42] and later refined by them in [43]. The basic concepts underlying their approach (ignoring for the moment some of the refinements needed to ensure convergence) can be summarized in the following steps:

(1) Let $\Lambda = \{\boldsymbol{\lambda} \in \mathbb{R}^m \,|\lambda_i \geq 0; \sum_{i=1}^m \lambda_i = 1\}$. Select a $\boldsymbol{\lambda} \in \Lambda$ and maximize $\sum_{i=1}^m \lambda_i z_i$ for $\boldsymbol{z} \in \mathcal{Z}$. Let $\boldsymbol{z}^*$ be the resultant attribute vector.

(2) Search for a Λ-efficient solution, say $\boldsymbol{z}^1$, distinct from that represented by $\boldsymbol{z}^*$, such that the decision maker prefers $\boldsymbol{z}^1$ to $\boldsymbol{z}^*$ (where $\boldsymbol{z}$ is Λ-efficient if it maximizes $\sum_{i=1}^m \lambda_i z_i$ over $\boldsymbol{z} \in \mathcal{Z}$ for some $\boldsymbol{\lambda} \in \Lambda$). In the case of multiple objective linear programming, the search is typically restricted to basic solutions adjacent to that represented by $\boldsymbol{z}^*$. If no such solution is found, then STOP.

(3) Restrict the set Λ to those weight vectors satisfying $\sum_{i=1}^m [z_i^1 - z_i^*]\lambda_i > 0$. If Λ is now empty, then start deleting the oldest constraints of this form until Λ is no longer empty.

(4) Select a $\boldsymbol{\lambda} \in \Lambda$ and maximize $\sum_{i=1}^m \lambda_i z_i$ for $\boldsymbol{z} \in \mathcal{Z}$. (Zionts and Wallenius suggest using the "middlemost" solution, viz. that which maximizes the minimum slack over all constraints.) Let $\boldsymbol{z}^0$ be the resultant attribute vector.

(5) Set $\boldsymbol{z}^* = \boldsymbol{z}^0$ if $\boldsymbol{z}^0$ is preferred by the decision maker to $\boldsymbol{z}^1$, or $\boldsymbol{z}^* = \boldsymbol{z}^1$ otherwise. Include an additional constraint as in step 3 consistent with the preference stated between $\boldsymbol{z}^0$ and $\boldsymbol{z}^1$, and return to step 2.

The above steps are intuitively appealing, and will in fact lead to the maximum value solution if the value function is linear and the decision maker expresses preferences consistent with this linear value function at all times. In practice, however, tradeoffs will vary across the decision space, with the result that the inequalities in the λ_i generated from local preferences at one point in the decision space might not apply at the true optimum point, leading to termination at a sub-optimal solution. With this and other practicalities in mind, Zionts and Wallenius [43] introduced the following refinements:

- If no distinctly different adjacent basic solutions are found in step 2, then search for tradeoffs corresponding to Λ-efficient edges emanating from $\boldsymbol{z}^*$ which the decision maker considers desirable. If such a preferred tradeoff is found, then a similar constraint can be added to that of step 3, and the procedure continues.

- If step 2 yields neither a preferred adjacent solution nor a preferred tradeoff, then the search is extended to edges and adjacent solutions which are efficient but not Λ-efficient.

- If only preferred tradeoffs are identified, then there will be no z^1 solution. In this case, the comparison in the last step is made between z^0 and z^*; if $z^* \succ z^0$, then the procedure terminates. The solution is, however, only locally optimal (the optimal vertex), but more preferred non-basic solutions exist. Strictly speaking, adjacent facets of the LP should be investigated to find a truly globally optimal solution, but it is not clear that this will give any great advantage in MOLPs of realistic size.

Since the refined procedure either moves to an improved solution at each step, or terminates with a basic solution which is locally optimal amongst adjacent basic solutions, the result must be the extreme point of the simplex which has the largest value of $V(z)$.

A more direct, but rather heuristic approach to the same problem of finding the "correct" weight vector λ has been suggested by Steuer and co-workers (see for example Steuer [29], Chapters 13 and 14). Steuer proposes a number of variations to the basic idea, perhaps the easiest of which to describe being that which he terms the "interactive weighted-sums/filtering approach" (Steuer [29], Section 13.5). The set of feasible weight vectors at any step of the procedure is defined in the form $\Lambda = \{\lambda \in \mathbb{R}^m \,|\ell_i \leq \lambda_i \leq \mu_i; \sum_{i=1}^m \lambda_i = 1\}$. Initially the bounds are $\ell_i = 0$ and $\mu_i = 1$, but the interval is systematically shrunk at each iteration. In outline, the procedure can be described in terms of the following steps:

(1) Initialize $\ell_i = 0$ and $\mu_i = 1$ for all criteria i; let $W = 1$ be the width of the interval for each i.

(2) Randomly generate a specified number of vectors from Λ, and then filter these to obtain a smaller number of widely dispersed vectors. For each vector generated, find the corresponding attribute vector maximizing $\sum_{i=1}^m \lambda_i z_i$ for $z \in \mathcal{Z}$. Filter these solutions again to generate a specified number of attribute vectors which are as widely dispersed as possible. Let these be $z^1, z^2, \ldots, z^P$ say.

(3) Let the decision maker select the most preferred of $z^1, z^2, \ldots, z^P$. Denote by z^0 the vector which is chosen, and let λ^0 be the weight vector for which z^0 is optimal.

(4) Replace W by rW, where r is a chosen reduction factor. Select new values for ℓ_i and μ_i such that $\mu_i - \ell_i = W$, and such that λ_i^0 is positioned as close as possible to the centre of the interval while ensuring that $\ell_i \geq 0$ and $\mu_i \leq 1$.

(5) Return to step 2.

The basic philosophy, of course, is that the correct weights for generating the optimal solution should be in the vicinity of the weights generating the most preferred of a sample set of solutions, than in the vicinity of weights generating the less preferred solutions. The approach is essentially heuristic, as no guarantee of convergence to the optimal solution can be given, but has a degree of

plausibility, is easy to explain to users, encourages systematic exploration of the decision space, and is generally quick and easy to apply.

Steuer and Choo [30] applied the same general idea to the context in which solutions are generated not by maximizing a linear approximation to the value function, but by minimizing the distance from the ideal solution according to an augmented weighted Tchebycheff norm. Thus, if $z_i^{**} = \max_{z \in \mathcal{Z}}\{z_i\}$, then the solution corresponding to a weight vector $\boldsymbol{\lambda}$ is obtained by minimizing $D - \epsilon \sum_{i=1}^{m} z_i$ for some small $\epsilon > 0$, subject to:

$$D \geq \lambda_i(z_i^{**} - z_i) \quad \text{for } i = 1, 2, \ldots, m.$$

This is perhaps closer to concept of aspiration level methods (cf. Section 10.5), and indeed Steuer et al. [31] have extended the idea to combine the augmented Tchebycheff approach with the reference point approach of Wierzbicki. It is worth noting that the weighted Tchebycheff and related procedures are directly applicable to non-linear and even discrete problems.

Another approach for multiple objective linear programming which shares some of the features of the above methods is "SIMOLP" (Simplified Interactive Multiple Objective Linear Programming) introduced by Reeves and Franz [22]. SIMOLP also makes use of linear weighted sums of attribute values to generate solutions, with adjustment of weights in the light of selections made by the decision maker, although the authors stress that the weights should not be viewed as approximating a value function, but simply as a device for generating new solutions. The method can be described in terms of the following steps:

(1) Maximize each objective in turn, to generate m solutions, say $\boldsymbol{z}^1, \ldots, \boldsymbol{z}^m$.

(2) Identify the m-dimensional hyperplane, described by $\sum_{i=1}^{m} \lambda_i z_i = c$, which passes through all m points. In order to ensure that the solution to be obtained in the next step is efficient, set any negative λ_i-values to zero.

(3) Maximize $\sum_{i=1}^{m} \lambda_i z_i$ subject to $z \in \mathcal{Z}$, in order to generate a solution which is a compromise between the previous m solutions, and let the solution be $\boldsymbol{z}^*$.

(4) If $\boldsymbol{z}^*$ is distinct from all of $\boldsymbol{z}^1, \boldsymbol{z}^2, \ldots, \boldsymbol{z}^m$, and is preferred to at least one of them, then replace the least preferred of the m solutions by $\boldsymbol{z}^*$, and return to step 2. Otherwise the procedure terminates.

The procedures of Zionts and Wallenius and of Steuer and co-workers described above systematically reduce the decision space *indirectly* by means of restrictions placed on the range of value functions. Comparisons between alternatives can also be used to place direct constraints on the decision space itself. This is particularly useful in the case of discrete alternatives where linear value functions can exclude efficient solutions (which are "convex dominated", i.e. dominated by a hypothetical alternative formed by the convex combination

of the attribute values of two other alternatives). The basic idea is relatively simple. For any pair of attribute vectors z^a and z^b such that $z^a \succ z^b$ we define:

$$C[z^a \succ z^b] = \{z \in \mathcal{Z} \,|\, z = z^b + \mu(z^b - z^a) \text{ for some } \mu > 0\}. \tag{10.5}$$

This set is commonly referred to as a "*convex cone*", although it is strictly speaking the differences $z - z^b$ which constitute the cone. If the underlying value function is quasi-concave (a slightly weaker assumption than that of pseudoconcavity), then it follows directly from the definition of quasi-concavity that $z^b \succ z$ for any z which either belongs to, or is dominated by an element of, $C[z^a \succ z^b]$. The corresponding alternative can in consequence be eliminated from further consideration.

A single preference statement as above does not usually provide much of a reduction in the decision space. The approach becomes much more powerful if the least preferred of a set of more than two alternatives can be provided. Thus suppose that z^q has been identified as the least preferred of q attribute vectors $z^1, z^2, \ldots, z^q$. We then define:

(10.6)

$$C[z^1, z^2, \ldots, z^{q-1} \succ z^q] = \{z \in \mathcal{Z} \,|\, z = z^q + \sum_{k=1}^{q-1} \mu_k(z^q - z^k) \text{ for some } \mu_i > 0\}$$

which is also referred to as a q-point convex cone (with the same caveat as above). The assumption of a quasi-concave value function then implies again that $z^q \succ z$ for any z either belonging to or dominated by an element of $C[z^1, z^2, \ldots, z^{q-1} \succ z^q]$, so that the corresponding alternative can be eliminated. It is simple to test whether any particular attribute vector z is in or dominated by $C[z^1, z^2, \ldots, z^{q-1} \succ z^q]$ by solving the linear programming problem:

$$\text{Maximize } \epsilon$$

subject to

$$\sum_{k=1}^{q-1} \mu_k(z_i^q - z_i^k) - \epsilon \geq z_i - z_i^q \quad \text{for } i = 1, 2, \ldots, m.$$

If the optimum value for this LP is non-negative then z can be eliminated.

The use of convex cones in the above manner was introduced by Korhonen et al. [13] as a method for extending the Zionts-Wallenius approach to discrete choice problems in MCDM. Although the approach was developed initially for the discrete problem, Ramesh et al. [21] have pointed out that it is a useful addition to the Zionts-Wallenius algorithm for the multiple objective linear programming as well. The advantage in the linear programming case is that even if earlier constraints on the weights λ need to be deleted, it is still valid to retain the convex cones derived from the same comparisons, so that the earlier preference information is not completely lost.

While the convex cone approach is potentially useful particularly for the case of a discrete set of alternatives, its value is limited if attention is restricted entirely to actual alternatives in $\mathcal{A}$. The problem is that if z^q is substantially less preferred than any of $z^1, z^2, \ldots, z^{q-1}$, then the preference statement may be relatively uninformative, and may not reduce the decision space to any great degree. In recognition of this problem, Köksalan et al. [8], Köksalan and Sagala [9], and Köksalan and Taner [10] have suggested the use of "*dummy alternatives*" to generate the cones. In essence, convex combinations of actual attribute vectors are generated in such a way that all are only slightly preferred to the vector z^q. It is shown that cones generated in this way are much more informative and eliminate larger portions of the decision space. The cited references suggest a number of practical means by which useful dummy alternatives may be generated.

The links between the convex cones and the Zionts-Wallenius procedures have been extended by Prasad et al. [20], who introduce the concept of "p-cone efficiency". For any specific attribute vector z, we minimize p subject to:

$$\sum_{k=1}^{q-1} \mu_k(z_i^q - z_i^k) - pz_i \geq z_i - z_i^q \quad \text{for } i = 1, 2, \ldots, m.$$

This is an alternative test for cone dominance of z in the sense that if the solution is zero then z is cone-dominated. If the optimizing value of p is greater than zero, then $z^q + \sum_{k=1}^{q-1} \mu_k(z^q - z^k) \geq (1-p)z$, i.e. $(1-p)z$ is cone dominated. The magnitude of p thus gives a measure of how nearly cone dominated the vector z is. Prasad et al. suggest using this idea in two ways. Firstly, the alternative giving the maximum value for p is selected for comparison with the current best solution, and generation of a new cone, as this appears to give a maximum rate of improvement in the interactive process. Secondly, the procedure can be terminated with little loss if all remaining alternatives have values for p less than a given threshold, i.e. are nearly cone dominated.

Malakooti [17] suggests another approach to selecting an alternative to be compared with the current best solution, also with the idea of providing maximum information and progress with each comparison. His proposal is to choose the next alternative to be compared as that which maximizes the number of alternatives in $\mathcal{A}$ which would be eliminated if the current best solution turns out to be preferred to it. It is claimed that this gives a high likelihood of improvement at each stage, while even if there is no immediate improvement the feasible space is maximally reduced. Malakooti also allows the use of gradient, i.e. tradeoff information, in which case the cones are equivalent to the tradeoff cuts which we have discussed earlier.

10.4 A GENERALIZED INTERACTIVE VALUE FUNCTION APPROACH

The methods discussed in the previous two sections all assume the existence of a value function with certain properties, typically pseudoconcavity. Preferences

expressed by the decision maker are translated either into explicit restrictions on the family of value functions best describing the decision maker's preferences, or into restrictions on the set of alternatives which are potentially optimal for value functions consistent with the expressed preferences. The preferences are expressed either in terms of choice between two or more actual or hypothetical (dummy) alternatives, or in terms of tradeoffs. Tradeoffs can, in fact, be viewed as statements of indifference between a real and hypothetical alternative, differing on two attributes only.

In previous papers [32, 36] we have suggested an integrated or generalized approach, in which both restrictions on the family of value functions and on the decision space can be generated from responses of either the direct comparison or tradeoff forms. The approach represents a more formal estimation of the underlying value function, based on the piecewise linear additive value function concept originally used as the basis for the "UTA" (Utilité Additive) model of Siskos [27]. This generalized interactive approach does assume the existence of an additive value function, so that the property of preferential independence is presumed to hold, and should in principle be checked. On the other hand, weaker assumptions than concavity or pseudoconcavity can be made.

The piecewise linear marginal value functions as used in the interactive methods described here is formulated as follows. The basic model is of the additive form:

$$V(z) = \sum_{i=1}^{m} v_i(z_i). \tag{10.7}$$

Let $z_{i(min)}$ and $z_{i(max)}$ represent the limits of attribute values which will require consideration (which may often, but need not necessarily, represent the range of outcomes amongst the specific alternatives which have to be compared). Since the preference ordering implied by (10.7) is unchanged by the addition or subtraction of a constant term, we can always define the partial value functions such that $v_i(z_{i(min)}) = 0$.

The partial value function $v_i(z_i)$ is often decomposed into a product of a standardized function (e.g. taking on values 0 and 100 at $z_{i(min)}$ and $z_{i(max)}$ respectively), and an importance "weight". It is convenient in our context to retain the unstandardized form, in which case $v_i(z_{i(max)})$ represents the importance of the increase in values for attribute i from $z_{i(min)}$ to $z_{i(max)}$, relative to the corresponding increases in the other attributes. We now represent each function $v_i(z_i)$ in piecewise linear form. In simulation studies (cf. Section 10.6) it has been found that the use of purely linear partial value functions can severely bias results, but that this problem can largely be overcome by subdividing the function into a small number (as few as 3 or 4) of linear segments. For ease of nomenclature, let us assume that the same number of segments (say ν) is used for each attribute (although this is not essential to the method). The ν linear segments are defined by $\nu + 1$ "breakpoints", including the endpoints of the range, say $z_{i(min)} = z_{i(0)} < z_{i(1)} < \cdots < z_{i(\nu)} = z_{i(max)}$. These could be defined by the decision maker directly, or could simply be equally spaced

across the interval. Note that the slope of the linear approximation for $v_i(z_i)$ over the j-th segment is given by $u_{ij}/(z_{i(j)} - z_{i(j-1)})$ which we shall represent by $t_{ij}u_{ij}$, where $t_{ij} = 1/(z_{i(j)} - z_{i(j-1)})$.

The marginal value function $v_i(z_i)$ is then fully defined by ν parameters u_{ij} $(j = 1, \ldots, \nu)$, where $u_{ij} = v_i(z_{i(j)}) - v_i(z_{i(j-1)})$, i.e. the increase in value corresponding to a change from $z_{i(j-1)}$ to $z_{i(j)}$. For an arbitrary value of z_i, not necessarily occurring at one of the breakpoints, we can express $v_i(z_i)$ in the form:

$$v_i(z_i) = \sum_{j=1}^{\nu} c_{ij}(z_i)u_{ij} \tag{10.8}$$

where $c_{ij}(z_i) = 0$ for each j such that $z_i \leq z_{i(j-1)}$, $c_{ij}(z_i) = 1$ for each j such that $z_i \geq z_{i(j)}$, while:

$$c_{ij}(z_i) = \frac{z_i - z_{i(j-1)}}{z_{i(j)} - z_{i(j-1)}}$$

for the segment j (unique if it exists) for which $z_{i(j-1)} < z_i < z_{i(j)}$.

The parameters u_{ij} for $i = 1, \ldots, m$, $j = 1, \ldots, \nu$ must be non-negative, and need to be standardized in some manner, for example such that $\sum_{i=1}^{m}\sum_{j=1}^{\nu} u_{ij} = 100$. Additional *a priori* constraints on the u_{ij} can be imposed in order to represent any prior knowledge of the shape of the function. For example, concavity would require that:

$$t_{i,j+1}u_{i,j+1} < t_{ij}u_{ij}. \tag{10.9}$$

A quite rich family of functional shapes for the marginal value functions, motivated in part by the empirical work of Kahnemann and Tversky [7], is obtained by restricting $v_i(z_i)$ to be convex, concave or "S-shaped". These three forms taken together imply that once $t_{i,j+1}u_{i,j+1} < t_{ij}u_{ij}$ for *some* j, then $t_{i,k+1}u_{i,k+1} < t_{ik}u_{ik}$ for all $k > j$. With a view to the LP estimation procedures to be described below, we note that this condition can be enforced by defining integer variables $\delta_{ij} \in \{0, 1\}$ for each attribute i and for $j = 2, \ldots, \nu$, and imposing the constraints:

$$-M(1 - \delta_{ij}) \leq t_{ij}u_{ij} - t_{i,j-1}u_{i,j-1} \leq M\delta_{ij} \quad \text{for } j = 2, \ldots, \nu \tag{10.10}$$

for suitably "big" M, and:

$$\delta_{i,j+1} \geq \delta_{ij} \quad \text{for } j = 2, \ldots, \nu - 1 \tag{10.11}$$

for each attribute.

It is interesting to note that a pure "S-shape" can be imposed without recourse to integer variables if $\nu = 4$, by simply adding the constraints

$$t_{i1}u_{i1} < t_{i2}u_{i2} \text{ and } t_{i3}u_{i3} > t_{i4}u_{i4}. \tag{10.12}$$

Any specific case would be based on one of the above sets of constraints, i.e. (10.9), or (10.10) and (10.11), or (10.12). We shall generally refer to these as "*shape constraints*", which we note are linear.

With this background, an approach essentially similar to that of Zionts-Wallenius can be implemented, with greater scope for modeling changing tradeoffs and less need for discarding earlier preference statements. Suppose that the decision maker states a preference for z^a over z^b. This implies the following constraint on the u_{ij}:

$$\sum_{i=1}^{m}\sum_{j=1}^{\nu}[c_{ij}(z_i^a) - c_{ij}(z_i^b)]u_{ij} \geq 0. \tag{10.13}$$

In principle, after any number of such preference statements have been made, a "middlemost" estimate of the u_{ij}'s can be obtained by maximizing the minimum slack across all inequalities of the form given by (10.13), subject to the non-negativity and standardization of the parameters, and to any desired shape constraints. In fact, by allowing the "slacks" to be negative, this approach will also generate the best fit parameter values when the preference statements are inconsistent with the model.

In practice, however, there is (as shown by the studies in [36]) a substantial advantage to be gained by searching for "dummy" alternatives between which the decision maker is approximately indifferent, much as discussed in the context of convex cones above. Some suggestions for achieving this end are given in [36]. Note that if the dummy alternatives differ on two criteria only, then this becomes equivalent to the assessment of tradeoffs. In any case, if we have two attribute vectors, say z^a and z^b (real or dummy), between which the decision maker is approximately indifferent, then we would replace (10.13) by:

$$\sum_{i=1}^{m}\sum_{j=1}^{\nu}[c_{ij}(z_i^a) - c_{ij}(z_i^b)]u_{ij} + d_{ab}^{+} - d_{ab}^{-} = 0. \tag{10.14}$$

In this case the same middlemost concept is achieved my minimizing the maximum of all the deviation variables (i.e. d_{ab}^{+} and d_{ab}^{-} for all pairs $\{a, b\}$ for which an indifference statement has been made).

The LPs described above are aimed at producing the currently most consistent value functions, and by implication the alternative which maximizes this value function can be identified (to serve as the current best estimate for the optimum alternative). It is also possible to eliminate any alternative a if there exists an attribute vector, say z^0 which is either an element of $\mathcal{Z}$ or is dominated by an element of $\mathcal{Z}$, such that a negative value is obtained for the LP maximizing:

$$\sum_{i=1}^{m}\sum_{j=1}^{\nu}[c_{ij}(z_i^a) - c_{ij}(z_i^0)]u_{ij} \tag{10.15}$$

subject to the relevant constraints (i.e. (10.13) or (10.14) and the non-negativity, scaling and shape constraints). In [32] we proposed that in the discrete alternative case, z^0 could be selected to be the attribute vector of the current optimal

solution, and that each remaining alternative could be tested against z^0 whenever the optimal solution changed.

With this background, the interactive algorithm can be summarized in the following steps for the case in which the judgements of the decision maker are expressed as tradeoffs. (Use of indifferences between more general dummy alternatives involves appropriate modifications to the third and fourth steps below, and replacement of the constraint in step (4) by the more general (10.14).)

(1) Select an arbitrary initial guess for the u_{ij} values (e.g. all values equal).

(2) Let $\hat{V}(z)$ be the value function corresponding to the current estimates of u_{ij}. Let z^0 be the attribute vector maximizing $\hat{V}(z)$ subject to $z \in \mathcal{Z}$. If (in the judgement of the DM) the results have appeared to stabilize, then STOP; otherwise go on to the next step.

(3) If tradeoffs have not yet been assessed at the point represented by z^0, then set $z^a = z^0$. Otherwise select a distinctly different attribute vector z^a at which tradeoffs have not been assessed. (If $\mathcal{A}$ is discrete, z^a might be the vector maximizing $\hat{V}(z)$ amongst those elements of $\mathcal{Z}$ for which tradeoffs have not been assessed; for other problems z^a may be selected by randomly selecting sets of u_{ij} satisfying the current constraints, and maximizing the corresponding value function.)

(4) Selecting one criterion (without loss of generality denoted as criterion 1) as a reference, obtain tradeoffs on the other criteria corresponding to a fixed increment on the reference criterion (i.e. the amount, say τ_i which the decision maker would give up on criterion i in order to obtain an increase of Δ on criterion 1). In this case the constraint given by (10.14) becomes for each i:

$$\sum_{j=1}^{\nu}[c_{1j}(z_1^a + \Delta) - c_{1j}(z_1^a)]u_{1j} + [c_{ij}(z_i^a - \tau_i) - c_{ij}(z_i^a)]u_{ij} + d_{ai}^+ - d_{ai}^- = 0.$$

Relaxations of the need to fix on to one reference criterion is discussed in [36].

(5) If $\mathcal{A}$ is discrete, then apply the LP maximizing (10.15) to eliminate all solutions which are no longer potentially optimal. If only z^0 remains, then stop.

(6) With the addition of the latest set of constraints on the u_{ij} generated in step (4), solve the LP minimizing the maximum over all deviational variables subject to the non-negativity, scaling and shape constraints on u_{ij}. Return to step 2.

The numerical results reported in [36], which were based on discrete sets with quite large numbers (50-100) of alternatives, indicated that the above procedure tended to converge after about 6 or 7 iterations. At this stage, results were

effectively identical to that obtained by *a priori* fitting of a value function in the standard manner. It is worth noting that in the case of multiple objective linear programming, the piecewise linear approximations generate many more basic solutions, reducing the need for detailed exploration of an optimal facet (as suggested by Zionts and Wallenius in [43]).

10.5 PROCEDURES BASED ON ASPIRATION LEVELS

The procedures discussed in this section are based on the concept of representing values and preferences of the decision maker in terms of "*aspiration levels*", i.e. desirable levels of performance in terms of each attribute. The distinction between aspiration level procedures and those previously described can be quite fuzzy. We have, for example, in discussing Steuer's interactive weighted sums and filtering approach (distinctly a method based on value function approximations), referred to enhancements ([30, 31]) which are related to reference point methods, and which should perhaps thus also be viewed together with the aspiration level methods discussed in this section. It seems, nevertheless, helpful to give specific consideration to aspiration level approaches.

In broad terms, an interactive aspiration level procedure would involve the following steps:

(1) Start with an initial vector of "aspiration levels" for each attribute, say $\mathbf{g} = \{g_1, g_2, \ldots, g_m\}$.

(2) Find a feasible solution which best approaches these aspiration levels in some sense.

(3) Present the resulting solution to the DM. If the DM is "satisfied" with this solution, then the procedure terminates; otherwise the aspiration levels (g_i) are adjusted (either implicitly, or directly by the DM) and the procedure returns to the previous step.

An immediate problem which arises is in the interpretation of the term "aspiration level", which is almost certainly understood differently by different users. For some, the aspiration level will be a largely unattainable ideal goal; for others it may represent a non-negotiable bottom line or reservation level, such that any alternatives which do not meet all aspirations are immediately rejected; for yet others, the aspirations may be an *a priori* assessment of the likely characteristics of a desirable solution. At this stage we still lack a fully formulated theory as to how decision makers may form and modify aspirations, so that the challenge to the interactive implementation particularly is do provide a methodology which is robust to different interpretations of the aspiration level concept.

The basic idea of an aspiration level is central to goal programming and to reference point methods which are described elsewhere in this book. The original definitions of goal programming were not framed in the context of interactive methods as we have defined them. It is nevertheless difficult to see how goal programming could be implemented in anything other than an interactive

sense in most applications, as decision makers are unlikely immediately to be able to express realistic goals in any but the most familiar problem settings. The interaction may, however, be carried out in a somewhat *ad hoc* or unstructured manner, rather than in the structured interactions discussed in this chapter, although Tamiz and Jones [39] do provide a more formal interactive framework for the application of goal programming.

In this section we shall briefly discuss some aspiration-based methodologies which have been developed with an explicitly interactive context in mind. The reader is referred to the chapters on goal programming and on reference point approaches for more details on the overall philosophies of these methods.

Some of the earliest approaches which can be seen to fall into the interactive aspiration level based category are those of Zeleny's concept of "compromise programming" linked to the theory of the "displaced ideal" (Zeleny [41]), and STEM (Benayoun et al. [1]). In both cases, the aspiration levels are not directly specified by the decision maker, but are simply the ideals relative to the currently specified decision space, i.e. $g_i = z_i^{**} = \max_{x \in \mathcal{Z}} z_i$. The distinctive feature of these two approaches is that the interaction with the decision maker is not directly aimed at modifying the goals, but at placing constraints on the decision space, which in turn modifies the set of alternatives, and hence the ideals, under consideration.

The concept of *compromise programming* is simply to minimize norms of the form:

$$\left[\sum_{i=1}^{m} [w_i(g_i - z_i)]^p\right]^{1/p}$$

for some $p \geq 1$, which tends to the Tchebycheff norm:

$$\max_i [w_i(g_i - z_i)]$$

as $p \to \infty$. Note that in this case $g_i \geq z_i$ by definition. The weights w_i appear primarily to serve the function of ensuring a comparable scaling for all criteria (e.g. to normalize all deviations to the [0,1] interval), leaving value judgements to be expressed by effective changes in the goals. By minimizing the above norm for a variety of values for the exponent p, a range of efficient solutions can be identified and presented to decision makers. In the light of the solutions generated, decision makers are encouraged to eliminate clearly undesirable options (either by eliminating specific alternatives in the discrete case, or by inserting lower bounds for achievement on certain criteria). This leads to a shift, or "displacement" of the ideal, after which the process is repeated with adjustment of the goals to the new ideals. The process ultimately terminates when the difference between the ideal and the compromise solutions are found to be acceptably small.

The *STEM* (or *Step Method*) approach of Benayoun et al. [1] was formulated in the multiple objective linear programming context, but can be generalized to other problems, and shares many features with the compromise programming approach. The first step is to evaluate the ideals z_i^{**} by maximizing z_i subject

to $z \in \mathcal{Z}$ for each attribute i. Let z_i^k be the value of attribute i achieved when maximizing z_k. Then $z_i^{**} = z_i^i$, while $z_{*i} = \min_{1 \leq k \leq m}\{z_i^k\}$ represents a pessimistic level of achievement for attribute i. Once again, the ideals are used as the initial goals for each criterion, while deviations are minimized in this case using the Tchebycheff norm only. Weights are not selected subjectively, but are automatically generated for each criterion as the product of $(z_i^{**} - z_{*i})/z_i^*$ (representing the relative ranges of values available on each criterion) and a term which standardizes the objective functions (in the case of linear programming by standardizing the objective function coefficients for each criterion to unit Euclidean norm). Suppose that the solution then generated has values $\hat{z}_i$ for each criterion i. Decision makers are required to classify these values into those which are "satisfactory" and "unsatisfactory" respectively; in addition, an amount (say Δz_i) is required to be specified for each "satisfactory" criterion, being the amount which the decision maker would be prepared to sacrifice on this criterion, in order to achieve gains in the "unsatisfactory" criteria. Decision alternatives are then constrained to satisfy:

$$z_i \geq \hat{z}_i - \Delta z_i \quad \text{for "satisfactory" criteria}$$

and

$$z_i \geq \hat{z}_i \quad \text{for "unsatisfactory" criteria.}$$

This eliminates some alternatives, and leads to a shift in the ideal, after which the process is repeated. As each iteration constrains at least one criterion to "satisfactory" solutions (unless there are none at that iteration), it follows that the process must terminate in at most m iterations, with either a solution in which all criteria are "satisfactory", or a definite conclusion that no such solution exists (in which case attempts are presumably needed to create or discover new alternatives).

The key to the success of both STEM and compromise programming is the ability of decision makers either to specify what constitutes a satisfactory level of performance or to identify alternatives which can be eliminated. As previously discussed, the process by which this is done is not well-understood, but requires at least a substantial degree of global understanding of the available tradeoffs, and there must always be some question of the extent to which this is true. In problem settings which are relatively familiar to the decision maker (for example, in the selection of investment portfolios, which is likely to be a repetitive task) the choices may well be justifiable, and it is in such contexts that methods such as STEM and compromise programming may well offer an efficient means of decision support.

An approach which has been termed *interactive multiple goal programming (IMGP)*, although it does not quite fit into the usual definition of goal programming, was introduced by Spronk [28]. As with compromise programming and STEM, IMGP is also based on a pruning of the decision space, and it is interesting to note that the primary applications of IMGP appear to have been in capital budgeting and financial planning, i.e. the type of familiar decision context to which we have referred. In essence, the IMGP approach is based

on presenting the decision maker at each stage of the iterative process with what is termed by Spronk a "potency matrix", which consists simply of the vectors of ideal and pessimistic values, based on the current decision space. In this case the pessimistic values represent effectively a set of guaranteed lower bounds on performance for each criterion. At each iteration, the decision maker is asked which criterion should be first improved, and a lower bound greater than its current pessimistic value is tentatively imposed, and a new set of ideals calculated. If the decision maker is satisfied that any consequent losses in the ideal values are worth the gain in the guaranteed performance bound for the chosen criterion, then the tentative lower bound is made permanent and the process repeats. If the decision maker is not satisfied, then provision is made for backtracking (i.e. relaxing the tentative lower bound until the decision maker is satisfied with the tradeoffs).

IMGP thus, to a greater extent than the previous two methods, does seek to assist the decision maker to explore the decision space, by providing direct feedback on the effects of increasing aspirations on one criterion on performance levels for the other criteria. This feedback is very approximate however, as the loss in ideal may be substantially different to the actual (local) tradeoffs which are available. Nevertheless, in some practical experience reported in Stewart [33], IMGP appeared to generate more satisfying results than other interactive goal programming methods, even though convergence was somewhat slower.

A more systematic form of interactive goal programming, and one which is closer to the spirit of the original goal programming, was proposed by Masud and Hwang [18]. Their approach was formulated in the context of a multiobjective mathematical programming framework, but the principles seem equally appropriate to discrete choice problems. The core of the approach is essentially based on the standard Archimedean goal programming formulation (although there appears to be no reason why the approach could not equally well be applied in the framework of a Tchebycheff formulation), but contains certain interesting variations:

- As in some of the other methods discussed above, weights are not subjectively assessed to reflect relative importances, but merely ensure comparable scaling. In this case, the weights have the effect of re-scaling the criterion values so that the difference between g_i and z_i^{**} is the same for each criterion.

- Apart from the standard goal programming solution, the method also generates a further m solutions by solving a sequence of additional goal programming problems, in each of which the goal for one of the criteria is replaced by a hard constraint (i.e. $z_i \geq g_i$, with no deviations allowed). This set of solutions is meant to inform the decision maker when reassessing the goals (see below).

- Non-dominated solutions are avoided by including maximization of over-achievement of goals as a second-order objective in a pre-emptive goal programming sense.

At each iteration of the ISGP process, the decision maker is presented with the $m + 1$ solutions. If one of the $m + 1$ solutions is satisfactory, the process terminates; otherwise the decision maker is asked to revise the goals in the light of the solutions presented. The idea is that the multiple solutions which are presented inform the decision maker concerning available tradeoffs, thus contributing to more realistic goal specifications. The value of the procedure nevertheless still depends fundamentally on the ability of the decision maker to specify meaningful goals.

The usual goal programming paradigm is based on minimization of the underachievement of stated goals. A problem can arise if the goals are unduly modest, to the extent that feasible solutions exist satisfying all goals simultaneously. If the goals are truly levels of universal and objective satisfaction, then this may be a pleasing result, but more typically the decision maker would not be satisfied with this outcome, especially since the solution generated by the algorithm may then be dominated. We have noted that the ISGP procedure does recognize this problem, by including maximization of over-achievements as a low priority objective (to be sought only after minimization of underachievement has been completed). The reference point approach described by Wierzbicki (Chapter 9 of this book) addresses the problem more directly. In this approach, the "reference" level, indicates levels of achievement currently viewed as a good starting point for further exploration of the decision space. Wierzbicki introduces the concept of a "*scalarizing function*", which may be viewed as a surrogate value function to be applied in the vicinity of the reference point. Optimization of the scalarizing function then produces an efficient solution which is in a sense closest to the reference point. In the spirit of interactive goal programming, the decision maker is required to judge whether the solution found is satisfactory (in which case the process terminates), and if not to revise the reference values.

The reference point approach can clearly and naturally be applied in the context of the interactive aspiration level approach, as defined at the start of this section. The key to applying the reference point procedure is the choice of the scalarizing function. Although a number of options have been proposed, the most commonly applied form is closely allied to the Tchebycheff norm for goal programming, but with two important variations:

(1) The deviational variables $\delta_i = g_i - z_i$ are permitted to be negative, in which case $-\delta_i$ becomes a measure of over-achievement.

(2) The scalarizing function to be *minimized* is:

$$\max_{i=1}^{m} w_i \delta_i \; + \; \epsilon \sum_{i=1}^{m} w_i \delta_i .$$

Clearly the min-max term dominates, and while any δ_i remains strictly positive, is equivalent to goal programming using a Tchebycheff norm. Only once all reference levels have exceeded does maximization of over-achievement come into play. The summation term is weighted by the small constant ϵ, and serves primarily to ensure that solutions are non-dominated in cases when the min-max solution is not unique.

Some of the concepts of the reference point approach were adapted by Korhonen and Laakso [12] in a multiple objective linear programming context, and by Korhonen [11] for the discrete choice problem, to provide a visual interactive graphical procedure for applying goal or aspiration level procedures. Starting from a particular efficient solution, say $\hat{\mathbf{z}}$, a "reference direction" (rather than a single reference point), say $\mathbf{d}$, is chosen by the chosen by the decision maker. A sequence of reference points of the form $\hat{\mathbf{z}} + \theta\mathbf{d}$ is then generated, and projected on to the efficient frontier by minimizing Wierzbicki's scalarizing function for each point. This generates a path along the efficient frontier which the decision maker can examine to find a best point. A new reference direction can be chosen at this point from which the process can restart. Korhonen and Laakso suggest a visualization of the process in terms of a "*Pareto race*", i.e. driving along the Pareto frontier, which has been incorporated into a software package called VIG (or VIMDA for discrete choice problems). The current solution at any point is represented by bar graphs representing levels of achievement for each attribute. By pressing "accelerator" or "brake" keys the user can see how these levels change as one moves along the current direction, and can reverse direction by changing "gears". At any stage, the user can change direction by requiring greater emphasis on a specified attribute.

Another practical implementation of the reference point approach in an interactive framework, with associated software, is AIM (Aspiration-level Interactive Model) as described in Lotfi et al. [15]. This implementation was originally designed for discrete choice problems, but a generalization to multiple objective linear programming problems has also been proposed by Lotfi et al. [16]. The initial reference level is generated as a median value for each criterion, and information is provided to the decision maker at each stage as to the feasibility of the current reference level (expressed in terms of proportions of the decision space satisfying the references levels singly and jointly), and the relative sizes of the increments which are available for each criterion. Overall, however, the problem remains in any implementation, that if the decision maker indicates shifts in reference level which are strongly at odds with available tradeoffs, then the solution generated at the next iteration may be perceived to be worse than before and the process may be terminated prematurely. The idea from ISGP, of generating a $m + 1$ feasible solution for comparison, could be incorporated here with some benefit.

An interactive implementation of a reference point approach in much the same philosophy as that of AIM, but in the context of selecting a portfolio of R&D projects (which is a multiple objective integer programming context), is described in Stewart [34].

One theoretical problem relevant to the development and use of interactive methods based on aspiration levels, is the lack of a comprehensive theory as to how decision makers form and modify their aspirations (or goals or reference levels). This points to the need for a programme of behavioural research to support the algorithmic developments, similar to the manner in which the work of Kahnemann and Tversky (see, for example, [7]) has supported the development of utility-based decision support. One potential problem which we have previously identified ([33, 38]) is that aspiration-based methodologies may tend to terminate too early, especially for discrete choice problems, when decision makers fail to perceive sufficient progress being made. For this reason, we have suggested in [38] that the following two enhancements be incorporated into any interactive aspiration level procedure:

- Check at each iteration whether the new solution generated is perceived by the decision maker to be an improvement on the previous solution, and, if not, adjust the direction of change in the aspiration levels towards the direction of known preference implied by the comparison between the previous and current solutions.
- Encourage the decision maker to persevere for a few iterations at least, even in the absence of improvement.

10.6 CONVERGENCE PROPERTIES OF INTERACTIVE PROCEDURES

It is almost strange (some may even say irrelevant) to discuss the concept of convergence within the context of interactive MCDM methods. The number of iterations of the process through which a decision maker may be prepared to go is likely to be too small (possibly around 6 or 8 at most) for any concept of mathematical convergence to have much meaning. Furthermore, inputs from the decision maker are unlikely to be sufficiently consistent with the assumptions of mathematical programming theory for any mathematical proofs of convergence to have validity. Some of the problems inherent in determining convergence properties are evident in the study of convergence of reference point methods reported by Bogetoft et al. [2], where they need to include procedures for "'guiding' the decision maker in his choice of new reference points". In addition, it can quite legitimately be argued that the primary purpose of interactive methods is for the decision maker to gain understanding of the space $\mathcal{Z}$ and of his/her preferences in that space, and thus that "convergence" to a particular solution is uninteresting. Even when termination of the process within a finite number of steps can be proven (as in STEM, for example), this may not promote efficient learning and understanding, and may not identify the decision maker's most preferred alternative.

In spite of the above conceptual problems, however, it remains important for there to be some assurance that the decision aiding methodologies do not of themselves bias the results towards particular types of solutions, and do lead to a point at which the decision maker has some confidence in the solu-

tion found. Attempts at seeking such assurance include empirical studies (e.g. Buchanan [3]) and numerical studies (e.g. Shin and Ravindran [26]).

In a series of studies ([36, 37, 38]), we have suggested the use of comprehensive simulation methods, in which various "non-idealities" of decision makers in responding to the interactive questions, and errors in the assumptions made concerning preference structures, can be modeled. The effects of such non-idealities and errors on the solutions obtained with various interactive methods can thus be assessed systematically. The simulation approach proposed in the above references (elaborated in greatest detail in [37]), starts with a hypothetical decision maker whose long-run goals are assumed to be consistent with a preference structure which satisfies assumptions of completeness, transitivity and preferential independence for an ideal set of m criteria (on which the performance measures z_i are defined). These assumptions imply the existence of an additive value function, defined on the z_i, representing these long-run goals, even though in the short-term (during the learning and preference construction process) the decision maker's perceived preferences and responses may not be consistent with any such additive function. It is important to note that this value function need not necessarily be explicitly knowable or accessible. In fact, the reason for adopting aspiration-based methods in practice may well be that the elicitation of an explicit value function is impractical in the light of the problem setting and time and resource constraints.

The above assumptions are consistent with the philosophy of decision aid aiming to support the construction of preferences consistently with desirable "rationality" properties. We note also that the ideal marginal (single criterion) value functions may in principle include sharp thresholds and changes in slope, and may be "S-shaped" (rather than purely concave, for example), properties which may be masked in the assessment process, leading to non-ideal behaviour even for interactive methods based on additive value function approaches, in spite of assumptions which seem to be biassed in favour of such approaches. In applying the simulation, it is also assumed that in practice the m ideally preferentially independent criteria may be corrupted by the inadvertent elimination of some criteria (so that alternatives are described in terms of $q \leq m$ attributes), and/or by mixing of the remaining criteria which destroys preferential independence. DM responses are assumed to be consistent with the underlying idealized preference structure, with unmodeled criteria assumed (unconsciously) to be fixed at some neutral reference level.

The reader is referred to the original references for details as to how the simulations were carried out for different types of procedure. Perhaps the key conclusions can be summarized as follows:

(1) The generalized interactive value function approach described in Section 10.4 requires approximately the same number of value judgements from the decision maker as that required to estimate a value function directly, in order to obtain the same level of precision of solution.

(2) It is extremely valuable to model non-linearities in preference structure directly (as with the piecewise linear models described in Section 10.4),

but there is little gain in modeling such non-linearities to greater precision than that obtained from a piecewise linear model with three or four segments.

(3) Violation of preferential independence assumptions when constructing a hierarchy of criteria can seriously bias solutions obtained from value function methods, but has little or no influence on the performance of aspiration level methods.

(4) Aspiration level methods may tend to terminate to early in the absence of enhancements such as those described at the end of Section 10.5.

(5) Apart from the previous two points, the quality and sensitivity of solutions obtained from either methods based on value functions or those based on aspiration levels are essentially the same.

Overall, the conclusions that can be drawn from the simulation studies cited above are that interactive methods do (if used with some care and understanding) lead to solutions to the decision problem which can be adopted with confidence. Choice between interactive methods, or between interactive and "non-interactive" methods, can and should be based on what is more comfortable for the decision maker, and on the (sometimes conflicting) needs for building understanding and for providing an adequate "audit trail" regarding how the final recommendation was reached.

References

[1] Benayoun, R., J. de Montgolfier, J. Tergny, O. Larichev. Linear programming with multiple objective functions: step method (STEM). Mathematical Programming 1(1971)366–375

[2] Bogetoft, P., Å. Hallefjord, M. Kok. On the convergence of reference point methods in multiobjective programming. European Journal of Operational Research 34(1988)56–68

[3] Buchanan, J.T. An experimental evaluation of interactive MCDM methods and the decision making process. Journal of the Operational Research Society 45(1994)1050–1059

[4] Gardiner, L. R., R. E. Steuer. Unified interactive multiple objective programming. European Journal of Operational Research 74(1994)391–406

[5] Gardiner, L. R., R. E. Steuer. Unified interactive multiple objective programming: an open architecture for accommodating new procedures. Journal of the Operational Research Society 45(1994)1456–1466

[6] Geoffrion, A. M., J. S. Dyer, A. Feinberg. An interactive approach for multicriterion optimization with an application to the operation of an academic department. Management Science 19(1972)357–368

[7] Kahneman, D., A. Tversky. Prospect theory: An analysis of decision under risk. Econometrica 47(1979)263–291

[8] Köksalan, M. M., M. H. Karwan, S. Zionts. An improved method for solving multiple criteria problems involving discrete alternatives. IEEE Transactions on Systems, Man and Cybernetics SMC-14(1984)24–34

[9] Köksalan, M. M., P. N. S. Sagala. An interactive approaches for choosing the best of a set of alternatives. Journal of the Operational Research Society 43(1992)259–263

[10] Köksalan, M. M., O. V. Taner. An approach for finding the most preferred alternative in the presence of multiple criteria. European Journal of Operational Research 60(1992)52–60

[11] Korhonen, P. A visual reference direction approach to solving discrete multiple criteria problems. European Journal of Operational Research 34(1988)152–159

[12] Korhonen, P., J. Laakso. A visual interactive method for solving the multiple criteria problem. European Journal of Operational Research 24(1986)277–287

[13] Korhonen, P., J. Wallenius, S. Zionts. Solving the discrete multiple criteria problem using convex cones. Management Science 30(1984)1336–1345

[14] Loganathan, G. V., H. D. Sherali. A convergent interactive cutting plane algorithm for multiobjective optimization. Operations Research 35(1987)365–377

[15] Lotfi, V., T. J. Stewart, S. Zionts. An aspiration-level interactive model for multiple criteria decision making. Computers and Operations Research 19(1992)671–681

[16] Lotfi, V., Y. S. Yoon, S. Zionts. Aspiration-based search algorithm (ABSALG) for multiple objective linear programming problems: theory and comparative tests. Management Science 43(1997)1047–1059

[17] Malakooti, B. A decision support system and a heuristic interactive approach for solving discrete multiple criteria problems. IEEE Transactions on Systems, Man and Cybernetics 18(1988)273–284

[18] Masud, A. S. M., C. L. Hwang. Interactive sequential goal programming. Journal of the Operational Research Society 32(1981)391–400

[19] Musselman, K., J. Talavage. A tradeoff cut approach to multiple objective optimization. Operations Research 28(1980)1424–1435

[20] Prasad, S. Y., M. H. Karwan, S. Zionts Use of convex cones in interactive multiple objective decision making. Management Science 43(1997)723–734

[21] Ramesh, R., M. H. Karwan, S. Zionts Interactive multicriteria linear programming: an extension of the method of Zionts and Wallenius. Naval Research Logistics 36(1989)321–335

[22] Reeves, G. R., L. S. Franz. A simplified interactive multiobjective linear programming procedure. Computers and Operations Research 12(1985)589–601

[23] Rosinger, E. E. Interactive algorithm for multiobjective optimization. Journal of Optimization Theory and Applications 35(1981)339–365

[24] Sakawa, M. Interactive multiobjective decision making by the sequential proxy optimization technique: SPOT. European Journal of Operational Research 9(1982)386–396

[25] Shin, W. S., A. Ravindran. Interactive multiple objective optimization: Survey I – continuous case. Computers and Operations Research 18(1991)97–114

[26] Shin, W. S., A. Ravindran. A comparative study of interactive tradeoff cutting plane methods for MOMP. European Journal of Operational Research 56(1992)380–393

[27] Siskos, J. Comment modéliser les préférences au moyen de fonctions d'utilité additives. RAIRO Recherche Opérationnelle/Operations Research 14(1980)53–82

[28] Spronk, J. *Interactive Multiple Goal Programming* Martinus Nijhoff Publishing, 1981.

[29] Steuer, R. E. *Multiple Criteria Optimization: Theory, Computation, and Application.* John Wiley & Sons, New York 1986.

[30] Steuer, R. E., E. U. Choo. An interactive weighted Tchebycheff procedure for multiple objective programming. Mathematical Programming 26(1983)326–344

[31] Steuer, R. E., J. Silverman, A. W. Whisman. A combined Tchebycheff/aspiration criterion vector interactive multiobjective programming procedure. Management Science 39(1993)1255–1260

[32] Stewart, T. J. Pruning of decision alternatives in multiple criteria decision making, based on the UTA method for estimating utilities. European Journal of Operational Research 28(1987)79–88

[33] Stewart, T. J. Experience with prototype multicriteria decision support systems for pelagic fish quota determination. Naval Research Logistics 35(1988)719–731

[34] Stewart, T. J. A multicriteria decision support system for R&D project selection. Journal of the Operational Research Society 42(1991)17–26

[35] Stewart, T. J. A critical survey on the status of multiple criteria decision making theory and practice. OMEGA: International Journal of Management Science 20(1992)569–586

[36] Stewart, T. J. Use of piecewise linear value functions in interactive multicriteria decision support: A monte carlo study. Management Science 39(1993)1369–1381

[37] Stewart, T. J. Robustness of additive value function methods in MCDM. Journal of Multi-Criteria Decision Analysis 5(1996)301–309

[38] Stewart, T. J. Evaluation and refinement of aspiration-based methods in MCDM. European Journal of Operational Research 1998. (To appear).

[39] Tamiz, M. D. F. Jones. Interactive frameworks for investigation of goal programming models: theory and practice. Journal of Multi-Criteria Decision Analysis 6(1997)52–60

[40] Vetschera, R. An interactive outranking system for multi-attribute decision making. Computers and Operations Research 15(1988)311–322

[41] Zeleny, M. *Multiple Criteria Decision Making.* McGraw-Hill, New York 1982.

[42] Zionts, S. J. Wallenius. An interactive programming method for solving the multiple criteria problem. Management Science 22(1976)652–663

[43] Zionts, S. J. Wallenius. An interactive multiple objective linear programming method for a class of underlying nonlinear utility functions. Management Science 29(1983)519–529

11 OUTRANKING APPROACH

Philippe Vincke

Contents

T. Gal et al. (eds.), *Multicriteria Decision Making*

Abstract: The purpose of this chapter is to present the so-called outranking approach, which was proposed in about 1970 by B. Roy as a complementary approach to multiattribute utility theory. After the description of the motivation and of the basic principles, we present about ten methods, give some comments on the determination of their parameters and illustrate how theoretical results can help to choose a method. The bibliography contains the most known references on the subject.

11.1 PRELIMINARY DEFINITIONS AND NOTATIONS

We denote by A the set of alternatives (decisions, solutions, candidates, ...) which has to be explored in the decision process. If not otherwise mentioned, we consider that A is finite.

A binary relation $\mathcal{R}$ on A is a subset of $A \times A$. It is said to be

- symmetric iff $a\mathcal{R}b \Rightarrow b\mathcal{R}a$, $\forall\ a, b \in A$,
- asymmetric iff $a\mathcal{R}b \Rightarrow b\not\mathcal{R}a$, $\forall\ a, b \in A$,
- complete iff $a\not\mathcal{R}b \Rightarrow b\mathcal{R}a$, $\forall\ a, b \in A$,
- transitive iff $a\mathcal{R}b,\ b\mathcal{R}c \Rightarrow a\mathcal{R}c$, $\forall\ a, b, c \in A$,
- a *complete preorder* iff it is complete and transitive,
- a *partial preorder* iff it is transitive and not complete.

Defining a *valued relation* on A consists in associating a real number with every ordered pair (a, b) of elements of A.

A *criterion* g is a real valued function defined on A and representing preferences according to a certain point of view, in the sense that

$$\begin{cases} g(a) > g(b) & \text{iff} \quad a \text{ is preferred to } b, \\ g(a) = g(b) & \text{iff} \quad a \text{ is indifferent to } b. \end{cases}$$

It is well known that the preferences can be represented by a criterion iff the relation $\mathcal{R}$, defined by

$$a\mathcal{R}b \text{ iff } a \text{ is preferred or indifferent to } b,$$

is a complete preorder.

A *pseudo-criterion* is (g, q, p) a triplet of real valued functions representing preferences according to a certain point of view, in the sense that

$$\begin{cases} g(a) > g(b) + p(g(b)) \text{ iff } a \text{ is strictly preferred to } b, \\ g(b) + p(g(b)) \geq g(a) > g(b) + q(g(b)) \text{iff } a \text{ is weakly preferred to } b, \\ a \text{ is indifferent to } b \text{ iff there is no strict or weak preference between them.} \end{cases}$$

The functions q and p are respectively called *indifference and preference thresholds*. They can be constant or not. If not, they are assumed to satisfy some coherence conditions. The underlying preference structure is called a *pseudo-order* (see Roy and Vincke [62]). In the particular case where $p \equiv q$, the structure is a *semiorder* and (g, q) is called a *quasi-criterion*.

11.2 WHY AN OUTRANKING APPROACH ?

Decision-aid can be based on preference models which are expressed by a unique numerical function: it is the case in utility (or value) theory, in econometrics, in many financial models, in mathematical programming, and so on. It is also the case, in the framework of MCDM, in multiattribute utility (or value) theory and in most of the methods of multiobjective programming. Two main advantages of such models are the fact that they exclude any incomparability and that the preferences are transitive, so that ranking the alternatives or choosing the so-called "best" one are trivial operations (even if they necessitate many calculations). These advantages have of course a counterpart: all the imprecisions, incertitudes, arbitrary aspects of the so-called "data" (which in general are not given but must be built) and of the parameters of the method are finally aggregated in one number for each alternative. The consequence is that this aggregated information is sometimes delicate (even if some sensivity analysis can help) is this approach mixes all the ambiguities and conflicts into one number and treats in the same way very well established and precise evaluations or very uncertain and fuzzy ones. The question then is to be sure that the final decision is really justified by the characteristics of the initial problem and not by the mathematical properties of the method.

Another approach consists in working on preference models which point out the more or less solid aspects of the information. This means that such models accept incomparabilities and do not impose any transitivity properties; this also means that the preferences cannot be expressed by a unique numerical function and, consequently, that ranking the alternatives or choosing the best one are not trivial problems any more. The outranking approach is based on this kind of model. Scientists are sometimes hindered by the fact that this approach is based on preference models including incomparabilities; however, *incomparability* has to be considered as interesting and important as "preference" or "indifference". To conclude that two alternatives are incomparable is also *decision-aiding*: it points out the conflicts or lack of information and invites the analyst and the decision-maker to go deeply (if they want to do so) into some aspects of the problem. The resolution of a decision problem is a dynamic process where the preferences evolve with the information: in this context, it is useful to have models which are able to represent the preferences at the interpretation of each step of the process and, particularly, at steps where some alternatives are (eventually temporarily) incomparable.

Here are some examples of situations where the outranking approach can be justified (Roy and Bouyssou [55]):

- when at least one criterion is not quantitative, so that preference intervals ratios have no sense,

- when the units of the different criteria are so heterogeneous that coding them into one common scale seems to be very difficult or artificial,

- when the compensations between gains on some criteria and losses on other criteria are not clear,

- when some preference or veto thresholds have to be taken into account.

Of course, the outranking approach is complementary to the other approaches, whose interest is not in question. Too often this approach (sometimes called French or European approach) has been contrasted with multiattribute utility theory (called American approach) or to multiobjective optimization. This opposition has no sense: each approach has advantages, disadvantages and pertinent fields of applications. They often can be mixed in the treatment of concrete situations. Confronted to a particular problem, the role of the scientist is to use appropriate tools or combinations of tools in order to progress in the resolution of the problem. The outranking methods are tools among others.

11.3 BASIC PRINCIPLES OF THE OUTRANKING APPROACH

The specialists of the outranking approach traditionally consider that the outranking methods consist of two steps:

i) the building of the outranking model, which represents the holistic preferences and which can be formed by one or several valued or crisp binary relations; the *outranking concept* was defined for the first time by B. Roy (see for example [48]) and can be schematized as follows: an alternative a outranks b if, given the information about the preferences of the decision maker, there are sufficient arguments to affirm that a is at least as good as b and there is no really important reason to refuse this assertion. The various outranking methods which can be found in the literature differ in the way they formalize this definition.

ii) the exploitation of the outranking model in function of the problem to solve; B. Roy [51] defines three main problems: choosing alternative(s), sorting them into categories or ranking them from the best to the worst, but many variants and combinations can occur in the practical applications (see Bana e Costa [3]). As mentioned in the previous section, the second step is usually not trivial, due to the fact that an outranking model does not satisfy, in general, nice mathematical properties like completeness and transitivity.

11.4 OUTRANKING METHODS

11.4.1 Introduction

Not all the outranking methods can be presented in detail here. We confine ourselves to eight of them: the oldest ones (in order to illustrate the basic principles), the most common ones (which are often mentioned in the literature) and some others which are rarely mentioned but which contain some particularities.

Section 11.4.11 presents an overview of these methods and section 11.4.12 gives a brief description of some other ones.

11.4.2 ELECTRE I (Roy [47])

ELECTRE I is already 30 years old: as it was the first historically and as the formulae are rather simple, giving a good illustration of the basic principles, we recall once more how it works.

The basic information is a set of n criteria $\{g_1, g_2, \ldots, g_n\}$ on A and, for each of them:

- a "weight" w_j, expressing the relative importance of criterion g_j (see section 11.5.1),
- a veto threshold $v_j(g_j) > 0$ (see section 11.5.2).

For each ordered pair (a, b), a *concordance index* $c(a, b)$ is calculated by

$$c(a,b) = \frac{1}{W} \sum_{j:g_j(a) \geq g_j(b)} w_j,$$

where

$$W = \sum_{j=1}^{n} w_j.$$

This index varies from 0 to 1 and can be considered as a measure of the arguments in favour of the assertion "a outranks b".

Then, choosing a *concordance level* s, a is declared to outrank b, denoted by aSb iff:

$$\left| \begin{array}{l} c(a,b) \geq s, \\ \\ \forall\, j \text{ such that } g_j(a) < g_j(b), \text{ the interval} \\ (g_j(a), g_j(b)) \text{ is smaller than } v_j(g_j(a)). \end{array} \right.$$

As we see, this formalization corresponds to the definition of outranking given in the basic principles.

Note also that the outranking relation is unchanged if we transform the weights in such a way that the ordering of their partial sums (obtained for all the subsets of criteria) is maintained (and assuming that the concordance level is modified in the same manner). We can for example multiply all the

weights by a same number; if the new weights are integers, the building of the outranking relation in ELECTRE I can be interpreted as a voting procedure with a special majority rule (characterized by the concordance level).

Having the outranking relation S, which can be represented by a graph where the vertices represent the actions, one seeks a subset N of actions such that:

$$\begin{cases} \forall\, b \in A \setminus N, \exists a \in N : aSb, \\ \forall\, a, b \in N,\ a \not S b. \end{cases}$$

One aims to find a subset N of actions such that any action which is not in N is outranked by at least one action of N and the actions of N are incomparable (the latter condition allows to render N minimal for inclusion). In graph theory, this type of set is called a *kernel* of the graph and there exist procedures to determine it. Let us also recall that if the graph has no circuit, the kernel exists and is unique. One possible technique consists in reducing the initial graph's circuits (i.e. replacing each circuit by a unique element, which is equivalent to considering the actions in the circuit as tied), but the latter operation may eliminate a great deal of the information contained in the outranking relation. Another technique consists in using the concept of minimum weakness quasi-kernel (see Hansen et al. [23]).

In order to advance towards the best possible compromise, a more refined analysis of the kernel's actions must be performed. Practically, it is advised to use fluctuations of the method's parameters and to study the sensitivity of the result with respect to those variations. The latter sensitivity analysis can also be used to break up ties between the kernel's actions.

11.4.3 ELECTRE IS (Roy and Skalka [60])

This method is an adaptation of ELECTRE I to the case where indifference and preference thresholds are defined for some criteria and must be taken into account in the building of the outranking relation. The interested reader will find in Roy and Bouyssou [55] all the details about that method.

11.4.4 ELECTRE II (Roy and Bertier [54])

We give here the original presentation of this method; the reader will find in Nadeau et al. [37] a more recent and more sophisticated variant, called ELECCALC.

The basic information is a set of criteria $\{g_1, g_2, \ldots, g_n\}$ on A and for each of them:

- a "weight" w_j, expressing the relative importance of criterion g_j,
- two veto thresholds v_j^1 and v_j^2 such that $v_j^1 \leq v_j^2$.

For each ordered pair (a, b), a concordance index $c(a, b)$ is calculated by

$$c(a,b) = \frac{1}{W} \sum_{j : g_j(a) \geq g_j(b)} w_j,$$

where

$$W = \sum_{j=1}^{n} w_j.$$

This index varies from 0 to 1 and can be considered as a measure of the arguments in favour of the assertion "a outranks b".

Then choosing two concordance levels s^1 and s^2, such that $s^1 > s^2$, two outranking relations S^1 and S^2 are determined as follows: for $i = 1, 2$, we have aS^ib iff :

$$\left|\begin{array}{l} c(a,b) \geq s^i, \\ \\ \displaystyle\sum_{j:g_j(a)>g_j(b)} w_j > \sum_{j:g_j(a)<g_j(b)} w_j, \\ \\ \forall\, j \text{ such that } g_j(a) < g_j(b), \text{ the interval } (g_j(a), g_j(b)) \text{ is} \\ \text{smaller than } v_j^i\,(g_j(a)). \end{array}\right.$$

It is clear that $S^1 \subset S^2$: S^1 is called *strong outranking* (and is only satisfied when outranking is not disputable) and S^2 is called *weak outranking* (in the sense that it is less reliable).

As in ELECTRE I, we can multiply all the weights by a same number in order to obtain integers and without changing S^1 and S^2, so that the building of these relations can be interpreted as the application of a voting procedure.

The class of best actions (the first class of the ranking) is obtained as follows: after reducing the circuits of S^1 (cf. ELECTRE I), one determines the set B of actions which are not strongly outranked by any other action; inside that set, the circuits of S^2 are reduced and one determines the set A^1 of actions which are not weakly outranked by any other action of B. The set A^1 is the first class of the ranking and the procedure is started again in the remaining set, thereby yielding a complete preorder.

A second complete preorder is built in an analogous way but by starting with the class of worst actions (those which outrank no other action) and "going up" toward the best actions.

The two preorders obtained are, in general, not the same: if they are close, the decision-maker is offered a "*median preorder*" (for details, see Roy and Bertier [54]). Otherwise, a more thorough study is required since it is possible that the data are too divergent to be able to build an acceptable complete preorder. In both cases, a robustness analysis is obviously necessary.

Note for example, that if a certain action doesn't outrank any other and is itself outranked by no other (in other words, if it is difficult to compare it with the others), it will appear as first in the first ranking and as last in the second: the comparison between the two complete preorders is thus quite useful to detect "problematic" actions. This is why it is advised to build the partial preorder resulting from the intersection of the two complete preorders.

Another way to obtain the complete preorders is based upon the degrees of the graph's vertices, i.e. the number of actions which strongly outrank and

which are strongly outranked by each action, ties being eliminated on the basis of the weak outranking relation (see the Section 11.6 for a comparison of these different ways for obtaining complete preorders).

11.4.5 ELECTRE III (Roy [49])

The basic information is a set of n pseudo-criteria $\{(g_j, q_j, p_j),\ j = 1, \ldots, n\}$ on A and, for each of them:

- a "weight" w_j, expressing the relative importance of criterion g_j,
- a veto threshold $v_j(g_j) > 0$.

For each ordered pair (a, b), a concordance index $c(a, b)$ and *discordance indices* $d_j(a, b)$ are calculated as follows:

$$c(a,b) = \frac{1}{W} \sum_{j=1}^{n} w_j c_j(a,b),$$

where

$$W = \sum_{j=1}^{n} w_j.$$

and

$$c_j(a,b) = \begin{cases} 1 \text{ if } g_j(a) + q_j\,(g_j(a)) \geq g_j(b), \\ 0 \text{ if } g_j(a) + p_j\,(g_j(a)) \leq g_j(b), \\ \text{linearly decreasing in the intermediary region;} \end{cases}$$

$$d_j(a,b) = \begin{cases} 0 \text{ if } g_j(a) + p_j\,(g_j(a)) \geq g_j(b), \\ 1 \text{ if } g_j(a) + v_j\,(g_j(a)) \leq g_j(b), \\ \text{linearly increasing in the intermediary region.} \end{cases}$$

The concordance index $c(a, b)$ can be considered as a measure of the arguments in favour of the assertion "a outranks b", while the discordance indices $d_j(a, b)$ measure the "strength" of the reasons for refusing this assertion. A *valued outranking relation* is then defined by calculating, for each ordered pair (a, b), the quantity

$$S(a,b) = \begin{cases} c(a,b) \text{ if } d_j(a,b) \leq c(a,b), \quad \forall\, j \\ \\ c(a,b). \displaystyle\prod_{j \in J(a,b)} \frac{1 - d_j(a,b)}{1 - c(a,b)}, \end{cases}$$

where $J(a, b)$ is the set of criteria j such that $d_j(a, b) > c(a, b)$. $S(a, b)$ can be interpreted as a measure of the credibility of the assertion "a outranks b". The advantage of working with a valued outranking relation is the fact that it is less sensitive to a variation of the necessarily somewhat arbitrary values of the parameters.

The value $\lambda = \max_{a,b \in A} S(a,b)$ is determined and only the arcs having values sufficiently close to λ are considered, i.e. more precisely, those which have a value larger or equal to $\lambda - s(\lambda)$ where $s(\lambda)$ is a threshold to be determined (it allows the values close enough to λ to be defined). The latter yields a non-valued outranking relation for which the qualification $Q(a)$ of each action a can be computed (that is the number of actions which are outranked by a minus the number of actions which outrank a). The set of actions having the largest qualification will be called the first distillate D_1.

If D_1 only contains one action, the previous procedure is started again in $A \setminus D_1$. Otherwise, the same procedure is applied inside D_1; if distillate D_2 which is thereby obtained is a singleton, the procedure is started again in $D_1 \setminus D_2$ (except if the latter set is empty); otherwise, it is applied inside D_2, and so forth until D_1 is used up entirely, before starting with $A \setminus D_1$. This procedure, which is called a descending distillation chain, yields a first complete preorder.

A second complete preorder is obtained by an ascending distillation chain, in which the actions having the smallest qualification are retained.

The information which can be drawn from these two preorders is analogous to that obtained by ELECTRE II. For an example of robustness analysis, we recommend the work of Roy, Present and Silhol [58].

11.4.6 ELECTRE IV (Roy and Hugonnard [56])

A characteristic of this method is the fact that it does not introduce any weight expressing the relative importance of the criteria.

The basic information is a set of n pseudo-criteria $\{(g_j, q_j, p_j), j = 1, \ldots, n\}$ on A and, for each of them:

- a veto threshold $v_j(g_j) > 0$.

Two embedded outranking relations, a strong one S^1 and a weak one S^2, are defined as follows:

$$aS^1b \quad \text{if} \quad \begin{cases} \bullet g_j(a) + p_j(g_j(a)) \geq g_j(b), \quad \forall\, j \\ \bullet ||\{j : g_j(a) + q_j(g_j(a)) < g_j(b)\}|| \leq \\ \qquad\qquad ||\{j : g_j(b) + q_j(g_j(b)) < g_j(a)\}||; \end{cases}$$

$$aS^2b \quad \text{if} \quad \bullet\ g_j(a) + p_j(g_j(a)) \geq g_j(b), \quad \forall\, j,$$

or

$$\text{if} \quad \begin{cases} \bullet \text{ there exists exactly one criterion } k \text{ such that} \\ \quad g_k(a) + v_k(g_k(a)) \geq g_k(b) > g_k(a) + p_k(g_k(a)), \\ \bullet \ ||\{j : g_j(b) + p_j(g_j(b)) < g_j(a)\}|| \geq \frac{n}{2}, \end{cases}$$

where $||A||$ denotes the number of elements contained in the set A.

The reader will find in Roy and Bouyssou [55] a more refined variant based on five embedded outranking relations instead of two. It is clear that a lot of variants can be considered according to the context of the application.

The fact that no weight is introduced to express the relative importance of the criteria does not imply that the method is only applicable when all the criteria have the same importance (even if the definitions of S^1 and S^2 are based on some counting of criteria). The reader will find in Roy and Bouyssou [55] some interesting comments on the underlying assumptions of the method with respect to the importance of the criteria.

The exploitation is performed as in ELECTRE III (distillations) but is made simpler by the fact that there are only two outranking levels. One determines the subset D_1 of actions which have the largest qualification in A for S^1 (remember that the qualification of a is the number of actions outranked by a, minus the number of actions which outrank a). If D_1 is a singleton, qualifications are computed again in $A \setminus D_1$ and the subset D_2 of actions which have the largest qualification in $A \setminus D_1$ for S^1 is determined, and so forth. When a D_h contains more than one action, the same procedure is applied inside D_h but on the basis of relation S^2. This descending procedure terminates when all the actions are ranked in a complete preorder. A second complete preorder is built by an ascending procedure (by determining each time the actions which have the smallest qualification). The information drawn from these two procedures is analogous to that obtained in ELECTRE II or ELECTRE III.

11.4.7 ELECTRE TRI (Roy and Bouyssou [55])

The basic information is a set of n pseudo-criteria $\{(g_j, q_j, p_j),\ j = 1, \ldots, n\}$ on A and, for each of them:

- a "weight" w_j expressing the relative importance of criterion g_j,
- a veto threshold $v_j(g_j) > 0$.

Moreover, as the purpose of the method is to assign the alternatives to k predefined ordered categories, the reference alternatives b^h $(h = 0, \ldots, k)$, considered as defining the limits between the categories, are defined by the vectors of their values for functions g_j, denoted by $(g_1^h, \ldots, g_n^h)$ in such a way that

$$g_j^h > g_j^{h-1}, \quad \forall\, j; \quad \forall\, h \in \{1, \ldots, k\}.$$

Giving an alternative a, an outranking relation is then built on the set

$$\{a\} \bigcup \{b^h, h = 0, \ldots, k\}.$$

One way is to calculate a valued outranking relation as in ELECTRE III method and to only consider the values greater or equal to a certain level, but any other way can be considered. The pessimistic assignment procedure consists in assigning a to the highest category c^h such that a outranks b^{h-1}. The optimistic assignment procedure consists in assigning a to the lowest category c^f such that b^f strictly outranks a.

The reader will find in Roy and Bouyssou [55] a detailed discussion on the theoretical properties and the practical aspects of this method (in particular a comparison of the pessimistic and optimistic procedures).

11.4.8 *PROMETHEE I and II (Brans et al. [15])*

The basic information is a set of n so-called generalized criteria (g_j, F_j) on A, and, for each of them:

- a "weight" w_j expressing the relative importance of criterion g_j.

A valued strict preference relation is defined by calculating, for each ordered pair (a, b), the quantity

$$\pi(a,b) = \frac{1}{W} \sum_{j=1}^{n} w_j F_j(a,b),$$

where

$$W = \sum_{j=1}^{n} w_j,$$

and $F_j(a,b)$ is the degree of preference of a over b for criterion j.

$F_j(a,b)$ is a number, between 0 and 1, which increases with the interval between $g_j(a)$ and $g_j(b)$. It is determined on the basis of a battery of functions which are discussed with the decision-maker and which are presented in Fig. 11.1. For each criterion g_j, a particular function F_j is chosen and the corresponding parameters (q_j, p_j or σ_j) are fixed. In the original method, these parameters are constant but there is no conceptual difficulty to consider that they vary with the position of the interval between $g_j(a)$ and $g_j(b)$.

Note that the valued relation defined here is a preference relation and not an outranking relation as, by construction, $\pi(a,a) = 0, \forall\ a$. Note also that no discordance aspect was introduced in the original method.

Just as in the previous methods, two complete preorders are built: one consists in ranking the actions following the decreasing order of the numbers

$$\phi^+(a) = \sum_{b \in A} \pi(a,b) \qquad \text{(outgoing flow)},$$

and the other following the increasing order of the numbers $\phi^-(a)$ such that

$$\phi^-(a) = \sum_{b \in A} \pi(b,a) \qquad \text{(ingoing flow)}.$$

Their intersection yields the partial preorder of the PROMETHEE I method. The PROMETHEE II method consists in ranking the actions following the decreasing order of the numbers $\phi(a)$ such that

$$\phi(a) = \phi^+(a) - \phi^-(a)$$

(generalizing the concept of qualification from the previous methods) and thus yields a unique complete preorder.

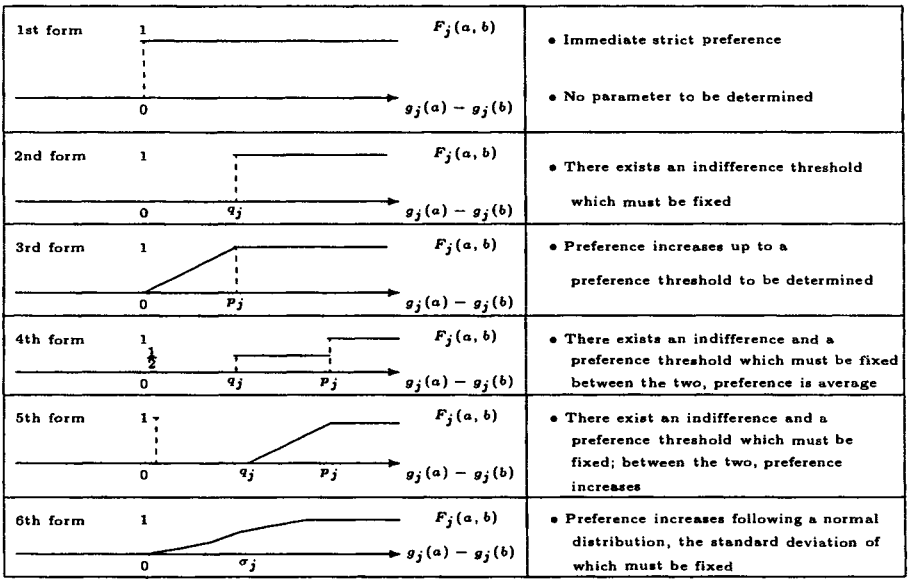


Figure 11.1 Preference functions

The PROMETHEE method gave birth to the development of a very user-friendly sofware called PROMCALC which includes, besides the interactive construction of functions F_j and the determination and exploitation of the valued outranking relation, a sensitivity analysis of the result with respect to the weights w_j (cf. Mareschal [29]). The GAIA software yields, on the basis of the results determined by PROMCALC, a geometrical representation of the actions and of the criteria by application of a principal components analysis (cf. Mareschal and Brans [30]).

11.4.9 MELCHIOR (Leclercq [26])

A family of n pseudo-criteria is at hand, provided with a relation T such that $i\,T\,j$ means: "criterion i is at least as important as criterion j". No assumption is made beforehand on the properties of T. The basic idea is to say that a outranks b if the criteria which are unfavourable to the latter assertion are "hidden" by those which are in its favour and if no criterion j exists such that $g_j(b) > g_j(a) + v_j$, where v_j is a veto threshold (no discordance).

It remains to define what we call:

- criteria which are in favour of the outranking of b by a,
- criteria which are unfavourable to the outranking of b by a,

- "to hide".

The author of the method proposes the following definitions:

- A criterion j will be said to be in favour of the outranking of b by a if

$$\begin{aligned} & aP_jb && \text{(1st definition)} \\ \text{or} & \\ & a\,(P_j \bigcup Q_j)\,b && \text{(2nd definition)} \\ \text{or} & \\ & g_j(a) > g_j(b) && \text{(3rd definition).} \end{aligned}$$

- A criterion j will be said to be unfavourable to the outranking of b by a if

$$\begin{aligned} & bP_ja && \text{(1st definition)} \\ \text{or} & \\ & b\,(P_j \bigcup Q_j)\,a && \text{(2nd definition)} \\ \text{or} & \\ & g_j(b) > g_j(a) && \text{(3rd definition).} \end{aligned}$$

- A subfamily G of criteria "hides" a subfamily H of criteria if, for any criterion j of H, there exists a criterion i of G such that

$$\begin{aligned} & i\ T\ j && \text{(1st definition)} \\ \text{or} & \\ & i\ T\ j \text{ or not } (j\ T\ i) && \text{(2nd definition),} \end{aligned}$$

  the same criterion i of G not being allowed to hide several criteria of H. Obviously, other definitions could be introduced.

By choosing two combinations of definitions, one stricter than the other, one obtains a strong and a weak outranking relations which are in turn exploited as in ELECTRE IV method (the latter in fact coincides with MELCHIOR in the particular case where T is empty). Let us note here that the choice of combinations of definitions is not arbitrary. Leclercq [26] gives examples of coherent combinations, a study of the properties of the resulting outranking relations and a numerical example.

11.4.10 TRICHOTOMIC SEGMENTATION (Moscarola and Roy [35])

The procedure described in this section was built in order to help a decision-maker who must, during the process of discovering the actions, decide to which category he will assign them among several ones. These categories are defined in respect of the treatment they will receive later; this kind of situation occurs, for example, in loan allocation problems, when launching new products or research projects, when awarding promotions, and so on. As indicated by its name, this procedure is limited to the case where there are three categories considered: K^+, K^- and $K^?$ (in the example of loan allocation, they correspond to "accepting", "refusing" and "awaiting extra information").

Assume the considered actions are evaluated through criteria $g_1, g_2, \ldots, g_n$. The procedure consists in fixing (with the decision-maker's help) some pairs $\{\bar{b}^k, \bar{c}^k\}$, $k = 1, 2, \ldots, l$, of n-dimensional vectors so that if there exists a value of k such that

$$g_j(a) \geq b_j^k, \ \forall\, j,$$

then a is assigned to K^+; if there exists a value of k such that

$$g_j(a) \leq c_j^k, \ \forall\, j,$$

then a is assigned to K^-; if there exists a value of k such that

$$b_j^k \geq g_j(a) \geq c_j^k, \ \forall\, j,$$

then a is assigned to $K^?$.

Intuitively, the $\bar{b}^k$'s correspond to "high profiles" for which the decision-maker chooses category K^+ without any hesitation (he agrees to award a loan) and the $\bar{c}^k$'s correspond to "low profiles" for which the decision-maker chooses category K^- without any hesitation (he refuses the loan). There is a one-to-one correspondence between the $\bar{b}^k$'s and the $\bar{c}^k$'s so that any action offering a profile between $\bar{b}^k$ and $\bar{c}^k$ (for the same value of k) is assigned to category $K^?$.

Of course, many actions fall into none of the three situations described above. The authors therefore recommend to compute, for each action a, the following outranking indexes

$$S(a, b^k), S(b^k, a), S(a, c^k), S(c^k, a), \qquad k = 1, 2, \ldots, l,$$

where b^k and c^k respectively denote fictitious actions such that

$$\begin{cases} g_j(b^k) = b_j^k & \forall\, j, \\ g_j(c^k) = c_j^k & \forall\, j, \end{cases}$$

and where outranking indices are computed as in the ELECTRE III method (Section 11.4.5).

The assignment of each action to one of the three categories is performed by using a decision tree of the type described in Fig. 11.2, where

$$\begin{aligned} S(a, \hat{b}) &= \max_k S(a, b^k), \\ S(b^*, a) &= \max\{S(b^k, a) : b^k \neq \hat{b}\}, \end{aligned}$$

$$\begin{aligned} S(\hat{c}, a) &= \max_k S(c^k, a), \\ S(a, c^*) &= \max\{S(a, c^k) : c^k \neq \hat{c}\}, \end{aligned}$$

and where s, t, s', t' are thresholds to be fixed in function of the considered application and in particular by taking into account the inconveniences due to an assignment error and those resulting from any assignment to category $K^?$ (loss of time, extra costs, etc.). See Roy [50] for some examples of applications.

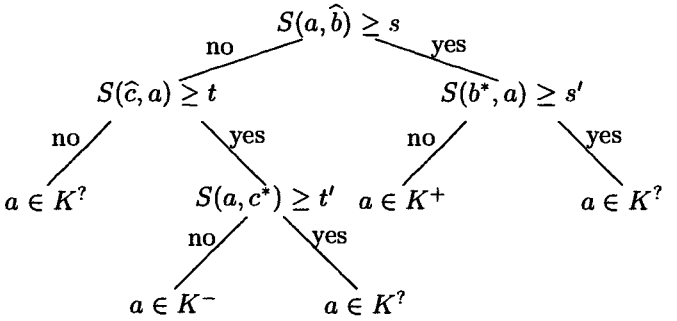


Figure 11.2 The assignment tree

11.4.11 Synoptic tableau

Table 11.1 below summarizes some characteristics of the outranking methods. It can be used as a guide to progress in the choice of a method, but not all the elements to be taken into account are considered in this tableau (see for example Guitouni and Martel [22], for a more complete tableau). In section 11.7, we will come back to this question of the choice of a method.

Table 11.1 Synoptic tableau

Method	References	Needed informations	Type of relation	Final output
ELECTRE I	Roy [47]	Criteria Veto thresholds Weights Concordance level	One outranking relation	Subset of A
ELECTRE IS	Roy, Skala [60] Roy, Bouyssou [55]	Pseudo-criteria Veto thresholds Weights concordance level	One outranking relation	Subset of A + some indices
ELECTRE II	Roy, Bertier [54]	Criteria Pairs of veto thresholds Weights 2 concordance levels	Two imbedded outranking relations	Partial preorder
ELECTRE III	Roy [49]	Pseudo-criteria Veto thresholds Weights	Valued outranking relation	Partial preorder
ELECTRE IV	Roy, Hugonnard [56]	Pseudo-criteria Veto thresholds	Imbedded outranking relations	Partial preorder
ELECTRE Tri	Roy, Bouyssou [55]	Pseudo-criteria Veto thresholds Weights Profiles	Valued outranking relation	Assignment to categories
PROMETHEE I,II	Brans et al. [15]	Generalized criteria Weights	Valued preference relation	I: partial preorder II: complete preorder
MELCHIOR	Leclercq [26]	Pseudo-criteria Veto thresholds Importance relation	Imbedded outranking relations	Partial preorder
Trichotomic Segmentation	Moscarola, Roy [35]	Pseudo-criteria Veto thresholds Weights Profiles Assignment thresholds	Valued outranking relation	Assignment to 3 categories

11.4.12 Some other methods

The QUALIFLEX method (Paelinck [38]) consists in exploring the set of weights which are compatible with the importance relation on the criteria (assumed to be a complete preorder) and, for each of them, to determine the ranking of the actions which is at minimum distance from the weighted average of the rankings yielded by the criteria.

The ORESTE method (Roubens [46]) combines, in a rather audacious way, the ranks of the actions on the criteria with the ranks of the criteria themselves (the importance relation on the criteria is also a complete preorder), in order to obtain a global rank for each ordered pair (action, criterion) to be used as the basis for comparing actions. Pastijn and Leysen [39] bring some light on the interpretation of the parameters of the method.

The TACTIC method (Vansnick [68]) is a way for building a global preference relation on the basis of a family of n quasi-criteria, weights and veto thresholds; it is rather similar to the first step of ELECTRE I, but adapted to quasi-criteria and to the building of a strict preference relation instead of an outranking relation.

Some outranking methods were also proposed for the treatment of uncertain data or decision under risk (see for example D'Avignon and Vincke [18]; Martel and Zaras [32]).

An outranking method can also be obtained when one uses a weighted sum with some imprecision on the "weights". This situation was treated for the first time in a concrete problem (see Montgolfier and Bertier [19]). New developments were proposed by Bana e Costa [2], Bana e Costa and Vincke [5].

11.5 SOME COMMENTS ON THE PARAMETERS OF THE METHODS

11.5.1 The "weights" of the criteria

The "*weights*" of the criteria are of course crucial parameters in every multicriteria method. The most important remark is that their meaning, hence the way to assess them, completely depends on the method where they are used. In other words, the concept of importance of the criteria has no intrinsic interpretation: it depends on the way it is manipulated. To take an example outside of the field of outranking methods, it is well-known that the so-called "weights" of the weighted sum are nothing more than that tradeoffs between the criteria, so that a good question to ask the decision-maker to assess them is: "How many units on criteria g_k are you ready to loose in order to obtain one unit more on criterion g_j" (the answer gives the ratio between the "weights" of the two criteria).

In that model, the fact that $w_j > w_k$ does not mean that criterion g_j is more important than criterion g_k: a simple change of units can reverse the inequality. So, if values are given to the weights independently of the scales used to express the utility functions, then these values cannot be used in an additive model.

A consequence of this connection between the meaning of the weights and the way they are manipulated in the methods is that it is generally a nonsense to compare two different methods by applying them to a same numerical example, and by using the same weights.

Another consequence is that it is hazardous to use, in a method, weights which are spontaneously given by a decision-maker (this situation is not so rare).

The question of the exact meaning of the weights in the outranking methods is not easy: let us examine what we can say on this subject (see also section 11.5.2 about the veto thresholds which are not completely disconnected from the relative importances of the criteria).

In ELECTRE I and II, as we have seen, the weight of a criterion plays the same role as a number of votes in a voting procedure. More rigorously, in a non-compensatory method (such as ELECTRE I), one can define a "more important than" relation on the set of coalitions of criteria. Given two coalitions of criteria G and H (two subsets of family $\mathcal{F}$), "G is more important than H"

if two (real or fictitious) actions a and b can be found such that

$$\begin{cases} a \text{ is better than } b \text{ for all the criteria in } G, \\ b \text{ is better than } a \text{ for all the criteria in } H, \\ a \text{ and } b \text{ are indifferent for all the other criteria,} \\ a \text{ is globally better than } b \end{cases}$$

(because of the non-compensation, one can be sure that the latter definition won't lead to any contradiction). If one assumes that this "more important than" relation can be represented using n constants $p_1, p_2, \ldots, p_n$ (associated to the n criteria) such that the comparison between G and H is equivalent to that between $\sum_{j \in G} p_j$ and $\sum_{j \in H} p_j$, one is taken back to well-known non-compensatory methods (on this subject, see Vansnick [68]). From a theoretical point of view, this allows some necessary and sufficient existence conditions for these "weights" to be established. In practice, their determination can be deduced from comparisons between real or fictitious actions.

Another way of characterizing the nature of the weights used in a method consists in seeking the transformations of these weights which are allowed without any modification of the method's results and finding inspiration in the results of measurement theory (see Roberts [45]). In a deterministic context, most methods permit a multiplication of the weights by a positive constant (this allows the weights to be normalized by making them sum to 1). In the weighted average method, the latter is the only transformation which is allowed (one says that the "weights" make up a ratio scale). In the ELECTRE I method, any transformation preserving the order of the coalitions' weights is authorized (the weight of a coalition being the sum of the weights of the criteria composing it). A great deal of research remains to be done in this direction.

In ELECTRE III, ELECTRE TRI and PROMETHEE I and II, the exact interpretation of the weights is not very easy: in fact they are partly numbers of votes (for the non-compensentory parts where $c_j(a,b)$ or $F_j(a,b)$ are equal to 0 or 1) and partly tradeoffs between "degrees of outranking or preference" (for the compensatory parts where $c_j(a,b)$ or $F_j(a,b)$ vary between 0 and 1).

A common aspect of all these methods is that they are not affected by a change of units of the criteria; this means that the inequality $w_j > w_k$ implies that the role of criterion g_j is really heavier than that of g_k, so that the term "weight" can be considered as well appropriate.

For the practical way of building the weights, many procedures have been proposed in the literature (see for example Mousseau [36] for a survey of these procedures). Roy and Bouyssou [55] propose (page 301) some principles to be respected in such a procedure and describe (page 303) a particular procedure in the framework of the outranking approach.

Finally, it is clear that sensitivity and robustness analysis are necessary to complete and validate the results of a method (see section 11.5.3). For the reader interested in a reflexion on the concept of importance, see also Roy and Mousseau [57].

11.5.2 The veto thresholds

The *veto threshold* $v_j\,(g_j(a))$ is the smallest interval $(g_j(a), g_j(b))$, where $g_j(a) < g_j(b)$, which is sufficient to refuse the assertion that a outranks b.

In a certain sense, the veto threshold is also connected to the notion of importance. To assess its value, the first question is to know whether the decision-maker wants to associate a veto threshold to criterion g_j (for each j). One way is to present two alternatives a and b such that

- a is significantly better than b for all the criteria except for g_j,
- $g_j(a)$ and $g_j(b)$ are respectively one of the worst and one of the best values for criterion g_j.

If the decision-maker prefers a over b without hesitation, there is no reason to introduce a veto threshold associated to g_j (it is generally the case when g_j is not very important).

If the decision-maker has some difficulties in comparing a and b, it is justified to introduce a veto threshold. The reader will find in Roy and Bouyssou [55] some comments on how to assess the value of this threshold.

11.5.3 Sensitivity and robustness analysis

As already mentioned, the analyst must often choose values for parameters which are not very well known (because information is never complete and reliable) or which are rather technical (because they are necessary to build the model): weights, veto thresholds, indifference and preference thresholds, concordance levels,... are necessarily a little bit arbitrary, as well as tradeoffs, utility functions or probability distributions in other approaches. One way is to choose rather central values in order to obtain a first solution and then to make a sensitivity analysis: the influence of each parameter on the solution is studied in order to detect the most critical ones and to inform the decision-maker on the neighbourhood of the proposed solution (see Mareschal [29], for an example about PROMETHEE method).

Another way is to try to find solutions which are robust in the sense that they are good for the different plausible sets of values of the parameters. With a robust solution, the decision-maker has some guarantee that the decision proposed by the method will not lead to a catastrophe if the values of the parameters are different from the values used in the model. This rather new concept of *robustness* should lead to very interesting developments in multicriteria analysis and operations research (see Roy and Bouyssou [55]; Roy [53]; Vincke [73]).

Just to illustrate this aspect, consider the problem of building an outranking relation on a finite set A of alternatives on basis of the following information:

- n complete preorders (complete and transitive relations) $R_1, R_2, \ldots, R_n$ defined on A (rankings of the alternatives according to n qualitative criteria or given by n experts),

- a "strictly more important than" relation defined in the set of coalitions of criteria, i.e. in the set of all subsets $N = \{1, 2, \ldots, n\}$; this relation, denoted by Q, is assumed to satisfy the following properties:
  - it is asymmetric and transitive,
  - $\forall\, C \subset N, C \neq \emptyset \Rightarrow C\ Q\ \emptyset$ (every coalition is strictly more important than the empty coalition),
  - $\forall\, C, D, E \subset N, C\ Q\ D \Rightarrow C \bigcup E\ Q\ D$ (adding criteria to a coalition reinforces its importance),
  - $\forall\, C, D, E \subset N$, with $C \bigcap E = \emptyset$ and $D \bigcap E = \emptyset$, $C\ Q\ D \Leftrightarrow C \bigcup E\ Q\ D \bigcup E$ (one can add criteria to or delete common criteria from two coalitions without changing the importance relation between them),
  - it is not necessarily complete.

We are looking for a method for building a relation R on A which should satisfy the following two conditions:

$$\begin{cases} -\ \forall\, a, b \in A : a\ R_i\ b,\ \forall\, i \Rightarrow a\ R\ b; \\ -\ \forall\, a, b \in A : \{i :\ a\ P_i\ b\}\ Q\ \{j :\ b\ P_j\ a\} \Rightarrow \text{not } b\ P\ a, \end{cases}$$

where P_i, P_j and P are the asymmetric parts of R_i, R_j and R respectively. The first condition is unanimity rule while the second condition expresses the fact that if the coalition of criteria for which a is strictly better than b is more important than the coalition of criteria for which b is strictly better than a, then b cannot strictly outrank a.

A traditional approach would consist to asociate, with each R_i, a "weight" w_i such that,

$$\forall\, C, D \subset N,\ C\ Q\ D \Rightarrow \sum_{i \in C} w_i > \sum_{j \in D} w_j.$$

Then, on basis of these weights, a rule should be introduced to define R. For a given rule different relations R can be obtained, depending on the choice of the weights. The idea of robustness is to try to define a rule giving a result which is as stable as possible with respect to variations of the weights.

For example if we decide that two results R and R' are contradictory if there is a pair $\{a, b\}$ of alternatives such that $a\ P\ b$ and $b\ P'\ a$ (where P and P' are the asymmetric part of R and R' respectively), we can define a robust rule as a rule which never leads to contradictory results.

In this example, it can be proved (see Vincke [69]) that rule 2 here below is robust while rule 1 is not.

Rule 1: Choose a value k in the interval $[.5, 1]$;

$$\forall\, a, b \in A, a\ R\ b \quad \text{iff} \quad \sum_{i : a\ R_i\ b} w_i \geq k \sum_{j=1}^{n} w_j$$

$$\text{or} \quad \{i :\ a\ R_i\ b\}\ Q \{j :\ b\ P_j\ a\}$$

Rule 2: Choose a value k in the interval $[.5, 1]$;

$$\forall\, a,b \in A, a\, R\, b \quad \text{iff} \quad \sum_{i:a\ R_i\ b} w_i \geq k \sum_{j=1}^{n} w_j$$

$$\text{and} \quad \{i:\ a\ P_i\ b\}\ Q\{j:\ b\ P_j\ a\}$$
$$\text{or} \quad a\ R_i\ b,\ \forall\, i$$

(Note that rule 2 is also robust for variations of k.)

This small example shows how a concept of robustness can help the scientist to analyze the characteristics and the differences between two a priori reasonable methods.

11.6 THEORETICAL ASPECTS OF OUTRANKING METHODS

The outranking methods have often been criticized for their lack of axiomatic foundations (see Bouysson et al. [11]). Since a few years several authors have attempted to progress in this way; this section proposes a brief overview of the main results obtained until now.

A first interesting question is to ask whether the outranking relations may have some structural properties (beside the fact that we know they are in general neither complete nor transitive). This question was studied by Bouyssou [9]; the conclusion is that the relations obtained by ELECTRE methods do not possess any particular property (in other words, any relation can be obtained). It is not the case of the relations obtained by PROMETHEE methods but their characterization is still an open problem. On the other hand, one can try to choose the parameters of the methods in order to obtain outranking relations with nice mathematical properties: such a problem is not so far from Arrow's problem and social choice theory. Generally speaking, the results show that the only way to obtain nice outranking relations is to accept the existence of coalitions of criteria which impose their preferences (see for example Perny [40]). Note also that all these considerations are connected to the difficult problem of how to define nice properties of valued (or fuzzy) relations (see, for example, Fodor and Roubens [20], or Perny and Roy [41]).

Another theoretical approach consists in defining potential fundamental properties of the methods and in verifying which properties are satisfied by which methods. This kind of research can lead to different types of results:

- impossibility theorems, showing that some sets of properties are incompatible and explaining why certain methods cannot satisfy certain properties;
- characterization theorems, giving, for a particular method, a list of properties which are simultaneously satisfied only by this method and improving so the understanding of the method;
- typologies of the methods based on fundamental properties, allowing to compare them and giving ideas for the development of methods.

To illustrate this last aspect let us consider two methods which can be found in the literature for building a ranking on basis of an outranking relation where the circuits have been reduced.

Let A be the set of alternatives after reduction of the circuits in the outranking relation S. We say that "a is better than b" or that "b is worse than a" when $a\ S\ b$, and we note $a\ P\ b$ when $a\ S\ b$ and not $b\ S\ a$.

First method

- Put in the first class the elements of A which are not worse than any other element.
- Remove the elements of the first class.
- Put in the second class the elements which are not worse than any other element in the remaining set.
- Remove the elements of the second class.
- Continue this procedure until all the elements are removed.

Second method

- Calculate the score of each element of A (the score of a is the number of elements which are worse than a minus the number of elements which are better than a).
- Rank the elements in the decreasing order of their score.

Let us now consider the following list of "reasonable" properties.

1. if S has already the structure of a ranking, than the procedure should give this ranking;
2. if it is possible to obtain a ranking from S without deleting any information contained in S, then the procedure should propose such a ranking; in particular,
   (a) if S is transitive, the procedure should give a ranking containing S;
   (b) if the transitive closure of S is a ranking the procedure should give this ranking;
   (c) if P has no cycle, the procedure should give a ranking containing P;
3. the procedure should be independent of the labels of the alternatives;
4. if a new information is added, which is favourable to alternative a, than the position of a in the final ranking should not be depreciated; more precisely, if a new relation S' is obtained from S by adding the fact that $a\ S'\ b$ (everything else being unchanged), than the place of a in the

ranking associated to S' should be at least as good as its place in the ranking associated to S.

5. if the class of best elements in the ranking associated to S is removed, and if the procedure is applied to the remaining elements, then the ranking of those remaining elements should be unchanged.

It can be proved (see Vincke [71]) that the first method satisfies all these properties, while the second one satisfies (1), (2b) and (3) but violates the other properties. With such an information, the choice of a method is not completly arbitrary: it can be justified on a sound basis.

In any case, this kind of research is a very efficient way for improving the understanding of the methods by the scientists (even if the obtained results are not always communicable to the practitioners).

A lot of work has still to be done in this way, but here are some references which illustrate this approach: Henriet [24], Bouyssou [7, 8], Bouyssou and Perny [10], Pirlot [42], Vincke [71, 69], Arrow and Raynaud [1], Perny [40], Marchant [27, 28], Tsoukias and Vincke [67], Bouyssou and Vincke [14].

Finally, let us also mention the suggestion of finding a common framework regrouping most of the methods for aggregating preferences and based on various definitions of compensation and independence of the criteria and on the numerical representation of not necessarily transitive preference structures (see Bouyssou and Pirlot [12], Pirlot and Vincke [43], Bouyssou, Pirlot and Vincke [13]).

11.7 HOW TO CHOOSE A METHOD ?

Many elements can be taken into account in order to choose a method, and several authors have presented comparative studies of multicriteria methods (including outranking methods and other approaches): the interested reader can consult Guitouni and Martel [22] for a good survey on these aspects.

Assuming that an outranking method will be used (see section 11.2 for the justification of such a choice), here are some questions to consider for choosing the method.

- What kind of output do we want to obtain ?
- What kind of information is it possible to obtain as input (different kinds of scales for the criteria, different kinds of thresholds, type of information about the relative importance of the criteria,...) ?
- Which properties are considered as important for the method ?

A tableau such as that in Section 11.4.11 and theoretical studies (as those mentioned in Section 11.6) can be helpful to cope with these questions. Beside them, other more pragmatic or subjective arguments can also be used as, for example, the access to a good software, the more or less great effort to be made by the analyst to understand the method, the more or less great effort to be made to explain the principle of the method to the decision-maker, the access to a good specialist of the method, the habits of the analyst, ...

However, we consider that the subjective aspects should not lead to completely neglecting the first questions, if we want to give a real scientific status to multicriteria decision-aid.

11.8 SOFTWARE

For most of the outranking methods, there exist software packages which are regularly updated. The best way to obtain them is to directly contact the authors of the methods: in general, demo versions can be obtained very easily.

11.9 APPLICATIONS

Outranking methods have been used in several concrete applications, although this is not really clear from the international journals. The main reason is probably the fact that describing a real decision process, with all the hesitations, discussions, tentatives, corrections, ..., is a tremendous work and does not really possess the good characteristics to be published in a scientific journal. On the other hand, reducing the description of the decision process to the strict application of a method to a set of data is not very interesting.

Siskos et al. [66] established a long list of applications. More recent examples can be found in Bana e Costa and D. Neves [4], Barda et al. [6], Briggs et al. [16], D'Avignon and Mareschal [17], Grassin[21], Hens et al. [25], Roy et al. [59], Roy et al. [61], Simos [63], Siskos and Hubert [64], Siskos et al. [65].

A lot of other examples in French can also be found in the Cahiers and the Documents of the Lamsade (collection directed by B. Roy).

References

[1] Arrow, K., H. Raynaud: *Social choice and multicriterion decision-making.* MIT Press, Cambridge, Mass, 1986

[2] Bana e Costa, C.: A multicriteria decision aid methodology to deal with conflicting situations on the weights. European Journal of Operational Research 26 (1986) pp. 22–34

[3] Bana e Costa, C.: Les problématiques de l'aide à la décision: vers l'enrichissement de la trilogie choix-tri-rangement. Revue de Recherche Opérationnelle 30 (1996) 2 pp. 191–216

[4] Bana e Costa, C., C.F. Dinis Das Neves: Describing and formalizing the evaluation process of portuguese navy officers. In: A.G. Lockett, G. Islei (eds.): *Improving Decision Making in Organisations*, Springer–Verlag, 1989 , pp. 355–369

[5] Bana e Costa, C., Ph. Vincke: Measuring credibility of compensatory preference statements when trade-offs are interval determined. Theory and Decision 39 (1995) pp. 127–155

[6] Barda, O.H., J. Dupuis, P. Lencioni: Multicriteria location of thermal power plants. European Journal of Operational Research 45 (1990) pp. 332–346

[7] Bouyssou, D.: Some remarks on the notion of compensation in MCDM. European Journal of Operational Research 26 (1986) pp. 150–160

[8] Bouyssou, D.: Ranking methods based on valued preference relations: a characterization of the net flow method. European Journal of Operational Research 60 (1992) pp. 61–67

[9] Bouyssou, D.: Outranking relations: do they have special properties ? Journal of Multicriteria Decision Analysis 5 (1996) pp. 99–111

[10] Bouyssou, D., P. Perny: Ranking methods for valued preference relations: a characterization of a method based on leaving and entering flows. European Journal of Operational Research 61 (1992) pp. 186–194

[11] Bouyssou, D., P. Perny, M. Pirlot, A. Tsoukias, Ph. Vincke: The manifesto of the new MCDA era. Journal of Multicriteria Decision Analysis 2 (1993) pp. 125–127

[12] Bouyssou, D., M. Pirlot: A general framework for the aggregation of semiorders. Tech. rep., ESSEC, Cergy-Pontoise, (1997)

[13] Bouyssou, D., M. Pirlot, Ph. Vincke: A general model of preference aggregation. In: M. Karwan, J. Spronk, J. Wallenius (eds.): *Essays in Decision-Making*, Springer Verlag, 1997, pp. 120–134

[14] Bouyssou, D., Ph. Vincke: Ranking alternatives on basis of preference relations: a progress report with special emphasis on outranking relations. Journal of Multicriteria Decision Analysis 6 (1997) pp. 77–85

[15] Brans, J.P., B. Mareschal, Ph. Vincke: Promethee: A new family of outranking methods in multicriteria analysis. In: J.P. Brans (ed.): *Operational*

Research'84, Elsevier Science Publishers B.V. (North-Holland), 1984, pp. 408–421

[16] Briggs, Th., P. Kunsch, B. Mareschal: Nuclear waste management: an application of the multicriteria Promethee methods. European Journal of Operational Research 44 (1990) pp. 1–10

[17] D'Avignon, G., B. Mareschal: Specialization of hospital services in Québec an application of the Promethee and Gaia methods. Mathematical and Computer Modelling 12 (1989) pp. 1393–1400

[18] D'Avignon, G., Ph. Vincke: An outranking method under uncertainty. European Journal of Operational Research 36 (1988) pp. 311–321

[19] de Montgolfier, J., P. Bertier: *Approche multicritère des problèmes de décision.* Hommes et Techniques, Paris, 1978

[20] Fodor, J., M. Roubens: *Aggregation, ranking and choice procedures with applications to multiple criteria decision making methods dealing with valued binary relations.* Kluwer, 1993

[21] Grassin, N.: *Aide à la décision pour le choix des lignes à très haute tension.* Ph.D. thesis, Université de Paris-Dauphines, (1986), Thèse de 3e cycle

[22] Guitouni, A., J.M. Martel: Tentative guidelines to help choosing an appropriate MCDA method. *European Journal of Operational Research*, forthcoming

[23] Hansen, P., M. Anciaux, Ph. Vincke: Quasi-kernels of outranking relations. In: H. Thiriez, S. Zionts (eds.): *Multiple Criteria Decision Making*, vol. 30, Springer-Verlag, 1976, pp. 53–63

[24] Henriet, D.: The Copeland choice function - an axiomatic characterization. Social Choice and Welfare 2 (1985) pp. 49–64

[25] Hens, L., H.Pastijn, W. Struys: The allocation of grants to subcentral authorities: a multicriteria decision making approach. Revue des Systèmes de Décision 1 (1992) pp. 155–174

[26] Leclercq, J.P.: Propositions d'extensions de la notion de dominance en présence de relations d'ordre sur les pseudo-critères : Melchior. Revue Belge de Recherche Operationnelle, de Statistique et d'Informatique 24 (1984) pp. 32–46

[27] Marchant, T.: Valued relations aggregation with the Borda method. Journal of Multiple Criteria Decision Analysis 5 (1996) pp. 127–132

[28] Marchant, T.: Cardinality and the Borda score. To appear in the *European Journal of Operational Research*

[29] Mareschal, B.: Weights stability intervals in multicriteria decision aid. European Journal of Operational Research 33 (1988) pp. 54–64

[30] Mareschal, B., J.P. Brans: Geometrical representations for MCDA. European Journal of Operational Research 34 (1988) pp. 69–77

[31] Martel, J.: Aggregating preference: utility function and outranking approaches. In: J. Climaco (ed.): *Multicriteria Analysis*, Springer Verlag, 1997, pp. 76–85

[32] Martel, J., K. Zaras: Stochastic dominance in multicriterion analysis under risk. Theory and Decision 35 (1995) pp. 31–49

[33] Matarazzo, B.: Multicriterion analysis of preferences by means of pairwise actions and criterion comparisons (Mappac). Applied Mathematics and Computation 18 (1986) pp. 119–141

[34] Matarazzo, B.: Preference ranking global frequencies in multicriterion analysis (Pragma). European Journal of Operational Research 36 (1988) pp. 36–50

[35] Moscarola, J., B. Roy: Procédure automatique d'examen de dossiers fondée sur une segmentation trichotomique en présence de critères multiples. RAIRO Recherche Opérationnelle 11 (1977) pp. 145–173

[36] Mousseau, V.: *Problèmes liés à l'évaluation de l'importance en aide multicritère à la décision: Réflexions théoriques et expérimentations.* Ph.D. thesis, Université de Paris-Dauphine, (1993)

[37] Nadeau, R., J.M. Martel, L. Kiss: Eleccalc: Modelling the decision maker's preference with Electre II. In: M. Cerny, D. Glückaufova, D. Loula (eds.): *Multicriteria Decision Making: Methods, Algorithms, Applications*, Liblice, Czechoslovakia, (March 18–22 1991), 1991, pp. 109–119

[38] Paelinck, J.: Qualiflex, a flexible multiple criteria method. Economics Letters 3 (1978) pp. 193–197

[39] Pastijn, H., J. Leysen: Constructing an outranking relation with Oreste. Mathematical and Computer Modelling 12 (1989) pp. 1255–1268

[40] Perny, P.: *Modélisation, agrégation et exploitation des préférences floues dans une problématique de rangement : Bases axiomatique, procédures et logiciels.* Ph.D. thesis, Université de Paris-Dauphine, (1992)

[41] Perny, P., B. Roy: The use of fuzzy outranking relations in preference modelling. Fuzzy Sets and Systems 49 (1992) pp. 33–53

[42] Pirlot, M.: A characterization of "min" as a procedure for exploiting valued preference relations and related results. Journal of Multiple Criteria Decision Analysis 4 (1995) pp. 37–56

[43] Pirlot, M., Ph. Vincke: *Semiorders: properties, representations, applications.* Kluwer, 1997

[44] Renard, F.: Utilisation d'Electre dans l'analyse des réponses à un appel d'offres : Le cas de la machine de tri paquets à la direction générale des postes. Cahier du LAMSADE 73 (1986), 36 p.

[45] Roberts, F.: *Measurement theory, with applications to decision making, utility and the social sciences.* Addison-Wesley, 1979

[46] Roubens, M.: Preference relations on actions and criteria in multicriteria decision making. European Journal of Operational Research 10 (1982) pp. 51–55

[47] Roy, B.: Classement et choix en présence de points de vue multiples (la méthode Electre). Revue Française d'Informatique et de Recherche Opérationnelle 8 (1968) pp. 57–75

[48] Roy, B.: How outranking relations helps multiple criteria decision making. In: J. Cochrane, M. Zeleny (eds.): *Multicriteria Decision Making*, University of South Carolina, 1973, pp. 179–201

[49] Roy, B.: Electre III: un algorithme de classement fondé sur une représentation floue des préférences en présence de critères multiples. Cahiers du CERO 20 (1978) pp. 3–24

[50] Roy, B.: A multicriteria analysis for trichotomic segmentation problems. In: P. Nijkamp, J. Spronk (eds.): *Multicriteria Decision Analysis: Operational Methods*, Gower Press, 1981, pp. 245–257

[51] Roy, B.: *Méthodologie Multicritère d'Aide à la Décision.* Economica, 1985

[52] Roy, B.: The outranking approach and the foundations of Electre methods. Theory and Decision 31 (1991) pp. 49–73

[53] Roy, B.: Un chaînon manquant en RO-AD : les conclusions robustes. Cahier du LAMSADE 44 (1997), 125 p.

[54] Roy, B., P. Bertier: La méthode Electre II – une application au média planning. In: M. Ross (ed.): *OR'72*, North-Holland Publishing Company, 1973, pp. 291–302

[55] Roy, B., D. Bouyssou: *Aide Multicritère à la Décision: Méthodes et cas.* Economica, 1993

[56] Roy, B., J.-Chr. Hugonnard: Ranking of suburban line extension projects on the Paris metro system by a multicriteria method. Transportation Research 16A (1982) pp. 301–312

[57] Roy, B., V. Mousseau: A theoretical framework for analysing the notion of relative importance of criteria. Cahier du LAMSADE 131 (1995)

[58] Roy, B., M. Présent, D. Silhol: Programmation de la rénovation des stations du métro parisien: un cas d'application de la méthode Electre III. Document du LAMSADE 24 (1983), 67 p.

[59] Roy, B., M. Présent, D. Silhol: A programming method for determining which Paris metro should be renovated. European Journal of Operational Research 24 (1986) pp. 318–334

[60] Roy, B., J.M. Skalka: Electre IS - aspects méthodologiques et guide d'utilisation. Document du LAMSADE 30 (1984), 125 p.

[61] Roy, B., R. Slowinski, W. Treichel: Multicriteria programming of water supply systems for rural areas. Water Resources Bulletin 28 (1992) pp. 13–31

[62] Roy, B., Ph. Vincke: Pseudo-orders: definition, properties and numerical representation. Mathematical Social Sciences 62 (1987) pp. 263–274

[63] Simos, J.: *Evaluer l'impact sur l'environnement.* Presses Polytechniques et Universitaires Romandes, 1990

[64] Siskos, J., Ph. Hubert: Multi-criteria analysis of the impacts of energy alternatives: a survey and a new comparative approach. European Journal of Operational Research 13 (1983) pp. 278–299

[65] Siskos, J., J. Lombard, A. Oudiz: The use of multicriteria outranking methods in the comparison of control options against a chemical pollutant. Journal of Operational Research Society 37 (1986) pp. 357–371

[66] Siskos, J., G. Wäscher, H.M. Winkels: A bibliography on outranking approaches (1966-1982). Cahier du LAMSADE 45 (1983), 31 p.

[67] Tsoukias, A., Ph. Vincke: A new axiomatic foundation of partial comparability. Theory and Decision 39 (1995) pp. 79–114

[68] Vansnick, J.C.: On the problems of weights in MCDM (the noncompensatory approach. European Journal of Operational Research 24 (1986) pp. 288–294

[69] Vincke, Ph.: Robust and neutral methods for aggregating preferences into an outranking relation. Forthcoming

[70] Vincke, Ph.: *L'aide multicritère à la décision.* Editions Ellipses (France) et Editions de l'Université de Bruxelles (Belgique), 1989

[71] Vincke, Ph.: Exploitation of a crisp relation in a ranking problem. Theory and Decision 32 (1992) pp. 221–240

[72] Vincke, Ph.: *Multicriteria decision aid.* J. Wiley and Sons, Chichester, 1992

[73] Vincke, Ph.: Robust solutions and methods in decision-aid. Forthcoming

[74] Yu, W.: *Aide multicritère à la décision dans le cadre de la problématique du tri : Concepts, méthodes et applications.* Ph.D. thesis, Université de Paris-Dauphine, (1992)

12 MULTI-CRITERIA PROBLEM STRUCTURING AND ANALYSIS IN A VALUE THEORY FRAMEWORK

Valerie Belton

CONTENTS

T. Gal et al. (eds.), *Multicriteria Decision Making*

Abstract: This chapter focuses on the use in practice of multi-attribute value theory (MAVT). MAVT is a simplification of multi-attribute utility theory (MAUT) in that, unlike MAUT, MAVT does not seek to model the decision maker's attitude to risk. As a consequence it rests on simpler elicitation procedures which are more widely accepted by practising decision makers. The most significant recent advances in this field relate not to the underlying theory, but to the way in which MAVT can be, and is, used in practice to support decision making. The chapter begins with a brief review of the concepts of value theory. An exemplary decision is then used to convey a sense of how the process of decision making may currently be facilitated through the use of MAVT. The final section reviews recent developments which are beginning to impact on practice. These relate to: the use of problem structuring methods: advances in technology: and organisational developments.

12.1 INTRODUCTION

This chapter will focus on the use in practice of multi-attribute value theory (MAVT). MAVT is a simplification of multi-attribute utility theory (MAUT) in that, unlike MAUT, MAVT does not seek to model the decision maker's attitude to risk. As a consequence it rests on simpler elicitation procedures which are more widely accepted by practising decision makers. The most significant recent advances in this field relate not to the underlying theory, but to the way in which MAVT can be, and is, used in practice to support decision making. The chapter begins with a brief review of the concepts of value theory. An exemplary decision is then used to convey a sense of how the process of decision making may currently be facilitated through the use of MAVT. The final section reviews recent developments which are beginning to impact on practice.

12.2 INTRODUCTION TO VALUE THEORY

12.2.1 Fundamental requirements of value theory

The principal tenet of value theory is simply that it is possible to represent an individual's preferences in a defined context by a value function, V( ), such that if option A is preferred to option B then V(A) > V(B). The word "option" is used in a broad sense; the available options may be simple objects, such as chocolate bars;

they may be complex action plans; they may be alternative futures defined by the interaction of different strategies and scenarios.
For this representation to be possible, the individual's preferences must satisfy the following two properties:

12.2.1.1 Transitivity. Consider any 3 possible options, A, B and C. If a decision maker prefers A to B and B to C, then she should prefer A to C. This property is clearly apparent in the world of physical measurement; if Alec is taller than Bob and Bob is taller than Charlie, then we know that Alec is taller than Charlie. It is also intuitively appealing in the context of preference measurement, however, it is argued by some that the condition is too strong. Thus, it is perhaps appropriate to delve a little more deeply into the issue.

It is clear that in a descriptive sense an individual's preferences can appear to be intransitive. For example, suppose I have depleted my fruit stores to the extent that I have available just one apple, one orange and one banana. I choose to eat a banana with my breakfast, leaving me to chose between the apple and orange for lunch. I choose the orange. Later that afternoon someone gives me another banana (just as good as the first one) to add to the remaining apple, broadening the choice available for tea. I choose the apple. So we have:

Banana *preferred to* orange *preferred to* apple *preferred to* banana

- that is, an intransitive set of preferences. Here it is easy to justify the situation by arguing that my preferences are changing during the day, thus the context of the decision is changing. For example, at breakfast time my objective is to maximise energy in preparation for the day at work, so the banana with its high carbohydrate level is the obvious choice. By lunchtime I am quite thirsty and looking for something refreshing, suggesting the juicy orange. And when it comes to teatime I want something crunchy to accompany soft sandwiches. So, the argument is not that intransitive preferences never occur, but that where they do occur it is because of an underlying change, in either, or both, of the decision maker's objectives and the decision context. Value theory demands that in a static situation preferences should be transitive.

12.2.1.2 Comparability. This second condition demands that, given any two options A and B in a specified decision context, a decision maker must be able to indicate whether she prefers A to B, or B to A, or is indifferent between the two options.

That is not to say that the decision maker must be able to do this immediately and unaided, the rationale of decision analysis is to help in this regard. Neither is it to deny that it might be very difficult to compare A and B.

If these two conditions are satisfied then it is possible to determine a value function V( ) such that:

$$A \; P \; B \text{ implies and is implied by } \quad V(A) > V(B).$$

A I B implies and is implied by $V(A) = V(B)$.

Where A P B means the decision maker strictly prefers A to B.
A I B means the decision maker is indifferent between A and B.

Such a function is an **ordinal** value function; it gives information only about **the order of preferences**, no more. Such a function is unique up to strictly increasing transformations. That is if $V_1(\ )$ and $V_2(\ )$ are two ordinal value functions which both describe an individual's preferences then there exists a strictly increasing function $\varphi(\ )$, such that: $V_1(\ ) = \varphi\ (\ V_2(\ ))$.

A **measurable value function** goes beyond an ordinal value function in that it captures **strength of preference**. Measurable value functions are founded on value difference measurement. Suppose we know that A P B P C and that V( ) is a measurable value function, then if V(A) - V(B) > V(B) - V(C), this tells us that the difference in preference between A and B is greater than the difference in preference between B and C. An alternative way of interpreting this is that the increase in preference obtained by moving from B to A (exchanging B for A), is greater than that obtained by moving from C to B (exchanging C for B). Such functions are unique up to positive affine transformations, that is:

$$V_1(\ ) = \alpha + \beta\ V_2(\ ), \text{ where } \beta > 0.$$

To be able to determine a measurable value function which describes an individual's preferences it is essential that the individual understands and is able to make judgements about the relative value of such value differences, or exchanges. This notion is fundamental to the use of value functions in multi-criteria decision support.

12.2.2 Multi-attribute value theory

The discussion so far has considered preferences only in a holistic sense, assuming that the decision maker is taking implicit account of all relevant factors in making preference judgements regarding two options. If the context and the elements of choice are simple, for example deciding which chocolate bar to have with afternoon tea, then this may be quite possible. However, as we mentioned in the introduction, the options may be a complex combination of actions and future scenarios. In such a situation the decision maker may try to make sense of the problem by breaking it down into component parts. The field of MCDA is a body of approaches which seek to support decision makers by making explicit, and modelling, the multi-faceted nature of preferences. Thus **multi-attribute value theory**, which is one approach to MCDA, extends the concept of value measurement to the development of models which explicitly incorporate multiple factors.

An important stage in the development of a multi-attribute value function (MAVF) is clearly the process of determining the factors which are to be included in the model. This will be discussed in some detail later in the chapter, but for the moment we assume that n such factors have been identified and that option A is

described by the vector $(a_1, a_2, \ldots\ldots a_n)$, where a_i describes the performance of option A with respect to factor, or attribute, i.

The principle underlying MAVF modelling is one of dis-aggregation / synthesis, sometimes referred to as "divide and conquer". Instead of assessing preferences over the available choices on a holistic basis the decision maker's overall preference is synthesised from individual building blocks, where each building block describes preferences with respect to one of the fundamental factors which have been identified. That is, instead of directly assessing V(A), the decision maker first focuses on assessing $v_i(a_i)$, the partial value function describing preferences with respect to factor i, for all factors.

The simplest form of multi-attribute value function, and most widely used in practice, is the additive form:

$$V(A) = \sum_i v_i(a_i),$$

This is generally expressed as follows,

$$V(A) = \sum_i w_i \; v_i(a_i) \quad \text{where} \quad 0 \le v_i(a_i) \le 1.$$

$$\sum_I w_i = 1, \qquad w_i \ge 0.$$

The w_i are scaling factors which define acceptable trade-offs between different factors. These are values on a ratio scale and thus are unique up to multiplication by a constant. The value of w_i/w_j defines the relative value of one unit measured on factor i in comparison with one unit on factor j. Note that this means that the values of the w_i depend on the units used in the definition of the individual factors, as will be seen later.

This form of value function is appropriate if and only if the decision maker's preferences satisfy a condition known as **mutual preferential independence**. Given a set of attributes, X, then a subset Y of X, is preferentially independent of its complement, Z, if preferences relating to the attributes contained in the subset Y, do not depend on the level of attributes in the complementary set, Z. Mutual preference independence requires that every subset of attributes is preferentially independent of its complement.

This notion of preferential independence is an important one that merits further illustration by an example. A marine engineer has been offered two jobs and in considering her choice between these two jobs has identified salary, holiday entitlement, location and job satisfaction as important factors. The two jobs are outlined in table 12.1.

The numbers in brackets by location indicate an assessment of the attractiveness of the location to the decision maker on a 0 (low) to 10 (high) scale.

Table 12.1 Example to illustrate preferential independence

	Location	Salary	Holiday	Job Satisfaction
A	North Scotland (2)	£50k	20 days	Very high
B	Mediterranean (8)	£45k	22 days	Very high

Consideration of the conditions of preferential independence highlights that the extent to which the decision maker is willing to trade off an increase in the number of days holiday entitlement against a reduction in salary depends on the location. She would be willing to sacrifice more salary for a one day increase in holiday entitlement if the job is located in the Mediterranean than if it is located in the North of Scotland. This lack of preferential independence should not, however, lead to the immediate abandoning of an additive value function. It may be possible to redefine, or restructure the attributes in a way which achieves preferential independence. In this example, we might begin by probing a little more deeply into why location is important. From this it emerges that there are two reasons for its importance; firstly, proximity to Brussels where the engineer's family is based; secondly the climate and the extent to which this allows her to pursue her passion for windsurfing in her spare time. It is this latter aspect which relates to the length of holiday allowed. A possible way of overcoming the dependence between attributes may be to redefine location as distance from family and to redefine holiday as opportunity to windsurf (accepting that both jobs have a sufficient basic holiday allowance).

It is common to structure the factors of relevance to the decision in the form or a value tree or criteria hierarchy, as illustrated in Fig. 12.1. Partial value functions and scaling factors are assessed with respect to bottom-level criteria (those on the right of the tree below) and progressively aggregated to reflect preferences at the intermediate levels and overall.

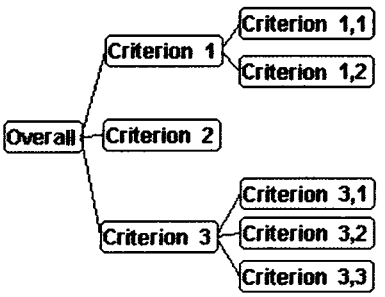


Figure 12.1 Illustrative Value Tree

This concludes the brief and informal introduction to value theory. More detailed and formal expositions, in the context of MCDA, can be found in French [32] and in Keeney and Raiffa [41].

The remainder of the chapter focuses on the use of MAVT to underpin decision support in practice, although many issues which will be discussed are equally relevant to the use of other approaches to MCDA.

12.3 USING MULTI-ATTRIBUTE VALUE THEORY

12.3.1 Example applications

The potential uses for MCDA in general, and for MAVT in particular, are wide ranging. Whenever a situation calls for a detailed evaluation of options or for a better understanding of preferences there is scope for the use of MAVT. It may be used, for example: by an individual to assess an aspect of their personal life, or to explore a personal perspective on an organisational issue; by a collaborative group to address an organisational issue: by a negotiative group seeking to identify a best compromise: or by a group of public policy makers seeking to identify a plan which best matches the needs and desires of society. The issue of concern may be a one-off choice between, or ranking of, a number of clearly identified options, for example, Butterworth [25] and Keeney [39] describe location decisions and Belton [8] discusses the selection of contractor and Buede [24] the design of a command and control system. On the other and it may be a recurrent decision about allocation of resources, or performance measurement, such as the case studies described by Islei et al [37] and Belton [11]. The reader is referred to special issues of the Journal of the Operational Research Society (April, 1982) and of Interfaces (November - December, 1991) for a range of applications.

12.3.2 The process of analysis

It is important to remember that the use of MAVT is part of the much wider MCDA process, which itself may be embedded in an even wider process of problem structuring and resolution, as illustrated in Fig. 12.2. It is not possible to appreciate the use of MAVT in practice without an awareness of this. Ever since its origins in the late 1960's concerns for the practical application of multi-attribute value theory (or, to be more precise, multi-attribute utility theory, MAUT) have influenced developments in the field. For example, concerns about difficulties of using the more complex MAUT models in practice led to the development of SMART [29,30], a simplified multi-attribute rating approach which now underpins much practical analysis. The field has benefited from the longstanding interests of psychologists, engineers, management scientists and mathematicians which has brought a continuing awareness of behavioural and social issues as well as underlying theory. For example, Decision Conferencing, described in more detail

later, recognises the importance of the social context in which analysis is conducted. In recent years these issues have become more widely embraced by the MCDM community as a whole as discussed Bouyssou et al. [20] and Korhonen & Wallenius [44].

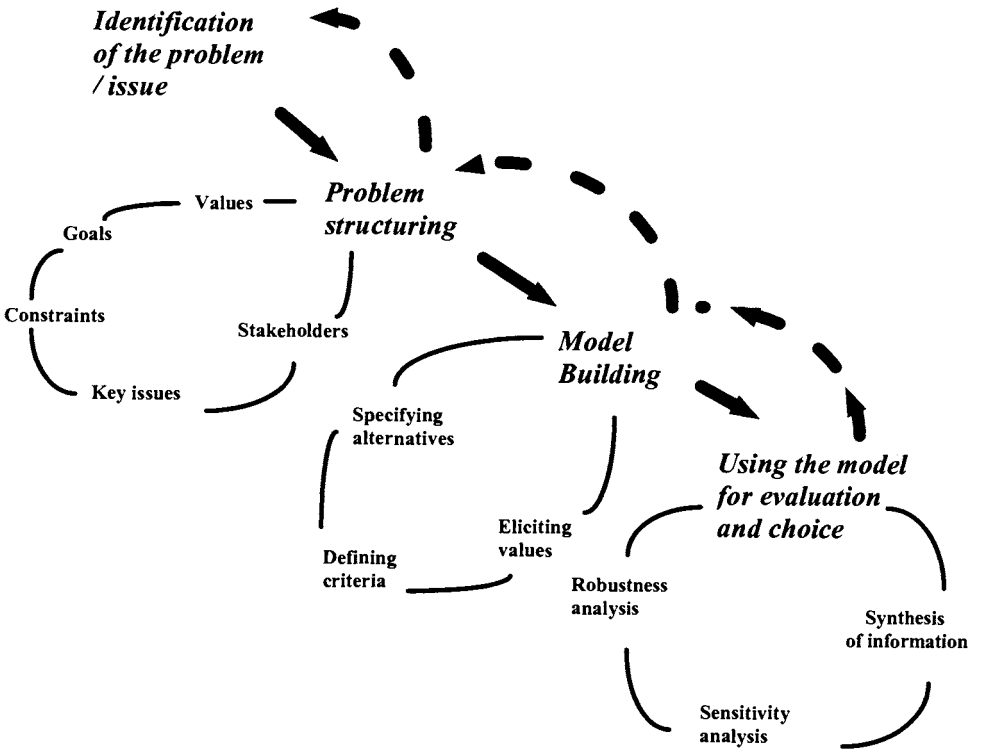


Figure 12-2 The process of MCDA

12.3.3 The nature of analyst / facilitator support

The majority of, if not all, multicriteria analyses reported in the literature are supported by one, or more, analysts or facilitators. The term analyst tends to be used when there is a strong emphasis on that person working independently to gather information and expertise; a facilitator is more commonly recognised as someone who also brings the skills of managing group processes. The analyst / facilitator may be an external or an in-house consultant, but in either case is recognised for their expertise in the approach to modelling.

This section explores in more detail the different analytical styles which may adopted. In section 12.4 we discuss the possibility of unsupported, or D.I.Y (do it yourself), analysis [16].

Buede [22] drew the distinction between the practice of decision analysis, including MCDA, as "engineering science" or "clinical art". From the engineering science perspective the role of decision analysis is to develop a model which, as far as possible, is an objective representation of reality based incorporating expert judgements together with the decision maker's values. Watson and Buede [62] comment that "an autonomous decision maker, receptive and willing to participate in an analytic process" would be best served by this approach, which they refer to as the "modelling" strategy for the conduct of a decision analysis. In contrast, the clinical art approach is characterised by the embedding of decision analysis in a facilitated social process which seeks to achieve a shared understanding of the issue and commitment to an agreed action plan. These models tend to be much less mathematically complex, focusing on providing a framework for explication of the decision makers values rather than a valid representation of an external reality. Watson and Buede refer to this as the "conferencing" strategy; it is epitomised by the Decision Conferencing approach pioneered by Cameron Peterson and colleagues which is described in detail by Phillips [52]. Of course these are not mutually exclusive approaches; a particular intervention may incorporate elements of both modelling and conferencing. Either approach can be used to support a one-off decision or to develop a decision support system to support recurrent decision making / monitoring.

Both approaches to MCDA continue to be practised to good effect. The RODOS project [33] and examples given by Von Winterfeldt [64] are recent and ongoing examples of the modelling approach. Quaddus et al. [54] and Belton et al. [13] describe applications in a decision conferencing environment.

The conferencing strategy has a number of parallels with the SODA (Strategic Options Development and Analysis) approach developed by Eden [27]. Potential synergies between these (suggested by Watson and Buede [62], Belton [9], and Ackermann and Belton, [2] are now beginning to be exploited (Belton, Ackermann and Shepherd [13]) and will be discussed later in the chapter.

12.4 AN EXAMPLE TO ILLUSTRATE THE USE OF MAVT

12.4.1 The problem

It is impossible to give a good sense of what is involved in building a multi-attribute value model without an illustrative example. However, it must be remembered that what follows is just one way of working through the process of building and using a multi-attribute value model, supported by a particular software tool. There are many possible variations on the way of working and a number of different tools available; it is impossible to describe them all here. To illustrate the process I have chosen an issue which is not too complex, but it is a real decision which is faced by the

International Society for MCDM every other year. The question is; where should the next Conference be? Imagine that a group of interested people have come together to resolve this issue.

The problem is already well defined; it is accepted that there will be a Conference and the problem is to decide on the location. Thus there is no need, in this case, to begin with a broad problem structuring process and we can move directly to model building. The aim of the process is to develop a value tree which captures the decision makers values relating to the issue. The process thus encompasses both elicitation and structuring of values. The extent to which the two stages of the process are distinguished and made explicit will depend on the analyst.

12.4.2 Problem structuring and model building

It is well recognised by practising analysts that "good problem structuring is the key to successful analysis" (Von Winterfeldt and Edwards [63]). However, building a value tree is an art which is informed by experience; there is no "right" tree waiting to be constructed. It may be possible to represent an issue by a number of different value trees which differ in structure and/or level of detail (see, for example, Keeney and Raiffa [41] (p 422), Brownlow and Watson [21]). Useful guidance on building value trees can be found in Keeney and Raiffa [41], Buede [23], Watson and Buede [62] and in Von Winterfeldt and Edwards [63]. Keeney's book, "Value-focused Thinking" [40] also has much to contribute on the subject.

Keeney [40] distinguishes between value-focused thinking and alternative-focused thinking. In the former, the process focuses on eliciting the decision makers values prior to identifying alternatives. In the latter, alternatives are identified at an early stage in the decision process and the focus is on distinguishing and choosing between the alternatives. As such, value-focused thinking is closer to problem structuring as described earlier. Keeney suggests that decision analysis should be driven by value-focused thinking as alternative focused thinking tends to " ..anchor the thought process, stifling creativity and innovation" (p48). However, in practice it is often the case that decision makers are faced with a situation in which well defined alternatives already exist and in such circumstances the alternatives can provide a useful stimulus for thinking about values (as Keeney himself describes - p57).

Von Winterfeldt and Edwards [63] and Buede [23] both describe two distinct approaches to the structuring of value trees, the top-down and bottom-up approaches. The top-down approach tends to be objective led, whilst the bottom-up approach is alternative led. It would be easy to equate the top-down approach with value-focused thinking and the bottom-up approach with alternative focused thinking. However, I think this is too simplistic - the alternative-led approach can be value-focused. It is better to view the two approaches as complementary ways of helping the decision makers think about the situation and to determine relevant values. It may be useful to build two separate value trees for a problem, one using

the bottom up approach, the other using top-down. Each approach may yield different insights, which can then be combined in an aggregated tree.

The desirable characteristics of a value tree are identified by Keeney and Raiffa [41] as:

- Complete - all important aspects of the problem are captured
- Operational - it can be used with reasonable effort
- Decomposable - allowing different parts of the tree to be analysed separately
- Nonredundant - to avoid double counting of impacts
- Minimal or concise - keeping the level of detail to the minimum required

In "Value-focused Thinking" [40] Keeney adds to this list in defining desirable properties for attributes which should be shared by fundamental objectives:

- Measurable - possible to specify in a precise way the degree to which objectives are achieved through the association of appropriate attributes
- Understandable - to facilitate generation and communication of insights

Clearly there is tension between the desire for completeness and conciseness - exactly how much detail should be included? The notion of a requisite model introduced by Phillips [51] nicely captures the ideal balance. A requisite model is one which has generated adequate insight into the problem, one which captures the shared understanding of a decision making group. However, it is only through using the model that the extent to which it possesses these characteristics will emerge.

We begin with an initial idea generation phase focusing on the question "What factors should be taken into account in deciding a conference venue?". Although we have not yet specified options for consideration in our example, it is likely that the decision makers will have in mind previous and potential conference venues. There are a number of group processes which might be adopted to facilitate this stage [49]. A nominal group approach which requires participants to work individually at first, thus capturing individual perspectives, and then to pool ideas with other members of the group, works well in practice. The use of post-its [1], [12], greatly facilitates this way of working. Individuals are asked to write down their ideas on the post-its (one idea per post-it), which are then collected on the wall (covered with flip-chart paper if necessary) for everyone to see. As participants review the ideas contributed by others they are encouraged to add to and expand on these. The facilitator, aided by the participants, should attempt to cluster similar ideas.

Fig. 12.3 shows the output of the initial brainstorming session. The ideas have been clustered and each cluster given a label which captures the theme which links ideas. Note that many of the ideas capture similar or overlapping concepts (for example, "don't go back to the same place" and "not been there before"), which is to be expected if several people are involved in generating the ideas. Keeping hold

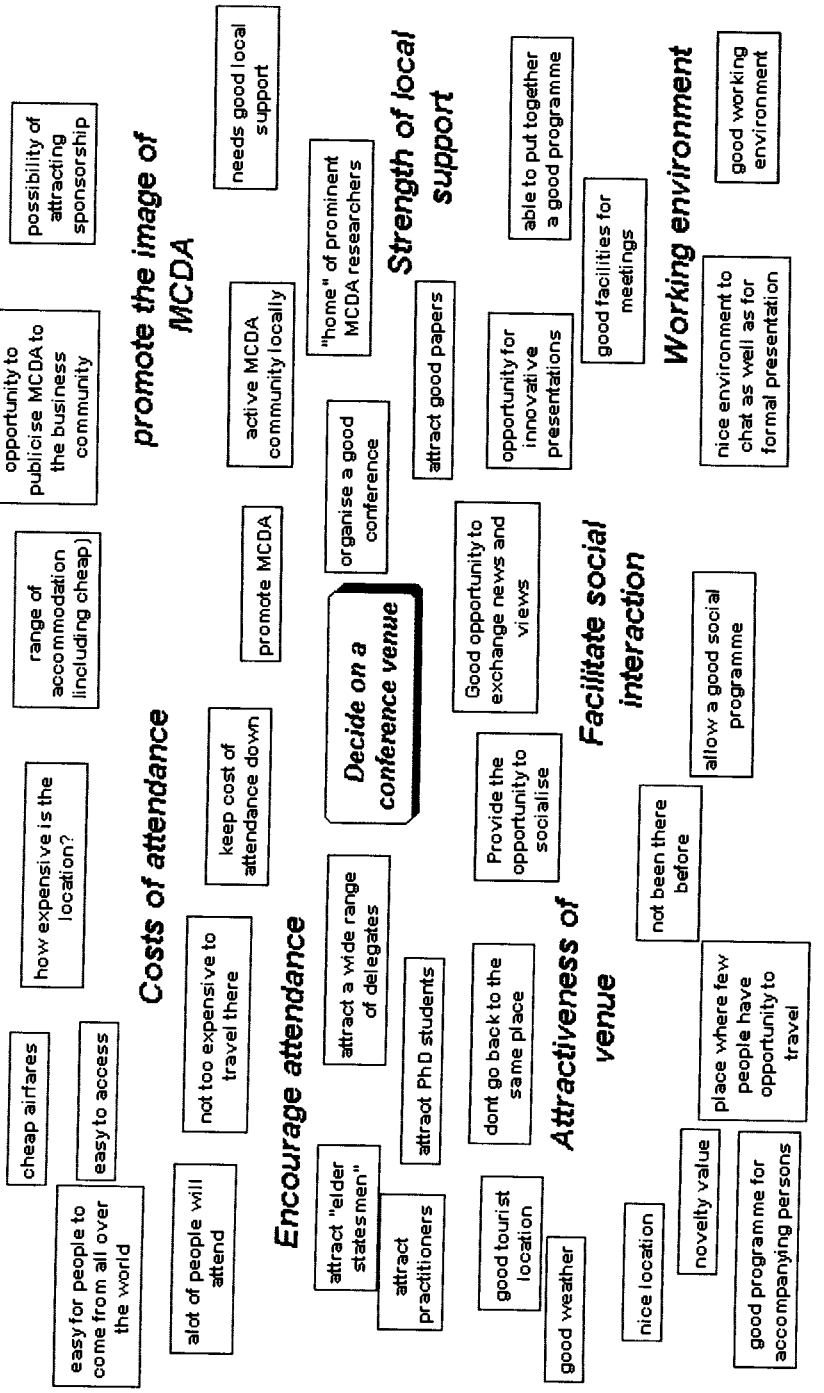


Figure 12.3 Post-It Session - Conference Venue problem

of all of these ideas at this stage increases the richness of the representation and increases ownership by maintaining the participants' own language.

The next stage is of analysis is to build a value tree which captures the issues which have emerged from the idea generation process. It is important to remember, as discussed above, that the value tree should reflect the decision makers' values, or objectives, rather than simply being a means of discriminating between options. In our example the post-its have been arranged so that the higher level objectives are positioned closer to the central question and these form the top level of criteria in the initial value tree, which is displayed in Fig. 12.4.

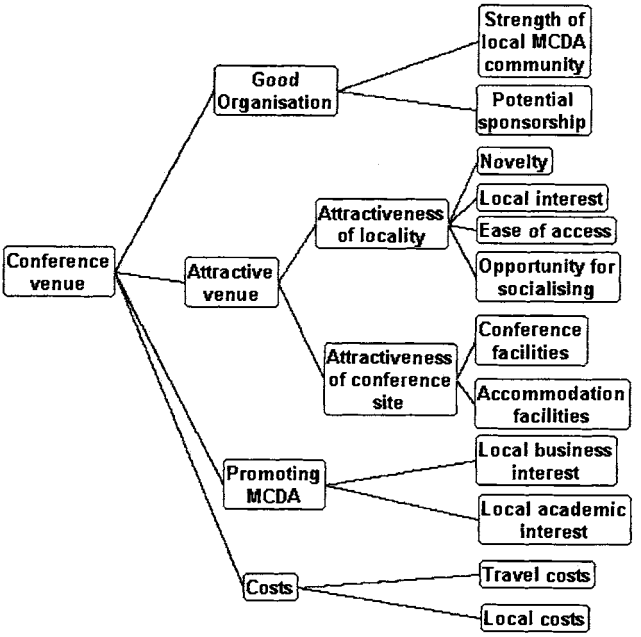


Figure 12.4 - Value tree for Conference Venue Problem

In practice it is unlikely that the initial value tree will "survive" the whole analysis. As the analysis proceeds it may emerge that certain criteria do not satisfy the preferential independence requirement and restructuring is required. Certain aspects may have been elaborated in too much detail, others in not enough detail. As indicated in Fig. 12.2 the whole process is an iterative one.

12.4.3 Eliciting values

12.4.3.1 Partial value functions. Once the value tree is structured and alternatives to be evaluated have been determined, the next stage is to assess the performance of each of the alternatives against those criteria which define the

bottom level of the tree. This is the process of determining the values $v_i(\ )$, the partial value functions, as indicated in section 12.2.2. Remember that the partial value functions define an interval scale of measurement, that is, a scale which focuses on the difference between points (the ratio of values has no meaning). To construct a scale it is necessary to define two reference points; these are often taken to be the bottom and top of the scale and assigned values of 0 and 100 respectively. For a local scale these points are defined by reference to the alternatives under consideration, assigning a score of 100 to the alternative which performs best on a particular criterion and 0 to the one which does least well. A global scale is defined by reference to the wider set of possibilities. The end points of a global scale are defined by the ideal and the worst conceivable performance on the particular criterion, or by the best and worst performance which could realistically occur. The definition of a global scale requires more work than a local scale. However, it has the advantages that it is more general and that it can be defined before consideration of specific alternatives.

Once the reference points have been determined consideration must be given to how intermediate scores are to be assessed. This can be done in the following three ways:

- *Definition of a value function*: relating value to performance against a measurable attribute reflecting the criterion of interest
- *Construction of a qualitative value scale*: against which the performance of alternatives can be assessed by reference to descriptive pointers
- *Direct rating of the alternatives*: no attempt is made to define a scale which characterises performance independently of the alternatives being evaluated

12.4.3.2 Defining a value function. The first step in defining a value function is to identify a measurable attribute scale which is closely related to the decision makers values. If this is not possible then it will be necessary to construct a value scale or directly assess the performance of alternatives, as described below. The value function reflects the decision makers preferences for different levels of achievement on the measurable scale. There are many ways of assessing value functions and the reader is referred to Keeney and Raiffa [41] or Watson and Buede [62] for a detailed exposition of some of the methods. A function may assessed directly, usually utilising a visual representation or by indirect questioning, utilising, for example, the bisection method.

Example 12.1

To illustrate the definition of a value function consider the criterion "strength of local MCDA community". This could be measured simply by the number of local persons active in the MCDA community, or we could attempt also to take account of seniority and experience. However, for purposes of illustration let us keep things

simple and measure strength by the number of members of the International Society. The minimum size is 1, as it would be considered unrealistic to organise a conference where there is no local representation. The maximum is less clearly defined, but 50 might be possible if there were a number of active research groups in the same city.

Fig. 12.5 shows a possible value function for the size of the local MCDA community. A value of 0 corresponds to a local community of 1 person, a value of 100 to a community of 50 people. As the graph illustrates, an increase from 1 to 5 people generates a significant increase in value and a further increase from 5 to 10 increases the assessment of value to a level of 80.

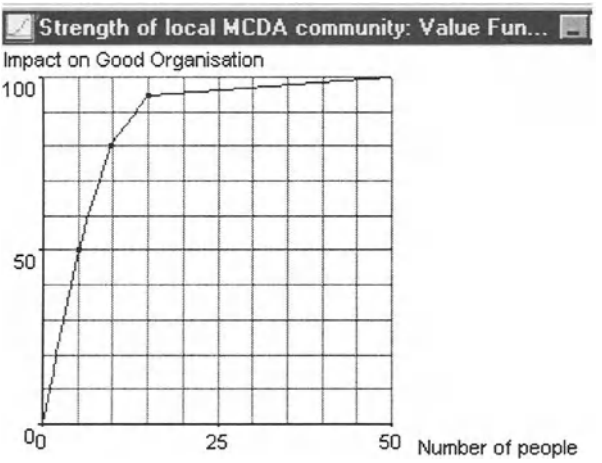


Figure 12.5 Partial value function for "Strength of local MCDA community"

Von Winterfeldt and Edwards [63] argue that if a problem has been well structured all value functions should be regular in form (i.e. no discontinuities) and further, that they should be linear or close to linear. However, Stewart [62] illustrates through experimental simulations that the results of analyses can be sensitive to such assumptions, hence care should be taken not to over-simplify a problem by inappropriate use of linear value functions. Non-monotonic value functions are often indicative of multiple underlying values and the analyst should consider restructuring a value tree to replace such by one or more monotonic functions.

12.4.3.3 Constructing a Qualitative Value Scale. If it is not possible to find a measurable attribute which captures a criterion an alternative approach is to build an appropriate qualitative scale. Examples of such scales in regular use are the well known Beaufort scale for measuring the strength of wind and the Richter scale for measuring the force of earthquakes. Points on these two scales are defined descriptively; an alternative approach is to associate specific alternatives, with which the decision makers are familiar, with points on the scale. The descriptors,

which correspond to a measurable attribute, must then be mapped onto a scale defining the decision maker's values. In constructing the scale the analyst and decision makers may start from the scale and develop descriptors, or word models, associated with specified value levels. Alternatively, the starting point may be a set of definitions, familiar to the decision makers, which are mapped on to points on the value scale. This may be done directly, or using a pairwise comparison procedure such as MACBETH [4]. A constructed scale should have the following attributes:

- Operational: allows decision makers to rate alternatives not used in the definition of the scale.
- Reliable: two independent ratings of an alternative should lead to the same evaluation.
- Value relevant: relates to the decision makers' objectives.
- Justifiable: an independent observer could be convinced that the scale is reasonable.

A constructed scale for accommodation facilities could be as detailed in Table 12.2.

Table 12.2 A constructed scale for the attribute "accommodation facilities"

Scale point	*Value*	*Descriptor*
Poor	0	No choice, poor quality student accommodation located at a distance from the conference site
	25	A limited choice (e.g. basic student residence or high quality hotel) but not conveniently located for the conference site or city
Middling	50	A range of types and quality of accommodation, but scattered around the city and somewhat distant from the conference site
	75	A wide range of types and standard of accommodation, but scattered around the city and not all conveniently located
Excellent	100	Wide range of types and standards of accommodation (basic student residences, executive student residences, budget hotels, high quality hotels) located close together within easy reach of the conference site and social amenities (restaurants, bars, sport)

12.4.3.4 Direct Rating of Alternatives. Direct rating can be viewed as the construction of a value scale, but defining only the two reference points. A local or a global scale, as described above, may be used, the former creating minimal work

for the decision makers. Alternatives are positioned directly on the scale to reflect their performance relative to the reference points. Even though no attempt is made to relate performance to a measurable scale, the positioning of alternatives can generate extensive discussion, yielding rich information on the decision makers values which should, ideally, be captured for future reference.

A process of pairwise comparisons is implicit in the direct rating approach - for example, in evaluating a particular alternative it is necessary to consider whether it should be positioned above, or below other alternatives. This process can be formalised in a manner which requires the decision makers to systematically consider each pair of alternatives. The pairwise comparison approach is one of the cornerstones of the Analytic Hierarchy Process (AHP) developed by Saaty [59]. However, the AHP treats responses as ratio judgements of preferences, which is not consistent with the value function approach. The underlying mathematics is easily modifiable to be consistent with difference measurement [31]. The MACBETH approach developed by Bana e Costa and Vansnick [3],[4], also based on pairwise comparisons, can be used to derive direct ratings.

One of the drawbacks of pairwise comparison methods is the large number of judgements required of the decision maker (N*(N-1)/2) for each criterion, where N is the number of alternatives. Nevertheless, the approach is a powerful one which can be effectively utilised if decision makers find the direct rating procedure difficult.

A scaling procedure must be defined for each bottom-level criterion. This is then employed to rate each of the alternative locations under consideration. It is beyond the scope of this chapter to consider each of the criteria in detail and to use the model to assess a number of real venues. In order to illustrate the rest of the process four hypothetical venues, based on the plausible scenarios outlined below, have been assessed.

- *Option A:* A North American industrial city, with a good university and strong MCDA support.
- *Option B:* An Australasian University city with good conference facilities and an active, but not large MCDA community
- *Option C:* An attractive and historic European city. One MCDA person locally working in a University with few conference facilities
- *Option D:* An exotic island in a remote location. A small local MCDA group. Good accommodation options and local facilities but poor conference facilities.

Fig. 12.6 shows the assessment of each of the alternatives against each of the bottom-level criteria in five profile graphs. It is clear from these graphs that all four locations have strengths and weaknesses. There is no dominating location, neither is any of the locations dominated by another.

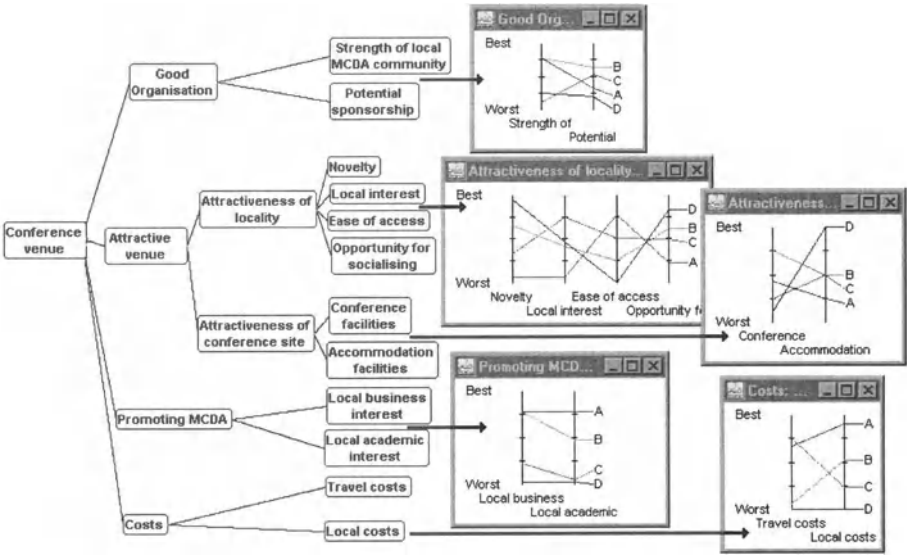


Figure 12.6 Assessment of alternatives against bottom-level criteria

We now consider the elicitation of scaling factors, or criteria weights. As these values are dependent on the scales used to assess the partial value functions, it is not possible to assess the weights until the partial value function scales have been defined.

12.4.3.5 Scaling factors. As we have seen already, the scaling factors, w_i, define acceptable trade-offs to the decision maker. How much would they be willing to give up on one factor in order to improve performance on another factor? There are many questioning procedures which seek to elicit this information. One which is widely used is the swing weights procedure.

The "swing" being considered is a swing from the worst value to the best value on each criterion. If the value tree is small, then the decision maker may be asked to consider all bottom-level criteria simultaneously and to assess which swing gives the greatest increase in overall value; this criterion will have the highest weight. The process is repeated on the remaining set of criteria, and so on, until the order of benefit resulting from a swing from worst to best on each criterion has been determined, thereby defining a ranking of the criteria weights. To assign values to the weights the decision maker must assess the relative value of the swings. For example, if a swing from worst to best on the most highly weighted criterion is assigned a value of 100, what is the relative value of a swing from worst to best on

the second ranked criterion? Note that it is not possible to assign swing weights until the scales for each criterion have been defined.

Fig. 12.7 shows the relative magnitude of swing weights assigned to the bottom-level criteria in the example value tree. The strength of the local MCDA community is judged to have the highest weight and other criteria are assessed relative to that.

The assessment of weights is also implicitly a process of pairwise comparison. This may be formalised by specifying a reference criterion against which all others are compared (requiring the minimal number of comparisons), or each criterion may be compared with every other one (requiring N (N-1)/2 comparisons) as in the AHP. For larger value trees it is easier first to consider families of criteria (i.e. those sharing the same parent criterion) and to compare across families.

Decision makers are generally very comfortable working with visual analogue scales such as the ones displayed in Fig. 12.6 and Fig. 12.7. These provide a means for communicating a good sense of the magnitude of judgements whilst removing the need for numerical precision. However, it is important that this degree of imprecision is not forgotten when information is aggregated.

The weights implied by the visual representation in Fig. 12.7 are translated to numerical values, normalised to sum to 1, as presented in Fig. 12.8. The figures in brackets are the "within family" weights normalised to sum to 1.

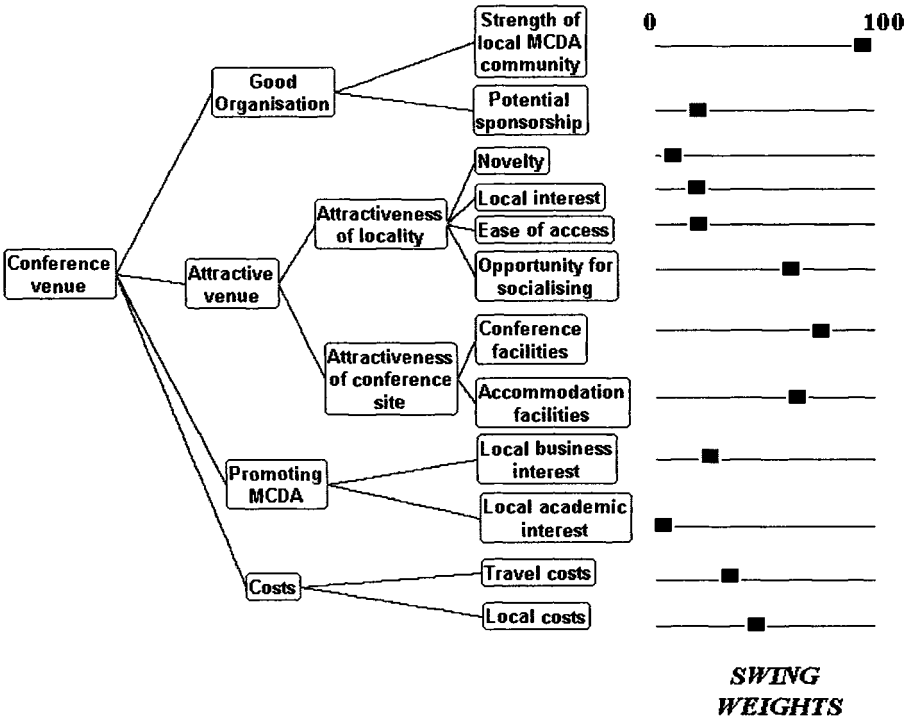


Figure 12.7 Swing weights for the Conference Venue value tree

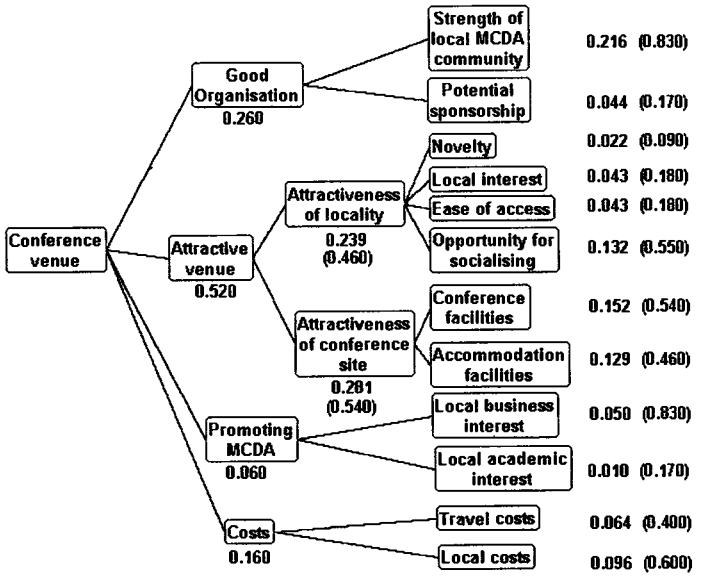


Figure 12.8 Numerical weights for the Conference Venue value tree

12.4.4 Using the information elicited

The elicited information can now be synthesised to give an overall value, V(A), for each of the alternatives, together with evaluations at intermediate levels of the value tree. Fig. 12.9 shows, on the right of the picture, the aggregation of values to the level of "Attractive venue" and on the left, the aggregation of values to the top of the value tree. It can be seen that, in terms of attractiveness of venue, locations B and D are preferred, followed by C and then A. Overall, B is slightly preferred to A with C and D rated some way behind. The overall profile graph emphasises the preference for A or B in most regards.

However, this should not be regarded as the end of the analysis. It may be that the result of this initial analysis is in conflict with the decision makers intuition. If this is the case, neither the model nor the intuitive preference should be discarded. Instead, the model should be further explored and intuition questioned with a view to better understanding the conflict. Has an important factor been omitted from the model? Should the specification of criteria weights be revisited? Or is the decision maker persuaded to change their view?

Even if there is no conflict between the model output and the decision makers intuition the robustness of the preferred option should be explored through sensitivity analysis. It is generally recognised (Rios Insua and French [57], Belton [9]) that sensitivity analysis is a very important part of any analysis. It enables the users to check out the implications of imprecise or uncertain judgements and to explore the implications of differing priorities. Furthermore, if the sensitivity

analysis is carried out interactively it serves as a vehicle for the decision makers to validate the underlying model, thereby increasing confidence and ownership.

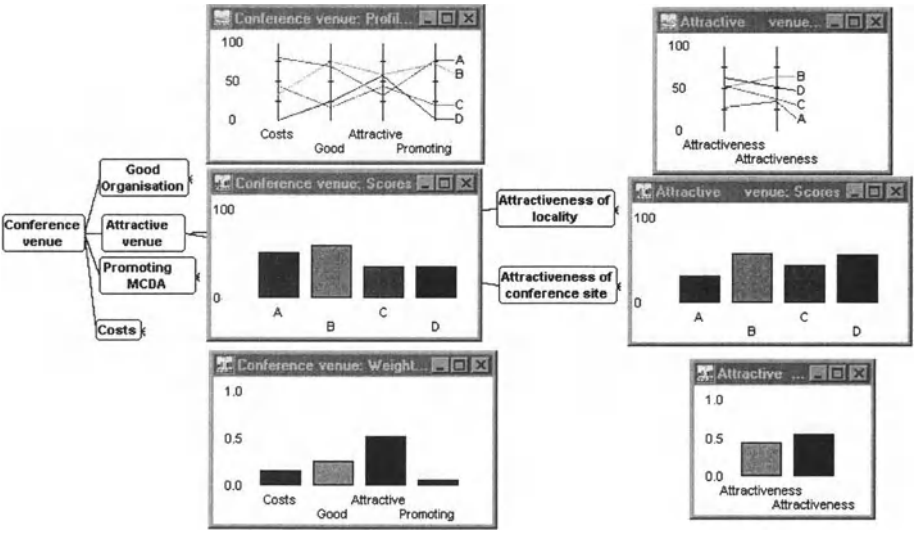


Figure 12.9 Aggregating the elicited information

Experience of the author in the use of MAVT in practice [8] prompted research into ways of quickly and effectively communicating the output of multicriteria sensitivity analyses. This led the author to develop V•I•S•A, a multi-criteria decision support system which supports visual interactive sensitivity analysis [19]. Motivated by developments in Visual Interactive Modelling [7] and inspired by Korhonen's implementation of visual interactive goal programming, the Pareto Race [43], V•I•S•A allows the user to explore the robustness of options through easy-to-conduct interactive sensitivity analysis. By dragging bars representing important parameters of the model, for example the criteria weights, the user can immediately see the implication of that change for, for example, the overall evaluation of the options. For example, increasing the weight given to the need for good organisational support and low costs would shift the balance of preference from location B to A in the above example, as seen in Fig. 12.10.

V•I•S•A allows the user (facilitator, analyst or decision maker) to carry out multi-dimensional sensitivity analysis in an exploratory manner - making ad-hoc changes to parameters of the model and immediately seeing the effect on the evaluation of alternatives. This is a valuable tool, but as it is completely user-driven it does not guarantee a systematic and comprehensive investigation: current work is exploring the provision of more systematic support for sensitivity analysis.

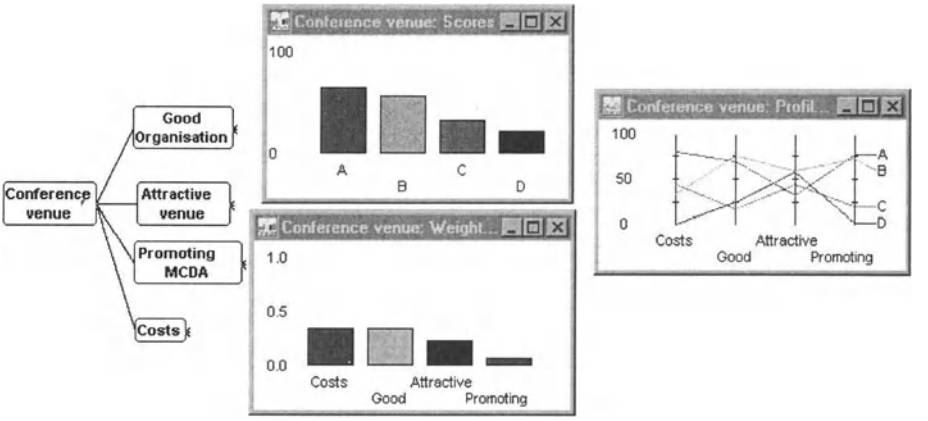


Figure 12.10 Visual interactive sensitivity analysis

12.5 WHAT'S NEW?

The theory supporting MAVF analysis has been in place since the late 1960's [55] and many accounts of the application of MAVT in practical decision making began to emerge in the 1970's, so you may well ask what place this chapter has in a book on advances in multiple criteria decision making. Although the underlying theory is substantially unchanged, the way in which it has been applied over the past 30 years has seen some changes. Many of these have been facilitated by rapidly developing technology, others have been driven by awareness of the importance of and increased attention to problem structuring, and some are a consequence of organisational factors.

12.5.1 Problem structuring for MCDA

As commented earlier, good problem structuring is the key to effective analysis. The shift in emphasis from modelling to problem structuring is reflected in the comparison between Keeney and Raiffa's, "Decisions with Multiple Objectives", published in 1976 [41], and "Value-focused Thinking" published by Keeney in 1992 [40]. In this respect the MCDA field can learn much from developments in so-called "soft operational research", or problem structuring methods, which has gained prominence in the UK [58]. The conferencing strategy has a number of parallels with the SODA (Strategic Options Development and Analysis) approach developed

by Eden [27] and potential synergies between these are now beginning to be exploited (Belton, Ackermann and Shepherd [13], Bana e Costa et al [5],[6]). In common with decision conferencing SODA is a process which uses modelling as part of a managed social process to achieve shared understanding. Cognitive mapping [26], which is a particular form of representing individual 'construct systems' based upon Kelly's Personal Construct Theory [42], is the underlying modelling approach. A cognitive map aims to represent a problem/issue as the decision maker (participant) perceives it, in the form of a means-ends network-like structure. Individual cognitive maps may be generated in one-to-one interviews and later merged to form a composite group map capturing all individuals perspectives. Alternatively a group map may be generated directly in a structured idea generation session. The shared map is used to facilitate conversation and discussion, leading to a shared understanding of the issue. The process is supported by the use of the Decision Explorer software (formerly known as COPE).

Ackermann and Belton [2] discuss the use of COPE and V•I•S•A for managing corporate knowledge, highlighting the parallels between the SODA and MCDA processes, and in Belton, Ackermann and Shepherd [13] they describe the integration of these two approaches to explore the development and implementation of strategy for the Supplies and Commercial Services Division of a hospital trust.

The embedding of MCDA within the broader framework for problem structuring and resolution provides a powerful vehicle for integrated decision support. The use of mapping allows the capture of a rich representation of the issue, which not only supports the decision making process but also provides for a longer term organisational memory. However, a multi-methodological approach raises a number of issues. Firstly, there are philosophical issues about the theoretical compatibility of approaches. Secondly, there are practical issues about the skills required of the analyst/facilitator. These are addressed by Belton, Ackermann and Shepherd [13] in the context of that particular intervention; and many of the issues are addressed in a general sense in the compilation of papers entitle "Multi-methodology" edited by Mingers and Gill [48].

12.5.2 Technology

Technological factors have influenced the way in which support is provided and the nature of support.

12.5.2.1 The way in which support is provided. The advent of mini-computers in the 1970's followed rapidly by the microcomputer in the 1980's enabled analysis to move out of the backroom and into the boardroom [10],[56]. This had a substantial impact on the practice of MCDA as evidenced by the introduction of the decision conferencing approach at DDE [45] in Washington, USA, followed by the development of the PODs at ICL and at London School of Economics in the UK [52]. The POD is a customised room, designed to provide the appropriate environment for a 2 day Decision Conference. The design of the room pays explicit

attention to the need to manage group dynamics as well as the need to provide appropriate, low key, technological support. With the introduction of portable microcomputers and display facilities an increased flexibility in working became possible - but it seemed that the more technology could provide, the more decision making groups demanded in terms of feedback which could be instantaneously conveyed to groups of decision makers, as described by Belton [8]. Thus, the way in which support could be provided prompted developments in the nature of the support. As described above, V•I•S•A was developed as a consequence of these experiences. Early work with V•I•S•A indicated that decision makers responded positively to the visual interactive implementation of the model when encouraged to "play", resulting in a better understanding of and increased confidence in the model [11].

12.5.2.2 Computer networked support. The advent of local area networks created new opportunities for group working. Pioneering systems such as Group Systems developed by Nunamaker et al. [50] focused on the provision of tools to enhance the creativity and productivity of groups in a workshop setting. Although the Group Systems software incorporates facilities for multicriteria evaluation, the procedures appear to have been developed in an ad hoc way rather than building on expertise in multicriteria modelling. The use of such systems was limited initially by the availability of suitable facilities - a laboratory with the requisite number of networked computers. The development of notebook computers and networking capabilities allowed a more flexible way of working - firstly, a customised laboratory was no longer required and secondly, the facilities were more portable. In very recent years we have seen the development of specialised group systems for multicriteria decision support (Group V•I•S•A and HIPRE are examples of two which are based on MAVT; Group Expert Choice, based on the AHP, and Group Promethee are two other examples). These systems allow decision makers simultaneously to input their individual values to a shared model; the information is then synthesised and aggregated to give a "group decision" or used as a basis for discussion aimed at achieving a consensus. Fig. 12.11 is a screenshot from Group V•I•S•A showing three "group views" in which each user is represented by a coloured dot (it may not be appropriate to distinguish users in this way, in which case all the dots are the same colour). The top left chart shows the overall evaluations of the four locations by each user; although A and B are preferred by all users, there is not a unanimous preference for one of these two. Looking at the bottom left chart showing the distribution of criteria weights reveals that the yellow user is less concerned about costs. The chart on the right looks in more detail at the evaluation of location A.

However, whilst group multicriteria decision support systems open up opportunities for new ways of working, this also generates many questions (see Belton and Pictet [17], Belton and Elder [14], Hämäläinen and Pöyhönen [36]). What is the best way of running such workshops - for example, how can manual, single-user and multi-user decision support technologies be best integrated? What aspects of a model should be common to all participants in a workshop and what

aspects are individual? How does the facilitator make sense of the mass of information generated and best use this to manage the group process? These questions will only be answered by action research to investigate the use of this new technology in practice.

The rapid development of the internet and worldwide web is pushing these developments even further. It is now possible for groups to work together to build and use multicriteria models over wide area networks, permitting dispersed and asynchronous MCDSS (see, for example, Web-HIPRE). This way of working opens up new issues of facilitation and user support which are the subject of further current research (Belton and Hodgkin,[16]).

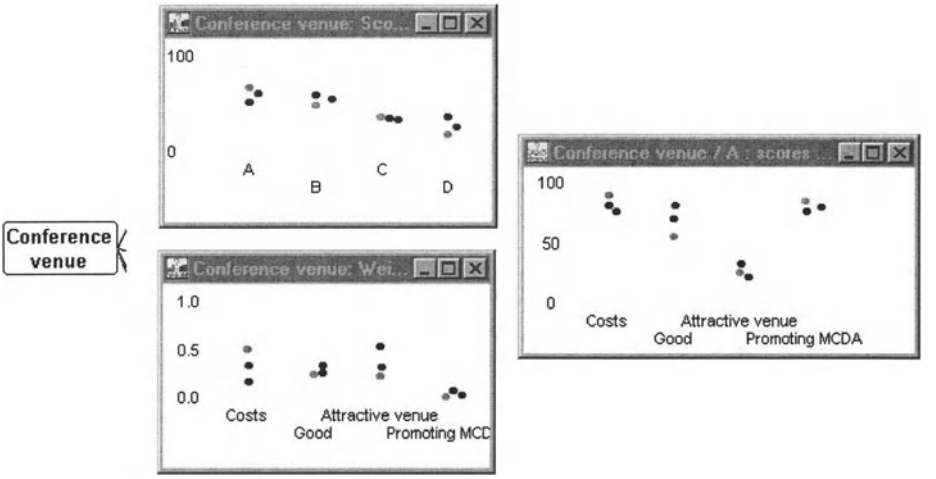


Figure 12.11 Group views

12.5.2.3 Nature of support. The widespread introduction of Windows technology from the early 1990's greatly facilitated the development and implementation of visual interactive interfaces and there are now a number of software tools incorporating such features which facilitate multi-attribute value analysis (for example, HiView, Logical Decisions, V•I•S•A).

Technological developments continue apace, continually providing new opportunities and challenges for the MCDA practitioner and researcher, for example:

Multimedia. The integration of materials which not so long ago were conceived of as distinct media, in a hyper-media environment is now common place. Graphics, animations, video, sound, text, interactive computer programmes, etc., are incorporated in multimedia teaching materials, multimedia advertising, multimedia

publishing, ... [34]. In what way can multimedia enhance multicriteria analysis? A few suggestions follow.

As discussed earlier, the process of problem structuring is a divergent one generating a wealth of information about the issue being investigated and about the values of the decision makers and other stakeholders. Much of this richness is set aside, if not lost completely, in the model building phase. A value tree is a very sparse representation of a problem; complex objectives and detailed alternatives or strategies may be captured by one word descriptors. We have already discussed the use of cognitive maps to act as a more detailed organisational memory of an analysis; multimedia can also play a role here. If the options are physical entities, then a visual record can provide a richer representation than a written descriptor. This may be a simple photograph or a video clip, or it may be a virtual reality representation of, for example, potential building designs. Fig. 12.12 shows a V•I•S•A model for the exploration of alternative sites for new office premises focusing on the differences between two possible sites.

If a model is to be used as part of an organisational memory, or to demonstrate and justify the basis for a decision to other parties, then audio could be used to capture and record the rationale for particular judgements.

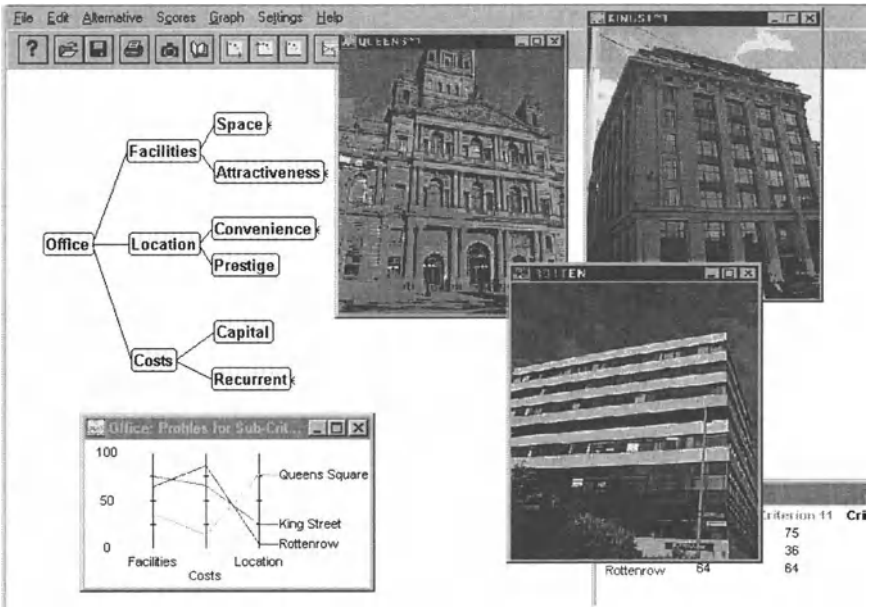


Figure 12.12 Multimedia in MCDA

Linking to other analytical tools. The ability to easily and dynamically link software tools opens up exciting opportunities for the integration of multicriteria analysis with other analytical approaches. It has been suggested elsewhere [53] that

if the field of MCDA is to achieve its full potential, then it must seek ways of achieving such integration - as the "classical" scenario of MCDA more often arises out of other investigations, than as a distinctly identifiable problem.

One such area of activity is the integration of Geographical Information Systems and MCDA as evidenced by the GIMDA (Geographic Information and Multicriteria Decision Analysis) group. Applications in the fields of land use planning, transport planning and environmental management particularly lend themselves to the integration of these two approaches.

Another potentially fruitful area for collaborative work is in the integration of simulation and MCDA. A number of surveys have indicated [47] that simulation is the most widely applied OR technique; it is used extensively in manufacturing and increasingly so to investigate operations in the public sector. Although such models invariably gather and report performance of the simulated system on multiple dimensions, simulation software does not generally extend its analytical capabilities to incorporate formal multicriteria analysis. This gap was bridged by Belton and Elder [15] who linked Simul8, a visual interactive system for discrete simulation, with V•I•S•A. The link is configured so that performance measures from a run of the simulation, defining a new option in the choice set, are automatically passed across to the multicriteria model where they are incorporated with other relevant information. Other examples of the integration of simulation are: Spengler and Penkuhn [60] describing a DSS which combines a flowsheet-based simulation with multicriteria analysis: work by Macharis [46] which incorporates a multicriteria choice rule within a system dynamics analysis of transport policy: and a study by Gravel et al [35] using multicriteria analysis to evaluate production plans developed using simulation.

12.5.3 Organisational issues

12.5.3.1 Nature of current applications. Applications of MAVT, at least those which are reported in the literature, tend to be focused in the public sector or large companies. Possible factors to explain this are:

- Awareness and access to expertise - large organisations are likely to have in-house consultants with the appropriate knowledge and skills
- Access to appropriate resources / facilities (decision support technology and facilitators)
- Need to be accountable to the public
- Need to take explicit account of multiple stakeholder views
- Enormity of the decisions merits detailed evaluation

Is there any evidence that MAVT, or MCDA in general, is becoming more widely used? It is interesting to note developments in two directions, which will be referred to as "D.I.Y." (do-it-yourself) analysis and "packaged" analysis.

12.5.3.2 "D.I.Y" MCDA. There is little to write about D.I.Y MCDA in practice as there is little evidence that any formal analysis takes place in this way. However, I believe that it is an area which deserves attention, for two main reasons. The first reason is a pool of middle and senior managers which is becoming increasingly educated in formal methods of management through MBA programmes. These managers are more technologically aware and have had greater exposure to modelling than their predecessors, two factors which should make them more receptive to the use of multicriteria analysis. This predisposition may make them more inclined to use MCDA without the support of a trained analyst/facilitator.

However, the notion of D.I.Y MCDA opens up new and interesting challenges, firstly in the education of potential users and secondly in the development of appropriate softwares. Edwards, Finlay and Wilson [28] discuss the general benefits of and difficulties in providing D.I.Y decision support and Belton and Hodgkin [16] focus on MCDA specific issues in an evaluation of the need and potential for intelligent decision support for D.I.Y users.

12.5.3.3 Packaged MCDA. A number of currently popular management methodologies, for example, Kaplan and Norton's Balanced Scorecard [38] and the EFQM (European Federation for Quality Management) model are founded on ad-hoc multicriteria analyses. In addition to these popular methodologies there are many "institutionalised" decision processes which require analysis of multiple factors. For example, the "Option Appraisal" process in the UK Health Service requires that before any capital project can be approved a full analysis of the costs and benefits of alternative proposals is conducted. COEIA (Combined Operational Effectiveness and Investment Analysis) is a similar process which underlies procurement decisions in the UK Ministry of Defence. There are many similar examples in other countries. These procedures are not themselves formal multicriteria analyses, but they point to the awareness of the need to explicitly model multiple factors and the existence of informal MCDA in organisations. This can open up substantial opportunities for supported [18] and for the development of tools to be used in an unsupported environment.

12.6 SUMMARY

This chapter has sought to describe the state of the art in the use of MAVT in practice and to highlight current developments and research which can be expected to have an impact on the nature and scope of applications in the near future. If MCDA is to achieve its full potential as a tool for decision support then it is my belief that the central MCDA community must take steps to encourage its wider use by OR/MS practitioners and more generally by practising managers. The current

mix of technological and organisational opportunities combine to facilitate this growth; but it is a challenge which the MCDA community must actively pursue.

REFERENCES

[1] Ackermann, F.: Using the 'domino technique' for problem structuring. University of Strathclyde, Working Paper Series in the Theory, Method and Practice of Management 93/23, 1993

[2] Ackermann, F., V. Belton: Managing corporate knowledge with SODA and V•I•S•A. British Journal of Management 5 (1994) 163-176

[3] Bana e Costa, C.A., J.C.Vansnick: MACBETH - an interactive path towards the construction of cardinal value functions. International Transactions in Operational Research 1 (1994) 489-500

[4] Bana e Costa, C.A., J.C. Vansnick: The MACBETH approach: basic ideas, software and an application. In: N. Meskens, M. Roubens (eds), *Advances in Decision Analysis*. Kluwer, Dordrecht (forthcoming, 1998)

[5] Bana e Costa, C.A., L. Ensslin, A.P. Costa: Structuring the process of choosing rice varieties at the south of Brazil. In: E. Beinat, P. Nijkamp (eds) *Multi-criteria Evaluation in Land-Use Management*. Kluwer, Dordrecht 1997, pp 33-45

[6] Bana e Costa, C.A., L. Ensslin, E.C. Correa, J.C. Vansnick: Decision support systems in action: integrated application in a multicriteria decision aid process. European Journal of Operational Research (forthcoming, 1999)

[7] Bell P.C. : Visual Interactive Modeling as an Operations Research technique. Interfaces 15 (1985) 26-33

[8] Belton, V.: The use of a simple multiple criteria model to assist in selection from a shortlist (Simpson Award winning paper, 1984). Journal of The Operational Research Society 36 (1985) 265-274

[9] Belton, V.: Multiple Criteria Decision Analysis, practically the only way to choose. In: L. Hendry, R. Eglese (eds), *Operational Research Tutorial Papers*. Operational Research Society, Birmingham 1990, pp 53-101

[10] Belton, V.: Operational Research and Information Technology - impacts and interactions. In: B. Spaul, B.C. Williams (eds) *IT and Accounting: The Impact of Information Technology*. Chapman & Hall, London 1991, pp 301-316

[11] Belton, V.: Project planning and prioritisation in the Social Services - an OR contribution. Journal of the Operational Research Society 44 (1993) 115-124

[12] Belton, V.: Implementing multicriteria decision analysis. University of Strathclyde, Working Paper Series in the Theory, Method and Practice of Management 97/10, 1997

[13] Belton, V., F. Ackermann, I. Shepherd: Integrated support from problem structuring through to alternative evaluation using COPE and V•I•S•A. Journal of Multiple Criteria Analysis 6 (1997) 115-130

[14] Belton V., M.D. Elder: Focusing discussion using Group V•I•S•A for multicriteria decision support. University of Strathclyde, Working Paper Series in the Theory, Method and Practice of Management Science (forthcoming 1998)

[15] Belton, V., M.D. Elder: Integrated MCDA: A simulation case study. In: T.J. Stewart, R. van Honert (eds) *Multiple Criteria Decision Making*, Springer-Verlag (forthcoming, 1998)
[16] Belton, V., J. Hodgkin: Facilitators, decision makers, D.I.Y Users: is intelligent multicriteria decision support for all feasible or desirable?. European Journal of Operational Research, Special issue on Intelligent MCDSS (forthcoming, 1998)
[17] Belton, V., J. Pictet: A Framework for group decision using a MCDA model: sharing, aggregating, or comparing individual information?. Journal of Decision Systems, Special Issue on Group Decision and Negotiation 6 (1997) 283-303
[18] Belton, V., R.L. Pratt: Multi-Criteria Decision Analysis: a structured approach to making choices. Journal of Defence Science (forthcoming, 1998)
[19] Belton V., S.P. Vickers: V•I•S•A - VIM for MCDA. In: A.G. Lockett A.G., G. Islei (eds) *Improving Decision Making in Organisations* Springer-Verlag, Berlin, Heidelberg 1989, pp 287-304
[20] Bouyssou, D., P. Perny, M. Pirlot, A. Tsoukias, P. Vincke: A manifesto for the new MCDA era. Journal of Multi-Criteria Analysis 2 (1993) 125-128
[21] Brownlow, S..A.,S.R. Watson: Structuring multi-attribute value hierarchies. Journal of The Operational Research Society 38 (1987) 309-318
[22] Buede, D.M.: "Applied decision analysis: engineering science versus clinical art" In: Proceedings of the International Conference on Cybernetics and Society, IEEE, New York, USA 1979
[23] Buede D.M.: Structuring value attributes. Interfaces 16 (1986) 52-62
[24] Buede D.M., R.W. Choisser: Providing an analytic structure for key system design choices. Journal of Multi-Criteria Decision Analysis 1 (1992) 17-27
[25] Butterworth, N.J.: Giving up 'The Smoke'; a major institution investigates alternatives to being sited in the City. Journal of the Operational Research Society 36 (1989) 711-718
[26] Eden, C.: Cognitive Mapping: a review. European Journal of Operational Research 36 (1988) 1-13
[27] Eden, C.: Using Cognitive Mapping for Strategic Options Development and Analysis (SODA). In: J. Rosenhead (ed) *Rational Analysis for a Problematic World.* Wiley, Chichester 1990, pp 21-42
[28] Edwards, J.S., P. Finlay, J.M. Wilson: "Do It Yourself" decision support - triumph or tragedy?. Paper presented at OR38, Kent, UK, 1996
[29] Edwards, W.: Social utilities. Engineering Economist, Summer Symposium Series, 6 (1971) 119-129
[30] Edwards, W., J.R. Newman: *Multiattribute Evaluation.* Sage, Beverly Hills, CA. 1982
[31] Freeling, A.N.S.: *Belief and Decision Aiding.* Ph.D. Thesis, Cambridge, UK 1983
[32] French, S.: *Decision Theory: An Introduction to the Mathematics of Rationality.* Ellis Horwood, Chichester 1988
[33] French, S., K.N. Papamichail, D.C. Ranyard, J.Q. Smith: Design of a decision support system for use in the event of a radiation accident. University of Leeds, School of Computer Studies, Working Paper 1996

[34] Grauer, M., U. Merten: Analysis of multimedia concepts for decision support. Arbeitsberichte Nr 20, Institut fuer Wirtschaftinsformatik, Universitaet Siegen 1995
[35] Gravel, M., J.M. Martel, R. Nadeau, W. Price, R. Tremblay: A Multicriterion View of Optimal Resource Allocation in Job-Shop Production. Presented to IFORS SPC1 on Decision Support Systems, Bruges, 1991
[36] Hämäläinen, R., M. Pöyhönen: On-line group decision support by preference programming in traffic planning. Group Decision and Negotiation 5 (1996) 485-500
[37] Islei, G., G.Lockett, B. Cox, S. Gisbourne, M. Stratford: Modelling strategic decision making and performance measurements at ICI Pharmaceuticals. Interfaces 21 (1991) 4-22
[38] Kaplan, R.S., D.P. Norton: *The Balanced Scorecard.* HBS Press, Boston, Mass. 1996
[39] Keeney, R.L.: *Siting Energy Facilities.* Academic Press. New York 1980
[40] Keeney, R.L.: *Value-focused Thinking: A Path to Creative Decision Making.* Harvard University Press, Cambridge, Mass. 1992
[41] Keeney, R.L., H. Raiffa: *Decisions with Multiple Objectives: Preferences and Value Tradeoffs* . Wiley, New York 1976
[42] Kelly, G.: *The Psychology of Personal Constructs.* Norton, New York 1955
[43] Korhonen, P., J. Laasko: A visual interactive method for solving the multiple criteria problem.. European Journal of Operational Research (1986) 277-287
[44] Korhonen, P., J. Wallenius: Behavioural issues in MCDM: neglected research questions. Journal of Multi-Criteria Decision Analysis 5 (1996) 178-182
[45] Kuskey, K.P.: Decision conferences, presented at ORSA/TIMS conference, Washington,1980
[46] Macharis, C.: Hybrid Modelling: System dynamics combined with multicriteria analysis. Vrije Universiteit Brussels, Centre for Statistics and Operational Research, Working Paper 1997
[47] Mingers, J.: The Contents of MSc Operational Research courses: results of a questionnaire to OR groups. Journal of the Operational Research Society 42 (1991) 375-382
[48] Mingers, J., A. Gill (eds) *: Multi-Methodology.* Wiley, Chichester 1997
[49] Moore, C.M.: *Group Techniques for Idea Building.* Sage, Thousand Oaks, Ca. 1994
[50] Nunamaker, J., A. Dennis, J, Valacich, D. Vogel, J. George: Electronic meeting systems that support group work. Communications of the ACM 34 (1991) 42-61
[51] Phillips, L.D.: A theory of requisite models. Acta Psychologica 56 (1984) 29-48
[52] Phillips, L.D.: Decision analysis for group decision support. In: C. Eden, J. Radford (eds) *Tackling Strategic Problems.* Sage, London 1990, pp 142-150
[53] Pictet, J., V. Belton: MCDA: What Message?. Opinion Makers Section, Newsletter of the European Working Group on Multicriteria Aid for Decisions (1997)
[54] Quaddus, M.A., D.J. Atkinson, M. Levy: An application of decision conferencing to strategic planning for a voluntary organisation. Interfaces 22 (1992) 61-71
[55] Raiffa, H.: *Decision Analysis.* Addison Wesley, Reading MA., USA 1968
[56] Ranyard, J. C.: A history of OR and computing. Journal of the Operational Research Society 39 (1988) 1073-1086.

[57] Rios Insua, D., S. French: A framework for sensitivity analysis in discrete multi-objective decision-making. European Journal of Operational Research 54 (1991) 176-190
[58] Rosenhead, J. (ed): *Rational Analysis for a Problematic World.* Wiley, Chichester 1989
[59] Saaty, T.L.: *The Analytic Hierarchy Process.* McGraw Hill, New York 1980
[60] Spengler, T., T. Penkuhn: KOSIMEUS - A combination of a flowsheeting program and a multicriteria decision support system. Accepted for publication in the Proceedings of SAMBA II Workshop (State of the Art Computer Programs for Material Balancing), Wuppertal Institut, 1996
[61] Stewart, T.J.: Robustness of additive value function methods in MCDM. Journal of Multi-Criteria Decision Analysis 5 (1996) 301-309
[62] Watson, S.R., D.M. Buede, *Decision Synthesis. The Principles and Practice of Decision Analysis.* Cambridge University Press, Cambridge 1987
[63] Von Winterfeldt, D., W. Edwards: *Decision Analysis and Behavioral Research.* Cambridge University Press, New York 1986
[64] Von Winterfeldt, D.: Major public policy applications of decision analysis in the USA: an overview and an example. Semi-plenary paper presented at SPUDM-16, Leeds 1997

13 FUNDAMENTALS OF INTERIOR MULTIPLE OBJECTIVE LINEAR PROGRAMMING ALGORITHMS

Ami Arbel

Contents

Abstract

The aim of the chapter is to expose its readers to some basic and generic ideas associated with interior algorithms and develop approaches for using these algorithms to address MOLP problems. In doing so, we discuss basic MOLP questions associated with interior algorithms, develop some specific interior MOLP approaches and illustrated them with examples.

13.1 Introduction

For many years the field of linear programming used the simplex algorithm, developed by G.B. Dantzig in 1947, as its major solution approach. With the introduction in 1984 of a new algorithm by N.K. Karmarkar, this situation has changed dramatically [10]. This algorithm had the desired property of polynomial-time convergence, had claims for great speed advantage over the simplex algorithm and, unlike the ellipsoid algorithm (which preceded it with the property of polynomial-time convergence), was practical to implement numerically. The class of algorithms which ensued Karmarkar's algorithm in subsequent years is generally referred to today as interior-point linear programming algorithms. This class of algorithms includes today a number of algorithmic variants (see, e.g., [1], [16]). Research activity in this new class of algorithms became so intensive that over 1500 papers were counted as of early 1990s.

T. Gal et al. (eds.), *Multicriteria Decision Making*

The difference between the class of interior algorithms and the simplex algorithm lies in their respective progress toward the optimal solution of the linear programming problem. The simplex algorithm makes its progress by moving the current solution point on the *exterior* of the constraint polytope and along its vertices until optimality is reached. In contrast, an interior algorithm makes its progress by moving the current solution point through the *interior* of the polytope.

Extensive numerical testing of these algorithms against sophisticated implementations of the simplex algorithm have caused the initial skepticism which greeted the introduction of this new class of algorithms to subside. It is now the general consensus that interior algorithms dominate the simplex algorithm when solving large-scale linear programming problems. Specifically, it is generally agreed today that for small to medium-sized problems, the simplex algorithm will have an edge over the interior algorithm, while the situation reverses when one moves to large-scale problems. Exact delineation of the boundary where one moves from one class of problems to another is not an easy task and depends on factors such as number of variables, density of the constraints matrix as well as its structure. In addition, it is the general consensus today that this new class does not replace the simplex algorithm but rather complements it and enhances solution capabilities of linear programming problems.

Currently, most interactive Multiple Objective Linear Programming (MOLP) procedures are simplex-based as they use the simplex algorithm and its pivoting mechanism to move from one facet to another one on the efficient surface (see, e.g., [8,13,14,19]). The pivoting operation may not pose any special computational difficulty, but the number of required operations increases rapidly when the size of the problem increases, making a simplex-based MOLP algorithm untenable for large-scale problems. These observations motivate the need to explore alternative search procedures for interactive algorithms and lead to the class of interior MOLP algorithms illustrated in this chapter. While the size of the MOLP problem has provided the initial motivation in this direction, it has been found that the use of interior algorithms allows new ways of interaction with a Decision Maker (DM) during the preference elicitation phase inherent with interactive MOLP algorithms.

First attempts at using interior-point single objective algorithms to MOLP problems were reported in [2-5]. These approaches used two different interior variants to generate an interior sequence of iterates. Recently, the use of the achievement scalarizing function has resulted in yet another interior approach for MOLP problems. This function was first proposed by Wierzbicki [17,18] in his reference point approach and led Korhonen and Laakso [11] to show how to use it in an MOLP context and how to explore the efficient frontier [12]. The combination of an interior algorithm with the achievement scalarizing function was recently demonstrated as well [6,7].

The aim of the chapter, therefore, is to present a modest state-of-the-art of interior-point algorithms as they pertain to MOLP problems. Specifically, we intend to present basic and generic ideas associated with interior algorithms and develop some approaches for using these algorithms in an MOLP context. In doing so, we discuss some basic MOLP questions associated with interior algorithms and develop specific interior MOLP approaches. These interior MOLP algorithms are then illustrated by simple examples. This chapter, therefore, is arranged as follows. In

section 2 we provide a brief review of the affine-scaling primal algorithm which is one simple variant of an interior linear programming algorithm. We use this algorithm to generate the interior solution trajectory for our MOLP algorithms. Section 3 initiates the departure from single objective linear programming problems and into MOLP problems. Section 4 develops our first interior MOLP algorithm that is based on performing convex combinations of individual step direction vectors based on local preference assessment. Section 6 reviews the fundamentals of achievement scalarizing functions and illustrates its use in developing an interior MOLP algorithm. Section 7 provides a summary and some suggestions for future research.

13.2 The Affine-Scaling Primal Algorithm

We consider a linear programming problem in *standard form,* given by

$$\begin{aligned} &\max \quad \boldsymbol{c}^T\boldsymbol{x} \\ &s.t. \quad \mathbf{A}\boldsymbol{x} = \boldsymbol{b} \\ &\qquad \boldsymbol{x} \geq \boldsymbol{0}, \end{aligned} \tag{13.1}$$

where $\boldsymbol{x}, \boldsymbol{c} \in \mathscr{R}^n$, $\boldsymbol{b} \in \mathscr{R}^m$ and where $\mathbf{A}$ is an $m \times n$ matrix of full row rank, m. To develop an interior algorithm, we have to provide for a stepping mechanism that starts from the interior of the feasible region and moves to another interior point in such a way that the objective is improved. This process repeats itself until no further improvement is possible. It stops when some termination conditions are met.

We concern ourselves in this section just with the stepping mechanism that allows us to move from one interior point to another while improving the objective function. To this end, let us assume that given the system in (13.1), and an initial feasible *and interior* point (i.e., one whose components are all strictly positive), we seek to proceed in an *ascent direction* to the next interior point. Therefore, denoting this starting interior feasible point by $\boldsymbol{x}_0$ ($\boldsymbol{x}_0 > \boldsymbol{0}$), then feasibility of (13.1) implies

$$\mathbf{A}\boldsymbol{x}_0 = \boldsymbol{b}. \tag{13.2}$$

Next, we seek a step direction vector, $\boldsymbol{dx}$, that takes us in an ascent direction to a new point $\boldsymbol{x}_{new}$ while maintaining feasibility. If the new point is obtained through

$$\boldsymbol{x}_{new} = \boldsymbol{x}_0 + \boldsymbol{dx}, \tag{13.3}$$

then feasibility requires

$$\mathbf{A}\boldsymbol{x}_{new} = \mathbf{A}(\boldsymbol{x}_0 + \boldsymbol{dx}) = \mathbf{A}\boldsymbol{x}_0 + \mathbf{A}\boldsymbol{dx} = \boldsymbol{b}$$

but since $\mathbf{A}\boldsymbol{x}_0 = \boldsymbol{b}$, this results in the first condition on the step direction vector, $\boldsymbol{dx}$, which is then given by

$$\mathbf{A}\boldsymbol{dx} = \boldsymbol{0}. \tag{13.4}$$

In addition, if the step direction vector, $\boldsymbol{dx}$, takes us in an ascent direction, then the following condition must hold: $\boldsymbol{c}^T\boldsymbol{x}_{new} \geq \boldsymbol{c}^T\boldsymbol{x}_0$. Using the expression for $\boldsymbol{x}_{new}$ from (13.3) we have

$$\boldsymbol{c}^T\boldsymbol{x}_{new} = \boldsymbol{c}^T(\boldsymbol{x}_0 + \boldsymbol{dx}) \geq \boldsymbol{c}^T\boldsymbol{x}_0,$$

which leads to the second condition on the step direction vector stated through

$$c^T dx \geq 0. \tag{13.5}$$

Thus we have derived two requirements that the step direction vector, dx, must satisfy. The first one requires that the step direction vector has to be orthogonal to every row of the constraints matrix $\mathbf{A}$. The second requires that the *inner product* of the step direction vector, dx, with the objective vector, c, be nonnegative.

We know that when searching for a solution to a maximization problem one should step along the direction of the gradient. For the linear case under consideration the gradient is simply the objective vector, c. Clearly, stepping along the objective vector is valid only if feasibility of the current point is maintained and if it satisfies the ascent condition. Therefore, if the gradient satisfies the conditions stated in (13.4) - (13.5) it provides a valid ascent direction, otherwise, an alternative direction is needed. Can we modify the gradient direction somehow so that (13.4) - (13.5) are both satisfied?

We start our search by noting that the first condition requires that the step direction vector, dx, belongs to the null space of the $m \times n$ matrix $\mathbf{A}$. The latter is defined through

$$\mathcal{N}(\mathbf{A}) = \{x: \mathbf{A}x = 0,\ x \in \mathcal{R}^n\}. \tag{13.6}$$

A *projection operator* for this space is defined by the $n \times n$ matrix $\mathbf{P}$, where

$$\mathbf{P} = \mathbf{I}_n - \mathbf{A}^T\left(\mathbf{A}\mathbf{A}^T\right)^{-1}\mathbf{A}. \tag{13.7}$$

With this projection operator, given any vector $z \in \mathcal{R}^n$, its projection on the null space of $\mathbf{A}$ given by $y = \mathbf{P}z$, satisfies $\mathbf{A}y = 0$. For this projection matrix, in addition to satisfying $\mathbf{AP} = \mathbf{O}$, the following two properties are readily verified

$$\mathbf{P} = \mathbf{P}^T, \quad \text{and} \quad \mathbf{P} = \mathbf{P}^2. \tag{13.8}$$

For the linear programming problem of (13.1) the gradient, c, should not be expected to satisfy (13.4)-(13.5) for each case. Can we modify this direction? Projecting it perhaps on the null space of $\mathbf{A}$? To check if this is a valid approach, what is needed is to check the projected gradient against the two conditions of (13.4)-(13.5). Projecting the gradient on the null space of $\mathbf{A}$, we find the step direction vector through $dx = \mathbf{P}c$. The feasibility condition of (13.4) is satisfied automatically by virtue of the projection itself. For the ascent condition we find, using (13.8), the following

$$c^T dx = c^T(\mathbf{P}c) = c^T(\mathbf{P}^2 c) = (\mathbf{P}c)^T(\mathbf{P}c) = \|\mathbf{P}c\|^2 \geq 0, \tag{13.9}$$

which indicates that we are stepping in an ascent direction as required by (13.5).

The discussion thus far indicates that stepping through the interior from one feasible point to another, while moving in an ascent direction, should be along the projected gradient. Note, however, that according to our result thus far we can do it only once, since wherever we stop, the next step will be along the same direction. Stepping again along the same direction will, sooner or later, terminate us at the boundary of the polytope. So, even though our step direction, dx, satisfies both (13.4) and (13.5) it provides us with a constant direction that does not change from

one iteration to another. To cause it to change we need to start each iteration from a somewhat different point. A way to do it is provided through the concept of *centering*. This concept and its use in modifying our stepping mechanism is discussed next.

Given a starting vector, $\boldsymbol{x} = [x_1 \quad x_2 \quad \cdots \quad x_n]^T$, its components measure the distance from the respective "walls" of the polytope where $x_i = 0$. If one wishes to place a point at the "center" of the polytope, a simple way of doing it is by choosing a point whose components are all equal. One such choice is a point $\boldsymbol{x}_1$ whose components are all of a *unit* distance from the walls. Therefore, given any starting point, $\boldsymbol{x}$, we translate it to a point that is centered in the polytope, and which is of unit distance from all the wall of the polytope by using the centering transformation defined through

$$\boldsymbol{x}_1 = \mathbf{D}^{-1}\boldsymbol{x}, \tag{13.10}$$

where the scaling operation and the resulting scaled point are given by

$$\mathbf{D} = \begin{pmatrix} x_1 & 0 & \cdots & 0 \\ 0 & x_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & x_n \end{pmatrix}, \quad \text{and } \boldsymbol{x}_1 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}. \tag{13.11}$$

Furthermore, since the original solution vector, $\boldsymbol{x}$, is interior to the polytope, the diagonal elements of the matrix $\mathbf{D}$ are strictly positive and, therefore, $\mathbf{D}$ is invertible. This allows performing the scaling operation of (13.10) for any given interior point. Next, since the original solution vector, $\boldsymbol{x}$, of the system in (13.1) is now scaled through the transformation given by (13.10), we replace it with the scaled (centered) vector, $\boldsymbol{x}_1$ by using $\boldsymbol{x} = \mathbf{D}\boldsymbol{x}_1$. This lead to

$$\begin{aligned} \max \quad & \boldsymbol{c}^T\mathbf{D}\boldsymbol{x}_1 \\ s.t. \quad & \mathbf{A}\mathbf{D}\boldsymbol{x}_1 = \boldsymbol{b} \\ & \boldsymbol{x}_1 \geq \boldsymbol{0}, \end{aligned}$$

which results in the *scaled linear programming problem* described through

$$\begin{aligned} \max \quad & \boldsymbol{c}_1^T\boldsymbol{x}_1 \\ s.t. \quad & \mathbf{A}_1\boldsymbol{x}_1 = \boldsymbol{b} \\ & \boldsymbol{x}_1 \geq \boldsymbol{0}, \end{aligned} \tag{13.12}$$

$$\mathbf{A}_1 = \mathbf{A}\mathbf{D}, \tag{13.13a}$$

$$\boldsymbol{c}_1 = \mathbf{D}\boldsymbol{c}. \tag{13.13b}$$

and where $\boldsymbol{x}_1, \boldsymbol{c}_1 \in \mathscr{R}^n$, $\boldsymbol{b} \in \mathscr{R}^m$.

As shown earlier, we should step in the direction of the projected gradient. Since we deal now with the scaled system of (13.12), we must distinguish between operations applied to the original system and those applied to the scaled system. The projection operator for the scaled linear programming problem of (13.12) is now given by $\mathbf{P}_1$ where

$$\mathbf{P}_1 = \mathbf{I}_n - \mathbf{A}_1^T\left(\mathbf{A}_1\mathbf{A}_1^T\right)^{-1}\mathbf{A}_1. \tag{13.14}$$

The (ascent) step direction vector for the scaled problem, dx_1, is found by projecting the gradient of the scaled problem, c_1, on the null space of the scaled matrix $\mathbf{A}_1$. That is

$$dx_1 = \mathbf{P}_1 c_1. \tag{13.15}$$

Using the expression for the scaled projection operator shown in (13.14) and the relations $\mathbf{A}_1 = \mathbf{AD}$ and $c_1 = \mathbf{D}c$, we derive the ascent step direction vector, dx_1, in terms of the original information which is given by

$$dx_1 = \mathbf{P}_1 c_1 = \mathbf{D}\left[c - \mathbf{A}^T\left(\mathbf{AD}^2\mathbf{A}^T\right)^{-1}\mathbf{AD}^2 c\right]. \tag{13.16}$$

Letting $\left(\mathbf{AD}^2\mathbf{A}^T\right)y = \mathbf{AD}^2 c$ leads to

$$y = \left(\mathbf{AD}^2\mathbf{A}^T\right)^{-1}\mathbf{AD}^2 c, \tag{13.17}$$

and this allows us to simplify the expression for dx_1 to the one given by

$$dx_1 = \mathbf{D}\left[c - \mathbf{A}^T y\right]. \tag{13.18}$$

Recalling that this step is still in the scaled space, requires us to go back to the original space, by a re-scaling operation. Reversing the scaling shown in (13.10), the step direction vector for the new iterate, dx, in the original space is now given by

$$dx = \mathbf{D}dx_1 = \mathbf{D}^2\left[c - \mathbf{A}^T y\right]. \tag{13.19}$$

It should be pointed out, however, that the vector, y, provides an estimate for the dual variable, associated with the dual problem of (13.1). In addition, note that the scaling and rescaling operations are now an implicit part of the sequence of algorithmic steps leading to the derivation of the step direction vector dx shown in (13.19), and are not taken explicitly.

With the step direction vector, dx, given by (13.19) we take a step in that direction and obtain the next iterate of the solution vector, x_{new}. We use an updating formula given by

$$x_{new} \leftarrow x_0 + dx \tag{13.20}$$

Since the step direction vector, dx, used in this update satisfies the feasibility condition of (13.5), the new iterate of the solution vector, x_{new}, satisfies the equality constraints $\mathbf{A}x = b$. Note, however, that these equality constraints are satisfied regardless of how far we step along that direction. To guard against violating the *nonnegativity constraints* of the solution vector, x, we must establish a *maximum allowable step* in that direction. We denote this maximum step by the parameter α, and modify (13.20) into

$$x_{new} = x_0 + \rho\alpha dx, \qquad \alpha > 0, \tag{13.21}$$

where $0 < \rho < 1$, is a *step size factor* that keeps the new iterate interior. Note that if all components of dx are positive, one can increase the value of the objective without violating any of the feasibility constraints. The problem then is clearly *unbounded* and the solution process terminates. Otherwise, similar to the simplex

algorithm, this maximum allowable step size (or simply the *step size*), α, is found from a *ratio test* shown in (13.22).

$$\alpha = \min_{1 \le i \le n} \{-x_0^i / dx_i : 1 \le i \le n, \ \textit{and} \ dx_i < 0\}. \tag{13.22}$$

For computer implementation reasons, however, it is more advantageous to perform the ratio test of (13.22) in its reciprocal form which is given by

$$\alpha = \max_{1 \le i \le n} \{-dx_i / x_0^i : 1 \le i \le n, \ \textit{and} \ dx_i < 0\}. \tag{13.23}$$

This ratio test, of course, leads to an update relation given by

$$\boldsymbol{x} = \boldsymbol{x}_0 + (\rho / \alpha)\boldsymbol{dx}, \qquad 0 < \rho < 1. \tag{13.24}$$

While the only constraint on the stepsize factor ρ is that shown in (13.24), values in practice range from 0.95 to 0.995. The sequence of steps described above continues until the problem is primal and dual feasible and when the duality gap falls below a pre-specified threshold.

13.3 Moving From Single to Multiple Objectives

The primal algorithm described earlier for single-objective linear programming (SOLP) problems goes through a series of steps in which the current iterate is first centered through a scaling operation, and then movement is made from the current position along the gradient projected on the null space of the scaled constraints matrix. Re-scaling the step direction vector allows us to derive the new iterate in the original space. When dealing with MOLP problems there are a few objective vectors and, therefore, the projected gradients usually point in more than one direction. Moving from a single-objective to a multiple-objective problem requires reconciling these different directions and finding a compromise direction. Arriving at such a compromise can be done in more ways than one as we show next.

In the previous section we considered a single-objective linear programming problem in standard form which is described through

$$\begin{aligned} &\max \ \boldsymbol{c}^T \boldsymbol{x} \\ &s.t. \quad \mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \\ &\qquad \boldsymbol{x} \ge \boldsymbol{0}, \end{aligned} \tag{13.25}$$

where $\boldsymbol{x}, \boldsymbol{c} \in \mathscr{R}^n$, $\boldsymbol{b} \in \mathscr{R}^m$ and $\mathbf{A}$ is an $m \times n$ constraints matrix. In this section we consider an MOLP *maximization* problem having q linear objective functions. Such a problem is described through

$$\begin{aligned} &\begin{cases} \max f_1(\boldsymbol{x}) = \boldsymbol{c}_1^T \boldsymbol{x} \\ \max f_2(\boldsymbol{x}) = \boldsymbol{c}_2^T \boldsymbol{x} \\ \quad \vdots \\ \max f_q(\boldsymbol{x}) = \boldsymbol{c}_q^T \boldsymbol{x} \end{cases} \\ &s.t. \quad \boldsymbol{x} \in \boldsymbol{S} = \left\{ \boldsymbol{x} \in \mathscr{R}^n : \mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \ \boldsymbol{x} \ge \boldsymbol{0} \right\}. \end{aligned} \tag{13.26}$$

Such a problem is written in matrix form as

$$\begin{aligned} &\text{"}max\text{"} \quad \mathbf{C}\boldsymbol{x} \\ &s.t. \quad \boldsymbol{x} \in \boldsymbol{S} \equiv \{\boldsymbol{x} \in \mathcal{R}^n \mid \mathbf{A}\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geq \boldsymbol{0}\} \end{aligned} \tag{13.27}$$

where the $q \times n$ matrix $\mathbf{C}$ has the q objective vectors of (13.26) as its rows.

Before moving on, however, it is useful to recall some important definitions. We say that a point $\boldsymbol{x}^* \in \boldsymbol{S}$, in decision space is *efficient* iff there does not exist another $\boldsymbol{x} \in \boldsymbol{S}$ such that $\mathbf{C}\boldsymbol{x} \geq \mathbf{C}\boldsymbol{x}^*$, and $\mathbf{C}\boldsymbol{x} \neq \mathbf{C}\boldsymbol{x}^*$. We say that $\boldsymbol{x}^* \in \boldsymbol{S}$ is *weakly efficient* iff there does not exist another $\boldsymbol{x} \in \boldsymbol{S}$ such that $\mathbf{C}\boldsymbol{x} > \mathbf{C}\boldsymbol{x}^*$. Letting $\boldsymbol{V} \equiv \{\boldsymbol{v} \in \mathcal{R}^q : \boldsymbol{v} = \mathbf{C}\boldsymbol{x},\ \boldsymbol{x} \in \mathcal{R}^n\}$ be the set of *feasible* criterion vectors (i.e., the feasible region in objective space), a vector $\boldsymbol{v} \in \boldsymbol{V}$ corresponding to an efficient point is termed a *nondominated* criterion vector and a vector $\boldsymbol{v} \in \boldsymbol{V}$ corresponding to a weakly efficient point is termed a *weakly nondominated* criterion vector. The set of all efficient points is called the *efficient set*, denoted by $\boldsymbol{E}$, and the set of all nondominated criterion vectors is called the *nondominated set*, denoted by $\boldsymbol{N}$. For weakly efficient solutions, we use $\boldsymbol{E}^w$ and $\boldsymbol{N}^w$, respectively.

Of the approaches devised at addressing the vector optimization problem of (13.27) the *naïve approach* is noteworthy as an erroneous way of looking at this problem. When using such an approach one is tempted to bypass the difficulty introduced through the $q \times n$ matrix of objective vectors, $\mathbf{C}$, by trying some static weighting scheme to reduce the q objectives into a single weighted objective amenable to the direct application of a linear programming algorithm. Let Δ^q denote a set of q positive and normalized weights defined by

$$\Delta^q \equiv \left\{ \lambda \in \mathcal{R}^q :\ \lambda_i > 0,\ \forall 1 \leq i \leq q, \quad \text{and} \sum_{i=1}^{n} \lambda_i = 1 \right\}. \tag{13.28}$$

Then, using such weights, the problem in (13.27) is reduced to a single-objective linear programming problem described through

$$\begin{aligned} &\max \quad \boldsymbol{\lambda}^T \mathbf{C}\boldsymbol{x} \\ &s.t. \quad \boldsymbol{x} \in \boldsymbol{S} \equiv \{\boldsymbol{x} \in \mathcal{R}^n \mid \mathbf{A}\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geq \boldsymbol{0}\} \\ &\qquad \boldsymbol{\lambda} \in \Delta^q. \end{aligned} \tag{13.29}$$

The rationale behind such an approach is that one can easily establish some ranking of the objectives (e.g., *the first objective is twice as important than the second*, etc.) which allows the establishment of this (static) weight vector, $\boldsymbol{\lambda}$. This rationale, however, is faulty. The reason is that, in general, the relative preference for the objectives and its underlying weighting scheme depends on where we are in solution space and changes from one location of the solution vector $\boldsymbol{x}$, to another. Such weights should, therefore, be *locally-relevant*, and require reassessment during the solution process. They should not be taken as static entities.

A more credible approach that attempts replacing the vector-valued optimization problem with a scalar-valued (single objective) optimization problem involves the use of a utility function, $u(\boldsymbol{x}) \in \mathcal{R}^1$. This is a real-valued function capturing the DM's preference for all outcomes in decision space. When such a

function is available, the problem of (13.27) is effectively replaced by the problem shown in

$$\begin{aligned} &\max \ u(\boldsymbol{x}) \\ &s.t. \ \ \boldsymbol{x} \in \boldsymbol{S}, \ \ u: \mathcal{R}^n \rightarrow \mathcal{R}^1 \end{aligned} \tag{13.30}$$

and the optimal solution to this – generally nonlinear – problem is the solution vector, $\boldsymbol{x}^*$, yielding the highest value of the utility function over all possible feasible outcomes. Conceptually then, when a utility function is known, one starts at a feasible point, generates a search direction that moves the current iterate along the gradient of the utility function to a new one with a higher utility value. When no utility function is available, one has to find a substitute to the gradient of the utility function. For any reasonable MOLP problem one cannot assume the availability of a utility function due to the size of the decision vector, $\boldsymbol{x}$, which makes the assessment of a utility function, $u(\boldsymbol{x})$, impractical. One has to assume, therefore, that the utility function is only *implicitly-known*. By this we mean that while a DM does not have access to an explicit utility function, preference statements can still be made. Most realistic MOLP algorithms use some sort of proxy measures to this implicitly-known utility function.

The two approaches outlined above represent the two extremes of solution methods to MOLP problems: from the erroneous inherent in the naïve approach to the ideal, and mostly impractical, situation inherent with the utility-based approach. Specific methods developed over the years for MOLP problems naturally fall between these two extremes. They attempt to combine the straightforward nature of the naïve approach on the one hand with the clean mathematical nature of the utility-based approach on the other.

In general then, moving from SOLP problems to MOLP problems poses, therefore, two questions. The first is concerned with preference elicitation and usage in the absence of a utility function, and the second is concerned with using this preference information in devising a step direction vector along which one steps from the current to the next iterate.. Let us consider first the latter question. An immediate suggestion that follows our basic development of the SOLP algorithm is to apply the basic steps to each of the q objective vectors on their own. That is, given a starting feasible and interior point $\boldsymbol{x}_0 > \boldsymbol{0}$, a search direction, $\boldsymbol{dx}_i$, is then found by projecting the i-th objective vector $\boldsymbol{c}_i$, $(1 \le i \le q)$. Using these projections, the q respective new iterates, $\{\boldsymbol{x}_i\}$, are then evaluated through

$$\boldsymbol{x}_i = \boldsymbol{x} + \rho\alpha_i \boldsymbol{dx}_i \quad \text{where } 1 \le i \le q \tag{13.31}$$

where the step sizes, α_i, are evaluated for each of the separate directions, $\boldsymbol{x}_0$ is the current iterate and ρ is the step size factor. The value of each of the q objective functions at *each* of these new iterates, is summarized in a vector of objective values, $\boldsymbol{v}_i$, where $\boldsymbol{v}_i = \mathbf{C}\boldsymbol{x}_i$. The description of the current iterate and its corresponding q new iterates in decision space is shown in Figure 3.1 below.

Starting at the current iterate, $\boldsymbol{x}_0$, the DM can now step along each of the q step direction vectors $\{\boldsymbol{dx}_i\}$ and end up at the respective new iterate $\boldsymbol{x}_i$ in decision space. Finding the corresponding points in objective space is easily evaluated through $\boldsymbol{v}_i = \mathbf{C}\boldsymbol{x}_i$. Since a step from the current iterate is taken along a single

direction, the q different directions should, somehow, be combined. One such way is offered by assessing relative preference to objective values at each of the q new iterates. We outline next such a procedure.

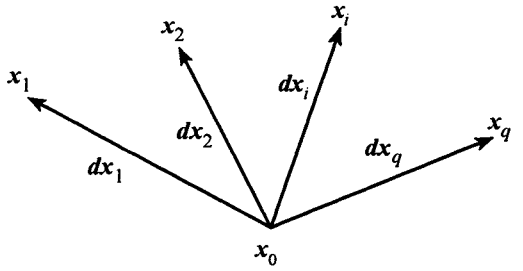


Figure 13.1: Current and new iterates in decision space

13.4 Combining Individual Step Direction Vectors

We consider again the MOLP problem described through

$$\text{"}max\text{"}\ \ \mathbf{C}\boldsymbol{x}$$

$$s.t.\ \ \boldsymbol{x} \in \boldsymbol{S} \equiv \{\boldsymbol{x} \in \mathcal{R}^n \mid \mathbf{A}\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geq \boldsymbol{0}\} \tag{13.32}$$

where the $q \times n$ matrix $\mathbf{C}$ has the q objective vectors of (13.26) as its rows.

Starting from an initial feasible and interior point in decision space, $\boldsymbol{x}_0$, each new interior iterate, corresponding to a specific objective function, $\boldsymbol{c}_i$ is derived through

$$\boldsymbol{x}_i = \boldsymbol{x}_0 + \rho\alpha_i \boldsymbol{dx}_i, \tag{13.33}$$

and the values of all the objectives at each of these new iterates are summarized in an objective vector, $\boldsymbol{v}_i$, evaluated through

$$\boldsymbol{v}_i = \mathbf{C}\boldsymbol{x}_i, \quad \text{where } 1 \leq i \leq q. \tag{13.34}$$

Next, we proceed to assess relative preference for objective vectors $\{\boldsymbol{v}_i\}$ at each of the new q iterates. This step results in a set of weights which are then applied to the q objectives to reduce them to a single objective. Note, however, that unlike the naïve approach, these weights are locally-relevant rather than static, and are re-assessed at each iteration.

Looking from the current iterate, $\boldsymbol{x}_0$, to each of the new iterates and their respective objective values, the DM is asked to assess relative preferences for each of these new iterates. Preference for, say, $\boldsymbol{x}_i$ over $\boldsymbol{x}_j$ implies that the DM assesses the implicitly-known utility at $\boldsymbol{x}_i$ to be higher than that at $\boldsymbol{x}_j$ and, therefore

$$\boldsymbol{x}_i \succ \boldsymbol{x}_j \quad \Leftrightarrow \quad u(\boldsymbol{x}_i) > u(\boldsymbol{x}_j). \tag{13.35}$$

In the absence of a utility function, this preference assessment is carried out using some sort of an evaluation methodology capable of assigning numerical values to the strength of preference for each of the q new iterates. Whatever methodology one uses should result in a vector of weights or priorities, $\boldsymbol{w}$, providing a measure of relative preference for each of the objective vectors $\{\boldsymbol{v}_i\}$. If the actual utility function was available, we could use the (normalized) utilities at the new iterates $\{\boldsymbol{x}_i\}$ as components for such a weighting vector. In fact, if the DM's answers truly reflect his utility function, then whenever $\boldsymbol{x}_i \succ \boldsymbol{x}_j$, this implies

$$u(\boldsymbol{x}_i) > u(\boldsymbol{x}_j) \quad \Leftrightarrow \quad w_i > w_j. \tag{13.36}$$

Normalizing the assessed weights allows us to maintain the proper ratio between the utility values at the respective points in objective space.

With the utility at each of the new iterates in objective space approximated through the respective component of the assessed weight vector, $\boldsymbol{w}$, we proceed to establish a *compromise* direction. Specifically, the components of the weight vector are now used as coefficients of a convex combination of the q single step direction vectors yielding a single step direction vector, $\boldsymbol{dx}$, along which we take the next step. The resulting new interior iterate is then found from

$$\boldsymbol{x}_{new} = \boldsymbol{x}_0 + \rho \boldsymbol{dx} = \sum_{i=1}^{q} w_i (\boldsymbol{x}_0 + \rho \alpha_i \boldsymbol{dx}_i). \tag{13.37}$$

Since the weight vector, $\boldsymbol{p}$, is normalized, we find

$$\boldsymbol{x}_{new} = \boldsymbol{x}_0 + \rho \boldsymbol{dx} = \boldsymbol{x}_0 + \rho \sum_{i=1}^{q} (w_i \alpha_i) \boldsymbol{dx}_i \tag{13.38}$$

and the combined step direction vector is, therefore, given by

$$\boldsymbol{dx} = \sum_{i=1}^{q} w_i (\alpha_i \boldsymbol{dx}_i), \tag{13.39}$$

The new iterate is now evaluated through

$$\boldsymbol{x}_{new} = \boldsymbol{x}_0 + \rho \boldsymbol{dx}, \qquad 0 < \rho < 1, \tag{13.40}$$

where $\boldsymbol{w} \in \Delta^q$, namely

$$w_i > 0, \ \forall 1 \le i \le q, \text{ and } \sum_{i=1}^{q} w_i = 1. \tag{13.41}$$

We see, therefore, that the set of weights, $\{w_i\}$, together with the step sizes $\{\alpha_i\}$, are used to combine the q individual step direction vectors, $\{\boldsymbol{dx}_i\}$, into a single combined direction, $\boldsymbol{dx}$ given by (13.39).

The procedure for using interior search directions and generating a trajectory of interior iterates, while using the affine-scaling primal algorithm is, therefore, as follows. We project each of the individual objective vectors, establish the corresponding step sizes, $\{\alpha_i\}$, and using a common step size factor, ρ, we step into q different interior new iterates, $\{\boldsymbol{x}_i\}$. Next, we evaluate the value of the q objective vectors, $\{\boldsymbol{v}_i\}$, at each of the iterates in decision space, $\{\boldsymbol{x}_i\}$. Assessing relative preference for each of these objective vectors, allows us to arrive at a set of

normalized weights, $\{w_i\}$, which are then used to form a convex combination that yields the combined step direction vector, $\boldsymbol{dx}$, as given in (13.39), along which we step to the next interior iterate. This general procedure is depicted below in Fig. 13.2a, where an MOLP problem with two objectives is considered.

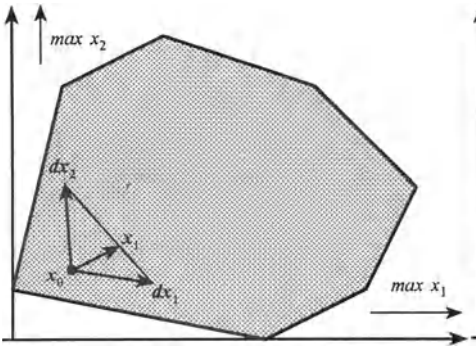


Figure 13.2a: Taking one Combined Step

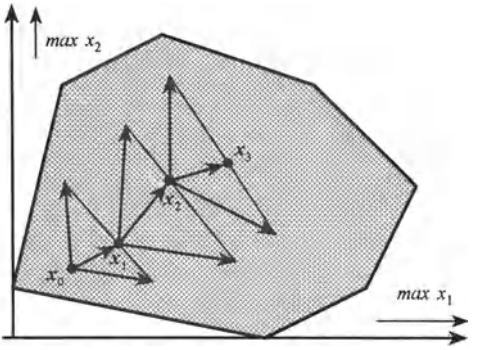


Figure 13.2b: Combining Step Directions

In a similar manner, the procedure repeats itself to generate an entire interior trajectory of solution iterates as shown in Fig. 13.2b. This procedure seems straightforward enough. Let us illustrate it with a numerical example before proceeding onward with our development.

Example 13.1: Consider an MOLP problem described by

$$\begin{cases} \max\ f_1 = x_1 \\ \max\ f_2 = x_2 \end{cases}$$

$$s.t.\quad x_1 + x_2 \le 10$$

$$x_1, x_2 \ge 0$$

$$x_0 = \begin{bmatrix} 1 & 1 & 8 \end{bmatrix}^T, \text{(for the problem in standard form)}$$

and where $u(x) = x_1 x_2$.

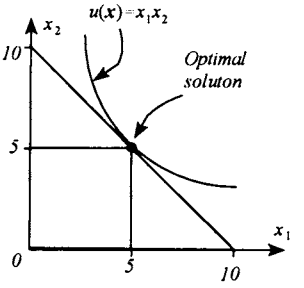


Figure 13.3: Feasible region

Note that with the utility function as stipulated above, the MOLP problem has a unique optimal solution given by $x_1^* = 5, x_2^* = 5,$ and where $u(x_1^*, x_2^*) = 25$. Using the affine-scaling primal algorithm, we proceed next to project each of the objective vectors. This results with the two step direction vectors given by

$$dx_1 = \begin{bmatrix} 0.9848 \\ -0.0152 \\ -0.9697 \end{bmatrix}, \qquad dx_2 = \begin{bmatrix} -0.0152 \\ 0.9848 \\ -0.9697 \end{bmatrix}.$$

Next, performing the required ratio test, and taking an interior step using a step size factor of $\rho = 0.5$ we arrive at the two new interior iterates given by

$$x_1 = x + \rho\alpha dx_1 = \begin{bmatrix} 5.0625 \\ 0.9375 \\ 4.0000 \end{bmatrix}, \qquad x_2 = x + \rho\alpha dx_2 = \begin{bmatrix} 0.9375 \\ 5.0625 \\ 4.0000 \end{bmatrix}.$$

The availability of a utility function allows us direct evaluation of the necessary weighting coefficients at the new iterates and avoid the need for an interaction with a DM. The availability of the utility function allows us, in addition, to *test* the algorithmic steps and verify that we arrive at the true optimal solution. The utility values and the normalized weighting coefficients are given by

$$u(x_1) = 4.7461, \qquad u(x_2) = 4.7461.$$

$$w_1 = \frac{u(x_1)}{u(x_1) + u(x_2)}, \qquad w_2 = \frac{u(x_2)}{u(x_1) + u(x_2)}.$$

The combined step direction vector, dx, is evaluated now through

$$dx = w_1 dx_1 + w_2 dx_2 = \begin{bmatrix} 0.4848 \\ 0.4848 \\ -0.9697 \end{bmatrix}$$

and the new iterate (corresponding to a step size factor of $\rho = 0.5$) with its respective utility value are then given by

$$x_{new} = x_{old} + \rho\alpha dx = \begin{bmatrix} 3.0000 \\ 3.0000 \\ 4.0000 \end{bmatrix},$$

and the value of the utility function at this point is given by

$$u(x_{new}) = 9.000.$$

The complete summary of the first ten iterations is shown in the table below

Table 13.1: Summary of iterations (ρ=0.5)

k	$x_1(k)$	$x_2(k)$	$u(x)$
0	1.0000	1.0000	1.0000
1	3.0000	3.0000	9.0000
2	4.0000	4.0000	16.0000
3	4.2500	4.2500	18.0625
4	4.3824	4.3824	19.2050
5	4.4694	4.4694	19.9756
6	4.5324	4.5324	20.5426
7	4.5806	4.5806	20.9822
8	4.6190	4.6190	21.3354
9	4.6505	4.6505	21.6267
10	4.6767	4.6767	21.8718

The value of the utility at the current iterate is depicted in Fig. 13.4 below.

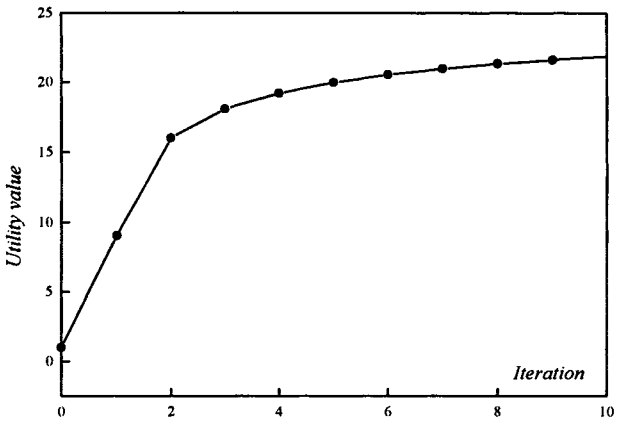


Figure 13.4: Summary of iterations

■

13.5 Anchoring Points

The procedure outlined above for combining the individual step direction vectors into a single compromise direction through a convex combination seems like a straightforward plausible approach that is likely to make good progress toward the optimal solution. *Unfortunately this is not always the case!* We discuss next the reason for this statement and the required corrective measures offered through the introduction of *anchoring points*. Before proceeding, however, we need to define the

notion of an *end point*. When stepping from the current iterate along a step direction vector to a new interior iterate, the updating formula we use is given by

(13.42) $$x = x_0 + \rho\alpha dx, \text{ where } 0 < \rho < 1.$$

The step size factor, ρ, determines how close we come to the boundary of the constraints polytope. Therefore, taking a *full* step by choosing the step size factor according to $\rho = 1$, we derive a new iterate that is located on the boundary of the constraints polytope. This point is therefore given by

(13.43) $$x = x_0 + \alpha dx,$$

where the step size, α, is determined in the usual way to ensure the nonnegativity of the new decision iterate, x. We term such an end point an *anchor point* for reasons discussed below.

Definition 13.1: An *anchor point*, x_a, to an MOLP problem is a point on the boundary of the constraints polytope found by taking a full step ($\rho = 1$) along a combined step direction vector, dx. Such a point is derived by $x_a = x + \alpha dx$.

With these definitions, consider now the situation depicted in Fig. 13.5.1 below. What is shown is a feasible region of a simple linear programming problem having two objective functions given by $f_1 = \max x_1$, and $f_2 = \max x_2$. We start from an interior feasible point denoted by $x(0)$, that is placed on a utility curve $u(x) = const.$, and the gradient shown in Fig. 13.5 indicates the direction of increased utility values. The optimal point where the utility function reaches its highest value is indicated as well. In addition, we show the two step direction vectors dx_1 and dx_2 generated by applying the algorithmic details of an interior point linear programming algorithm.

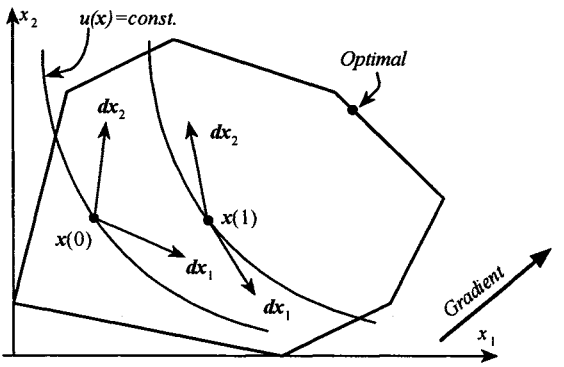


Figure 13.5: Moving Along Gradients

By assessing local preferences for the two directions indicated above and forming the necessary convex combination, we arrive at the new interior iterate, $x(1)$, as indicated above. Note that at x_0, whatever the assessed weights are for the two directions, the new interior iterate, x_1, leads to a utility curve of higher values.

Furthermore, at $x(1)$, we see that if one is to move to a utility curve of higher value, the step direction vector dx_2 should be weighted much heavier than that of dx_1 which points to a direction of lower utility.

We proceed by moving from $x(1)$, to a new interior iterate $x(2)$, at which the step direction vectors dx_1 and dx_2 point in the direction as indicated in Fig. 13.6 below. Clearly, at this point there does not exist a convex combination of the two step direction vectors, dx_1 and dx_2 capable of moving the current iterate to a higher utility value. Whatever convex combination one forms at the current iterate points in such a way that the next iterate has a lower utility value. Is there a way to move from the current iterate to a point of higher utility value rather than decreasing it? This is where we make use of the concept of an *anchor* point.

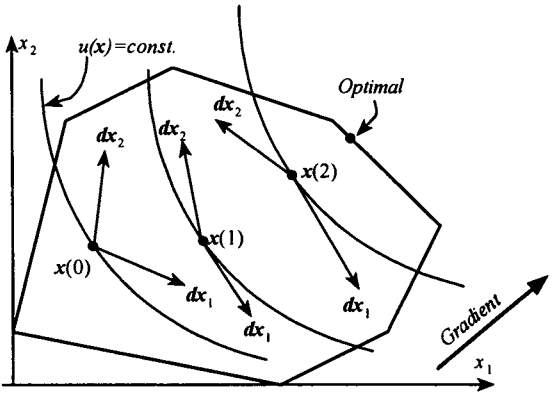


Figure 13.6: The Need for Anchoring

Note that the iterate denoted by $x(1)$, is derived by stepping away from $x(0)$ while using a certain step size factor $0 < \rho < 1$. Taking a full step (i.e., $\rho = 1$) along dx moves us all the way to the boundary where we establish a *boundary point*, x_{end}. Having the boundary point x_{end}, however, introduces an additional step direction vector derived through $dx_{end} = x_{end} - x(1)$, which points from $x(1)$, towards the boundary point. One proceeds now to form a convex combination that includes dx_1, dx_2, as well as dx_{end}. With the aid of an anchor point, when we are at the iterate denoted by $x(1)$ we have an additional step direction vector pointing to a point on the boundary. With the aid of this additional boundary point – defined earlier as an anchor point – we are now capable of moving the current iterate to higher utility value. We refer to such a boundary point as an anchor point as it provides an anchor capable of lifting us to higher values of utility.

Let us illustrate a conceptual sequence of iteration using anchor points together with the step direction vectors. Consider the situation depicted in Fig. 13.7 below. The two objectives are given by $f_1 = \max x_1$, and $f_2 = \max x_2$. The optimal solutions for each of the two objectives are denoted by x_1^* and x_2^* respectively. Starting the

solution process from an interior feasible iterate denoted by $\boldsymbol{x}_0$, we project the two objective vectors and derive the two step direction vectors denoted by $\boldsymbol{dx}_1$, and $\boldsymbol{dx}_2$, respectively. Using a certain step size factor, ρ, we arrive at two new interior iterates. We proceed next to assess relative preference for the objective vectors at these two new iterates, and forming the required convex combination we establish our next interior iterate denoted by $\boldsymbol{x}_1$, below. Note that by taking a full step from the starting iterate, $\boldsymbol{x}_0$, along $\boldsymbol{dx}$ we establish an anchor point on the boundary. Such a point is denoted as $\boldsymbol{A}$ below.

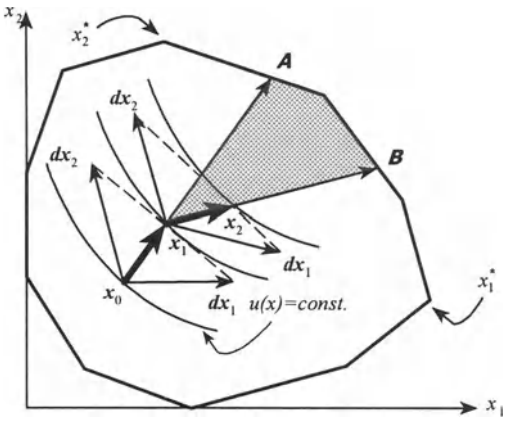


Figure 13.7: Successive Anchors

Next, we move from $\boldsymbol{x}_1$, to the next iterate. We start by deriving the step direction vectors $\boldsymbol{dx}_1$, and $\boldsymbol{dx}_2$, but now we add an additional step direction that points toward the anchor point, **A**. Stepping along these three directions by using a step size factor, ρ, we arrive at three new interior iterates. Assessing preference and taking a convex combination we derive a step direction vector, $\boldsymbol{dx}$, along which we move to the next iterate denoted by $\boldsymbol{x}_2$. Taking a full step along the derived step direction, $\boldsymbol{dx}$, establishes a new anchor point denoted by $\boldsymbol{B}$ in *Figure 13.7* above.

Clearly then, each iteration adds another anchor to the existing set of points generated thus far. This raises an interesting question of how many of them to keep and how to trim the set to a manageable number. The simplest way of handling this issue is to keep all anchor points. This, however, increases the set of anchors by one at each iterate. The drawback with such an approach is the increased burden on the DM. Recall that when forming the convex combination, relative preference of objective vectors is assessed. Since these objective vectors include also those corresponding to anchor points, increasing the number of anchor points increases the number of vectors compared. After a few iterations this may become an unruly process and quite tedious for the DM.

A simpler approach is offered by keeping a constant number of anchors. Once this number is reached, each additional anchor point is compared against those

already in the set. If the value of the objective vector at the new anchor point is preferred to one of the existing anchors, we bring the new point into the set, and remove the least preferred anchor from the set. Doing so, keeps the most preferred anchors encountered during the solution process. In addition, since solving single-objective LP problems is cheap, while interacting with a DM is not, this process allows an added benefit. We may start the solution process by solving the MOLP problem for a few weighted combinations of its objectives. That is, solve the following problem for q different sets of $\lambda \in \Delta^q$

$$\begin{aligned} &\max \ \lambda^{\mathrm{T}}\mathbf{C}\boldsymbol{x} \\ &s.t. \ \ \boldsymbol{x} \in \boldsymbol{S} \equiv \{\boldsymbol{x} \in \mathscr{R}^n \mid \mathbf{A}\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \geq \boldsymbol{0}\} \\ &\lambda \in \Delta^q. \end{aligned} \tag{13.44}$$

It is well known [9] that a solution to this single-objective parametric linear programming problem results in an *efficient* solution. We therefore, start our solution with a set of efficient anchor points which are generated *off-line*, and do not require any interaction with a DM. A simple way to start, is to generate an approximate ideal point. Recall that an ideal point is that where each of the objectives reaches its highest value. Therefore, by solving (13.44) q times where at each run we favor heavily one of the objectives, we derive a set of anchors each favoring one of the objectives. The solution process that ensues serves to bring this starting set closer to each other until a final compromise solution is arrived.

The importance of anchor points was illustrated graphically, we now provide the required algorithmic modifications and a summary of the proposed MOLP algorithm. Recall that an interior iterate, $\boldsymbol{x}_{new}$, and its corresponding anchor point (derived and the end of a full step along $\boldsymbol{dx}$), $\boldsymbol{x}_{end}$, are derived from similar updating formulas differing only in the use of a step size factor, ρ. Specifically, we have

$$\boldsymbol{x}_{new} = \boldsymbol{x}_0 + \rho\alpha\boldsymbol{dx}, \ \text{where} \ 0 < \rho < 1.$$

$$\boldsymbol{x}_{end} = \boldsymbol{x}_0 + \alpha\boldsymbol{dx}.$$

Assuming that we keep on hand a full set of q anchor points, $\{\boldsymbol{x}^i_{anchor}\}$, where $1 \leq i \leq q$, a step direction vector, $\boldsymbol{dx}_{end}$, pointing from the current iterate, $\boldsymbol{x}_0$, toward a particular end point, $\{\boldsymbol{x}_{end}\}$, is readily derived through

$$\boldsymbol{dx}_{end} = \boldsymbol{x}_{end} - \boldsymbol{x}_0. \tag{13.45}$$

Using a step size factor, ρ, the interior point along this direction, $\bar{\boldsymbol{x}}_{end}$, as well at the value of the objectives at this point, $\bar{\boldsymbol{v}}_{end}$, are evaluated through

$$\bar{\boldsymbol{x}}_{end} = \boldsymbol{x}_0 + \rho\boldsymbol{dx}_{end}, \ \text{and} \ \bar{\boldsymbol{v}}_{end} = \mathbf{C}\bar{\boldsymbol{x}}_{end}. \tag{13.46}$$

Assessing relative preference for each of the q interior iterates, $\{\boldsymbol{x}^i\}$, together with the q anchor points, $\{\bar{\boldsymbol{x}}^i_{anchor}\}$, allows the derivation of the combined step direction vector, $\boldsymbol{dx}$, which is now derived through

$$\boldsymbol{dx} = \sum_{i=1}^{q} (w_i\alpha_i)\boldsymbol{dx}_i + \sum_{i=1}^{q} \bar{w}_i\boldsymbol{dx}_{end}. \tag{13.47}$$

where $\{w_i\}$ is the set of weights expressing assessed relative preference for the interior iterates, $\{\boldsymbol{x}^i\}$ generated by stepping along $\{\boldsymbol{dx}^i\}$ and using a step size factor, ρ, and the set $\{\overline{w}_i\}$ expresses relative preference for the interior iterates, $\{\overline{\boldsymbol{x}}^i_{anchor}\}$, generated according to (13.46). Note that these sets of weights are assessed together and satisfy

$$\sum_{i=1}^{q} w_i + \sum_{i=1}^{q} \overline{w}_i = 1, \quad \text{where } w_i > 0,\ \overline{w}_i > 0, \quad \forall 1 \le i \le q. \tag{13.48}$$

With the modified step direction vector, $\boldsymbol{dx}$, given as in (13.47), the new interior iterate is readily derived according to

$$\boldsymbol{x}_{new} = \boldsymbol{x}_0 + \rho \boldsymbol{dx}. \tag{13.49}$$

It should be noted, however, that by taking a full step along the combined direction, $\boldsymbol{dx}$, we establish a new boundary point given by

$$\boldsymbol{x}_{new} = \boldsymbol{x}_0 + \boldsymbol{dx}. \tag{13.50}$$

Evaluating the value of the objectives at this point through $\boldsymbol{v}_{new} = \mathbf{C}\boldsymbol{x}_{new}$, allows us to compare DM's preference for this point against the existing anchor points. If this point is preferable to any of the existing anchor points, we replace that anchor point with the newly derived anchor points. The current set of anchor points continues to contain the most preferred boundary points encountered during the solution process. Before demonstrating the modified interior-point algorithm through an example, we provide a summary of its algorithmic steps below.

Summary of the Affine-Scaling Interior MOLP Algorithm:

Consider a MOLP problem in standard form given by

$$\text{"max"}\ \mathbf{C}\boldsymbol{x}$$

$$s.t.\ \ \boldsymbol{x} \in \boldsymbol{S} \equiv \{\boldsymbol{x} \in \mathcal{R}^n \mid \mathbf{A}\boldsymbol{x} = \boldsymbol{b},\ \boldsymbol{x} \ge \boldsymbol{0}\}$$

for which a starting feasible and strictly interior solution vector, $\boldsymbol{x}_0$, is given. That is, $\mathbf{A}\boldsymbol{x}_0 = \boldsymbol{b}$, and $\boldsymbol{x}_0 > \boldsymbol{0}$. The affine-scaling interior-point MOLP algorithm (*ASIMOLP*) proceeds as follows:

Step 1: Initialize iteration counter, $k := 0$, and solution vector through $\boldsymbol{x}(k) = \boldsymbol{x}_0$. Choose q different weight vectors, $\boldsymbol{\lambda} \in \Delta^q$, and solve – for each set – the resulting weighted MOLP problem as shown in (13.44). The resulting q solutions provide the initial set of anchor points, $\{\overline{\boldsymbol{x}}^i\}$.

Step 2: Increment the iteration counter, $k := k+1$, and define the scaling matrix, $\mathbf{D}$, as $\mathbf{D} = diag[x_1(k)\ \ x_2(k)\ \ \cdots\ \ x_n(k)]^T$. Solve for the m dimensional vector $\boldsymbol{y}^i(k)$ from the q symmetric systems of equations given by

$$\left(\mathbf{A}\mathbf{D}^2\mathbf{A}^T\right)\boldsymbol{y}^i(k) = \mathbf{A}\mathbf{D}^2\boldsymbol{c}^i, \quad \forall 1 \le i \le q,$$

where the i-th objective, $\boldsymbol{c}^i$, is the i-th row of the $q \times n$ objective matrix

C. The i-th step direction vector at the k-th iteration, $\boldsymbol{dx}^i(k)$, is then derived through

$$\boldsymbol{dx}^i(k) = \mathbf{D}^2\left[c^i - \mathbf{A}^T y^i(k)\right], \quad \forall 1 \le i \le q.$$

Step 3: Find the set of q new interior iterates, $x^i(k+1)$, through

$$x^i(k+1) = x^i(k) + \rho\alpha_i \boldsymbol{dx}^i(k), \quad \forall 1 \le i \le q, \text{ and } 0 < \rho < 1.$$

Proceed to evaluate the value vectors $v^i(k+1)$, at these points, through

$$v^i(k+1) = \mathbf{C}x^i(k+1), \quad \forall 1 \le i \le q.$$

Step 4: Use an assessment methodology to determine the relative preference, $\{w_i\}$, for each of the q objective vectors at the end of the new interior iterates, as well as preference, $\{\overline{w}_i\}$, for the objective vectors at the end of the step directions pointing toward the anchor points. Use these weights to derive the step combined direction vector, $\boldsymbol{dx}$, described through

$$\boldsymbol{dx} = \rho\sum_{i=1}^{q}(w_i\alpha_i)\boldsymbol{dx}_i + \rho\sum_{i=1}^{q}\overline{w}_i\left(x_{end} - x_0\right).$$

Step 5: Find the next iterate through the updating equation given by

$$x(k+1) = x(k) + \rho\boldsymbol{dx}, \quad \text{where } 0 < \rho < 1.$$

Take a full step along the direction $\boldsymbol{dx}$ and find the new boundary point:

$$x_{new} = x(k) + \boldsymbol{dx},$$

Evaluate the objective vector, v_{new}, at this point through $v_{new} = \mathbf{C}x_{new}$. If the value vector v_{new} is preferred to any of the objective vectors of the anchor points, remove the least preferred anchor point from the current list and replace it with x_{new}. If termination conditions are met, ***STOP***; otherwise, Goto *Step 2*.

■

Example 13.2: Consider the problem [14] described through

$$\begin{cases} \max\ f_1 = x_1 \\ \max\ f_2 = x_2 \end{cases}$$

$$s.t.$$

$$\begin{cases} 8x_1 + 6x_2 \ge 112 \\ 5x_1 + 7x_2 \ge 96 \\ x_1 + x_2 \le 18 \\ 5x_1 + 7x_2 \ge 0 \end{cases}$$

Where: $x_0 = [13 \quad 4.5 \quad 19 \quad 0.5 \quad 0.5]^T$, and $u(x) = x_1 x_2$. The feasible region for this problem is depicted in Fig. 13.8.

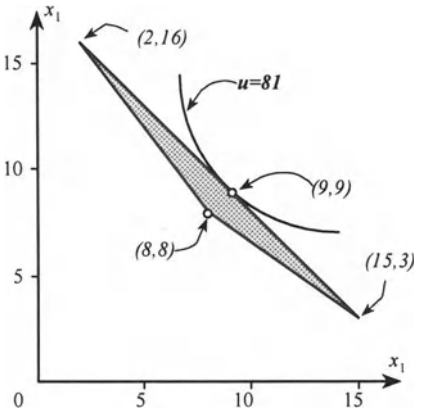


Figure 13.8: Feasible region

Performing the steps of the proposed MOLP algorithm, we generate a sequence of interior iterates summarized in Table 13.2 below. Note that in lieu of an interaction with a DM we have used the postulated utility function.

Table 13.2, Summary of iterations (ρ = 0.4)

k	$x_1(k)$	$x_2(k)$	$u_{current}$	u_{anchor}
0	13.0000	4.5000	58.5000	45.0001
1	11.1799	6.3796	71.3235	45.0001
2	9.9388	7.7369	76.8955	74.6191
3	9.4939	8.3303	79.0869	80.9989
4	8.4007	9.5149	79.9322	80.9989
5	8.3498	9.6099	80.2414	80.9989
6	8.5744	9.4064	80.6548	80.9989
7	8.7535	9.2374	80.8594	80.9989
8	8.8328	9.1629	80.9336	80.9989
9	8.8518	9.1461	80.9596	80.9989
10	8.8566	9.1424	80.9706	80.9989

The plot of the utility value at the current iterate and that of the best anchor point is shown in Fig. 13.9 below. In addition, we show the plot of the utility value at the current iterate when no anchor points are used.

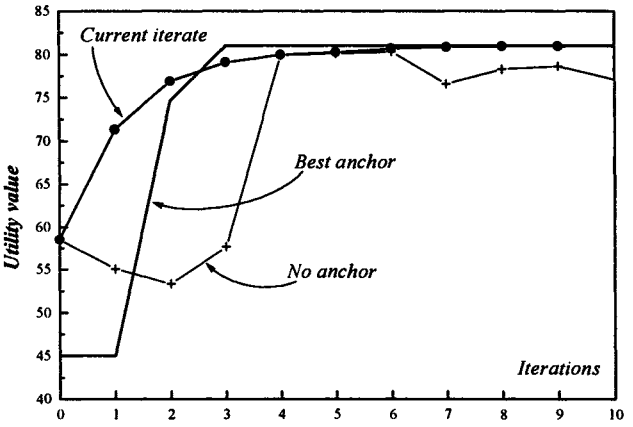


Figure 13.9: Results with Anchors (ρ=0.4)

■

This concludes the development of our first interior-point MOLP algorithm. The resulting approach uses the DM to generate local preference information to guide the generation of interior solution iterates. No restriction was placed on the step size taken when moving from one interior iterate to another. It can be as small or as large as desired. Paying closer attention to the issue of step size and its implications, another MOLP approach can be developed. In this approach one steps along an estimate to the gradient of the utility function at the current iterate. For the approximation to be valid, small steps have to be taken. We do not pursue the development of this variant here. Interested readers should refer to [3].

13.6 Projecting Aspirations

The MOLP approach presented thus far relies on the DM to generate local preference information that guides the interior trajectory of solution iterates toward the final compromise solution. There is another way for using the DM and that is by stating aspirations for values of the objectives. Doing so, allows one to generate an interior solution of iterates that moves toward the projection of the stated aspiration in objective space on the efficient frontier in decision space. This is the subject of discussion in this section.

Searching the set of nondominated solutions is greatly aided by means of the so-called *achievement scalarizing function* suggested by Wierzbicki [17-18]. Such a function projects any given (feasible or infeasible) point in objective space onto the set of (weakly) nondominated solutions. The simplest form proposed by Wierzbicki is the one given by

$$s(g, v, w) = \max_{1 \le i \le q} \left\{ \frac{g_i - v_i}{w_i} \right\}, \tag{13.51}$$

where $\boldsymbol{w} > \boldsymbol{0}$, $(\boldsymbol{w} \in \mathcal{R}^q)$ is a vector of weights, $\boldsymbol{g} \in \mathcal{R}^q$, is a vector of stated aspirations in objective space, $\boldsymbol{v} \in \boldsymbol{V} \equiv \{\boldsymbol{v} \in \mathcal{R}^q | \boldsymbol{v} = \boldsymbol{C}\boldsymbol{x},\ \boldsymbol{x} \in \boldsymbol{S}\}$. We show next that by minimizing $s(\boldsymbol{g}, \boldsymbol{v}, \boldsymbol{w})$ subject to $\boldsymbol{v} \in \mathbf{V}$, we obtain a solution $\boldsymbol{v}^*$, which is weakly nondominated. We start by noting that for all $\boldsymbol{v} \in \mathbf{V}$, we have by definition

$$s(\boldsymbol{g}, \boldsymbol{v}^*, \boldsymbol{w}) \leq s(\boldsymbol{g}, \boldsymbol{v}, \boldsymbol{w})$$

from which it follows that

$$\left\{\frac{g_i - v_i^*}{w_i}\right\} \leq \left\{\frac{g_i - v_i}{w_i}\right\}, \text{ for at least one } i \in \{1,2,\cdots,q\}.$$

This, in turn, implies that $v_i \leq v_i^*$, for at least one $i \in \{1,2,\cdots,q\}$. Hence, it follows that there exists no other $\boldsymbol{v} \in \mathbf{V}$, such that $\boldsymbol{v} > \boldsymbol{v}^*$, implying that $\boldsymbol{v}^*$ is weakly nondominated. If the given aspiration point $\boldsymbol{g} \in \mathcal{R}^q$ is feasible, then for a solution $\boldsymbol{v}^* \in \mathbf{N}^w$, we have $\boldsymbol{v}^* \geq \boldsymbol{g}$. Note that if the solution is unique, $\boldsymbol{v}^* \in \mathbf{N}$, otherwise it may be dominated. To generate only nondominated (instead of weakly nondominated) solutions, more complicated forms for the achievement scalarizing function should be used to guarantee uniqueness. For simplicity, we assume in the discussion that follows that the solution $\boldsymbol{v}^*$ is always nondominated.

To illustrate the use of an achievement scalarizing function, let us assume that we have a two-criterion problem with a feasible region having three extreme points located at {(0,0), (0,8), (16,0)}, as shown in Fig. 13.10 below. Let us illustrate the projection of different aspiration points in objective space. We consider both the case of feasible aspirations as well as that of infeasible aspirations. Let us assume that the DM first specifies a feasible aspiration level point $\boldsymbol{g}^A = [2 \quad 2]^T$, as denoted by point A below. Using a weight vector $\boldsymbol{w} = [2 \quad 1]^T$, the minimum of the achievement scalarizing function is reached at a point $\boldsymbol{v}^A = [7 \quad 4.5]^T$.

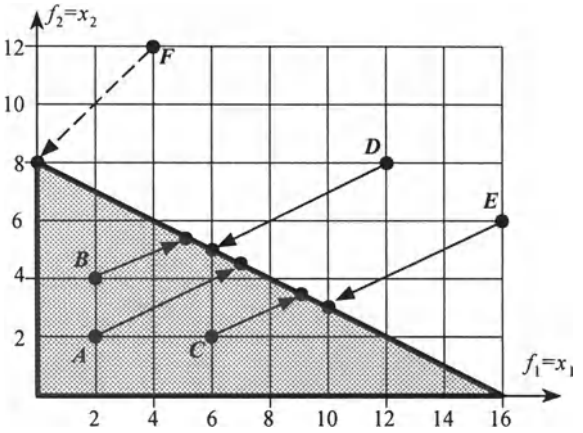


Figure 13.10: Projecting feasible and infeasible aspiration points.

Similarly, if an aspiration point is infeasible, say $g^D = [12 \quad 8]^T$, then the minimum of the achievement scalarizing function is reached at point $v^A = [6 \quad 5]^T$. The value of the scalarizing function is -1 in the first case and +1 in the second. If the aspiration point is dominated by a feasible point, then the value of scalarizing function is always negative; otherwise it is non-negative. It is zero, if an aspiration level point is weakly-nondominated. Additional aspiration points in objective space and their respective projections are shown in Fig. 13.10.

To project an aspiration point requires the minimization of an achievement scalarizing function such as described in (13.51). Doing so results in the following min-max problem

$$\min_{x \in S}\{s(g, v, w)\} = \min_{x \in S}\left\{\max_{1 \le i \le q}\left(\frac{g_i - v_i}{w_i}\right)\right\} \tag{13.52}$$

It is readily shown that the problem in (13.52) is equivalent to the problem described through

$$\begin{aligned} &min \quad \alpha \\ &s.t. \quad \mathbf{C}x + \alpha w \ge g \\ &\qquad x \in S, \quad \alpha \, free. \end{aligned} \tag{13.53}$$

Subtracting a surplus vector, z, we find

$$\begin{aligned} &min \quad \alpha \\ &s.t. \quad \mathbf{C}x + \alpha w - z = g \\ &\qquad x \in S, \quad \alpha \, free, \quad z \ge 0. \end{aligned} \tag{13.54}$$

Letting an augmented solution vector be defined through $\bar{x} = [x \quad \alpha \quad z]^T$, the problem in (13.54) is now compactly described through

$$\begin{aligned} &min \quad \alpha \\ &s.t. \quad \bar{\mathbf{A}}\bar{x} = \bar{b} \\ &\qquad x \ge 0, \; z \ge 0, \\ &\qquad\qquad \alpha \, free. \end{aligned} \tag{13.55}$$

where

$$\bar{\mathbf{A}} = \left[\begin{array}{c|c|c} \mathbf{A} & 0 & \mathbf{O} \\ \hline \mathbf{C} & w & -\mathbf{I}_q \end{array}\right], \qquad \bar{b} = \left[\begin{array}{c} b \\ \hline g \end{array}\right]$$

The original MOLP-problem is now reduced to a single-objective optimization problem. The augmented system is amenable now to the application of a regular single-objective interior-point algorithm such as the affine-scaling primal algorithm described earlier. The procedure starts by asking the DM to state an aspiration. Starting from an interior solution, the algorithm generates a few interior iterates that moves the interior trajectory of iterates closer to the boundary. As we move closer,

the DM has the option of varying the original stated aspiration. Doing so, changes the structure of the augmented system in (13.55) and, as a result, the interior iterates will deviate from their starting trajectory. If the DM keeps the current aspiration unchanged, the solution process continues unhindered until it reaches a nondominated solution. Using an interior algorithm for the augmented system of (13.55) obviates the need for many pivot operations and going through many intermediate solution steps requiring DM's intervention. In addition, starting deep in the interior of the feasible region, allows one to generate a win-win trajectory of iterates along which all objective values improve as we move through the interior toward the boundary. It is our belief that presenting the DM with dominated solutions allows easier elicitation of preference that steers the solution trajectory in an overall desired direction. In fact, we believe that this class of algorithms where one moves through the interior may provide an easier setting for the DM to respond to when asked for the preference information inherent with every MOLP procedure.

Example 13.3:

We demonstrate our proposed approach with an example whose solution was reported earlier using simplex-based approaches [12]. Consider the MOLP problem described through:

$$\begin{aligned} &\max f_1(x) = x_1 \\ &\max f_2(x) = x_2 \\ &\max f_3(x) = x_3 \end{aligned}$$

subject to:

$$\begin{aligned} 3x_1 + 2x_1 + 3x_3 &\le 18 \\ x_1 + 2x_2 + x_3 &\le 10 \\ 9x_1 + 20x_2 + 7x_3 &\le 96 \\ 7x_1 + 20x_2 + 9x_3 &\le 96 \\ x_1, \quad x_2, \quad x_3 &\ge 0 \end{aligned}$$

Adding the necessary slack variables, this problem becomes

$$\begin{aligned} &\text{"maximize"}\ \mathbf{C}\boldsymbol{x} \\ &s.t. \quad \mathbf{A}\boldsymbol{x} = \boldsymbol{b} \\ &\qquad \boldsymbol{x} \ge \boldsymbol{0}, \end{aligned}$$

where:

$$\mathbf{A} = \begin{bmatrix} 3 & 2 & 3 & 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 1 & 0 & 0 \\ 9 & 20 & 7 & 0 & 0 & 1 & 0 \\ 7 & 20 & 9 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 18 \\ 10 \\ 96 \\ 96 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Converting this system of constraints to the form required by our proposed approach results in the augmented system given by

$$\begin{aligned} &min \quad \alpha \\ &s.t. \quad \bar{\mathbf{A}}\bar{\boldsymbol{x}} = \bar{\boldsymbol{b}} \\ &\qquad \boldsymbol{x} \geq \boldsymbol{0},\ \boldsymbol{z} \geq \boldsymbol{0}, \\ &\qquad\quad \alpha \text{ free}. \end{aligned}$$

where $\bar{\boldsymbol{x}} = [\boldsymbol{x} \quad \alpha \quad \boldsymbol{z}]^T$, $\boldsymbol{x} \in \mathcal{R}^7, \boldsymbol{z} \in \mathcal{R}^3$, $\boldsymbol{b} \in \mathcal{R}^7$, and the matrix $\bar{\mathbf{A}}$ and the vector $\bar{\boldsymbol{b}}$ are given by

$$\bar{\mathbf{A}} = \left[\begin{array}{c|c|c} \mathbf{A} & 0 & \mathbf{O} \\ \hline \mathbf{C} & \boldsymbol{w} & -\mathbf{I}_3 \end{array}\right], \quad \bar{\boldsymbol{b}} = \left[\begin{array}{c} \boldsymbol{b} \\ \hline \boldsymbol{g} \end{array}\right]$$

The optimization problem is now *min* α, subject to the augmented system of constraints. Assuming a utility function given by $u(\boldsymbol{x}) = \min(3x_1, 5x_2, 3x_3)$, leads to a unique solution given by $\boldsymbol{x} = [2.5\ 1.5\ 2.5]^T$. Note that this solution is on a face of a polytope and not at an extreme point. To test our algorithm, we set our aspiration level at $\boldsymbol{g} = [3.5\ 2.5\ 3.5]^T$, and run the proposed algorithm with a step size factor of $\rho = 0.9$. The results of the iterative process, for an initial vector defined by $\boldsymbol{x}_0 = [1\ 1\ 1\ 10\ 6\ 60\ 60]^T$, are shown in Table 13.3. For brevity we show only the original decision vector's components and exclude the slack and surplus variables. We also evaluate the relative duality gap during the iterative process and show it in the last column of the table.

Table 13.3, Solution results ($\rho = 0.9$)

k	x_1	x_2	x_3	$\boldsymbol{c}^T\boldsymbol{x}$	*Gap*
0	1.0000	1.0000	1.0000	10.000	–
1	1.0851	1.0431	1.0851	3.1649	0.2958
2	2.0440	1.1160	2.0440	1.5548	0.2554
3	2.4455	1.4883	2.4455	1.1424	0.0541
4	2.4924	1.5053	2.4924	1.0355	0.0213
5	2.4967	1.4989	2.4967	1.0061	0.0032
6	2.4994	1.5006	2.4994	1.0024	0.0012
7	2.4998	1.4998	2.4998	1.0005	0.0002
8	2.5000	1.5000	2.5000	1.0002	0.0001
9	2.5000	1.5000	2.5000	1.0000	0.0000
10	2.5000	1.5000	2.5000	1.0000	0.0000

We should point out, however, that the primal algorithm generates a sequence of iterates that provide a descent direction for the primal objective.

The plot of the decision variables for the first three components of the sequence of iterates is shown in Fig. 13.11.

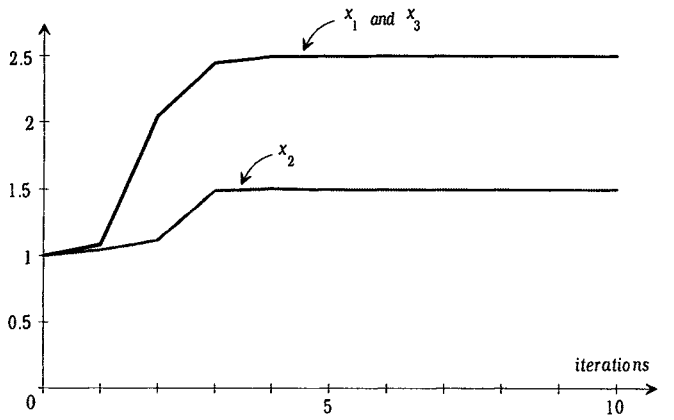


Figure 13.11: Solution trajectories for the decision vector's components ($\rho = 0.9$)

Next we run the algorithm with a series of aspiration levels which terminate with the one leading to the true optimal solution. We run 5 iterations with each aspiration level vector before asking the DM to reconsider new aspiration levels. We assume that the DM is willing to specify aspiration levels three times, i.e., at the first, 6-*th*, and 11-*th* iterations. The specific levels used are given by

$$\boldsymbol{g}_1 = [0.5\ 2.5\ 3.5]^T,$$

$$\boldsymbol{g}_6 = [3.25\ 1.25\ 3.75]^T,$$

$$\boldsymbol{g}_{11} = [3.5\ 2.5\ 3.5]^T.$$

Solution results for this series of aspirations is summarized in Table 13.4, and plotted in Fig. 13.12.

Note how the solution trajectory reacts to the change in aspirations at the proper intervals. At the 16-th iteration the DM is not willing anymore to specify new aspiration levels. Therefore, we continue the procedure until the duality gap falls below a pre-specified threshold (usually 10^{-6}). A good estimate for a nondominated solution is found. In this example, we use $\rho = 0.6$. By varying ρ, we can control how close to the boundary we allow the DM to move, before new aspiration levels are asked. If ρ is close to one, the system may first make a "U-turn" which moves the solution trajectory away from the boundary, when new aspiration levels are specified, before it approaches the boundary. Observing this kind of "U-turn" may be confusing to the DM. On the other hand, small ρ increases the number of iterations. At the moment, the specification of the step size factor, ρ, is one possible direction for future research.

Table 13.4, Solution results ($\rho = 0.6$)

k	x_1	x_2	x_3	$c^T x$	*Gap*
0	1.0000	1.0000	1.0000	10.0000	—
1	1.0135	1.0670	1.0253	6.3330	0.5936
2	1.0052	1.2624	1.0393	4.6976	0.3822
3	0.9247	1.8693	0.9974	3.5147	0.2186
4	0.7062	3.0969	0.8775	2.1274	0.2187
5	0.5012	3.6569	0.9006	1.5013	0.3023
6	0.5166	3.6314	0.9508	5.6595	0.5725
7	0.5609	3.5515	1.0960	3.8595	0.2931
8	0.6960	3.3530	1.4170	3.0222	0.0776
9	1.1961	2.8923	1.9538	2.2412	0.0999
10	1.6528	2.4902	2.3183	1.7548	0.0111
11	1.6740	2.4867	2.3078	1.4556	0.7030
12	1.7631	2.4692	2.2512	1.0925	0.5052
13	1.9773	2.4058	2.1296	0.8859	0.2888
14	2.1517	2.2409	2.1542	0.7980	0.1260
15	2.3365	1.8484	2.3360	0.7525	0.0109
16	2.4314	1.5953	2.4312	5.1409	0.0257
17	2.4697	1.5115	2.4696	3.1660	0.0392
18	2.4865	1.4988	2.4865	2.2714	0.0210
19	2.4944	1.5002	2.4944	1.5680	0.0086
20	2.4977	1.5000	2.4977	1.2599	0.0034

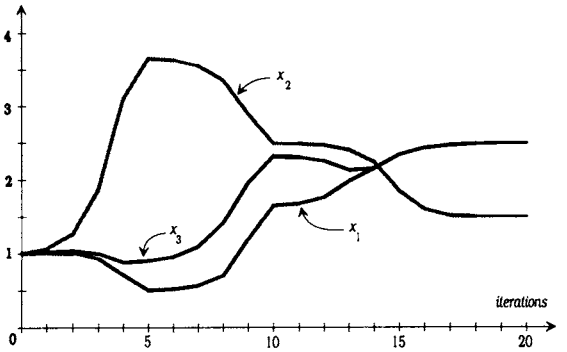


Figure 13.12: Solution trajectories for the decision vector's components ($\rho = 0.6$)

13.7 Summary

We have discussed in this paper the modification needed for adopting single-objective interior-point algorithm to MOLP problems. Specifically, we have described in considerable detail one variant of an interior-point algorithm termed the affine-scaling primal algorithm. We have discussed the basic questions one faces

when moving from a single-objective setting into a multiple-objective setting. We have developed two rather complete approaches for MOLP problems using an interior algorithm and have demonstrated these algorithms with numerical examples. Future work in this area should continue to explore ways of adopting interior algorithms to MOLP problems, compare relative performance issues of simplex-based MOLP algorithms and interior MOLP. Regardless of the specific topic one chooses to explore, we are certain that the use of interior-point algorithm will command a prominent position in the future of MOLP algorithm similar to that experience in the linear programming research literature.

References

[1] Arbel, A., *Exploring Interior Point Linear Programming: Algorithms and Software*, MIT Press, Cambridge, MA, 1993

[2] Arbel, A.: An Interior Multiobjective Linear Programming Algorithm. Computers and Operations Research 20(1993)723-735

[3] Arbel, A., S.S. Oren: Generating Interior Search Directions for Multiobjective Linear Programming. Journal of Multicriteria Decision Analysis 2(1993)73-86

[4] Arbel, A.: Anchoring Points and Cones of Opportunities in Interior Multiobjective Linear Programming Problems. Journal of the Operational Research Society 45(1994)83-96

[5] Arbel, A.: A Multiobjective Interior Primal-Dual Linear Programming Algorithm. Computers and Operation Research, 21(1994)433-445

[6] Arbel, A., P. Korhonen: Using Aspiration Levels in an Interactive Interior Multiobjective Linear Programming Algorithm. European Journal of Operational Research 89(1996)193-201

[7] Arbel, A., P. Korhonen: Using Aspiration Levels in an Interior Primal-Dual Multiobjective Linear Programming Algorithm. Journal of Multicriteria Decision Analysis 5(1996)61-71

[8] Chankong, V., Y.Y. Haimes: *Multiobjective Decision Making: Theory and Methodology*, North-Holland, 1983

[9] Gal, T.: *Postoptimal Analyses, Parametric Programming, and Related Topics*, 2nd/3rd edition, de Gruyter, Berlin 1996, Chapter 9

[10] Karmarkar, N.K.: A new polynomial time algorithm for linear programming. Combinatorica 4(1984)373-395

[11] Korhonen, P., J. Laakso: A Visual Interactive Method for Solving the Multiple Criteria Problem. European Journal of Operational Research 24(1986)277-287

[12] Korhonen, P., J. Wallenius: A Pareto Race. Naval Research Logistics 35(1988)615-623

[13] Shin, W.S., A. Ravindran: Interactive Multiobjective Optimization Survey I - Continuous Case. Computers and Operations Research 18(1991)97-114

[14] Steuer, R.E.: *Multiple Criteria Optimization: Theory, Computation, and Application*, Wiley and Sons, New York, 1986

[15] Vanderbei, R.J., M.S. Meketon, B.A. Freedman: A modification of Karmarkar's linear programming algorithm. Algorithmica 1(1986)395-407

[16] Vanderbei, R. J.: *Linear Programming*, Kluwer, 1996

[17] Wierzbicki, A.: "The Use of Reference Objectives in Multiobjective Optimization." In: G. Fandel, T. Gal, (Eds.): *Multiple Criteria Decision Making: Theory and Application.* Springer-Verlag, Berlin, 1980, pp. 468-486

[18] Wierzbicki, A.: On the Completeness and Constructiveness of Parametric Characterizations to Vector Optimization Problems, OR-Spectrum, 8(1986)73-87

[19] Zeleny, M.: *Multiple Criteria Decision Making*, McGraw-Hill Book Co., New York, 1982

14 THE USE OF ROUGH SETS AND FUZZY SETS IN MCDM

Salvatore Greco, Benedetto Matarazzo, Roman Slowinski

CONTENTS

T. Gal et al. (eds.), *Multicriteria Decision Making*

Abstract: The rough sets theory has been proposed by Z. Pawlak in the early 80's to deal with inconsistency problems following from information granulation. It operates on an information table composed of a set U of objects (actions) described by a set Q of attributes. Its basic notions are: indiscernibility relation on U, lower and upper approximation of a subset or a partition of U, dependence and reduction of attributes from Q, and decision rules derived from lower approximations and boundaries of subsets identified with decision classes. The original rough sets idea has proved to be particularly useful in the analysis of multiattribute classification problems; however, it was failing when preferential ordering of attributes (criteria) had to be taken into account. In order to deal with problems of multicriteria decision making (MCDM), like sorting, choice or ranking, a number of methodological changes to the original rough sets theory were necessary. The main change is the substitution of the indiscernibility relation by a dominance relation (crisp or fuzzy), which permits approximation of ordered sets in multicriteria sorting. In order to approximate preference relations in multicriteria choice and ranking problems, another change is necessary: substitution of the information table by a pairwise comparison table, where each row corresponds to a pair of objects described by binary relations on particular criteria. In all those MCDM problems, the new rough set approach ends with a set of decision rules, playing the role of a comprehensive preference model. It is more general than the classic functional or relational model and it is more understandable for the users because of its natural syntax. In order to workout a recommendation in one of the MCDM problems, we propose exploitation procedures of the set of decision rules. Finally, some other recently obtained results are given: rough approximations by means of similarity relations (crisp or fuzzy) and the equivalence of a decision rule preference model with a conjoint measurement model which is neither additive nor transitive.

14.1 A general view of rough sets

14.1.1 Introduction

The rough sets theory introduced by Z. Pawlak [58, 61] has often proved to be an excellent mathematical tool for the analysis of a vague description of objects (called actions in decision problems). The adjective vague, referring to the quality of information, means inconsistency or ambiguity which follows from information granulation. The rough sets philosophy is based on the assumption that with every object of the universe there is associated a certain amount of information (data, knowledge), expressed by means of some attributes used for object description. For example, if the objects are firms applying for a bank mortgage, the information given concerns their financial, economic and technical characteristics, that constitute their description. Objects having the same description are indiscernible (similar) with respect to the available information. The *indiscernibility relation* thus generated constitutes a mathematical basis of the rough sets theory; it induces a partition of the universe into blocks of indiscernible objects, called elementary sets, that can be used as "bricks" to build knowledge about a real or abstract world. The use of the indiscernibility relation results in information granulation.

Any subset X of the universe may be expressed in terms of these bricks either precisely (as a union of elementary sets) or approximately only. In the latter case, the subset X may be characterized by two ordinary sets, called *lower* and *upper approximations*. A rough set is defined by means of these two approximations, which coincide in the case of an ordinary set. The lower approximation of X is composed of all the elementary sets included in X (whose elements, therefore, certainly belong to X), while the upper approximation of X consists of all the elementary sets which have a non-empty intersection with X (whose elements, therefore, may belong to X). Obviously, the difference between the upper and lower approximations constitutes the boundary region of the rough set, whose elements cannot be characterized with certainty as belonging or not to X, using the available information. The information about objects from the boundary region is, therefore, inconsistent or ambiguous. Clearly, in ordinary sets the boundary region is empty. The cardinality of the boundary region states, moreover, to what extent it is possible to express X in exact terms, on the basis of the available information. For this reason, this cardinality may be used as a measure of vagueness of the information about X.

The rough sets theory, dealing with representation and processing of vague information, presents a series of intersections and complements with respect to many other theories and mathematical techniques dealing with imperfect information, like probability theory, evidence theory of Dempster-Shafer, fuzzy sets theory, discriminant analysis and mereology (see [7, 8, 43, 59, 60, 64, 76, 86]).

Some important characteristics of the rough set approach make of this a particularly interesting tool in a number of problems and concrete applications. With respect to the input information, it is possible to deal with both quantitative and qualitative data and inconsistencies need not to be removed prior to the

analysis. With reference to the output information, it is possible to acquire a posteriori information regarding the relevance of particular attributes and their subsets to the quality of approximation considered in the problem at hand, without any additional inter-attribute preference information. Moreover, the final result in the form of "*if..., then...*" decision rules, using the most relevant attributes, is easy to interpret.

14.1.2 The information table and indiscernibility relation

For algorithmic reasons, the information regarding the objects is supplied in the form of an *information table*, whose separate rows refer to distinct *objects* (actions), and whose columns refer to the different *attributes* considered. Each cell of this table indicates, therefore, an *evaluation* (quantitative or qualitative) of the object placed in that row by means of the attribute in the corresponding column. In the case of quantitative evaluations on an attribute q, the domain of this attribute is suitably divided into subintervals and then codified, e.g. by natural numbers. This pre-processing of data, called *discretization*, is commonly used in machine learning in order to get a description of the phenomenon studied without noisy details. The problem of discretization is rather delicate, since the results of the analysis depend, in general, on the way of discretization. To this purpose suitable techniques and methodologies have been proposed (see, for example, [81, 10, 5, 56, 95]).

Formally, an *information table* is the 4-tuple $S = \langle U,Q,V,f \rangle$, where U is a finite set of *objects* (universe), $Q=\{q_1,q_2,...,q_m\}$ is a finite set of *attributes*, V_q is the domain of the attribute q, $V = \bigcup_{q \in Q} V_q$ and $f:U \times Q \rightarrow V$ is a total function such that $f(x,q) \in V_q$ for each $q \in Q$, $x \in U$, called *information function*.

Therefore, each object x of U is described by a vector (string) $Des_Q(x)=[f(x,q_1),f(x,q_2),...,f(x,q_m)]$, called *description* of x in terms of the evaluations of the attributes from Q; it represents the available information about x. Obviously, $x \in U$ can be described in terms of any non-empty subset $P \subseteq Q$.

To every (non-empty) subset of attributes P is associated an *indiscernibility relation* on U, denoted by I_P:

$$I_P = \{(x,y) \in U \times U: f(x,q) = f(y,q), \ \forall q \in P\}.$$

If $(x,y) \in I_P$, it is said that the objects x and y are P-indiscernible. Clearly, the indiscernibility relation thus defined is an equivalence relation (reflexive, symmetric and transitive). The family of all the equivalence classes of the relation I_P is denoted by $U|I_P$ and the equivalence class containing an element $x \in U$ by $I_P(x)$. The equivalence classes of the relation I_P are called P-*elementary sets*. If P = Q, the Q-elementary sets are called *atoms*.

14.1.3 Approximations

Let S be an information table, X a non-empty subset of U and $\varnothing \neq P \subseteq Q$. The P-*lower approximation* and the P-*upper approximation* of X in S are defined, respectively, by:

$$\underline{P}(X) = \{x \in U : I_P(x) \subseteq X\},$$
$$\overline{P}(X) = \bigcup_{x \in X} I_P(x).$$

The elements of $\underline{P}(X)$ are all and only those objects $x \in U$ which belong to the equivalence classes generated by the indiscernibility relation I_P, *contained* in X; the elements of $\overline{P}(X)$ are all and only those objects $x \in U$ which belong to the equivalence classes generated by the indiscernibility relation I_P, *containing at least one* object x belonging to X. In other words, $\underline{P}(X)$ is the largest union of the P-elementary sets included in X, while $\overline{P}(X)$ is the smallest union of the P-elementary sets containing X.

The P-*boundary* of X in S, denoted by $Bn_P(X)$, is

$$Bn_P(X) = \overline{P}(X) - \underline{P}(X).$$

The following relation holds: $\underline{P}(X) \subseteq X \subseteq \overline{P}(X)$.

Therefore, if an object x belongs to $\underline{P}(X)$, it is certainly also an element of X, while if x belongs to $\overline{P}(X)$, it may belong to the set X. $Bn_P(X)$ constitutes the "doubtful region" of X: nothing can be said with certainty about the belonging of its elements to the set X.

The following relation, called *complementarity property*, is satisfied: $\underline{P}(X) = U - \overline{P}(U - X)$.

If the P-boundary of X is empty, $Bn_P(X) = \varnothing$, then the set X is an ordinary (exact) set with respect to P, that is, it may be expressed as the union of a certain number of P-elementary sets; otherwise, if $Bn_P(X) \neq \varnothing$, the set X is an approximate (rough) set with respect to P and may be characterized by means of the approximations $\underline{P}(X)$ and $\overline{P}(X)$. The family of all the sets $X \subseteq U$ having the same P-lower and P-upper approximations is called a *rough set.*

We define the following ratio as *accuracy* of the approximation of X, $X \neq \varnothing$, by means of the attributes from P

$$\alpha_P(X) = \frac{|\underline{P}(X)|}{|\overline{P}(X)|},$$

where |Y| indicates the cardinality of a (finite) set Y. The result is, obviously, $0 \le \alpha_P(X) \le 1$; if $\alpha_P(X)=1$, X is an ordinary (exact) set with respect to P; if $\alpha_P(X)<1$, X is a *rough* (vague) set with respect to P.

We also define a *quality* of the approximation of X by means of the attributes from P as

$$\gamma_P(X) = \frac{|\underline{P}(X)|}{|X|}.$$

The quality $\gamma_P(X)$ represents the relative frequency of the objects correctly classified using the attributes from P. Moreover, we have $0 \le \alpha_P(X) \le \gamma_P(X) \le 1$, $\gamma_P(X)=0$ iff $\alpha_P(X)=0$ and $\gamma_P(X)=1$ iff $\alpha_P(X)=1$.

The definition of approximations of a subset $X \subseteq U$ can be extended to a classification, i.e. a partition $Y = \{Y_1,...,Y_n\}$ of U. Subsets Y_i, i=1,...,n, are disjunctive classes of *Y*. By P-lower (P-upper) approximation of *Y* in S we mean sets $\underline{P}Y = \{\underline{P}Y_1,...,\underline{P}Y_n\}$ and $\overline{P}Y = \{\overline{P}Y_1,...,\overline{P}Y_n\}$, respectively. The coefficient

$$\gamma_P(Y) = \frac{\sum_{i=1}^{n} |\underline{P}Y_i|}{|U|}$$

is called *quality of the approximation of classification Y* by set of attributes P, or in short, *quality of classification*. It expresses the ratio of all P-correctly classified objects to all objects in the system.

The main preoccupation of the rough sets theory is approximation of subsets or partitions of U, representing a knowledge about U, with other sets or partitions built up using available information about U. From the viewpoint of a particular object $x \in U$, it may be interesting, however, to use the available information to assess the degree of its membership to a subset X of U. The subset X can be identified with a concept of knowledge to be approximated. Using the rough set approach one can calculate the membership function $\mu_X^P(x)$ (*rough membership function*) as

$$\mu_X^P(x) = \frac{|X \cap I_P(x)|}{|I_P(x)|}.$$

The value of $\mu_X^P(x)$ may be interpreted analogously to conditional probability and may be understood as the *degree of certainty* (credibility) to which x belongs to X. Observe that the value of the membership function is calculated from the available data, and not subjectively assumed, as it is the case of membership functions of fuzzy sets.

Between the rough membership function and the approximations of X the following relationships are valid:

$$\underline{P}(X)=\{x\in U:\ \mu_X^P(x)=1\},$$
$$\overline{P}(X)=\{x\in U:\ \mu_X^P(x)>0\},$$
$$Bn_p(X)=\{x\in U:\ 0<\mu_X^P(x)<1\},$$
$$\underline{P}(U-X)=\{x\in U:\ \mu_X^P(x)=0\}.$$

In the rough sets theory there is, therefore, a close link between vagueness (granularity) connected with rough approximation of sets and uncertainty connected with rough membership of objects to sets.

14.1.4 Dependence and reduction of attributes

A very important concept for concrete applications is that of dependence of attributes. Intuitively, a set of attributes $T\subseteq Q$ *totally depends* on a set of attributes $P\subseteq Q$ (notation $P\rightarrow T$) if all the values of the attributes from T are uniquely determined by the values of the attributes from P, that is, if a functional dependence exists between evaluations by the attributes from P and by the attributes from T. In other words, the partition generated by the attributes from P is "finer" than that generated by the attributes from T, so that it is sufficient to use the attributes from P to build the partition $U|I_T$. Formally, T totally depends on P iff $I_P\subseteq I_T$.

Therefore, T is totally (partially) dependent on P if all (some) elements of the universe U may be univocally assigned to classes of the partition $U|I_T$, using only the attributes from P.

Another issue of great practical importance is that of "superfluous" data in an information table. Superfluous data can be eliminated, in fact, without deteriorating the information contained in the original table.

Let $P\subseteq Q$ and $p\in P$. It is said that attribute p is *superfluous* in P if $I_P=I_{P-\{p\}}$; otherwise, p is *indispensable* in P.

The set P is *independent* (orthogonal) if all its attributes are indispensable. The subset P' of P is a *reduct* of P (denotation Red(P)) if P' is independent and $I_{P'}=I_P$.

A reduct of P may also be defined with respect to an approximation of a partition Y of U. It is then called *Y-reduct* of P (denotation $Red_Y(P)$) and specifies a minimal subset P' of P which keeps the quality of classification unchanged, i.e. $\gamma_{P'}(Y)=\gamma_P(Y)$. In other words, the attributes that do not belong to Y-reduct of P are superfluous with respect to the classification Y of objects from U.

More than one Y-reduct (or reduct) of P may exist in an information table. The set containing all the indispensable attributes of P is known as the *Y-core*. Formally,

$$Core_Y(P)=\bigcap Red_Y(P).$$

Obviously, since the Y-core is the intersection of all the Y-reducts of P, it is included in every Y-reduct of P. It is the most important subset of attributes from P, because none of its elements can be removed without deteriorating the quality of classification.

The calculation of all the reducts is fairly complex (see [2, 44, 78, 96]). Nevertheless, in many practical applications it is not necessary to calculate all the reducts, but only some of them. For example, in [82], the following heuristic procedure has been used to obtain the most satisfactory reduct. Starting from single attributes, the one with the greatest quality of classification is chosen; then to the chosen attribute, another attribute is appended that gives the greatest increase to the quality of classification for the pair of attributes; then yet another attribute is appended to the pair giving the greatest increase to the quality of classification for the triple, and so on, until the required quality is reached by a subset of attributes. Thus, for further analysis, it is often sufficient to take into consideration a reduced information table, where the set Q of attributes is confined to the most satisfactory reduct.

14.1.5 Decision table and decision rules

If in an information table the attributes of set Q are divided into *condition* attributes (set $C\neq\emptyset$) and *decision* attributes (set $D\neq\emptyset$), $C\cup D=Q$ and $C\cap D=\emptyset$, such a table is called a *decision table*. The decision attributes induce a partition of U deduced from the indiscernibility relation I_D in a way that is independent of the condition attributes. D-elementary sets are called *decision classes*. There is a tendency to reduce the set C while keeping all important relationships between C and D, in order to make decisions on the basis of a smaller amount of information. When the set of condition attributes is replaced by one of its reducts, the quality of approximation of the classification induced by the decision attributes is not deteriorating.

Since the tendency is to underline the functional dependencies between condition and decision attributes, a decision table may also be seen as a set of *decision rules*. These are logical statements (implications) of the type "*if*..., *then*...", where the antecedent (condition part) specifies values assumed by one or more condition attributes (description of C-elementary sets) and the consequence (decision part) specifies an assignment to one or more decision classes (description of D-elementary sets). Therefore, the syntax of a rule is the following:

"*if* $f(x,q_1)$ is equal to r_{q1} *and* $f(x,q_2)$ is equal to r_{q2} *and* ... $f(x,q_p)$ is equal to r_{qp}, *then* x belongs to Y_{j1} *or* Y_{j2} *or* ... Y_{jk}",

where $\{q_1,q_2,...,q_p\}\subseteq C$, $(r_{q1},r_{q2},...,r_{qp})\in V_{q1}\times V_{q2}\times...\times V_{qp}$ and Y_{j1}, $Y_{j2},...,Y_{jk}$ are some decision classes of the considered classification (D-elementary sets). If the consequence is univocal, i.e. $k=1$, then the rule is *exact*, otherwise it is *approximate* or *uncertain*.

An object $x\in U$ *supports* decision rule r if its description is matching both the condition part and the decision part of the rule. We also say that decision rule r *covers* object x. Each decision rule is characterized by its *strength*, defined as the

number of objects covered by the rule. In the case of approximate rules, the strength is calculated for each possible decision class separately.

Let us observe that exact rules are supported only by objects from the lower approximation of the corresponding decision class. Approximate rules are supported, in turn, only by objects from the boundaries of the corresponding decision classes.

Procedures for generation of decision rules from a decision table use an inductive learning principle. The objects are considered as examples of decisions. In order to induce decision rules with a unique consequent assignment to a D-elementary set, the examples belonging to the D-elementary set are called *positive* and all the others *negative*. A decision rule is *discriminant* if it is consistent, i.e. distinguishes positive examples from negative ones, and *minimal* if removing any attribute from a condition part gives a rule covering also negative objects. It may be also interesting to look for *partly discriminant* rules. These are rules that, besides positive examples, could cover a limited number of negative ones. They are characterized by a coefficient, called *level of discrimination*, telling to what extent the rule is consistent, i.e. what is the ratio of positive examples to all examples covered by the rule.

Generation of decision rules from decision tables is a complex task and a number of procedures have been proposed to solve it (see, for example, [36, 37, 51, 75, 77, 88, 94, 105]). The existing induction algorithms use one of the following strategies:

(a) generation of a minimal set of rules covering all objects from a decision table,
(b) generation of an exhaustive set of rules consisting of all possible rules for a decision table,
(c) generation of a set of 'strong' decision rules, even partly discriminant, covering relatively many objects each but not necessarily all objects from the decision table.

14.1.6 Fuzzy measures and rough sets

Let $N=\{1,2,...,n\}$ be a finite set, whose elements could be players in a game, criteria in a multicriteria decision problem, attributes in an information table, etc., and let $P(N)$ denote the power set of N, i.e. the set of all subsets of N. A *fuzzy measure* on N is a set function μ: $P(N)\rightarrow[0,1]$ satisfying the following axioms:

1) $\mu(\varnothing)=0$, $\mu(N)=1$,
2) $A\subseteq B$ implies $\mu(A)\leq\mu(B)$, for all $A,B\in P(N)$.

In the following, the first axiom is relaxed by considering the condition $\mu(N)\leq 1$ instead of $\mu(N)=1$.

Within game theory, the function $\mu(A)$ is called characteristic function and represents the payoff obtained by the coalition $A\subseteq N$ in a cooperative game ([73,

1]); in a multicriteria decision problem, $\mu(A)$ can be interpreted as the conjoint importance of the criteria from $A \subseteq N$ [14].

Some indices have been introduced in game theory as specific solutions of cooperative games. The most important are the Shapley value and the Banzhaf value. The Shapley value [73] for every $i \in N$ is defined by

$$\phi_S(i) = \sum_{K \subseteq N-\{i\}} \frac{(n-|K|-1)!|K|!}{n!} [\mu(K \cup \{i\}) - \mu(K)].$$

The Banzhaf value ([1]) for every $i \in N$ is defined by

$$\phi_B(i) = \frac{1}{2^{n-1}} \sum_{K \subseteq N-\{i\}} [\mu(K \cup \{i\}) - \mu(K)].$$

The Shapley value and the Banzhaf value can be interpreted as specific kinds of weighted average contribution of element i alone to all coalitions. Let us remind that in the case of $\phi_S(i)$ the value of $\mu(N)$ is shared among the elements of N, i.e. $\sum_{i=1}^{n} \phi_S(i) = 1$, while an analogous property does not hold for $\phi_B(i)$.

The Shapley and the Banzhaf indices have also been proposed to represent the average importance of particular criteria within multicriteria decision analysis, when for the conjoint importance of criteria fuzzy measures are used [54]. In addition to the indices concerning particular criteria, other indices have been proposed to measure the interaction between pairs of criteria. *Interaction indices* have been suggested by Murofushi and Soneda [55] and by Roubens [66] with respect to Shapley index and Banzhaf index, respectively.
The Murofushi-Soneda interaction index for elements $i,j \in N$ is defined by

$$I_{MS}(i,j) = \sum_{K \subseteq N-\{i,j\}} \frac{(n-|K|-2)!|K|!}{(n-1)!} [\mu(K \cup \{i,j\}) - \mu(K \cup \{i\}) - \mu(K \cup \{j\}) + \mu(K)].$$

The Roubens interaction index for elements $i,j \in X$ is defined by

$$I_R(i,j) = \frac{1}{2^{n-2}} \sum_{K \subseteq N-\{i,j\}} [\mu(K \cup \{i,j\}) - \mu(K \cup \{i\}) - \mu(K \cup \{j\}) + \mu(K)].$$

The interaction indices $I_{MS}(i,j)$ and $I_R(i,j)$ can be interpreted as specific kinds of average added values resulting from putting i and j together in each coalition. The following cases can happen:

- $I_{MS}(i,j)>0$ ($I_R(i,j)>0$): i and j are complementary,
- $I_{MS}(i,j)<0$ ($I_R(i,j)<0$): i and j are substitutive,
- $I_{MS}(i,j)=0$ ($I_R(i,j)=0$): i and j are independent.

The definition of interaction indices can be extended from non-ordered pairs $i,j \in N$ to any subset $A \subseteq N$, $A \neq \emptyset$. Extensions of interaction indices in this sense

have been proposed by Grabisch [14] and Roubens [66], with respect to Shapley index and Banzhaf index, respectively.

The Shapley interaction index of elements from $A \subseteq N$ is defined by

$$I_S(A) = \sum_{K \subseteq N-A} \frac{(n-|K|-|A|)!|K|!}{(n-|A|+1)!} \sum_{L \subseteq A} (-1)^{|A|-|L|} \mu(L \cup K).$$

The Banzhaf interaction index of elements from $A \subseteq N$ is defined by

$$I_B(A) = \frac{1}{2^{n-|A|}} \sum_{K \subseteq N-A} \sum_{L \subseteq A} (-1)^{|A|-|L|} \mu(K \cup L).$$

In addition to the interaction indices, another concept useful for the interpretation of the fuzzy measures is the *Möbius representation* of μ, i.e. the set function $m: P(N) \rightarrow R$ defined by

$$m(A) = \sum_{B \subseteq A} (-1)^{|A-B|} \mu(B)$$

for any $A \subseteq N$. Within Dempster-Shafer theory of evidence [72], m(A) is interpreted as basic probability assignment.

The relations between fuzzy measures μ, interaction indices I_S and I_B and Möbius representation m have been extensively studied in [15, 6, 66, 16].

Interaction indices I_S and I_B and Möbius representation m can be used within rough set analysis to study the relative value of the information supplied by different attributes [23]. Considering the quality of classification as a fuzzy measure, we conclude that

1) the Shapley index $\phi_S(i)$ and the Banzhaf index $\phi_B(i)$ can be interpreted as measures of the contribution of attribute $i=1,...,n$ to the quality of approximation of the considered classification,
2) the Murofushi-Soneda interaction index $I_{MS}(i,j)$ and Roubens interaction index $I_R(i,j)$ can be interpreted as the average conjoint contribution of the non-ordered pair of attributes $i,j=1,...,n$, $i \neq j$, to the quality of classification when adjoined to all sets $K \subset C$ such that $K \cap \{i,j\} = \emptyset$,
3) the Shapley interaction index $I_S(A)$ and the Banzhaf interaction index $I_B(A)$ can be interpreted as the average conjoint contribution of the subset of attributes $A \subset C$ to the quality of classification when adjoined to all sets $B \subset C$ such that $B \cap A = \emptyset$,
4) the Möbius representation m(A) of μ can be interpreted as the conjoint contribution of the subset of attributes $A \subseteq C$ to the quality of classification.

All of these indices can be useful to study the informational dependence among the considered attributes and to choose the best reducts.

14.1.7 An example

The following example (based on Pawlak [62]) illustrates the concepts introduced above. In Table 14.1, six warehouses are described by means of four attributes:

- A_1, capacity of the sales staff,
- A_2, perceived quality of goods,
- A_3, high traffic location,
- A_4, warehouse profit or loss.

The components of the information table S are: $U=\{1,2,3,4,5,6\}$, $Q=\{A_1,A_2,A_3,A_4\}$, $V_1=\{\text{high, medium, low}\}$, $V_2=\{\text{good, medium}\}$, $V_3=\{\text{no, yes}\}$, $V_4=\{\text{profit, loss}\}$, the information function f(x,q), taking values $f(1,A_1)=\text{high}$, $f(1,A_2)=\text{good}$, and so on.

Table 14.1 Information table of the illustrative example

Warehouse	A_1	A_2	A_3	A_4
1	high	good	no	profit
2	medium	medium	no	loss
3	medium	medium	no	profit
4	low	medium	no	loss
5	medium	good	yes	loss
6	high	medium	yes	profit

Observe that each warehouse has a different description in terms of the attributes A_1, A_2, A_3 and A_4, so that they can be distinguished (discerned) by means of the information supplied by the attributes considered. Formally, the indiscernibility relation based on all four attributes is $I_Q=\{(1,1),(2,2),(3,3),(4,4),(5,5),(6,6)\}$ and, therefore, there is no two distinct warehouses x and y such that $(x,y)\in I_Q$. However, warehouses 2 and 3 are indiscernible in terms of the attributes from $P=\{A_1,A_2,A_3\}$, since they have the same values on the three attributes. Formally, the indiscernibility relation based on P is, thus, $I_P=\{(1,1),(2,2),(2,3),(3,2),(3,3),(4,4),(5,5),(6,6)\}$. Similarly, warehouses 2, 3 and 4 are indiscernible with reference to the attributes of $P'=\{A_2, A_3\}$, and so on, considering all the possible subsets of attributes from Q.

Each $P\subseteq Q$ determines a partition $U|I_P$ that groups in the corresponding equivalence classes the objects having the same description in terms of the attributes from P: e.g., for $P'=\{A_2,A_3\}$, $U|I_{P'}=\{\{1\},\{2,3,4\},\{5\},\{6\}\}$, and thus, $\{1\},\{2,3,4\},\{5\},\{6\}$ are the P'-elementary sets.

Suppose that, using the set of attributes $P=\{A_1,A_2,A_3\}$, we wish to approximate the set X of warehouses which have made a profit, i.e., $X=\{1,3,6\}$. Since $U|I_P=\{\{1\},\{2,3\},\{4\},\{5\},\{6\}\}$, the result is

$$\underline{P}(X) = \{1,6\},\ \overline{P}(X) = \{1,2,3,6\},\ Bn_P(X) = \{2,3\}.$$

The answer to the question whether it is possible to describe X by means of the information supplied by the attributes from P is not unique. Observe that the P-boundary $Bn_P(X)$ is not empty: warehouses 2 and 3, belonging to the P-boundary have the same description in terms of attributes considered, but warehouse 2 has suffered a loss while warehouse 3 has made a profit. Nevertheless, the P-lower approximation of X, $\underline{P}(X)$, is also not empty and it consists of warehouses 1 and 6, whose descriptions are different from those of all the warehouses not belonging to X. Summing up, in intuitive terms, it may be said that, on the basis of the information supplied by the attributes from P:

- warehouses 1 and 6, from the P-lower approximation of X, certainly belong to the set of warehouses that make a profit,
- warehouses 1,2,3 and 6, from the P-upper approximation of X, could belong to the set of warehouses that make a profit,
- warehouses 2 and 3, from the P-boundary of X, represent cases of uncertain membership to the set of warehouses that make a profit.

Approximating the set Y of warehouses which have made a loss, i.e. $Y=\{2,4,5\}$, using again the set of attributes $P=\{A_1,A_2,A_3\}$, the result is

$$\underline{P}(Y) = \{4,5\},\ \overline{P}(Y) = \{2,3,4,5\},\ Bn_P(Y) = \{2,3\}.$$

Let us consider now the following subsets of Q: $P=\{A_1,A_2,A_3\}$, $R=\{A_1,A_2\}$, $T=\{A_1,A_3\}$, $W=\{A_2,A_3\}$. It is easy to observe that $I_R=I_P$, $I_T=I_P$, while $I_W \neq I_P$. This means that R and T are reducts of P, while W is not. In other words, R and T are minimal subsets of P that induce the same partition of the elements of U as the set of attributes P. It can also be observed that in the core of P, defined by $R \cap T$, there is attribute A_1, which is then indispensable for the approximation of the class of warehouses that make a profit (and also for the class of warehouses that make a loss), while other attributes from R and T may be mutually exchanged.

If in the set of attributes Q, condition attributes $C=\{A_1,A_2,A_3\}$ and decision attribute $D=\{A_4\}$ were distinguished, the information table could be seen as a decision table. In order to explain the evaluations of the decision attribute by means of the evaluations of the condition attributes, one can represent the information table as a set of decision rules. Such a representation of Table 14.1 gives the following rules:

1) *if* $f(x,A_1)$=high *and* $f(x,A_2)$=good *and* $f(x,A_3)$=no, *then* $f(x,A_4)$=profit (or, in linguistic terms, "*if* the capacity of the sales staff is high *and* the perceived quality of goods is good *and* the location is not in high traffic conditions, *then* the warehouse makes a profit"),
2) *if* $f(x,A_1)$=medium *and* $f(x,A_2)$=medium *and* $f(x,A_3)$=no, *then* $f(x,A_4)$=loss,
3) *if* $f(x,A_1)$=medium *and* $f(x,A_2)$=medium *and* $f(x,A_3)$=no, *then* $f(x,A_4)$=profit,
4) *if* $f(x,A_1)$=low *and* $f(x,A_2)$=medium *and* $f(x,A_3)$=no, *then* $f(x,A_4)$=loss,
5) *if* $f(x,A_1)$=medium *and* $f(x,A_2)$=good *and* $f(x,A_3)$=yes, *then* $f(x,A_4)$=loss,
6) *if* $f(x,A_1)$=high *and* $f(x,A_2)$=medium *and* $f(x,A_3)$=yes, *then* $f(x,A_4)$=profit.

The above set of rules may then be reduced by induction, obtaining a more concise representation of the decision table (within parantheses there are the objects supporting the corresponding rules):

1') *if* $f(x,A_1)$=high, *then* $f(x,A_4)$=profit, (1, 6)

2') *if* $f(x,A_1)$=low, *then* $f(x,A_4)$=loss, (4)

3') *if* $f(x,A_1)$=medium *and* $f(x,A_2)$=good, *then* $f(x,A_4)$=loss, (5)

4') *if* $f(x,A_1)$=medium *and* $f(x,A_2)$=medium, *then* $f(x,A_4)$=profit *or* loss. (2,3)

Observe that rules 1'), 2') and 3') have a univocal consequence and therefore these are exact rules, while rule 4') does not have a univocal consequence and for this reason it is an approximate rule.

Finally, quality of approximation, interaction indices I_S and I_B and Möbius representation of all subsets of attributes in C were calculated. Their values are presented in Table 14.2.

Table 14.2 Quality of approximations, Möbius representation and interaction indices

Attributes	Quality	Möbius	Shapley	Banzhaf
$\{A_1\}$	0.5	0.5	0.44	0.5
$\{A_2\}$	0	0	0.11	0.17
$\{A_3\}$	0	0	0.11	0.17
$\{A_1,A_2\}$	0.67	0.17	-0.17	-0.17
$\{A_1,A_3\}$	0.67	0.17	-0.17	-0.17
$\{A_2,A_3\}$	0.5	0.5	0.17	0.17
$\{A_1,A_2,A_3\}$	0.67	-0.67	-0.67	-0.67

The results represented in Table 14.2 can be interpreted as follows:

1) the second column shows the quality of approximation for the considered subset of attributes;
2) the third column presents the Möbius representation and gives a measure of the conjoint contribution of the corresponding set of attributes to the quality of classification. The negative value corresponding to $\{A_1,A_2,A_3\}$ should be read as a measure of the redundancy in the information from conjoint contribution of the three attributes;
3) the fourth column shows the Shapley interaction index: more precisely the first three values are the Shapley values and can be interpreted as measures of the importance of the corresponding attributes in the rough approximation. One can notice a relatively great importance of A_1 with respect to A_2 and A_3. Furthermore A_2 and A_3 are complementary, while A_1 and A_2, as well as A_1 and A_3, are substitutive. Finally there is redundancy between A_1, A_2 and A_3 as pointed out by the negative value of the corresponding interaction index;
4) the fifth column presents the Banzhaf interaction index, which has an interpretation analogous to the Shapley interaction index.

14.1.8 A comparison with the fuzzy sets theory and statistical analysis

A number of relations exist between the rough sets theory and other mathematical theories dealing with particular types of "uncertainty" or, more generally, with "imperfect" data. The rough sets have been compared with discriminant analysis, fuzzy sets and evidence theory (see [43, 7, 8, 76]). Tables 14.3 and 14.4 below give a brief synoptic comparison of the rough set approach with the classic statistical analysis and with the fuzzy set approach, respectively (Table 14.3 quotes some summary observations of Stefanowski [93]).

Often the rough set approach is not offered as an alternative, but as a complement to other approaches based on different theories or techniques. A number of concrete applications have been made using different approaches; the use of rough sets has very often proved to be particularly interesting, both for its ability of handling raw and even inconsistent data (the notable "poorness" of information required) and for the readable results obtained by this approach (reducts, core, relevance of the attributes, decision rules) which are useful for decision aiding.

Let us point out the particularly fruitful complementarity of fuzzy sets theory and rough sets theory, which has been acknowledged by a large number of studies (see e.g. [7, 8, 60, 86, 89, 90, 103]). Indeed, both theories deal with different types of imperfect information which can be encountered together. As to fuzzy sets, they deal with a type of imprecision which arises when the boundaries of a class of objects are not sharply defined. Informally, a fuzzy set may be regarded as a class in which there is a graduality of progression from membership to non-membership or, more precisely, in which an object may have a grade of membership intermediate between unity (full membership) and zero (non-membership). Three different semantics can be associated with the use of fuzzy sets [9]:

- a first semantics expressing closeness, proximity, similarity and the like; under this semantics, objects with membership one are viewed as prototypical (referent) objects of the fuzzy set, while the other membership grades estimate the closeness of objects (subjects) to the prototypical ones,
- a second semantics expressing an incomplete or vague state of information under the form of possibility distributions; this view of fuzzy sets enables imprecise or uncertain information to be processed,
- a third semantics expressing preferences between pairs of objects; the gradedness introduced by the use of fuzzy sets refines the classic crisp preference structures.

In many situations it is required to consider more than one semantics at a time. In the following, when using fuzzy sets in conjunction with rough sets, we will mention the type of semantics considered.

Table 14.3 Statistical analysis versus the classic rough set approach

Issue	Statistical methods	Rough set approach
Objectives	Identification and estimation of parameters of some structural equations in view of explaining a classification	Approximation of a classification using an indiscernibility relation, reduction of superfluous attributes, generation of decision rules
Representation of data	Two-entry table representing a sample	Information table
Types of attributes	Quantitative attributes and "binarized" qualitative attributes	Qualitative attributes and discretized quantitative attributes
Requirements of data	The sample must be statistically representative and the distribution of objects in decision classes must be well-balanced; normal multivariate distribution of attribute values is assumed	No requisite; possibility of analysing information tables of reduced dimensions
Operators for data aggregation	Average values, covariance matrices, statistical tests	No operator; the indiscernibility relation operates on original data
Reduction of data	Selection of attributes with the highest discriminating potential; typical instrument: statistical tests	Reducts of the set of attributes ensuring the same quality of classification as the whole set of attributes
Final results	Discriminant function or probabilistic tree classifier	Decision rules in the form of logical statements "*if..., then...*"

Table 14.4 Fuzzy sets versus classic rough sets

Issue	Fuzzy sets	Rough sets
Semantics of uncertainty (imperfect knowledge)	Closeness (similarity), incomplete state of information (imprecision) or degree of satisfaction (preference)	Inconsistency or ambiguity following from granularity of knowledge
Additional information	Context-dependent membership functions specifying the degree of membership of an object to a set	Discretization (if necessary) of quantitative attributes; the degree of membership of an object to a set can be calculated from available information
Mathematical modeling	Sets with soft boundaries; generalization of characteristic function of sets, of binary relation and of logical operators to continuum	Family of partitions consisting of classes of indiscernible objects; use of these partitions for approximation of a given set or classification; dependence and reduction of attributes
Processing of uncertainty	"Exact", using the membership functions	"Approximate", using the lower and upper approximations
Main preoccupation, for example in image processing	Levels of grey (degrees of membership)	Size of the pixels (granularity)

14.2 Generalization of the indiscernibility relation

As mentioned above, the classic definitions of lower and upper approximations have been introduced with reference to the binary indiscernibility relation, which is an equivalence relation. In this case, both the sets to be approximated and the relation used are ordinary (*crisp*).

A generalization, consisting in approximation of *fuzzy* sets with a *fuzzy* indiscernibility relation, was proposed by Dubois and Prade [7, 8]. Nevertheless, this approach is still based on the use of the indiscernibility relation.

Of particular interest are the proposals to use, instead of the indiscernibility relation, *similarity* relation, weaker than that of indiscernibility, since in the least demanding case it requires reflexivity only, relaxing the assumptions of symmetry and transitivity (see [9, 92]).

14.2.1 Similarity

As observed above, indiscernibility implies an impossibility to distinguish two objects of U having the *same* description in terms of the attributes from Q. This relation induces equivalence classes on U, which constitute the basic granules of knowledge. In reality, due to the imprecision of data describing the objects, small differences are often not considered significant for the purpose of discrimination. This situation may be formally modeled by considering similarity or tolerance relations (see e.g. [57, 46, 49, 65, 79, 91, 92, 104]).

In general, the similarity relations R do not generate partitions on U; the information regarding similarity may be represented using *similarity classes* for each object $x \in U$. Precisely, the similarity class of x, denoted by R(x), consists of the set of objects which are similar to x:

$$R(x) = \{y \in U: yRx\}.$$

It is obvious that an object $y \in R(x)$ may be similar to another object $z \in U$, and $z \notin R(x)$. The similarity relation is of course reflexive (each object is similar to itself). Slowinski and Vanderpooten [91, 92] have proposed a *similarity* relation which is only *reflexive*, relaxing therefore the properties of symmetry and transitivity. The abandon of the transitivity requirement is easily justifiable, remembering – for example – Luce's paradox of the cups of tea [47]. As for the symmetry, one should notice that yRx, which means "y is similar to x", is directional; there is a *subject* y and a *referent* x, and in general this is not equivalent to the proposition "x is similar to y", as maintained by Tversky [100]. This is quite immediate when the similarity relation is defined in terms of a percentage difference between evaluations of the actions compared on the attribute at hand, calculated with respect to the evaluation of the referent action. Therefore, the symmetry of the similarity relation should not be imposed and it makes sense to consider the inverse relation of R, denoted by R^{-1}, where $xR^{-1}y$ means again "y is similar to x"; $R^{-1}(x)$, $x \in U$, is then the class of referent objects to which x is similar:

$$R^{-1}(x) = \{y \in U: xRy\}.$$

Given a subset $X \subseteq U$ and a similarity relation R on U, an object $x \in U$ is said to be *non-ambiguous* in each of the two following cases:

- x belongs to X without ambiguity, that is $x \in X$ and $R^{-1}(x) \subseteq X$; such objects are also called *positive*;
- x does not belong to X without ambiguity (x clearly does not belong to X), that is $x \in U-X$ and $R^{-1}(x) \subseteq U-X$ (or $R^{-1}(x) \cap X = \emptyset$); such objects are also called *negative*.

The objects which are neither positive nor negative are said to be *ambiguous*.

A more general definition of lower and upper approximation may thus be offered (see [92]). Let $X \subseteq U$ and R a reflexive binary relation defined on U; the lower approximation of X, denoted by $\underline{R}(X)$, and the upper approximation of X, denoted by $\overline{R}(X)$, are defined, respectively, as:

$$\underline{R}(X) = \{x \in U: R^{-1}(x) \subseteq X\},$$

$$\overline{R}(X) = \bigcup_{x \in X} R(x).$$

It may be demonstrated that the key property: $\underline{R}(X) \subseteq X \subseteq \overline{R}(X)$, still holds and that

$$\underline{R}(X) = U - \overline{R}(U-X) \text{ (complementarity property) and}$$
$$\overline{R}(X) = \{x \in U: R^{-1}(x) \cap X \neq \varnothing\}.$$

Moreover, the definitions proposed are the only ones which correctly characterize the set of positive objects (lower approximation) and the set of positive or ambiguous objects (upper approximation) when a similarity relation is used which is reflexive, but not necessarily symmetric nor transitive.

Using similarity relation one is able to induce decision rules from a decision table. The syntax of a rule is the following:

"*if* $f(x,q_1)$ is similar to r_{q1} *and* $f(x,q_2)$ is similar to r_{q2} *and* ... $f(x,q_p)$ is similar to r_{qp}, *then* x belongs to Y_{j1} *or* Y_{j2} *or* ... Y_{jk}",

where $\{q_1,q_2,...,q_p\} \subseteq C$, $(r_{q1},r_{q2},...,r_{qp}) \in V_{q1} \times V_{q2} \times ... \times V_{qp}$ and $Y_{j1},Y_{j2},...,Y_{jk}$ are some classes of the considered classification (D-elementary sets). If k=1 then the rule is *exact*, otherwise it is *approximate* or *uncertain*. Procedures for generation of decision rules adapt the scheme described in Section 14.1.5. One such procedure has been proposed by Krawiec, Slowinski and Vanderpooten [42].

14.2.2 Fuzzy similarity

A further step towards generalization of the rough approximations consists in considering *fuzzy* reflexive binary relation R(x,y) defined on U, that is a relation of *fuzzy similarity*; using it one can define the lower and upper approximations of a *fuzzy set* [19, 28]. This use of fuzzy sets is concordant with the first semantics, according to the typology given in Section 14.1.8.

To this purpose, negation and the classic connectives of fuzzy logic are used in a suitable way, in particular those of the t-norm T as conjunction, of the t-conorm T^* as disjunction and of the appropriate fuzzy negation N (for a brief but thorough introduction to fuzzy logic see the first chapter of [13]). The set of the positive objects and that of the negative objects with respect to X are fuzzy sets, whose membership functions express the credibility of the respective statement: "for every $y \in U$, x is not similar to y and/or y belongs to X" and "for every $y \in U$, x is not similar to y and/or y does not belong to X".

The lower and upper approximations of X are fuzzy sets $\underline{R}(X)$ and $\overline{R}(X)$ in U with membership functions expressing the credibility of the respective statement: "for every $y \in U$, x is not similar to y and/or y belongs to X" and "at least one $y \in U$ exists such that x is similar to y and y belongs to X". Formally, the two respective membership functions are defined as:

$$\mu(x, \underline{R}(X)) = \underset{y \in U}{T} (T^*(N(R(x,y)), \mu_X(y))),$$

$$\mu(x, \overline{R}(X)) = \underset{y \in U}{T^*} (T(R(x,y), \mu_X(y))).$$

Let us remark that using the definition of the T^*-implication (i.e. $I^{\rightarrow}_{T^*,N}(a,b)=T^*(N(a),b)$), $\forall a,b \in [0,1]$), it is possible to rewrite the definition of $\mu(x,\underline{R}(X))$ and $\mu(x,\overline{R}(X))$ in the following way:

$$\mu(x, \underline{R}(X)) = \underset{y \in U}{T} (I^{\rightarrow}_{T^*,N}(R(x,y), \mu_X(y))),$$

$$\mu(x, \overline{R}(X)) = \underset{y \in U}{T^*} (N(I^{\rightarrow}_{T^*,N}(R(x,y), N(\mu_X(y))))).$$

Therefore, $\mu(x, \underline{R}(X))$ can be interpreted as the credibility of the statement "for each $y \in U$, the similarity of x to y implies that y belongs to X", while $\mu(x,\overline{R}(X))$ can be interpreted as the credibility of the statement "for at least one $y \in U$, the similarity of x to y does not imply that y does not belongs to X".

The following definitions are necessary to introduce the next results. A strict negation is a strictly decreasing continuous function $N:[0,1] \rightarrow [0,1]$ satisfying $N(0)=1$, $N(1)=0$. (T, T^*, N) is a De Morgan triplet iff $N(T^*(x,y))=T(N(x),N(y))$, where N is a strict negation. A negation N is involutive iff for all $x \in [0,1]$, $N(N(x))=x$. It may be demonstrated, in particular, that the following properties are valid:

1) $\mu(x,\underline{R}(X)) \leq \mu_X(x) \leq \mu(x,\overline{R}(X))$,

2) if (T, T^*, N) is a De Morgan triplet and N is involutive, then

$$\mu(x,\underline{R}(X)) = N(\mu(x,\overline{R}(U-X)),$$

where U-X represents the fuzzy set whose membership function, $\forall x \in U$, has the form $\mu_{U-X}(x) = N(\mu_X(x))$.

The above results can be read as follows in the sense of fuzzy sets: 1) means that X includes its lower approximation and is included in its upper approximation; 2) means that the lower approximation of X is the complement of the upper approximation of its complementary set (complementarity property).

Given set C of attributes, let us consider a fuzzy binary relation R_q for each attribute $q \in C$, i.e. a function $R_q: U \times U \rightarrow [0,1]$, where, $\forall x,y \in U$, $R_q(x,y)$ represents the intensity or degree of similarity of x to y with respect to the attribute q. More precisely, for $q \in C$ and $\forall x,y,w,z \in U$:

$R_q(x,y) = 0$ means absence of similarity of x to y,
$R_q(x,y) = 1$ means that x is absolutely similar to y, $(R_q(x,x)=1)$,
$R_q(x,y) \geq R_q(w,z)$ means that the similarity of x to y is at least as credible as the similarity of w to z.

To model the comprehensive similarity of x to $y \in U$ with respect to a subset of attributes $P=\{q_1,q_2,...,q_p\} \subseteq C$, denoted by $R_P(x,y)$, we consider the credibility of the statement "x is similar to y with respect to q_1 and x is similar to y

with respect to q_2 and ... x is similar to y with respect to q_p". This credibility is calculated as:

$$R_P(x,y) = \underset{q \in P}{T} (R_q(x,y)).$$

Making suitable use of such a fuzzy similarity relation and of the approximations defined above, it is possible to induce *decision rules* having the same syntax as the rules considered in Section 14.2.1. Contrary to the previous case, decision rules have, in general, different *degrees of credibility*.

14.3 Rough sets and multiattribute decisions

As mentioned above, a decision table contains all the information relative to a set of objects, described by a certain number of attributes. The traditional rough set analysis of such a table consists in approximating the classifications induced by decision attributes by means of the classifications induced by condition attributes. These two kinds of classifications are built independently, i.e. they are not deduced one from the other.

The aim of the decision analysis is to answer the following two basic questions. The first question is to *explain* decisions in terms of the circumstances in which they were made. The second is to give a *recommendation* how to make a decision under specific circumstances. Recommendation is mainly based on decision rules induced from a decision table. In this sense, the rough set approach is similar to the inductive learning approach; however, the former one is going far beyond the latter, because, in the rough set approach, the recommendation task is preceded by the explanation, which gives pertinent information useful for decision support (reducts, core, quality of approximation, relevance of attributes).

According to Roy [67], it is possible to distinguish the following three, most frequent decision problems:

- *classification*,
- *choice*,
- *ranking*.

In general, decisions are based on some characteristics of actions (objects). For example, when buying a car, the decisions can be based on such characteristics as price, maximum speed, fuel consumption, color, country of production, etc. We refer to these characteristics calling them *attributes*. Let us observe that, depending on interpretation given to the attributes by the DM, some of them may have ordinal properties expressing preference scales, while others may not. The former attributes are called *criteria*, while the latter ones keep the name of attributes. In the above example, price, maximum speed and fuel consumption are criteria, because, usually, a low price is better than a high price; most probably, color and country of production are not criteria but simple attributes, because, usually, red is not better than green, or a car is not better than another, simply because the former is produced in a country and the latter in another country. However, one can imagine that also those attributes could become criteria, because a DM could consider, for

instance, red better than green, or the country of production could actually orientate his/her preferences.

Moreover, decisions may be ordinal, because of expressing a preference or may not be ordinal. For example, a classification of cars for a catalogue does not impose any preference order among the classes (sport cars, family cars, utility cars, etc.); however, choice of the best car, or ranking of a set of cars from the best to the worst surely impose a preference order. Let us also observe that, depending on interpretation given to the classification by the DM, the classes may express a preference, so also a classification may be ordinal. For instance, the DM could be interested in a classification of cars in three categories: acceptable, probably acceptable, non-acceptable. This type of classification is called *sorting*.

In the case of any multicriteria and/or multiattribute decision problem, no recommendation can be elaborated before the DM provides some preferential information suitable to the preference model assumed.

There are two major models used until now in multicriteria decision analysis: functional and relational ones. The functional model has been extensively used within the framework of multiattribute utility theory [40]. The relational model has its most widely known representation in the form of an outranking relation [68] and a fuzzy relation [13]. These models require specific preferential information more or less explicitly related with their parameters. For example, in the deterministic case, the DM is often asked for pairwise comparisons of actions, from which we can assess the substitution rates in the functional model or importance weights in the relational model (see [11, 39, 53]). This kind of preferential information seems to be close to the natural reasoning of the DM. He/she is typically more confident exercising his/her comparisons than explaining them. The representation of this information by functional or relational models seems, however, less natural. According to Slovic [80], people make decisions by searching for *rules* that provide good justification of their choices. So, after getting the preferential information in terms of exemplary comparisons, it would be natural to build the preference model in terms of "*if*..., *then*..." rules. Then, these rules can be applied to a set of potential actions in order to obtain specific preference relations. From the exploitation of these relations, a suitable recommendation can be obtained to support the DM in decision problem at hand.

The induction of rules from examples is a typical approach of artificial intelligence. It is concordant with the principle of posterior rationality by March [48] and with aggregation-disaggregation logic by Jacquet-Lagrèze [38]. The rules explain the preferential attitude of the DM and enable his/her understanding of the reasons of his/her preference. As pointed out by Langley and Simon [45], the recognition of the rules by the DM justifies their use for decision support. So, the preference model in the form of rules derived from examples, fulfils both explanation and recommendation tasks mentioned above.

In Sections 14.3.2-14.3.4, we are presenting the main extensions of the rough set approach, resulting in a new methodology of modeling and exploitation of preferences in terms of decision rules. The rules are induced from the preferential information given by the DM in the form of examples of decisions. More precisely,

for A being a finite set of actions (real or fictitious, potential or not) considered in a multicriteria problem, the examples of decisions are confined to a subset of actions $B \subseteq A$, relatively well known to the DM, called *reference actions*. Depending on the type of the multicriteria problem, the examples concern either assignment of reference actions to decision classes (sorting problem) or pairwise comparisons of reference actions (choice and ranking problems).

14.3.1 Problems of multiattribute classification

Up to now, the rough set approach to decision analysis has been limited to problems of multiattribute classification, consisting in assigning a set of actions described by a set of attributes (not criteria) to one of pre-defined categories [63]. Rough set analysis is naturally adapted to this type of problems, because the set of classification examples may be represented directly in the decision table and it is possible to extract all the essential knowledge contained in the table using indiscernibility or similarity relations.

The rough sets theory has been successfully applied to a number of real classification problems in different fields, such as medicine, pharmacology, engineering, credit management, market research, financial analysis, environmental economics, linguistics, databases and other important sectors. The interesting results obtained have recently induced experts in various disciplines to become involved in the study of the theory and its implementation. For a collection of studies on the application of the rough set approach to real-world problems see [84]. A brief but thorough review of the most important applications has recently been made by Pawlak [62].

14.3.2 Problems of multicriteria sorting

As pointed out by Greco, Matarazzo and Slowinski [18], the original rough set approach cannot extract all the essential knowledge contained in the decision table of multicriteria sorting problems, i.e. problems of assigning a set of actions described by a set of criteria to one of pre-defined and ordered categories. Notwithstanding, in many real problems it is important to consider the *ordinal properties* of the considered criteria. For example, in bankruptcy risk evaluation, if the debt index (total debts/total activity) of company A has a modest value, while the same index of company B has a significant value, then, within the rough set approach, the two firms are just discernible, but no preference is given to one of them with reference to the attribute "debt ratio". In reality, from the point of view of the bankruptcy risk evaluation, it would be advisable to consider firm A better than firm B, and not simply different (discernible). Therefore, the attribute "debt ratio" is a criterion. Let us observe that the rough set approach based on the use of indiscernibility or similarity relations is not able to capture a specific kind of inconsistency which may occur when in the decision table there is at least one criterion. For instance, in the bankruptcy risk evaluation, which is a sorting problem, if firm A is better than B with respect to all the considered criteria (e.g.

debt ratio, return on equity, etc.) but firm A is assigned to a class of higher risk than B, then there is an inconsistency which cannot be captured by the original rough set approach, because these firms are discernible. In order to detect this inconsistency, the rough approximation should handle the ordinal properties of criteria. This can be made by replacing the indiscernibility or similarity relation by the dominance relation, which is a very natural concept within multicriteria decision making.

On the basis of these considerations, Greco, Matarazzo and Slowinski [24] have proposed a new rough set approach to multicriteria sorting problems, which is described in the next Sections. Let also mention that it is sometimes reasonable to consider both criteria and attributes (without ordered domains) in a sorting problems. More precisely a rough set approximation based on a binary relation which is partly of dominance (with respect to considered criteria) and partly of indiscernibility (with respect to considered attribute) has been proposed [26]. More generally, a binary relation which is partly of dominance, partly of indiscernibility and partly of similarity can be considered [33].

14.3.2.1 Approximation by means of dominance relations

Let S_q be an *outranking* relation [67] on U with reference to criterion $q \in C$, such that xS_qy means "x is at least as good as y with respect to criterion q". Suppose that S_q is a complete preorder, that is a strongly complete and transitive binary relation. Moreover, let $\mathbf{Cl} = \{Cl_t, t \in T\}$, $T = \{1,...,n\}$, be a set of classes of U, such that each $x \in U$ belongs to one and only one class $Cl_t \in \mathbf{Cl}$. We assume that $\forall r,s \in T$, such that $r>s$, each element of Cl_r is preferred (strictly or weakly [67]) to each element of Cl_s. More formally, if S is a comprehensive outranking relation on U, i.e. xSy means: "x is at least as good as y" $\forall x,y \in U$, then it is supposed that

$$[x \in Cl_r,\ y \in Cl_s,\ r>s] \Rightarrow [xSy \text{ and not } ySx].$$

Let us also consider the following upward and downward cumulated sets, respectively,

$$Cl_t^{\geq} = \bigcup_{s \geq t} Cl_s,$$

$$Cl_t^{\leq} = \bigcup_{s \leq t} Cl_s.$$

Observe that $Cl_1^{\geq} = Cl_n^{\leq} = U$, $Cl_n^{\geq} = Cl_n$ and $Cl_1^{\leq} = Cl_1$.

It is said that x *dominates* y with respect to $P \subseteq C$ (denotation xD_Py) if xS_qy $\forall q \in P$. Since the intersection of complete preorders is a partial preorder and S_q is a complete preorder for each $q \in P$, and $D_P = \bigcap_{q \in P} S_q$, then the dominance relation D_P is a partial preorder. Given $P \subseteq C$ and $x \in U$, let

$$D_P^+(x) = \{y \in U : yD_Px\},$$

$$D_P^-(x) = \{y \in U : xD_Py\}.$$

We can define the P-lower and the P-upper approximation of $Cl_t^{\geq}$, $t \in T$, with respect to $P \subseteq C$ (denotation $\underline{P}(Cl_t^{\geq})$ and $\overline{P}(Cl_t^{\geq})$, respectively), as:

$$\underline{P}(Cl_t^{\geq}) = \{x \in U: D_P^+(x) \subseteq Cl_t^{\geq}\},$$
$$\overline{P}(Cl_t^{\geq}) = \bigcup_{x \in Cl_t^{\geq}} D_P^+(x).$$

Analogously, we define the P-lower and the P-upper approximation of $Cl_t^{\leq}$, $t \in T$, with respect to $P \subseteq C$ (denotation $\underline{P}(Cl_t^{\leq})$ and $\overline{P}(Cl_t^{\leq})$, respectively), as:

$$\underline{P}(Cl_t^{\leq}) = \{x \in U: D_P^-(x) \subseteq Cl_t^{\leq}\},$$
$$\overline{P}(Cl_t^{\leq}) = \bigcup_{x \in Cl_t^{\leq}} D_P^-(x).$$

The P-lower and P-upper approximations so defined satisfy the following properties $\forall t \in T$ and $\forall P \subseteq C$:

$$\underline{P}(Cl_t^{\geq}) \subseteq Cl_t^{\geq} \subseteq \overline{P}(Cl_t^{\geq}),$$
$$\underline{P}(Cl_t^{\leq}) \subseteq Cl_t^{\leq} \subseteq \overline{P}(Cl_t^{\leq}).$$

Furthermore, the following specific complementarity properties hold:

$$\underline{P}(Cl_t^{\geq}) = U - \overline{P}(Cl_{t-1}^{\leq}),\ t=2,\ldots,n,$$
$$\underline{P}(Cl_t^{\leq}) = U - \overline{P}(Cl_{t+1}^{\geq}),\ t=1,\ldots,n-1,$$
$$\overline{P}(Cl_t^{\geq}) = U - \underline{P}(Cl_{t-1}^{\leq}),\ t=2,\ldots,n,$$
$$\overline{P}(Cl_t^{\leq}) = U - \underline{P}(Cl_{t+1}^{\geq}),\ t=1,\ldots,n-1.$$

The P-boundaries (P-doubtful regions) of $Cl_t^{\geq}$ and $Cl_t^{\leq}$ are defined as

$$Bn_P(Cl_t^{\geq}) = \overline{P}(Cl_t^{\geq}) - \underline{P}(Cl_t^{\geq}),$$
$$Bn_P(Cl_t^{\leq}) = \overline{P}(Cl_t^{\leq}) - \underline{P}(Cl_t^{\leq}).$$

We define the *accuracy of approximation* of $Cl_t^{\geq}$ and $Cl_t^{\leq}$, $\forall t \in T$ and $\forall P \subseteq C$, respectively, as:

$$\alpha_P(Cl_t^{\geq}) = \frac{\left|\underline{P}(Cl_t^{\geq})\right|}{\left|\overline{P}(Cl_t^{\geq})\right|}, \quad \alpha_P(Cl_t^{\leq}) = \frac{\left|\underline{P}(Cl_t^{\leq})\right|}{\left|\overline{P}(Cl_t^{\leq})\right|}.$$

The ratio

$$\gamma_P(\mathbf{Cl}) = \frac{\left|U - \left(\left(\bigcup_{t \in T} Bn_P(Cl_t^{\geq})\right) \cup \left(\bigcup_{t \in T} Bn_P(Cl_t^{\leq})\right)\right)\right|}{|U|}$$

defines the *quality of approximation of the partition* **Cl** by means of the set of attributes P, or, briefly, *quality of sorting*. This expresses the relation between all the P-correctly classified objects and all the objects in the table.

Every minimal subset $P \subseteq C$ such that $\gamma_P(\mathbf{Cl}) = \gamma_C(\mathbf{Cl})$ is called a reduct of C with respect to **Cl** and is denoted by $Red_{Cl}(P)$. Again, an information table may

have more than one reduct. The intersection of all the reducts is known as the core, denoted by $Core_{Cl}$.

14.3.2.2 Decision rules

On the basis of the approximations obtained by means of the dominance relations, it is possible to induce a generalized description of the preferential information contained in the decision table in terms of decision rules.

Assuming that for each $q \in C$, $V_q \subseteq \mathbf{R}$ (i.e. V_q is quantitative) and for each $x,y \in U$ $f(x,q) \geq f(y,q)$ implies xS_qy (i.e. V_q is ordered), the following three types of decision rules can be considered:

1) decision rules of the type $D_{\geq}$, which have the following form:

   if $f(x,q_1) \geq r_{q1}$ *and* $f(x,q_2) \geq r_{q2}$ *and*... $f(x,q_p) \geq r_{qp}$, *then* $x \in Cl_t^{\geq}$,

   where $P=\{q_1, q_2,...q_p\} \subseteq C$, $(r_{q1},r_{q2},...,r_{qp}) \in V_{q1} \times V_{q2} \times ... \times V_{qp}$ and $t \in T$; these rules are supported only by actions from the P-lower approximations of the classes $Cl_t^{\geq}$;

2) decision rules of the type $D_{\leq}$, which have the following form:

   if $f(x,q_1) \leq r_{q1}$ *and* $f(x,q_2) \leq r_{q2}$ *and*... $f(x,q_p) \leq r_{qp}$, *then* $x \in Cl_t^{\leq}$,

   where $P=\{q_1, q_2,...q_p\} \subseteq C$, $(r_{q1},r_{q2},...,r_{qp}) \in V_{q1} \times V_{q2} \times ... \times V_{qp}$ and $t \in T$; these rules are supported only by actions from the P-lower approximations of the classes $Cl_t^{\leq}$;

3) decision rules of the type $D_{\geq\leq}$, which have the following form:

   if $f(x,q_1) \geq r_{q1}$ *and* $f(x,q_2) \geq r_{q2}$ *and*... $f(x,q_k) \geq r_{qk}$ *and* $f(x,q_{k+1}) \leq r_{qk+1}$ *and*... $f(x,q_p) \leq r_{qp}$, *then* $x \in Cl_t^{\leq}$ *or* $x \in Cl_s^{\geq}$,

   where $O'=\{q_1, q_2,...q_k\} \subseteq C$, $O''=\{q_{k+1}, q_{k+2},...q_p\} \subseteq C$, $P=O' \cup O''$, O' and O'' not necessarily disjoint, $(r_{q1},r_{q2},..., r_{qp}) \in V_{q1} \times V_{q2} \times ... \times V_{qp}$, $s,t \in T$ such that $t<s$; these rules are supported only by actions from the P-boundaries of the classes $Cl_t^{\leq}$ and $Cl_s^{\geq}$.

If for some $q \in C$ V_q is not quantitative then the syntax of the above rules can easily extended, referring directly to the order induced by S_q on V_q.

Let us observe that the set of decision rules induced from the approximations defined using dominance relations gives, in general, a more synthetic representation of knowledge contained in the decision table than the set of rules induced from classic approximations defined using indiscernibility relations. The minimal sets of rules thus obtained have a smaller number of rules and use a smaller number of conditions. Moreover, the application of these rules to new objects gives better results, in general. This is due to the more general syntax of the rules ("≥" and "≤" are used instead of "=").

14.3.2.3 An example

Let us apply the rough approximation by dominance relation to the same decision table as considered in Section 14.1.7. Within this approach we approximate the class $Cl_1^{\leq}$ of the warehouses "at most making loss" and the class $Cl_2^{\geq}$ of the warehouses "at least making profit". Since only two classes are considered, we have $Cl_1^{\leq}=Cl_1$ and $Cl_2^{\geq}=Cl_2$. As previously, $C=\{A_1, A_2, A_3\}$ and $D=\{A_4\}$. In this case we must observe, however, that A_1, A_2 and A_3 are criteria. It means that

- with respect to A_1,"high" is better than "medium" and "medium" is better than "low",
- with respect to A_2, "good" is better than "medium",
- with respect to A_3,"yes" is better than "no".

The domain of A_4 is trivially ordered too.

The C-lower approximations, the C-upper approximations and the C-boundaries of sets $Cl_1^{\leq}$ and $Cl_2^{\geq}$ are equal, respectively, to:

$\underline{C}(Cl_1^{\leq})=\{4\}$, $\overline{C}(Cl_1^{\leq})=\{2, 3, 4, 5\}$, $Bn_C(Cl_1^{\leq})=\{2, 3, 5\}$,

$\underline{C}(Cl_2^{\geq})=\{1, 6\}$, $\overline{C}(Cl_2^{\geq})=\{1, 2, 3, 5, 6\}$, $Bn_C(Cl_2^{\geq})=\{2, 3, 5\}$.

Therefore, the accuracy of the approximation is 0.25 for $Cl_1^{\leq}$ and 0.4 for $Cl_2^{\geq}$, while the quality of sorting is equal to 0.5. There is only one reduct which is also the core, i.e. $Red_{Cl}(C) = Core_{Cl}(C) = \{A_1\}$.

The following minimal set of decision rules can be obtained from the considered decision table (within parentheses there are the actions supporting the corresponding rules):

1) *if* f(x, A_1) is at least high, *then* $x \in Cl_2^{\geq}$ (1, 6)

2) *if* f(x, A_1) is at most low, *then* $x \in Cl_1^{\leq}$ (4)

3) *if* f(x, A_1) is at least medium *and* f(x, A_1) is at most medium (i.e. f(x, A_1) is medium), *then* $x \in Cl_1^{\leq}$ *or* $x \in Cl_2^{\geq}$ (2, 3, 5)

14.3.2.4 Comparison of the results

The advantages of the rough set approach based on the dominance relation over the original rough set approach based on the indiscernibility relation are summarized below.

The results of the approximation are more satisfactory. This improvement is represented by a single reduct ($\{A_1\}$), while from the indiscernibility approach we obtain two reducts ($\{A_1, A_2\}$ and $\{A_1, A_3\}$). Let us observe that, even if the quality of the approximation obtained by dominance (0.5) is smaller than the quality of

approximation by indiscernibility (0.67), this is another point in favor of the new approach. In fact, this difference is due to the warehouses 3 and 5. Let us notice that warehouse 5 dominates warehouse 3, i.e. warehouse 5 is at least as good as warehouse 3 with respect to all the three criteria, however, 5 has a comprehensive evaluation worse than 3. Therefore, this can be interpreted as an inconsistency revealed by the approximation based on dominance, that cannot be captured by the approximation based on indiscernibility.

Moreover, let us remark that the decision rules induced from approximations defined using dominance relations give a more synthetic representation of knowledge contained in the decision table. The minimal set of decision rules obtained from the dominance approach has a smaller number of stronger rules (3 against 4) and uses a smaller number of conditions (3 against 6). Furthermore, let us observe that some rules obtained from the original rough set approach make problems with their interpretation. For example, rule 3') obtained by the original rough set approach says that "if the capacity of the sale staff is medium and the perceived quality of goods is good, then the warehouse makes loss". One would expect that a warehouse of lower quality, e.g. a warehouse with the same capacity of the sales staff but with a medium quality of goods, should still make loss. Surprisingly, warehouse 3 has these characteristics and, nevertheless, it makes profit.

14.3.2.5 Approximations by means of fuzzy dominance relations

The concept of dominance can be refined by introducing gradedness through the use of fuzzy sets in the sense of the third semantics, according to the typology given in 1.8 [29].

Let S_q be a fuzzy outranking relation on U with respect to criterion $q \in C$, i.e. $S_q: U \times U \rightarrow [0,1]$, such that $S_q(x,y)$ represents the credibility of the proposition "x is at least as good as y with respect to criterion q". Suppose that S_q is a fuzzy partial T-preorder, i.e. that it is reflexive ($S_q(x,x)=1$ for each $x \in U$) and T-transitive ($T(S_q(x,y),S_q(y,z)) \leq S_q(x,z)$, for each $x,y,z \in U$) (see [13]). Using the fuzzy outranking relations S_q, $q \in C$, a *fuzzy dominance relation* on U (denotation $D_P(x,y)$) can be defined for each $P \subseteq C$ as follows:

$$D_P(x,y) = \underset{q \in P}{T} (S_q(x,y)).$$

Given $(x,y) \in U \times U$, $D_P(x,y)$ represents the credibility of the proposition "x outranks y with respect to each criterion q from P".

Since the fuzzy outranking relations S_q are supposed to be partial T-preorders, then also the fuzzy dominance relation D_P is a partial T-preorder.

Furthermore, let $\mathbf{Cl}=\{Cl_t, t \in T\}$, $T=\{1,...,n\}$, be a set of fuzzy classes in U, such that for each $x \in U$, $Cl_t(x)$ represents the membership function of x to Cl_t. We suppose that the classes of **Cl** are increasingly ordered, i.e. that $\forall r,s \in T$, such that $r>s$, the elements of Cl_r have a better comprehensive evaluation than the elements of Cl_s. For example, in a problem of bankruptcy risk evaluation, Cl_1 is the set of

unacceptable risk firms, Cl_2 is the set of high risk firms, Cl_3 is the set of medium risk firms, and so on.

On the basis of the membership functions of the fuzzy class Cl_t, we can define fuzzy membership functions of two other sets:

1) the upward cumulated fuzzy set $Cl_t^{\geq}$, whose membership function $Cl_t^{\geq}(x)$ represents the credibility of the proposition "x is at least as good as the objects in Cl_t",

$$Cl_t^{\geq}(x)=\begin{cases} 1 & \text{if } \exists s \in T : Cl_s(x) > 0 \text{ and } s > t \\ Cl_t(x) & \text{otherwise (i.e. if } \forall s \in T \text{ such that } s > t \quad Cl_s(x) = 0) \end{cases}$$

2) the downward cumulated fuzzy set $Cl_t^{\leq}$, whose membership function $Cl_t^{\leq}(x)$ represents the credibility of the proposition "x is at most as good as the objects in Cl_t",

$$Cl_t^{\leq}(x)=\begin{cases} 1 & \text{if } \exists s \in T : Cl_s(x) > 0 \text{ and } s < t \\ Cl_t(x) & \text{otherwise (i.e. if } \forall s \in T \text{ such that } s < t \quad Cl_s(x) = 0) \end{cases}$$

The P-lower and the P-upper approximations of $Cl_t^{\geq}$ with respect to $P \subseteq C$ are fuzzy sets in U, whose membership functions (denotation $\underline{P}[Cl_t^{\geq}(x)]$ and $\overline{P}[Cl_t^{\geq}(x)]$) are defined as:

$$\underline{P}[Cl_t^{\geq}(x)] = \underset{y \in U}{T} (T^*(N(D_P(y,x)), Cl_t^{\geq}(y))),$$

$$\overline{P}[Cl_t^{\geq}(x)] = \underset{y \in U}{T^*} (T(D_P(x,y), Cl_t^{\geq}(y))).$$

$\underline{P}[Cl_t^{\geq}(x)]$ represents the credibility of the proposition "for all $y \in U$, y does not dominate x with respect to criteria from P and/or y belongs to $Cl_t^{\geq}$", while $\overline{P}[Cl_t^{\geq}(x)]$ represents the credibility of the proposition "there is at least one $y \in U$ dominated by x with respect to criteria from P which belongs to $Cl_t^{\geq}$".

The P-lower and P-upper approximations of $Cl_t^{\leq}$ with respect to $P \subseteq C$ (denotation $\underline{P}[Cl_t^{\leq}(x)]$ and $\overline{P}[Cl_t^{\leq}(x)]$) can be defined, analogously, as:

$$\underline{P}[Cl_t^{\leq}(x)] = \underset{y \in U}{T} (T^*(N(D_P(x,y)), Cl_t^{\leq}(y))),$$

$$\overline{P}[Cl_t^{\leq}(x)] = \underset{y \in U}{T^*} (T(D_P(y,x), Cl_t^{\leq}(y))).$$

$\underline{P}[Cl_t^{\leq}(x)]$ represents the credibility of the proposition "for all $y \in U$, x does not dominate y with respect to criteria from P and/or y belongs to $Cl_t^{\leq}$", while $\overline{P}[Cl_t^{\leq}(x)]$ represents the credibility of the proposition "there is at least one $y \in U$ dominating x with respect to criteria from P which belongs to $Cl_t^{\leq}$".

Let us remark that using the definition of the T^*-implication, it is possible to rewrite the definition of $\underline{P}[Cl_t^{\geq}(x)]$, $\overline{P}[Cl_t^{\geq}(x)]$, $\underline{P}[Cl_t^{\leq}(x)]$ and $\overline{P}[Cl_t^{\leq}(x)]$ in the following way:

$$\underline{P}[Cl_t^{\geq}(x)] = \underset{y \in U}{T} (I_{T^*,N}^{\rightarrow}(D_P(y,x), Cl_t^{\geq}(y))),$$

$$\overline{P}[Cl_t^{\geq}(x)] = \underset{y \in U}{T^*} (N(I_{T^*,N}^{\rightarrow}(D_P(x,y), N(Cl_t^{\geq}(y))))).$$

$$\underline{P}[Cl_t^{\leq}(x)] = \underset{y \in U}{T} (I_{T^*,N}^{\rightarrow}(D_P(x,y), Cl_t^{\leq}(y))),$$

$$\overline{P}[Cl_t^{\leq}(x)] = \underset{y \in U}{T^*} (N(I_{T^*,N}^{\rightarrow}(D_P(y,x), N(Cl_t^{\leq}(y))))).$$

The following results can be proved:

1) $\underline{P}[Cl_t^{\geq}(x)] \leq Cl_t^{\geq}(x) \leq \overline{P}[Cl_t^{\geq}(x)]$, for each $x \in U$ and for each $t \in T$,

2) $\underline{P}[Cl_t^{\leq}(x)] \leq Cl_t^{\leq}(x) \leq \overline{P}[Cl_t^{\leq}(x)]$, for each $x \in U$ and for each $t \in T$,

3) if (T, T^*, N) constitute a De Morgan triplet, if negation N is involutive and if $N(Cl_t^{\geq}(x)) = Cl_{t-1}^{\leq}(x)$ for each $x \in U$ and $t=2,\ldots,n$, then

$$\underline{P}[Cl_t^{\geq}(x)] = N(\overline{P}[Cl_{t-1}^{\leq}(x)], \quad t=2,\ldots n,$$

$$\underline{P}[Cl_t^{\leq}(x)] = N(\overline{P}[Cl_{t+1}^{\geq}(x)]), \quad t=1,\ldots,n-1,$$

$$\overline{P}[Cl_t^{\geq}(x)] = N(\underline{P}[Cl_{t-1}^{\leq}(x)]), \quad t=2,\ldots n,$$

$$\overline{P}[Cl_t^{\leq}(x)] = N(\underline{P}[Cl_{t+1}^{\geq}(x)]), \quad t=1,\ldots,n-1.$$

Results 1) to 3) can be read as the fuzzy counterparts of the following results well-known within the classic rough set approach: 1) says that $Cl_t^{\geq}$ includes its P-lower approximation and is included in its P-upper approximation; 2) has an analogous interpretation with respect $Cl_t^{\leq}$; 3) (complementarity property) says that the P-lower (P-upper) approximation of $Cl_t^{\geq}$ is the complement of the P-upper (P-lower) approximation of its complementary set $Cl_{t-1}^{\leq}$ (analogous property holds for $Cl_t^{\leq}$).

14.3.3 Multicriteria choice and ranking problems

As pointed out above, the use of rough sets in the past has been limited to problems of multiattribute classification only. In Section 14.3.2, we presented an extension of the rough set approach to the multicriteria sorting problem. In the case of

multicriteria choice and ranking problems we need further extensions, because the decision table in its original form does not allow the representation of preference relations between actions.

To handle binary relations within the rough set approach, Greco, Matarazzo and Slowinski [17] proposed to operate on, so-called, *pairwise comparison table* (*PCT*), i.e., with respect to a choice or ranking problem, a decision table whose objects are pairs of actions for which multicriteria evaluations and a comprehensive preference relation are known.

The use of an indiscernibility relation on the PCT makes problems with interpretation of the approximations of the preference relation and of the decision rules derived from these approximations. Indiscernibility permits handling *inconsistency*, which arrives when two pairs of actions have preferences of the same strength on considered criteria, however, the comprehensive preference relations established for these pairs are not the same. When we deal with criteria, there may arrive also another type of inconsistency connected with the dominance principle: on a given set of criteria, one pair of actions is characterized by some preferences and another pair has all preferences at least of the same strength, however, for the first pair we have a comprehensive preference and for the other – an inverse comprehensive preference. This is why the indiscernibility relation is not able to handle all kinds of inconsistencies connected with the use of criteria. For this reason, another way of defining the approximations and decision rules has been proposed, which is based on the use of *graded dominance relations.*

14.3.3.1 The pairwise comparison table

Let C be the set of criteria used for evaluation of actions from A. For any criterion $q \in C$, let T_q be a finite set of binary relations defined on A on the basis of the evaluations of actions from A with respect to the considered criterion q, such that $\forall (x,y) \in A \times A$ exactly one binary relation $t \in T_q$ is verified. More precisely, given the domain V_q of $q \in C$, if $v'_q, v''_q \in V_q$ are the respective evaluations of $x,y \in A$ by means of q and $(x,y) \in t$, with $t \in T_q$, then for each $w,z \in A$ having the same evaluations v'_q, v''_q by means of q, $(w,z) \in t$. For interesting applications it should be $\text{card}(T_q) \geq 2$, $\forall q \in C$. Furthermore, let T_d be a set of binary relations defined on set A (comprehensive pairwise comparisons) such that at most one binary relation $t \in T_d$ is verified $\forall (x,y) \in A \times A$.

The preferential information has the form of pairwise comparisons of reference actions from $B \subseteq A$, considered as examples of decision. The *pairwise comparison table* (PCT) is defined as information table $S_{PCT} = \langle B, C \cup \{d\}, T_C \cup T_d, g \rangle$, where $B \subseteq B \times B$ is a non-empty *set of exemplary pairwise comparisons of reference actions*, $T_C = \bigcup_{q \in C} T_q$, d is a decision corresponding to the comprehensive pairwise comparison (comprehensive preference relation), and $g: B \times (C \cup \{d\}) \rightarrow T_C \cup T_d$ is a total function such that $g[(x,y),q] \in T_q$ $\forall (x,y) \in A \times A$

and $\forall q \in C$, and $g[(x,y),d] \in T_d$ $\forall (x,y) \in B$. It follows that for any pair of reference actions $(x,y) \in B$ there is verified one and only one binary relation $t \in T_d$. Thus, T_d induces a partition of B. In fact, information table S_{PCT} can be seen as decision table, since the set of considered criteria C and decision d are distinguished.

We assume that the exemplary pairwise comparisons provided by the DM can be represented in terms of *graded preference relations* (for example "very weak preference", "weak preference", "strict preference", "strong preference", "very strong preference") P_q^h: $\forall q \in C$ and $\forall (x,y) \in AxA$,

$$T_q = \{P_q^h, h \in H_q\},$$

where H_q is a particular subset of the relative integers and

$xP_q^h y$, $h>0$, means that action x is preferred to action y by degree h with respect to the criterion q,

$xP_q^h y$, $h<0$, means that action x is not preferred to action y by degree h with respect to the criterion q,

$x P_q^0 y$ means that action x is similar (asymmetrically indifferent) to action y with respect to the criterion q.

Let us remark that P_q^0 is the same similarity relation as presented in Section 14.2.1 in very general terms, i.e. without any specific reference to preference modeling. Within the preference context, the similarity relation, even if not symmetric, resembles indifference relation. Thus, in this case, we call this similarity relation "asymmetric indifference".

Of course, $\forall q \in C$ and $\forall (x,y) \in A \times A$, $[xP_q^h y, h \geq 0] \Leftrightarrow [yP_q^k x, k \leq 0]$.

The set of binary relations T_d may be defined in a similar way, but $x P_d^h y$ means that action x is comprehensively preferred to action y by degree h.

Technically, the modeling of the binary relation P_q^h, i.e. the assessment of h, can be organized as follows:

- first, it is observed that for any $q \in C$ there exists a function c_q: $A \rightarrow \mathbf{R}$ which is increasing with respect to the preferences on q (the evaluations of c_q depend on the evaluations of the total function f(x,q), more precisely $f(x,q) = f(y,q)$ implies $c_q(x)=c_q(y)$)
- then, it is possible to define a function k_q: $\mathbf{R}^2 \rightarrow \mathbf{R}$ which measures the *strength of the preference* (positive or negative) of x over y (e.g. $k_q[c_q(x),c_q(y)]=c_q(x)-c_q(y)$); it should satisfy the following properties $\forall x,y,z \in A$:

$$\text{i) } c_q(x)>c_q(y) \Leftrightarrow k_q[c_q(x),c_q(z)] > k_q[c_q(y),c_q(z)],$$

$$\text{ii) } c_q(x)>c_q(y) \Leftrightarrow k_q[c_q(z),c_q(x)] < k_q[c_q(z),c_q(y)],$$

iii) $c_q(x)=c_q(y) \Leftrightarrow k_q[c_q(x),c_q(y)]=0$,

- next, the domain of k_q can be divided into intervals, using a suitable set of thresholds Δ_q, $\forall q \in C$; these intervals are numbered in such a way that $k_q[c_q(x),c_q(y)]=0$ belongs to interval no. 0,
- the value of h in the relation $xP_q^h y$ is then equal to the number of interval including $k_q[c_q(x),c_q(y)]$, for any $(x,y) \in A \times A$.

Actually, property iii) can be relaxed in order to obtain a more general preference model which, for instance, does not satisfy preferential independence [40].

We are considering a PCT where the set T_d is composed of two binary relations defined on A:

1) x outranks y (denotation xSy or $(x,y) \in S$), where $(x,y) \in B$,
2) x does not outrank y (denotation $xS^c y$ or $(x,y) \in S^c$), where $(x,y) \in B$,

and $S \cup S^c = B$, where "x outranks y " means "x is at least as good as y" [67]; observe that the binary relation S is reflexive, but neither necessarily transitive nor complete [68, 3].

14.3.3.2 Approximation by means of graded dominance relations

Let $H_P = \bigcap_{q \in P} H_q$, $\forall P \subseteq C$. Given $P \subseteq C$ and $h \in H_P$, $\forall (x,y) \in A \times A$ it is said that x *positively dominates* y by degree h with respect to criteria from P iff $x P_q^{fq} y$ with $fq \geq h$, $\forall q \in P$. Analogously, $\forall (x,y) \in A \times A$, x *negatively dominates* y by degree h with respect to criteria from P iff $x P_q^{fq} y$ with $fq \leq h$, $\forall q \in P$. Therefore, each $P \subseteq C$ and $h \in H_P$ generate two binary relations (possibly empty) on A called *positive P-dominance by degree h* (denotation D_{+P}^h) and *negative P-dominance by degree h* (denotation D_{-P}^h), respectively. They satisfy the following conditions:

(P1) if $(x,y) \in D_{+P}^h$, then $(x,y) \in D_{+R}^k$ for each $R \subseteq P$ and $k \leq h$;

(P2) if $(x,y) \in D_{-P}^h$, then $(x,y) \in D_{-R}^k$ for each $R \subseteq P$ and $k \geq h$.

Greco, Matarazzo, Slowinski [18] have proposed to approximate the outranking relation S by means of the dominance relation D_{+P}^h. Therefore, S is considered a *rough binary relation* (see [17]).

The P-lower approximation of S (denotation $\underline{P}(S)$) and the P-upper approximation of S (denotation $\overline{P}(S)$) are defined, respectively, as:

$$\underline{P}(S) = \bigcup_{h \in H_P} \left\{ \left(D_{+P}^h \cap B \right) \subseteq S \right\},$$

$$\overline{P}(S) = \bigcap_{h \in H_P} \left\{ \left(D_{+P}^h \cap B \right) \supseteq S \right\}.$$

Remembering (P1), $\underline{P}(S)$ may be interpreted as the dominance relation D^h_{+P} having the largest intersection with B included in the outranking relation S, and $\overline{P}(S)$ as the dominance relation D^h_{+P} including S and having the smallest intersection with B.

Analogously, it is possible to approximate the relation S^c by means of the dominance relation D^h_{-P}. Observe that, in general, the definitions of the approximations of S and S^c do not satisfy the condition of complementarity, i.e. it is not true, in general, that $\underline{P}(S)$ is equal to $B-\overline{P}(S^c)$ and that $\underline{P}(S^c)$ is equal to $B-\overline{P}(S)$; this is because S and S^c are approximated using two different relations, D^h_{+P} and D^h_{-P}, respectively. Nevertheless, the approximations thus obtained constitute a good basis for the generation of simple decision rules.

14.3.3.3 Decision rules

It is possible to represent the preferential information contained in a given PCT in terms of decision rules. Since approximations of S and S^c were made using graded dominance relations, it is possible to induce decision rules being propositions of the following type:

1) D_{++}-decision rule, which is a proposition of the type: *if* $xD^h_{+P}y$, *then* xSy,
2) D_{+-}-decision rule, which is a proposition of the type: *if not* $xD^h_{+P}y$, *then* xS^cy,
3) D_{-+}-decision rule, which is a proposition of the type: *if not* $xD^h_{-P}y$, *then* xSy,
4) D_{--}-decision rule, which is a proposition of the type: *if* $xD^h_{-P}y$, *then* xS^cy,

where P is a non-empty subset of C. Therefore, for example, a D_{++}-decision rule is a proposition of the type: "*if* x positively dominates y by degree h with respect to criteria from P, *then* x outranks y".

A *constructive definition* of these rules may be given, being a logical implication supported by the existence of at least one pair of actions from B satisfying one of the four propositions listed above, and by the absence of pairs from B contradicting it. Thus, for example, if

- there exists at least one pair $(w,z)\in B$ such that $wD^h_{+P}z$ and wSz and
- there does not exist any pair $(v,u)\in B$ such that $vD^h_{+P}u$ and vS^cu,
- then "*if* $xD^h_{+P}y$, *then* xSy" is accepted as a D_{++}- decision rule.

A D_{++}-decision rule "*if* $xD^h_{+P}y$, *then* xSy" is said to be *minimal* if there does not exist any rule "*if* $xD^k_{+R}y$, *then* xSy" such that $R\subseteq P$ and $k\leq h$. Analogous definitions hold for the other cases. In other words, a minimal decision rule is an implication for which there is no other implication whose premise is of at least the same weakness and whose consequence is of at least the same strength.

The following results show connections of the decision rules with the P-lower and P-upper approximations of S and S^c [18]:

1) for "*if* $xD^h_{+P}y$, *then* xSy" being a D_{++}-minimal decision rule, $\underline{P}(S) = D^h_{+P}\cap B$,

2) for "*if* $x D^{h}_{-P} y$, *then* $xS^{c}y$" being a D_{--}-minimal decision rule, $\underline{P}(S^{c}) = D^{h}_{-P} \cap B$,

3) for "*if not* $xD^{h}_{+P}y$, *then* $xS^{c}y$" being a D_{+-}-minimal decision rule, $\overline{P}(S) = D^{h}_{+P} \cap B$,

4) for "*if not* $x D^{h}_{-P} y$, *then* xSy" being a D_{-+}-minimal decision rule, $\overline{P}(S^{c}) = D^{h}_{-P} \cap B$.

14.3.3.4 Application of the decision rules and final recommendation

In order to obtain a recommendation in the choice or ranking problems with respect to a set of actions $M \subseteq A$, the decision rules induced from the approximations of S and S^{c} (defined with respect to reference actions from B) should be applied on set $M \times M$. The application of the rules to any pair of actions $(u,v) \in M \times M$ establishes the presence (uSv) or the absence ($uS^{c}v$) of outranking with respect to (u,v). More precisely,

1) from D_{++}-decision rule "*if* $xD^{h}_{+P}y$, *then* xSy" and from $uD^{h}_{+P}v$, one concludes uSv,

2) from D_{+-}-decision rule "*if not* $xD^{h}_{+P}y$, *then* $xS^{c}y$" and from *not* $uD^{h}_{+P}v$, one concludes $uS^{c}v$,

3) from D_{-+}-decision rule "*if not* $x D^{h}_{-P} y$, *then* xSy" and from *not* $u D^{h}_{-P} v$, one concludes uSv,

4) from D_{--}-decision rule "*if* $x D^{h}_{-P} y$, *then* $xS^{c}y$" and from $u D^{h}_{-P} v$, one concludes $uS^{c}v$.

After the application of the decision rules to each pair of actions $(u,v) \in M \times M$, one of the following four situations may occur:

- uSv and *not* $uS^{c}v$, that is *true* outranking (denotation $uS^{T}v$),
- $uS^{c}v$ and *not* uSv, that is *false* outranking (denotation $uS^{F}v$),
- uSv and $uS^{c}v$, that is *contradictory* outranking (denotation $uS^{K}v$),
- *not* uSv and *not* $uS^{c}v$, that is *unknown* outranking (denotation $uS^{U}v$).

The four above situations, which together constitute the so-called four-valued outranking (see [97, 98]), have been introduced to underline the *presence* and *absence* of *positive* and *negative* reasons for the outranking. Moreover, they make it possible to distinguish contradictory situations from unknown ones.

The following theorem underlines the operational importance of the minimal decision rules [18]: the application of *all* the decision rules obtained for a given S_{PCT} to a pair $(u,v) \in M \times M$ results in the same outranking relations S and S^{c} as those obtained from the application of the *minimal* decision rules *only*. Therefore, the set of the minimal decision rules totally characterizes the preferences of the DM contained in S_{PCT}.

A final *recommendation* can be obtained upon a suitable exploitation of the presence and the absence of outranking S and S^{c} on M. A possible exploitation procedure consists in calculating a specific score, called Net Flow Score, for each action $x \in M$:

$$S_{nf}(x) = S^{++}(x) - S^{+-}(x) + S^{-+}(x) - S^{--}(x),$$

where

$S^{++}(x) = |\{y \in M$: there is at least one decision rule which affirms $xSy\}|$,
$S^{+-}(x) = |\{y \in M$: there is at least one decision rule which affirms $ySx\}|$,
$S^{-+}(x) = |\{y \in M$: there is at least one decision rule which affirms $yS^c x\}|$,
$S^{--}(x) = |\{y \in M$: there is at least one decision rule which affirms $xS^c y\}|$.

The recommendation in ranking problems consists of the total preorder determined by $S_{nf}(x)$ on M; in choice problems it consists of the action(s) $x^* \in M$ such that $S_{nf}(x^*) = \max_{x \in M} S_{nf}(x)$.

The procedure described above has been recently characterized with reference to a number of desirable properties [34, 35]; moreover, a thorough axiomatic analysis of this and other exploitation procedures used to obtain a recommendation in choice and ranking problems has been carried out by Greco, Matarazzo and Slowinski [21].

14.3.3.5 Multigraded dominance

The graded dominance relation introduced in Section 14.3.3.2 assumes a common grade of preference for all the considered criteria. While this permits a simple calculation of the approximations and of the resulting decision rules, it is lacking in precision. A dominance relation allowing a different degree of preference for each considered criterion (multigraded dominance) gives a far more accurate picture of the preferential information contained in the pairwise comparison table S_{PCT}.

More formally, given $P \subseteq C$ ($P \neq \emptyset$), $(x,y),(w,z) \in A \times A$, (x,y) is said to dominate (w,z) taking into account the criteria from P (denotation $(x,y)D_P(w,z)$) if x is preferred to y at least as strongly as w is preferred to z with respect to each $q \in P$. Precisely, "at least as strongly as" means "by at least the same degree", i.e. $h_q \geq k_q$, where $h_q, k_q \in H_q$, $x P_q^{h_q} y$ and $w P_q^{k_q} z$, $\forall q \in P$. Let $D_{\{q\}}$ be the dominance relation confined to the single criterion $q \in P$. The binary relation $D_{\{q\}}$ is reflexive ($(x,y)D_{\{q\}}(x,y)$, $\forall (x,y) \in A \times A$), transitive ($(x,y)D_{\{q\}}(w,z)$ and $(w,z)D_{\{q\}}(u,v)$ imply $(x,y)D_{\{q\}}(u,v)$, $\forall (x,y),(w,z),(u,v) \in A \times A$), and complete ($(x,y)D_{\{q\}}(w,z)$ and/or $(w,z)D_{\{q\}}(x,y)$, $\forall (x,y),(w,z) \in A \times A$). Therefore, $D_{\{q\}}$ is a complete preorder on $A \times A$. Since the intersection of complete preorders is a partial preorder and $D_P = \bigcap_{q \in P} D_{\{q\}}$, $P \subseteq C$, then the dominance relation D_P is a partial preorder on $A \times A$.

Let $R \subseteq P \subseteq C$ and $(x,y),(u,v) \in A \times A$; then the following implication holds:

$$(x,y)D_P(u,v) \Rightarrow (x,y)D_R(u,v).$$

Given $P \subseteq C$ and $(x,y) \in A \times A$, let us introduce the positive dominance set (denotation $D_P^+(x,y)$) and the negative dominance set (denotation $D_P^-(x,y)$):

$$D_P^+(x,y) = \{(w,z) \in A \times A: (w,z)D_P(x,y)\},$$
$$D_P^-(x,y) = \{(w,z) \in A \times A: (x,y)D_P(w,z)\}.$$

Using the dominance relation D_P, it is possible to define P-lower and P-upper approximations of the outranking relation S with respect to $P \subseteq C$, respectively, as:

$$\underline{P}(S) = \{(x,y) \in B:\ D_P^+(x,y) \subseteq S\},$$
$$\overline{P}(S) = \bigcup_{(x,y)\in S} D_P^+(x,y).$$

Analogously, it is possible to define the approximations of S^c:

$$\underline{P}(S^c) = \{(x,y) \in B:\ D_P^-(x,y) \subseteq S^c\},$$
$$\overline{P}(S^c) = \bigcup_{(x,y)\in S^c} D_P^-(x,y).$$

It may be proved that

$$\underline{P}(S) \subseteq S \subseteq \overline{P}(S),$$
$$\underline{P}(S^c) \subseteq S^c \subseteq \overline{P}(S^c).$$

Furthermore, the following complementarity properties hold:

$$\underline{P}(S) = B - \overline{P}(S^c),\ \ \overline{P}(S) = B - \underline{P}(S^c),$$
$$\underline{P}(S^c) = B - \overline{P}(S),\ \ \overline{P}(S^c) = B - \underline{P}(S).$$

The P-boundaries (P-doubtful regions) of S and S^c are defined as

$$Bn_P(S) = \overline{P}(S) - \underline{P}(S),\ \ Bn_P(S^c) = \overline{P}(S^c) - \underline{P}(S^c).$$

It is possible to prove that $Bn_P(S) = Bn_P(S^c)$.

The concepts of accuracy, quality of approximation, reducts and core can be extended also to the approximation of the outranking relation by multigraded dominance relations. In particular, the *accuracy of approximation* of S and S^c by $P \subseteq C$ are defined, respectively, by the ratios:

$$\alpha_P(S) = \frac{|\underline{P}(S)|}{|\overline{P}(S)|},\ \ \alpha_P(S^c) = \frac{|\underline{P}(S^c)|}{|\overline{P}(S^c)|}.$$

The coefficient

$$\gamma_P = \frac{|\underline{P}(S) \cup \underline{P}(S^c)|}{|B|}$$

defines the *quality of approximation of* S and S^c by $P \subseteq C$. It expresses the ratio of all pairs of actions $(x,y) \in B$ correctly assigned to S and S^c by the set P of criteria to all the pairs of actions contained in B. Each minimal subset $P' \subseteq P$ such that $\gamma_{P'} = \gamma_P$ is called a *reduct* of P (denotation $Red_S(P)$). Let us remark that S_{PCT} can have more than one reduct. The intersection of all reducts is called the *core* (denotation $Core_S(P)$).

Using the approximations defined above, it is then possible to induce a generalized description of the preferential information contained in a given S_{PCT} in terms of suitable decision rules. The syntax of these rules is based on the concept of

upward cumulated preferences (denotation $P_q^{\geq h}$) and downward cumulated preferences (denotation $P_q^{\leq h}$), having the following interpretation:

- $xP_q^{\geq h}y$ means "x is preferred to y with respect to q by at least degree h",
- $xP_q^{\leq h}y$ means "x is preferred to y with respect to q by at most degree h".

Exact definition of the cumulated preferences, for each $(x,y)\in A\times A$, $q\in C$ and $h\in H_q$, is the following:

- $xP_q^{\geq h}y$ if $xP_q^k y$, where $k\in H_q$ and $k\geq h$,
- $xP_q^{\leq h}y$ if $xP_q^k y$, where $k\in H_q$ and $k\leq h$.

Using the above concepts, three types of decision rules can be obtained:

1) $D_{\geq}$-decision rules, being statements of the type:
   if $xP_{q1}^{\geq h(q1)}y$ *and* $xP_{q2}^{\geq h(q2)}y$ *and* ... $xP_{qp}^{\geq h(qp)}y$, *then* xSy,
   where $P=\{q1,q2,...,qp\}\subseteq C$ and $(h(q1),h(q2),...,h(qp))\in H_{q1}\times H_{q2}\times...\times H_{qp}$; these rules are supported by pairs of actions from the P-lower approximation of S only;
2) $D_{\leq}$-decision rules, being statements of the type:
   if $xP_{q1}^{\leq h(q1)}y$ *and* $xP_{q2}^{\leq h(q2)}y$ *and* ... $xP_{qp}^{\leq h(qp)}y$, *then* $xS^c y$,
   where $P=\{q1,q2,...,qp\}\subseteq C$ and $(h(q1),h(q2),...,h(qp))\in H_{q1}\times H_{q2}\times...\times H_{qp}$; these rules are supported by pairs of actions from the P-lower approximation of S^c only;
3) $D_{\geq\leq}$-decision rules, being statements of the type:
   if $xP_{q1}^{\geq h(q1)}y$ *and* $xP_{q2}^{\geq h(q2)}y$ *and* ... $xP_{qk}^{\geq h(qk)}y$ *and* $xP_{qk+1}^{\leq h(qk+1)}y$ *and* ... $xP_{qp}^{\leq h(qp)}y$, *then* xSy *or* $xS^c y$,
   where $O'=\{q1,q2,...,qk\}\subseteq C$, $O''=\{qk+1,qk+2,...,qp\}\subseteq C$, $P=O'\cup O''$, O' and O'' not necessarily disjoint, $(h(q1),h(q2),...,h(qp))\in H_{q1}\times H_{q2}\times...\times H_{qp}$; these rules are supported by actions from the P-boundary of S and S^c only.

14.3.3.6 Dominance without degrees of preference

The degree of graded preference considered in Section 14.3.3.1 is defined on a *quantitative* scale of the strength of preference k_q, $q\in C$. However, in many real world problems, the existence of such a quantitative scale is rather questionable. Roy [70] distinguishes the following cases:

- preferences expressed on an *ordinal* scale: this is the case where the difference between two evaluations has no clear meaning;
- preferences expressed on a *quantitative* scale: this is the case where the scale is defined with reference to a unit clearly identified, such that it is meaningful to consider an origin (zero) of the scale and ratios between evaluations (ratio scale);
- preferences expressed on a *numerical non-quantitative* scale: this is an intermediate case between the previous two; there are two well-known particular cases:

- interval scale, where it is meaningful to compare ratios between differences of pairs of evaluations,
- scale for which a complete preorder can be defined on all possible pairs of evaluations.

The preference scale has also been considered within economic theory (e.g. [74]), where cardinal utility is distinguished from ordinal utility: the former deals with a strength of preference, while, for the latter, this concept is meaningless. From this point of view, preferences expressed on an ordinal scale refer to ordinal utility while preferences expressed on a quantitative scale or a numerical non-quantitative scale deal with cardinal utility.

The strength of preference k_q and, therefore, the graded preference considered in Section 14.3.3.1, is meaningful when the scale is quantitative or numerical non-quantitative. If the information about k_q is non-available, then it is possible to define a rough approximation of S and S^c using a specific dominance relation between pairs of actions from $A\times A$, defined on an ordinal scale represented by evaluations $c_q(x)$ on criterion q, for $x\in A$ [30]. Let us explain this latter case in more details.

Let C^O be the set of criteria expressing preferences on an ordinal scale, and C^N, the set of criteria expressing preferences on a quantitative scale or a numerical non-quantitative scale, such that $C^O\cup C^N=C$ and $C^O\cap C^N=\emptyset$. Moreover, for each $P\subseteq C$, we denote by P^O the subset of P composed of criteria expressing preferences on an ordinal scale, i.e. $P^O=P\cap C^O$, and P^N the subset of P composed of criteria expressing preferences on a quantitative scale or a numerical non-quantitative scale, i.e. $P^N=P\cap C^N$. Of course, for each $P\subseteq C$, we have $P=P^N\cup P^O$ and $P^O\cap P^N=\emptyset$.

If $P=P^N$ and $P^O=\emptyset$, then the definition of dominance is the same as in the case of multigraded dominance (Section 14.3.3.5). If $P=P^O$ and $P^N=\emptyset$, then, given $(x,y),(w,z)\in A\times A$, the pair (x,y) is said to dominate the pair (w,z) with respect to P if, for each $q\in P$, $c_q(x)\geq c_q(w)$ and $c_q(z)\geq c_q(y)$. Let $D_{\{q\}}$ be the dominance relation confined to the single criterion $q\in P^O$. The binary relation $D_{\{q\}}$ is reflexive ($(x,y)D_{\{q\}}(x,y)$, $\forall(x,y)\in A\times A$), transitive ($(x,y)D_{\{q\}}(w,z)$ and $(w,z)D_{\{q\}}(u,v)$ imply $(x,y)D_{\{q\}}(u,v)$, $\forall(x,y),(w,z),(u,v)\in A\times A$), but non-complete (it is possible that *not* $(x,y)D_{\{q\}}(w,z)$ and *not* $(w,z)D_{\{q\}}(x,y)$ for some $(x,y),(w,z)\in A\times A$). Therefore, $D_{\{q\}}$ is a partial preorder. Since the intersection of partial preorders is also a partial preorder and $D_P=\bigcap_{q\in P} D_{\{q\}}$, $P=P^O$, then the dominance relation D_P is also a partial preorder.

If some criteria from $P\subseteq C$ express preferences on a quantitative or a numerical non-quantitative scale and others on an ordinal scale, i.e. if $P^N\neq\emptyset$ and $P^O\neq\emptyset$, then, given $(x,y),(w,z)\in A\times A$, the pair (x,y) is said to dominate the pair (w,z) with respect to criteria from P, if (x,y) dominates (w,z) with respect to both P^N and P^O. Since the dominance relation with respect to P^N is a partial preorder on $A\times A$ (because it is a multigraded dominance) and the dominance with respect to P^O is also a partial preorder on $A\times A$ (as explained above), then also the dominance D_P, being the intersection of these two dominance relations, is a partial preorder. In

consequence, all the concepts introduced in the previous point can be restored using this specific definition of dominance relation.

Using the approximations of S and S^c based on the dominance relation defined above, it is possible to induce a generalized description of the available preferential information in terms of decision rules. These decision rules are of the same type as the rules already introduced in the previous Section; however, the conditions on criteria from C^O are expressed directly in terms of evaluations belonging to domains of these criteria. Let $C_q=\{c_q(x), x\in A\}$, $q\in C^O$. The decision rules have in this case the following syntax:

1) $D_{\geq}$-decision rule, being a statement of the type:
   if $xP_{q1}^{\geq h(q1)}y$ *and* ... $xP_{qe}^{\geq h(qe)}y$ *and* $c_{qe+1}(x)\geq r_{qe+1}$ *and* $c_{qe+1}(y)\leq s_{qe+1}$ *and* ... $c_{qp}(x)\geq r_{qp}$ *and* $c_{qp}(y)\leq s_{qp}$, *then* xSy,
   where $P=\{q1,...,qp\}\subseteq C$, $P^N=\{q1,...,qe\}$, $P^O=\{qe+1,...,qp\}$, $(h(q1),...,h(qe))\in H_{q1}\times...\times H_{qe}$ and $(r_{qe+1},...,r_{qp})$, $(s_{qe+1},...,s_{qp})\in C_{qe+1}\times...\times C_{qp}$; these rules are supported by pairs of actions from the P-lower approximation of S only;
2) $D_{\leq}$-decision rule, being a statement of the type:
   if $xP_{q1}^{\leq h(q1)}y$ *and* ... $xP_{qp}^{\leq h(qp)}y$ *and* $c_{qe+1}(x)\leq r_{qe+1}$ *and* $c_{qe+1}(y)\geq s_{qe+1}$ *and* ... $c_{qp}(x)\leq r_{qp}$ *and* $c_{qp}(y)\geq s_{qp}$, *then* xS^cy,
   where $P=\{q1,...,qp\}\subseteq C$, $P^N=\{q1,...,qe\}$, $P^O=\{qe+1,...,qp\}$, $(h(q1),...,h(qe))\in H_{q1}\times...\times H_{qe}$ and $(r_{qe+1},...,r_{qp})$, $(s_{qe+1},...,s_{qp})\in C_{qe+1}\times...\times C_{qp}$; these rules are supported by pairs of actions from the P-lower approximation of S^c only;
3) $D_{\geq\leq}$-decision rule, being a statement of the type:
   if $xP_{q1}^{\geq h(q1)}y$ *and* ... $xP_{qe}^{\geq h(qe)}y$ *and* $xP_{qe+1}^{\leq h(qe+1)}y$ ... $xP_{qf}^{\leq h(qf)}y$ *and* $c_{qf+1}(x)\geq r_{qf+1}$ *and* $c_{qf+1}(y)\leq s_{qf+1}$ *and* ... $c_{qg}(x)\geq r_{qg}$ *and* $c_{qg}(y)\leq s_{qg}$ *and* $c_{qg+1}(x)\leq r_{qg+1}$ *and* $c_{qg+1}(y)\geq s_{qg+1}$ *and* ... $c_{qp}(x)\leq r_{qp}$ *and* $c_{qp}(y)\geq s_{qp}$, *then* xSy *or* xS^cy,
   where $O'=\{q1,...,qe\}\subseteq C$, $O''=\{qe+1,...,qf\}\}\subseteq C$, $P^N=O'\cup O''$, O' and O'' not necessarily disjoint, $P^O=\{qf+1,...,qp\}$, $(h(q1),...,h(qf))\in H_{q1}\times...\times H_{qf}$ and $(r_{qf+1},...,r_{qp}),(s_{qf+1},...,s_{qp})\in C_{qf+1}\times...\times C_{qp}$; these rules are supported by pairs of actions from the P-boundary of S and S^c only.

14.3.3.7 An example

For the illustration of results concerning the problems of multicriteria choice and ranking we will use the same example as considered in Section 14.1.7. Firstly, we assume that the DM accepts to express preferences with respect to criteria A_1, A_2, A_3 on a *numerical non-quantitative scale* for which a complete preorder can be defined on all possible pairs of evaluations. According to this assumption, in order to build the PCT, as described in Section 14.3.3.1, the DM specifies sets of possible degrees of preference; for example, $H_1=\{-2,-1,0,1,2\}$, $H_2=\{-1,0,1\}$, $H_3=\{-1,0,1\}$. Therefore, with respect to A_1, we have the following preference relations P_1^h:

- $x P_1^2 y$ (and $y P_1^{-2} x$), meaning that x is preferred to y with respect to A_1, if $f(x,A_1)$=high and $f(y,A_1)$=low,

- $x P_1^1 y$ (and $y P_1^{-1} x$), meaning that x is weakly preferred to y with respect to A_1, if $f(x,A_1)$=high and $f(y,A_1)$=medium or $f(x,A_1)$=medium and $f(y,A_1)$=low,

- $x P_1^0 y$ (and $y P_1^0 x$), meaning that x is indifferent to y with respect to A_1, if $f(x,A_1)=f(y,A_1)$.

Analogously, with respect to A_2 and A_3, we have the following preference relations P_2^h and P_3^h:

- $x P_2^1 y$ (and $y P_2^{-1} x$), meaning that x is weakly preferred to y with respect to A_2, if $f(x,A_2)$=good and $f(y,A_2)$=medium,

- $x P_2^0 y$ (and $y P_2^0 x$), meaning that x is indifferent to y with respect to A_2, if $f(x,A_2)=f(y,A_2)$,

- $x P_3^1 y$ (and $y P_3^{-1} x$), meaning that x is weakly preferred to y with respect to A_3, if $f(x,A_3)$=yes and $f(y,A_3)$=no,

- $x P_3^0 y$ (and $y P_3^0 x$), meaning that x is indifferent to y with respect to A_3, if $f(x,A_3)=f(y,A_3)$.

As to the comprehensive preference relation, the DM considers that, given two different warehouses $x,y \in U=\{1,2,3,4,5,6\}$, if x makes profit and y makes loss, then xSy and $yS^c x$. Moreover, the DM accepts xSx for each warehouse x. As to warehouses x and y which both make profit or both make loss, the DM abstains from judging whether xSy or $xS^c y$. Therefore, the set of exemplary pairwise comparisons supplied by the DM is B={(1,1), (1,2), (1,4), (1,5), (2,1), (2,2), (2,3), (2,6), (3,2), (3,3), (3,4), (3,5), (4,1), (4,3), (4,4), (4,6), (5,1), (5,3), (5,5), (5,6), (6,2), (6,4), (6,5), (6,6)}.

At this stage, we are able to build the PCT, shown in Table 14.5.

Table 14.5 Pairwise comparison table

Pairs	P_1^h	P_2^h	P_3^h	Outranking
(1,1)	0	0	0	S
(1,2)	1	1	0	S
(1,4)	2	1	0	S
(1,5)	1	0	-1	S
(2,1)	-1	-1	0	S^c
(2,2)	0	0	0	S
(2,3)	0	0	0	S^c

(2,6)	-1	0	-1	S^c
(3,2)	0	0	0	S
(3,3)	0	0	0	S
(3,4)	1	0	0	S
(3,5)	0	-1	-1	S
(4,1)	-2	-1	0	S^c
(4,3)	-1	0	0	S^c
(4,4)	0	0	0	S
(4,6)	-2	0	-1	S^c
(5,1)	-1	0	1	S^c
(5,3)	0	1	1	S^c
(5,5)	0	0	0	S
(5,6)	-1	1	0	S^c
(6,2)	1	0	1	S
(6,4)	2	0	1	S
(6,5)	1	-1	0	S
(6,6)	0	0	0	S

The C-lower approximations, the C-upper approximations and the C-boundaries of S and S^c obtained by means of multigraded dominance relations are as follows:

$\underline{C}(S) = \{(1,2),(1,4),(1,5),(3,4),(6,2),(6,4),(6,5)\}$,

$\overline{C}(S)=\{(1,1),(1,2),(1,4),(1,5),(2,2),(2,3),(3,2),(3,3),(3,4),(3,5),(4,4),(5,3),(5,5),(6,2),(6,4),(6,5),(6,6)\}$,

$\underline{C}(S^c) = \{(2,1),(2,6),(4,1),(4,3),(4,6),(5,1),(5,6)\}$,

$\overline{C}(S^c)=\{(1,1),(2,1),(2,2),(2,3),(2,6),(3,2),(3,3),(3,5),(4,1),(4,3),(4,4),(4,6),(5,1),(5,3),(5,5),(5,6),(6,6)\}$,

$Bn_C(S) = Bn_C(S^c) = \{(1,1),(2,2),(2,3),(3,2),(3,3),(3,5),(4,4),(5,3),(5,5),(6,6)\}$.

Therefore, the accuracy of the approximation is equal to 0.41 for S and also for S^c, while the quality of approximation is equal to 0.58. There is only one reduct which is also the core, i.e. $Red_S(C)=Core_S(C)=\{A_1\}$.

Finally, we obtain the following decision rules (within parentheses there are the pairs of actions supporting the rule):

1) *if* $x P_1^{\geq 1} y$, *then* xSy (or, in simple words, *if* x is at least weakly preferred to y with respect to A_1, *then* x outranks y), ((1,2),(1,4),(1,5),(3,4),(6,2),(6,4),(6,5)),
2) *if* $x P_1^{\leq -1} y$, *then* xS^cy (or, in simple words, *if* y is at least weakly preferred to x with respect to A_1, *then* x does not outrank y), ((2,1),(2,6),(4,1),(4,3),(4,6),(5,1),(5,6)),
3) *if* $x P_1^{\geq 0} y$ *and* $x P_1^{\leq 0} y$ (i.e. *if* $x P_1^0 y$), *then* xSy *or* xS^cy (or, in simple words, *if* x is indifferent with y with respect to A_1, *then* x outranks y *or* x does not outrank y), ((1,1),(2,2),(2,3),(3,2),(3,3),(3,5),(4,4),(5,3),(5,5),(6,6)).

Let us assume now that the DM accepts to express preferences with respect to criteria A_1, A_2, A_3 on an *ordinal scale* of preference for which there is only information about a partial preorder on all possible pairs of evaluations. In this case, S and S^c can be approximated in the way described in Section 14.3.3.6, i.e. without considering degrees of preference. The C-lower approximations, the C-upper approximations and the C-boundaries of S and S^c are as follows:

$\underline{C}(S) = \{(1,1),(1,2),(1,4),(1,5),(3,4),(4,4),(6,2),(6,4),(6,5),(6,6)\}$,

$\overline{C}(S)=\{(1,1),(1,2),(1,4),(1,5),(2,2),(2,3),(3,2),(3,3),(3,4),(3,5),(4,4),(5,3),(5,5),(6,2),(6,4),(6,5),(6,6)\}$,

$\underline{C}(S^c) = \{(2,1),(2,6),(4,1),(4,3),(4,6),(5,1),(5,6)\}$,

$\overline{C}(S^c)=\{(2,1),(2,2),(2,3),(2,6),(3,2),(3,3),(3,5),(4,1),(4,3),(4,6),(5,1),(5,3),(5,5),(5,6)\}$,

$Bn_C(S) = Bn_C(S^c) = \{(2,2),(2,3),(3,2),(3,3),(3,5),(5,3),(5,5)\}$.

Let us observe that the pairs (1,1), (4,4) and (6,6) belong now to the C-lower approximation of S and are not contained in the C-boundaries. Therefore, the accuracy of the approximation is equal to 0.59 for S and to 0.5 for S^c, while the quality of approximation is equal to 0.71. There is still only one reduct which is also the core, i.e. again $Red_S(C)=Core_S(C)=\{A_1\}$. The dominance relation with respect to the reduct $\{A_1\}$, i.e. the partial preorder on set B, is presented graphically by a Hasse diagram, together with approximations of S and S^c, in Figure 14.1.

The following decision rules are induced from the above approximations and boundaries (within parentheses there are the pairs of objects supporting the rule):

1) *if* $f(x,A_1)$ is at least high *and* $f(y,A_1)$ is at most high, *then* xSy, ((1,1),(1,2),(1,4),(1,5),(6,2),(6,4),(6,5),(6,6)),
2) *if* $f(x,A_1)$ is at least low *and* $f(y,A_1)$ is at most low, *then* xSy, ((1,4),(3,4),(4,4),(6,4)),
3) *if* $f(x,A_1)$ is at most medium *and* $f(y,A_1)$ is at least high, *then* xS^cy, ((2,1),(2,6),(4,1),(4,6)(5,1),(5,6)),
4) *if* $f(x,A_1)$ is at most low *and* $f(y,A_1)$ is at least medium, *then* xS^cy, ((4,1),(4,3),(4,6))
5) *if* $f(x,A_1)$ is at least medium *and* $f(y,A_1)$ is at most medium *and* $f(x,A_1)$ is at most medium *and* $f(y,A_1)$ is at least medium, (i.e. *if* $f(x,A_1)$ is equal to medium *and* $f(y,A_1)$ is equal to medium), *then* xSy *or* xS^cy, ((2,2),(2,3),(3,2),(3,3),(3,5),(5,3),(5,5)).

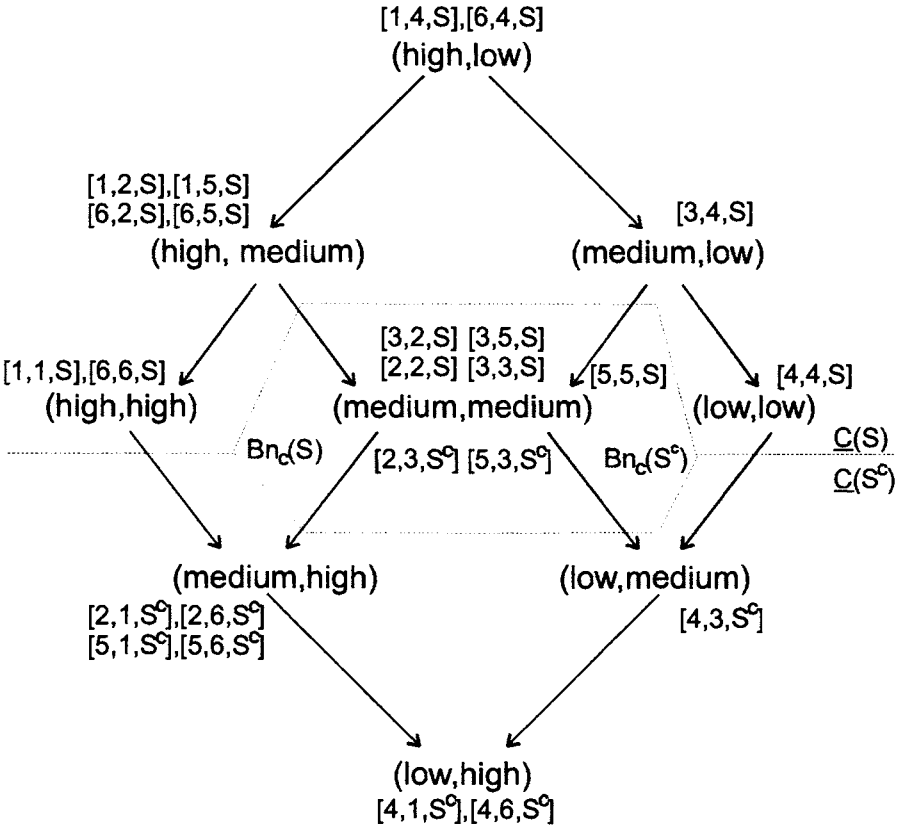


Fig. 14.1 Hasse diagram of the dominance relation with respect to the reduct $\{A_1\}$, i.e. partial preorder on B, and approximations of S and S^c

14.3.3.8 Fuzzy multigraded dominance

The rough set approximation of an outranking relation can be extended to a fuzzy context [27]. Let S be a comprehensive fuzzy outranking relation defined on A. As the knowledge of S is confined to the set $B \subseteq B \times B$, where $B \subseteq A$ is the set of reference actions, S is considered a function, $S:B \rightarrow [0,1]$, such that $\forall (x,y) \in B$, S(x,y) is the credibility of the statement "x is comprehensively at least as good as y". Of course, $S(x,x)=1$ for each $x \in B$ (reflexivity). The fuzzy relation S^c is the fuzzy complement of S, i.e. $S^c(x,y)=N(S(x,y))$.

As in the crisp case, for each $q \in C$, let H_q be a set of preference degrees with respect to q. Then, for each $q \in C$ and $h \in H_q$, a fuzzy graded preference relation on the set of actions A, $P_q^h: A \times A \rightarrow [0,1]$, can be introduced, where $P_q^h(x,y)$ is the

credibility of the statement "x is preferred to y with respect to criterion q by degree h".

In this case, according to the typology given in Section 14.1.8, the use of fuzzy sets for representation of graded preference relations is concordant with the second and the third semantics together, because we are in the presence of a vague state of information about the strength (degree) of preference. As to the use of fuzzy sets for representation of outranking relation, it is concordant with the third semantics.

Using an appropriate definition of the fuzzy multigraded dominance relation, built on the basis of fuzzy graded preference P_q^h, fuzzy outranking relations S and S^c can be approximated. More precisely, rough approximations of S and S^c are defined using a fuzzy positive dominance (denotation $D_P^+((x,y),(w,z))$ and a fuzzy negative dominance (denotation $D_P^-((x,y),(w,z))$, respectively, which are the fuzzy counterparts of sets $D_P^+(x,y)$ and $D_P^-(x,y)$ introduced in Section 14.3.3.5 to approximate crisp relations S and S^c.

Let us present step by step the process of building a fuzzy multigraded dominance satisfying some desirable properties.

First, on the basis of the fuzzy graded preference $P_q^h(x,y)$, $q\in C$, the fuzzy upward cumulated preference (denotation $P_q^{\geq h}(x,y)$) and the fuzzy downward cumulated preference (denotation $P_q^{\leq h}(x,y)$) are defined as follows:

1) the fuzzy upward cumulated preference $P_q^{\geq h}: A\times A\rightarrow[0,1]$, such that $\forall(x,y)\in A\times A$, $P_q^{\geq h}(x,y)$ is the credibility of the statement "x is preferred to y with respect to q by at least degree h", i.e.

$$P_q^{\geq h}(x,y)=\begin{cases} 1 & \text{if } \exists k\in H_q: P_q^k(x,y)>0 \text{ and } k>h \\ P_q^h(x,y) & \text{otherwise (i.e. if } \forall k>h \ P_q^k(x,y)=0) \end{cases}$$

2) the fuzzy downward cumulated preference $P_q^{\leq h}: A\times A\rightarrow[0,1]$, such that $\forall(x,y)\in A\times A$, $P_q^{\leq h}(x,y)$ is the credibility of the statement "x is preferred to y with respect to q by at most degree h", i.e.

$$P_q^{\leq h}(x,y)=\begin{cases} 1 & \text{if } \exists k\in H_q: P_q^k(x,y)>0 \text{ and } k<h \\ P_q^h(x,y) & \text{otherwise (i.e. if } \forall k<h \ P_q^k(x,y)=0) \end{cases}$$

Then, the comparison of degrees of fuzzy preference should be defined. This comparison represents the fuzzy counterpart of the statement "x is preferred to y by at least the same degree as w is preferred to z, with respect to $q\in C$". In formal terms, for $q\in C$, this corresponding crisp statement is: "$h\geq k$, where $x P_q^h y$ and $w P_q^k z$, $h,k\in H_q$" or, in short, "$h\geq k$, $h,k\in H_q$". From a semantic point of view, the latter comparison is equivalent to the following set of implications: "$k\geq r$ implies

$h \geq r$, for each $r \in H_q$". Thus, the statement "$h \geq k$, where $x P_q^h y$ and $w P_q^k z$, $h,k \in H_q$" is equivalent to: "$w P_q^{\geq r} z$ implies $x P_q^{\geq r} y$ for each $r \in H_q$". This set of implications is finally expressed in fuzzy terms and its credibility (denotation $P_q^+((x,y),(w,z))$) is equal to:

$$P_q^+((x,y),(w,z)) = \underset{r \in H_q}{T} (I^{\rightarrow}(P_q^{\geq r}(w,z), P_q^{\geq r}(x,y)).$$

Analogously, given $(x,y),(w,z) \in A \times A$, the comparison statement "x is preferred to y by at most the same degree as w is preferred to z, with respect to $q \in C$" is equivalent, in crisp terms, to the set of implications "$w P_q^{\leq r} z$ implies $x P_q^{\leq r} y$ for each $r \in H_q$", whose credibility (denotation $P_q^-((x,y),(w,z))$) is equal to:

$$P_q^-((x,y),(w,z)) = \underset{r \in H_q}{T} (I^{\rightarrow}(P_q^{\leq r}(w,z), P_q^{\leq r}(x,y))).$$

The following properties are desirable for $P_q^+((x,y),(w,z))$ and $P_q^-((x,y),(w,z))$, $q \in C$:

1) $[P_q^{\geq r}(x,y) \geq P_q^{\geq r}(w,z), \forall r \in H_q] \Leftrightarrow P_q^+((x,y),(w,z)) = 1$,

2) $[P_q^{\leq r}(x,y) \geq P_q^{\leq r}(w,z), \forall r \in H_q] \Leftrightarrow P_q^-((x,y),(w,z)) = 1$,

3) $P_q^+((x,y),(w,z)) = P_q^-((w,z),(x,y)), \forall (x,y),(w,z) \in A \times A$.

The first two properties represent monotonicity of comparisons with respect to cumulated preferences. The third property represents a kind of symmetry of comparisons. Property 1) says that for each $q \in C$, x is certainly (with credibility equal to 1) preferred to y by at least the same degree as w is preferred to z, iff for each $r \in H_q$ the fuzzy upward cumulated preference of x over y is not smaller than the fuzzy upward cumulated preference of w over z. Property 2) has an analogous interpretation with respect to the fuzzy downward cumulated preferences. Property 3) says that for each $q \in C$, the credibility of the comparison statement "x is preferred to y by at least the same degree as w is preferred to z" equals the credibility of the symmetric comparison statement "w is preferred to z by at most the same degree as x is preferred to y".

A sufficient condition for properties 1) and 2) is that the fuzzy implication $I^{\rightarrow}$ considered in the definition of $P_q^+((x,y),(w,z))$ and $P_q^-((x,y),(w,z))$ satisfies the following requirement, $\forall a,b \in [0,1]$:

$$I^{\rightarrow}(a,b) = 1 \Leftrightarrow a \leq b.$$

The following two conditions are sufficient for property 3):

- for each $h \in H_q - \{1\}$, $P_q^{\geq h}(x,y) = N(P_q^{\leq h-1}(x,y))$, $P_q^{\leq h-1}(x,y) = N(P_q^{\geq h}(x,y))$,
- for each $a,b \in [0,1]$, $I^{\rightarrow}(a,b) = I^{\rightarrow}(N(b),N(a))$.

The implications of Łukasiewicz and Fodor satisfy the above conditions concerning implications (see e.g. [13]).

Now, it is possible to define fuzzy positive and fuzzy negative dominance. Given $(x,y),(w,z)\in A\times A$ and $P\subseteq C$, let

$$D_P^+((x,y),(w,z)) = \underset{q\in P}{T} P_q^+((x,y),(w,z)),$$

$$D_P^-((x,y),(w,z)) = \underset{q\in P}{T} P_q^-((x,y),(w,z)).$$

Let us observe that $D_P^+((x,y),(w,z))$ represents the credibility of the statement "x is preferred to y by at least the same degree as w is preferred to z, taking into account all criteria q from P", while $D_P^-((x,y),(w,z))$ represents the credibility of the statement "x is preferred to y by at most the same degree as w is preferred to z, taking into account all criteria q from P".

Finally, the P-lower and the P-upper approximations of S are fuzzy sets defined on B whose membership functions (denotation $\underline{P}[S(x,y)]$ and $\overline{P}[S(x,y)]$) are, respectively:

$$\underline{P}[S(x,y)] = \underset{(w,z)\in B}{T} (T^*(N(D_P^+((w,z),(x,y))), S(w,z))),$$

$$\overline{P}[S(x,y)] = \underset{(w,z)\in B}{T^*} (T(D_P^+((x,y),(w,z)), S(w,z))).$$

$\underline{P}[S(x,y)]$ represents the credibility of the statement "for all $(w,z)\in B$, (w,z) does not dominate (x,y) and/or w outranks z", while $\overline{P}[S(x,y)]$ represents the credibility of the statement "there is at least one $(w,z)\in B$ such that (x,y) dominates (w,z) and w outranks z".

Analogously, the P-lower and the P-upper approximations of S^c are fuzzy sets defined on B whose membership functions (denotation $\underline{P}[S^c(x,y)]$ and $\overline{P}[S^c(x,y)]$) are, respectively:

$$\underline{P}[S^c(x,y)] = \underset{(w,z)\in B}{T} (T^*(N(D_P^-((w,z),(x,y))), S^c(w,z))),$$

$$\overline{P}[S^c(x,y)] = \underset{(w,z)\in B}{T^*} (T(D_P^-((x,y),(w,z)), S^c(w,z))).$$

Let us remark that using the definition of the T^*-implication, it is possible to rewrite the definition of $\underline{P}[S(x,y)]$, $\overline{P}[S(x,y)]$, $\underline{P}[S^c(x,y)]$ and $\overline{P}[S^c(x,y)]$ in the following way:

$$\underline{P}[S(x,y)] = \underset{(w,z)\in B}{T} (I^{\rightarrow}_{T^*,N}(D_P^+((w,z),(x,y)), S(w,z))),$$

$$\overline{P}[S(x,y)] = \underset{(w,z)\in B}{T^*} (N(I^{\rightarrow}_{T^*,N}(D_P^+((x,y),(w,z)), N(S(w,z))))),$$

$$\underline{P}[S^c(x,y)] = \underset{(w,z)\in B}{T} (I^{\rightarrow}_{T^*,N}(D_P^-((w,z),(x,y)), S^c(w,z))),$$

$$\overline{P}[S^c(x,y)] = \underset{(w,z)\in B}{T^*} (N(I^{\rightarrow}_{T^*,N}(D_P^-((x,y),(w,z)), N(S^c(w,z))))).$$

It may be proved that if properties 1) and 2) above hold, then

$$\underline{P}[S(x,y)] \leq S(x,y) \leq \overline{P}[S(x,y)],$$

$$\underline{P}[S^c(x,y)] \leq S^c(x,y)) \leq \overline{P}[S^c(x,y)],$$

i.e., in terms of the fuzzy inclusion, S and S^c include their lower approximation and are included in their upper approximation.

If (T,T^*,N) constitutes a De Morgan triplet, N is involutive and property 3) above is satisfied, then the following complementarity properties hold:

i) $\underline{P}[S(x,y)] = N(\overline{P}[S^c(x,y)])$,

ii) $\overline{P}[S(x,y)] = N(\underline{P}[S^c(x,y)])$,

iii) $\underline{P}[S^c(x,y)] = N(\overline{P}[S(x,y)])$,

iv) $\overline{P}[S^c(x,y)] = N(\underline{P}[S(x,y)])$.

Property i), expressed in terms of a complement of a fuzzy set, means that the lower approximation of S is the complement of the upper approximation of the complement of S, that is S^c. Properties ii) to iv) have an analogous interpretation.

14.3.4 Formal equivalence of decision rule preference models and conjoint measurement models

Traditionally, preferences are modeled using a value function $u(\cdot)$ such that action a is at least as good as action b, i.e. aSb, iff $u(a) \geq u(b)$. This implies that the relation S is complete (for each couple of actions $a,b \in A$, aSb and/or bSa) and transitive (for each triple of actions $a,b,c \in A$, aSb and bSc imply aSc). In a multicriteria context, each action a is generally seen as a vector $c(a)=(c_1(a),c_2(a),...,c_m(a))$ of evaluations with reference to the m criteria $c_1(a),c_2(a),...,c_m(a)$. It is often assumed that the value function is additive (see, e.g., [40] and, for an axiomatic characterization, [41, 102]), i.e.

$$u(a) = \sum_{q=1}^{m} u_q[c_q(a)],$$

where u_q, $q=1,...,m$, are non-decreasing functions.

The additive and transitive model represented by the additive value function is inappropriate in many situations, because:

- the indifference (the symmetric part of S) may not be transitive,
- S may not be complete, that is some actions may be incomparable,
- the compensation between conflicting criteria and the interaction between concordant criteria are far more complex than the capacity of representation by the additive value function.

To take these limitations into account, a variety of extensions have been proposed (e.g. [99, 12]). Bouyssou and Pirlot [4] have recently proposed a model generalizing the previous ones and creating an axiomatic basis to many multicriteria decision methods presented in the literature (see, e.g., [71, 101]). This model drops additivity, transitivity and completeness properties, and may be written as

i) aSb iff $F[\Psi_q(c_q(a), c_q(b))_{q=1,\ldots,m}] \geq 0$,

where Ψ_q: $\mathbf{R}^2 \rightarrow \mathbf{R}$ is a non-decreasing function in its first argument and a non-increasing function in its second argument, for q=1,...,m, and F: $\mathbf{R}^m \rightarrow \mathbf{R}$ is a non-decreasing function in every one of its arguments. Observe that the values assumed by the function Ψ_q may be interpreted as a measure of the strength of preference of a over b with respect to criterion q, q=1,...,m. Thus, Ψ_q plays the same role as the function k_q in the definition of the PCT.

Recently, Greco, Matarazzo and Slowinski [31, 32] have proposed some more general models of conjoint measurement. The first model may be written as [31]:

ii) aSb iff $G[\Psi_q(c_q(a), c_q(b))_{q=1,\ldots,k}, c_q(a), c_q(b)_{q=k+1,\ldots,m}] \geq 0$,

where the indices of the considered criteria are reordered such that {1,...,k} is the set of criteria for which the preference is expressed on a quantitative or a numerical non-quantitative scale and {k+1,...,m} is the set of criteria for which the preference is expressed on an ordinal scale; Ψ_q is defined as above for q=1,...,k, and G: $\mathbf{R}^{k+2(m-k)} \rightarrow \mathbf{R}$ is a non-decreasing function in its first k arguments, non-decreasing in each (k+'odd') argument ('odd'=1,3,...,2(m-k)-1), and non-increasing in each (k+'even') argument ('even'=2,4,...,2(m-k)). Greco, Matarazzo and Slowinski proved that model ii) is based on the same axioms as the model i), with the exception of an axiom which, for each q=1,...,m, introduces a total preorder on the set of pairs $(c_q(a), c_q(b))$, representing the preference strength, with respect to the considered criterion, in pairwise action comparisons. More precisely, this axiom is accepted only for q=1,...,k, i.e. for the set of criteria with a quantitative or a numerical non-quantitative preference scale.

Moreover, Greco, Matarazzo and Slowinski [32] have proposed a model of conjoint measurement to represent also some *inconsistencies in the preferences*. This model is based on the concepts of C-lower and C-upper approximation and of C-boundary of S and S^c, where C={1,...,m}. This model can be written as

iiia) $(a,b) \in \underline{C}(S)$ iff $G[\Psi_q(c_q(a), c_q(b))_{q=1,\ldots,k}, c_q(a), c_q(b)_{q=k+1,\ldots,m}] \geq t_2$,

iiib) $(a,b) \in \underline{C}(S^c)$ iff $G[\Psi_q(c_q(a), c_q(b))_{q=1,\ldots,k}, c_q(a), c_q(b)_{q=k+1,\ldots,m}] \leq t_1$,

iiic) $(a,b) \in Bn_C(S)$ (or, equivalently, $(a,b) \in Bn_C(S^c)$) iff $t_1 < G[\Psi_q(c_q(a), c_q(b))_{q=1,\ldots,k}, c_q(a), c_q(b)_{q=k+1,\ldots,m}] < t_2$,

where Ψ_q and G are defined as above and $t_1, t_2 \in \mathbf{R}$ such that $t_1 < t_2$. With respect to the model iiia)-iiic), Greco, Matarazzo and Slowinski proved that it is always possible to obtain such a representation, i.e. S should not satisfy any specific axiom.

It is interesting to compare the above models of conjoint measurement with the *decision rule preference models* resulting from the rough set approach (Sections 14.3.3.5 and 14.3.3.6). The following results have been proved:

1) the outranking relation S may be represented by means of the non-additive, non-transitive and non-complete model i) *if and only if* it may be

represented by means of a set of $D_{\geq}$-decision rules having a syntax defined in Section 14.3.3.5 [22],

2) the outranking relation S may be represented by means of the non-additive, non-transitive and non-complete model ii) *if and only if* it may be represented by means of a set of $D_{\geq}$-decision rules having a syntax defined in Section 14.3.3.6 [31].

Greco, Matarazzo and Slowinski [32] have also pointed out the clear equivalence between the representation of S and S^c obtained using the rough set approach proposed in Section 14.3.3.6 and the model of conjoint measurement iiia)-iiic). Furthermore, they observed that the rough set representation proposed in Section 14.3.3.5 can be viewed as a particular case of that one presented in 3.3.6, when the set of criteria with an ordinal preference scale is empty.

14.4 Conclusions

In this chapter, we made a synthesis of the contribution of the rough sets theory, and of related with it fuzzy sets theory, to multiattribute and multicriteria decision making. Classic use of the rough set approach, and more generally, of machine learning, data mining and knowledge discovery, deals with problems of multiattribute classification, i.e. problems where neither the attributes describing the objects, nor the classes to which the objects are assigned, are ordered. On the other hand, MCDM deals with problems where descriptions (evaluations) of objects (actions) by means of attributes (criteria), as well as decisions in sorting, choice and ranking problems, are ordered. The generalization of the rough set approach to problems in which order properties are important is possible upon two important methodological extensions: approximation by dominance relations, which allows to deal with order properties of criteria, and pairwise comparison table, which allows to handle preference relations for choice and ranking problems. Making these extensions, it was possible to keep the main advantage of the rough set approach that consists in its ability of handling inconsistency. The notion of inconsistency has been even enlarged, because the extended rough set approach can deal with inconsistency also in the sense of dominance principle. As the available information may not only be inconsistent but also imperfect in the sense of graduality of similarity, uncertainty and preference, the rough set approach can be fruitfully complemented by the fuzzy set approach.

Let us point out the main advantages of the extended rough set approach for MCDM:

- preferential information necessary to deal with any multicriteria decision problem is asked to the DM just in terms of exemplary decisions,
- the rough set analysis of preferential information supplies some useful elements of knowledge about the decision situation; these are: the relevance of attributes and/or criteria, information about their interaction (from quality of approximation and its analysis using fuzzy measures theory), minimal subsets of

attributes or criteria (reducts) conveying the relevant knowledge contained in the exemplary decisions, the set of the non-reducible attributes or criteria (core),

- the preference model induced from the preferential information is expressed in a natural and comprehensible language of "*if...*, *then...*" decision rules,
- the rules based on the use of dominance have a more general syntax than the rules based on indiscernibility,
- suitable procedures have been proposed to exploit the results of application of the decision rule preference model on a set of actions in order to workout a recommendation within choice, ranking, classification or sorting problem,
- no prior discretization of quantitative condition attributes and/or criteria is necessary,
- heterogeneous information (qualitative and quantitative, ordered and non-ordered, crisp and fuzzy evaluations, and ordinal, quantitative and numerical non-quantitative scales of preferences) can be processed within the extended rough set approach, while classic MCDM methods consider only quantitative ordered evaluations with rare exceptions,
- the proposed methodology fulfils some desirable properties for both rough set approach (the approximated sets include lower approximation and are included in upper approximation, and the complementarity property is satisfied), and for MCDM (the decision rule preference model is formally equivalent to the non-additive, non-transitive and non-complete conjoint measurement model, and to a more general model for preferences defined on all kinds of scales),
- the decision rule preference model resulting from the rough set approach is more general than all existing models of conjoint measurement, due to its capacity of handling inconsistent preferences (a new model of conjoint measurement is formally equivalent to the decision rule preference model handling inconsistencies),
- the proposed methodology is based on elementary concepts and mathematical tools (sets and set operations, binary relations), without recourse to any algebraic or analytical structures; main ideas are very natural and, in a certain sense, even objective: indiscernibility, similarity and dominance.

Acknowledgement

The research of the first two authors has been supported by grant no. 97.01387.CT10 from Italian National Council for Scientific Research (CNR). The third author wishes to acknowledge financial support from State Committee for Scientific Research, KBN research grant no. 8 T11C 013 13, and from CRIT 2 Esprit Project no. 20288.

References

[1] Banzhaf, J. F.: Weighted voting doesn't work: A mathematical analysis. Rutgers Law Review 19 (1965) 317-343

[2] Bazan, J., Skowron, A., Synak, P.: "Dynamic reducts as a tool for extracting laws from decision tables". In: M. Zemankowa, Z. Ras (eds): *Methodologies for Intelligent Systems*. LNAI, Vol. 869, Springer-Verlag, Berlin 1994, pp. 346-355

[3] Bouyssou, D.: Outranking relations: Do they have special properties? Journal of Multi-Criteria Decision Analysis 5 (2) (1996)99-111

[4] Bouyssou, D., Pirlot, M.: *A general framework for the aggregation of semiorders*. Technical Report, ESSEC, Cergy-Pontoise, 1997

[5] Chmielewski, M., Grzymala-Busse, J.: "Global discretization of continuous attributes as preprocessing for machine learning". In: Lin, T.Y., Wildberger, A. (eds): *Soft Computing: Rough Sets, Fuzzy Logic, Neural Networks, Uncertainty Management*. Simulation Councils Inc., San Diego, CA 1995, pp. 294-301

[6] Dennemberg, D., Grabisch, M.: Shapley value and interaction index. 1996, Working paper

[7] Dubois, D., Prade, H.: Rough fuzzy sets and fuzzy rough sets. Int. J. of General Systems 17(1990)191-200

[8] Dubois, D., Prade, H.: "Putting rough sets and fuzzy sets together". In: R. Slowinski (ed): *Intelligent Decision Support, Handbook of Applications and Advances of the Rough Sets Theory*. Kluwer, Dordrecht 1992, pp.203-233

[9] Dubois, D., Prade, H., Yager, R.R.: "A manifesto: Fuzzy information engineering". In: D. Dubois, H., Prade, R. R. Yager (ed): *Fuzzy Information Engineering*. Wiley, New York 1997, pp.1-8

[10] Fayyad, U.M., Irani, K.B.: On the handling of continuous-valued attributes in decision tree generation. Machine Learning 8(1992)87-102

[11] Fishburn, P. C.: Methods for estimating additive utilities. Management Science 13(1967)435-453

[12] Fishburn, P.C.: Nontransitive additive conjoint measurement. Journal of Mathematical Psychology 35(1991)1-40

[13] Fodor, J., Roubens, M.: *Fuzzy Preference Modelling and Multicriteria Decision Support*. Kluwer, Dordrecht 1994

[14] Grabisch, M.: The application of fuzzy integrals in multicriteria decision making. European Journal of Operational Research 89(1996) 445-456

[15] Grabisch, M.: k-order additive discrete fuzzy measures and their representation. Fuzzy Sets and Systems 92(1997)167-189

[16] Grabisch, M., Roubens, M.: "Equivalent representations of a set function with application to decision making", paper presented at *FUZZ-IEEE '97 Conference*, Barcelona 1997

[17] Greco S., Matarazzo, B., Slowinski, R.: *Rough set approach to multi-attribute choice and ranking problems*. ICS Research Report 38/95, Warsaw University of Technology, Warsaw 1995, and in Proceedings of the Twelfth International

Conference, Hagen (Germany); G. Fandel, T. Gal (eds): *Multiple Criteria Decision Making*, Berlin 1997, pp. 318-329

[18] Greco S., Matarazzo, B., Slowinski, R.: *Rough approximation of a preference relation by dominance relations.* ICS Research Report 16/96, Warsaw University of Technology, Warsaw 1996 and in print in European Journal of Operational Research (1998)

[19] Greco, S., Matarazzo, B., Slowinski, R.: *Rough approximations by fuzzy similarity relations.* Working paper, University of Catania, Catania 1997

[20] Greco, S., Matarazzo, B., Slowinski, R.: *Rough approximation of a preferential information.* Working paper, University of Catania, Catania 1997

[21] Greco, S., Matarazzo, B., Slowinski, R.: *Exploitation procedures for rough set analysis of multicriteria decision problems.* Working paper, University of Catania, Catania 1997

[22] Greco, S., Matarazzo, B., Slowinski, R.: *Preference modeling by decision rules.* Working paper, University of Catania, Catania 1997

[23] Greco, S., Matarazzo, B., Slowinski, R.: *Fuzzy measures representation as technique for rough set analysis.* Working paper, University of Catania, Catania 1997

[24] Greco, S., Matarazzo, B., Slowinski, R.: "A new rough set approach to evaluation of bankruptcy risk". In: C. Zopounidis (ed): *Operational Tools in the Management of Financial Risks.* Kluwer, Dordrecht, 1998, pp. 121-136

[25] Greco, S., Matarazzo, B., Slowinski, R.: "Rough approximation of a preference relation in a pairwise comparison table". In: L. Polkowski, A. Skowron (eds): *Rough Sets in Knowledge Discovery.* Physica-Verlag, Heidelberg 1998, to appear

[26] Greco, S., Matarazzo, B., Slowinski, R.: "A new rough set approach to multicriteria and multiattribute classification" In: L. Polkowski, A. Skowron (eds): *Proceedings of the First International Conference on Rough sets and Current Trends in Computing (RSTCTC '98)*, Warsaw, June 22-26, 1998; Springer-Verlag, 1998, 60-67

[27] Greco, S., Matarazzo, B., Slowinski, R.: *Rough approximation of a fuzzy preference relation.* Working paper, University of Catania, Catania 1998

[28] Greco, S., Matarazzo, B., Slowinski, R.: "Fuzzy similarity relation as a basis for rough approximation". In: L. Polkowski, A. Skowron (eds): *Proceedings of the First International Conference on Rough sets and Current Trends in Computing (RSTCTC '98)*, June 22-26, Warsaw 1998, Springer-Verlag, 1998, 283-289

[29] Greco, S., Matarazzo, B., Slowinski, R.: *Fuzzy dominance as a basis for rough approximations.* Working paper, University of Catania, Catania 1998

[30] Greco, S., Matarazzo, B., Slowinski, R.: *Rough approximation of a preference relation using ordinal scales of preference.* Working paper, University of Catania, Catania 1998

[31] Greco, S., Matarazzo, B., Slowinski, R.: *A conjoint measurement model to represent preference on ordinal scales.* Working paper, University of Catania, Catania 1998

[32] Greco, S., Matarazzo, B., Slowinski, R.: *A general model of conjoint measurement to represent preference inconsistencies.* Working paper, University of Catania, Catania 1998

[33] Greco, S., Matarazzo, B., Slowinski, R.: *A rough approximation using indiscernbility, similarity and dominance relations.* Working paper, University of Catania, Catania 1998

[34] Greco, S., Matarazzo, B., Slowinski, R., Tsoukias, A.: *Exploitation of a rough approximation of the outranking relation.* Cahier du LAMSADE no. 152, Université de Paris-Dauphine, Paris 1997

[35] Greco, S., Matarazzo, B., Slowinski, R., Tsoukias, A.: "Exploitation of a rough approximation of the outranking relation in multicriteria choice and ranking". In: T. Stewart (ed): *Multiple Criteria Decision Making.* Proceedings of the Thirteenth International Conference, Cape Town (South Africa), January 1997; Springer-Verlag, Berlin, 1998, to appear

[36] Grzymala-Busse, J.W.: "LERS - a system for learning from examples based on rough sets". In: R. Slowinski, (ed) *Intelligent Decision Support. Handbook of Applications and Advances of the Rough Sets Theory.* Kluwer, Dordrecht, 1992, pp. 3-18

[37] Grzymala-Busse, J.W.: A new version of the rule induction system LERS. Fundamenta Informaticae. 31(1997)27-39

[38] Jacquet-Lagrèze, E.: *Systèmes de décision et acteurs multiples - Contribution à une théorie de l'action pour les sciences des organisations.* Thèse d'Etat, Université de Paris-Dauphine, Paris 1981

[39] Jacquet-Lagrèze, E., Siskos, J.: Assessing a set of additive utility functions for multicriteria decision-making, the UTA method. European Journal of Operational Research 10(1982)151-164.

[40] Keeney, R. L., Raiffa, H.: *Decision with Multiple Objectives - Preferences and value Tradeoffs.* Wiley. New York 1976

[41] Krantz, D.M., Luce, R.D., Suppes, P., Tversky, A.: *Foundations of Measurements I.* Academic Press. 1978

[42] Krawiec, K., Slowinski, R., Vanderpooten, D.: "Learning of decision rules from similarity based rough approximations". In: A. Skowron, L. Polkowski (eds): *Rough Sets in Knowledge Discovery.* Physica-Verlag, Heidelberg 1998, to appear

[43] Krusinska, E., Slowinski, R., Stefanowski, J.: Discriminant versus rough set approach to vague data analysis. Applied Stochastic Models and Data Analysis 8(1992)43-56

[44] Kryszkiewicz, M., Rybinski, H.: Computation of reducts of composed information systems. Fundamenta Informaticae 27(1996)183-195

[45] Langley, P., Simon, H. A.: "Fielded applications of machine learning". In: R. S. Michalski, I. Bratko, M. Kubat (eds): *Machine Learning and Data Mining*, Wiley, New York 1998, pp. 113-129

[46] Lin, T.: "Neighborhood systems and approximation in database and knowledge base systems", in Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems, 1989

[47] Luce, R.D.: Semi-orders and a theory of utility discrimination. Econometrica 24(1956)178-191

[48] March, J. G.: "Bounded rationality, ambiguity, and the engineering of choice". In: D. E. Bell, H. Raiffa, A. Tversky (eds): *Decision Making, Descriptive, Normative and Prescriptive Interactions*. Cambridge University Press, New York 1988, pp. 33-58

[49] Marcus, S.: Tolerance rough sets, Cech topologies, learning processes. Bull. of the Polish Academy of Sciences, Technical Sciences 42 (3)(1994)471-487

[50] Michalski, R.S., Bratko, I. Kubat, M. (eds): Machine Learning and Data Mining – Methods and Applications. Wiley, New York 1998

[51] Mienko, R., Slowinski, R., Stefanowski, J., Susmaga, R.: "Rough family – software implementation of rough set based data analysis and rule discovery techniques". In: Shusaku Tsumoto *et al.* (eds), *Proceedings of the Fourth International Workshop on Rough Sets, Fuzzy Sets and Machine Discovery*, Tokyo University Press, Tokyo, November 6-8 1996, pp. 437-440

[52] Mienko, R., Stefanowski, J., Toumi, K., Vanderpooten, D.: *Discovery-oriented induction of decision rules*. Cahier du LAMSADE no. 141, Université de Paris Dauphine, Paris 1996

[53] Mousseau, V., *Problèmes liés à l'évaluation de l'importance en aide multicritère à la décision: Réflexions théoriques et expérimentations*, Thèse, Université de Paris-Dauphine, Paris 1993

[54] Murofushi, T.: "A technique for reading fuzzy measures (i): the Shapley value with respect to a fuzzy measure". Proc. 2nd Fuzzy Workshop, Nagaoka, Japan, October 1992, pp. 39-48, in Japanese

[55] Murofushi, T., Soneda, S.: "Techniques for reading fuzzy measures (iii): interaction index". Proc. 9th Fuzzy Systems Symposium, Sapporo, Japan, May 1993, pp. 693-696, in Japanese

[56] Nguyen, S.H., Skowron, A.: "Quantization of real value attributes: rough set and Boolean reasoning approach". Proc. 2nd Joint Annual Conference on Information Sciences, Wrightsville Beach, NC, 1995, pp. 34-37

[57] Nieminen, J.: Rough tolerance equality. Fundamenta Informaticae 11(3)(1988)289-296

[58] Pawlak, Z.: Rough sets. International Journal of Information & Computer Sciences 11(1982)341-356

[59] Pawlak, Z.: Rough probability. Bull. Polish Acad.. Scis., Technical Sci. 33(1985) 9-10

[60] Pawlak, Z.: Rough sets and fuzzy sets. Fuzzy Sets and Systems 17(1985)99-102

[61] Pawlak, Z.: *Rough Sets. Theoretical Aspects of Reasoning about Data*. Kluwer, Dordrecht 1991

[62] Pawlak, Z: Rough set approach to knowledge-based decision support. European Journal of Operational Research 99(1997)48-57

[63] Pawlak, Z., Slowinski, R.: Rough set approach to multi-attribute decision analysis. European Journal of Operational Research 72(1994)443-459

[64] Polkowski, L., Skowron, A.: "Rough mereology". *Proc. Symp. on Methodologies for Intelligent Systems*, Lecture Notes in Artificial Intelligence, vol. 869, Springer-Verlag, Berlin, 1994, pp. 85-94

[65] Polkowski, L., Skowron, A., Zytkow, J.: "Rough foundations for rough sets". In: T.Y. Lin, A. Wildberger (eds): *Soft Computing: Rough Sets, Fuzzy Logic, Neural Networks, Uncertainty Management, Knowledge Discovery*. Simulation Councils, Inc., San Diego, CA, 1995, pp. 142-149

[66] Roubens, M.: *Interaction between criteria through the use of fuzzy measures*, Report 96.007, Institut de Mathématique Université de Liège, Liège 1996

[67] Roy, B.: *Méthodologie Multicritère d'Aide à la Décision*. Economica, Paris 1985

[68] Roy, B.: The outranking approach and the foundation of ELECTRE methods. Theory and Decision 31(1991) 49-73

[69] Roy, B.: Decision science or decision aid science? European Journal of Operational Research, Special Issue on Model Validation in Operations Research 66(1993)184-203

[70] Roy, B.: *L'aide à la décision ajourd'hui: Que devrait-on en attendre?* Document du LAMSADE no. 104, Université de Paris Dauphine, Paris 1997. English version: "Aiding the Decision Maker", chapter in the present volume

[71] Roy, B., Bouyssou, D.: *Aide Multicritère à la Décision: Méthodes et Cas*. Economica, Paris 1993

[72] Shafer, G.: *A Mathematical Theory of Evidence*. Princeton University Press, Princeton 1976

[73] Shapley, L. S.: "A value for n-person games". In: H. W. Kuhn, A. W. Tucker (eds): *Contributions to the Theory of Games II*. Princeton University Press, Princeton 1953, pp. 307-317

[74] Shoemaker, P.J.H.: The expected utility model: Its variants, purposes, evidence and limitations. Journal of Economic Literature, XX(1982)529-562

[75] Skowron, A.: "Boolean reasoning for decision rules generation". In: J. Komorowski, Z.W. Ras (eds): *Methodologies for Intelligent Systems*. Lecture Notes in Artificial Intelligence, Vol. 689, Springer-Verlag, Berlin 1993, pp. 295-305

[76] Skowron, A., Grzymala-Busse, J.W.: "From the rough set theory to the evidence theory". In: M. Fedrizzi, J. Kacprzyk, R.R. Yager (eds): *Advances in the Dempster-Shafer Theory of Evidence*. John Wiley and Sons, New York 1994, pp. 193-236

[77] Skowron, A., Polkowski, L.: Decision algorithms: a survey of rough set-theoretic methods. Fundamenta Informaticae 27(3/4)(1997)345-358

[78] Skowron, A., Rauszer C.: "The discernibility matrices and functions in information systems". In: R. Slowinski (ed): *Intelligent Decision Support, Handbook of Applications and Advances of the Rough Set Theory*, Kluwer Academic Publishers, Dordrecht 1992, pp. 331-362

[79] Skowron, A., Stepaniuk, J.: "Generalized approximation spaces". In: T.Y. Lin, A. Wildberger (eds): *Soft Computing: Rough Sets, Fuzzy Logic, Neural Networks, Uncertainty Management, Knowledge Discovery*. Simulation Councils, Inc., San Diego, CA, 1995, pp. 18-21

[80] Slovic, P.: Choice between equally-valued alternatives. Journal of Experimental Psychology: Human Perception Performance 1(1975)280-287

[81] Slowinski, K., Slowinski, R.: Sensitivity analysis of rough classification. Int. J. Man-Machine Studies 32(1990)693-705

[82] Slowinski, K., Slowinski, R., Stefanowski, J.: Rough set approach to analysis of data from peritoneal lavage in acute pancreatitis. Medical Informatics 13(1988)143-159

[83] Slowinski, K., Stefanowski, J.: "On limitations of using rough set approach to analyse non-trivial medical information systems". In: Shusaku Tsumoto *et al.* (eds): *Proceedings of the Fourth International Workshop on Rough Sets, Fuzzy Sets and Machine Discovery*, Tokyo University Press, Tokyo 1996, pp. 176-183

[84] Slowinski, R. (ed): *Intelligent Decision Support. Handbook of Applications and Advances of the Rough Sets Theory*. Kluwer Academic Publishers, Dordrecht 1992

[85] Slowinski, R.: A generalization of the indiscernibility relation for rough set analysis of quantitative information. Rivista di matematica per le scienze economiche e sociali 15(1993)65-78

[86] Slowinski, R.: "Rough set processing of fuzzy information", In: T.Y.Lin, A.Wildberger (eds): *Soft Computing: Rough Sets, Fuzzy Logic, Neural Networks, Uncertainty Management, Knowledge Discovery*. Simulation Councils, Inc., San Diego, CA, 1995, pp. 142-145

[87] Slowinski, R., Stefanowski, J.: Rough classification in incomplete information systems. Mathl. Comput. Modelling 12 (10/11) (1989)1347-1357

[88] Slowinski, R., Stefanowski, J.: "RoughDAS and RoughClass software implementations of the rough sets approach". In: R. Slowinski (ed): *Intelligent Decision Support. Handbook of Applications and Advances of the Rough Sets Theory*. Kluwer Academic Publishers, Dordrecht 1992, pp. 445-456

[89] Slowinski, R., Stefanowski, J.: "Handling various types of uncertainty in the rough set approach". In: W. P. Ziarko (ed): *Rough Sets, Fuzzy Sets and Knowledge Discovery*, Springer-Verlag, London 1994, pp. 366-376

[90] Slowinski, R., Stefanowski, J.: Rough set reasoning about uncertain data. Fundamenta Informaticae 27(1996)229-243

[91] Slowinski, R., Vanderpooten, D.: *Similarity relation as a basis for rough approximations*, ICS Research Report 53/95, Warsaw University of Technology, Warsaw 1995. Also in: P.P. Wang (ed): *Advances in Machine Intelligence & Soft-Computing*, vol.IV, Duke University Press, Durham, NC, 1997, pp. 17-33

[92] Slowinski, R., Vanderpooten, D.: A generalized definition of rough approximations based on similarity. IEEE Transactions on Data and Knowledge Engineering 1998, to appear

[93] Stefanowski, J.: Rough sets theory and discriminant methods as tools for analysis of information systems. A comparative study. Foundations of Computing and Decision Sciences 17(1992)81-98

[94] Stefanowski, J.: "On rough set based approaches to induction of decision rules". In: A. Skowron, L. Polkowski (eds): *Rough Sets in Data Mining and Knowledge Discovery*, Physica-Verlag, Heidelberg 1998, to appear

[95] Susmaga, R.: Analysing discretizations of continuous attributes given a monotonic discrimination function. Intelligent Data Analysis 1 (3)(1997) on-line journal, http://www-east.elsevier.com

[96] Susmaga, R.: "Experiments in incremental computation of reducts". In: A. Skowron, L. Polkowski (eds): *Rough Sets in Data Mining and Knowledge Discovery*. Physica-Verlag, Heidelberg 1998, to appear

[97] Tsoukias, A., Vincke, Ph.: A new axiomatic foundation of the partial comparability theory. Theory and Decision 39(1995)79-114

[98] Tsoukias, A., Vincke, Ph.: "Extended preference structures in MCDA". In: J. Climaco (ed): *Multicriteria Analysis*, Springer-Verlag, Berlin 1997, pp. 37-50

[99] Tversky, A.: Intransitivity of preferences. Psychological Review 76(1969)31-48

[100] Tversky, A.: Features of similarity. Psychological Review 84 (4)(1977)327-352

[101] Vincke, Ph.: *Multicriteria Decision-Aid*. John Wiley and Sons, New York 1992

[102] Wakker, P.P.: *Additive representions of preferences. A new foundation of decision analysis*. Kluwer Academic Publishers, Dordrecht 1989

[103] Yao, Y.: "Combination of rough sets and fuzzy sets based on α-level sets". In: T. Y. Lin, N. Cercone (eds): *Rough Sets and Data Mining*, Kluwer, Boston 1996, pp. 301-321

[104] Yao, Y., Wong, S.: "Generalization of rough sets using relationships between attribute values". *Proceedings of the 2nd Annual Joint Conference on Information Sciences*, Wrightsville Beach, NC, 1995, pp. 30-33

[105] Ziarko, W., Shan, N.: "An incremental learning algorithm for constructing decision rules". In: W.P. Ziarko (ed): *Rough Sets, Fuzzy Sets and Knowledge Discovery*. Springer-Verlag, London 1994, pp. 326-334

15 USE OF ARTIFICIAL INTELLIGENCE IN MCDM

Patrice Perny
Jean-Charles Pomerol

CONTENTS

Abstract. We survey the different representations, including non-classical logic, issued from artificial intelligence and used in MCDM. On the one hand, we mainly focus on rule-based and object representations, and applications of non-classical logic. On the other hand, we examine the contribution of heuristic search ideas to interactive MCDM.

T. Gal et al. (eds.), *Multicriteria Decision Making*

FOREWORD

Even if multicriteria analysis can be proud of going back to the end of the 18th century (Borda, Condorcet), far before the emergence of AI, the two domains eventually begun to be really recognized during the sixties. Moreover, it is only during the eighties that authors explicitly referred to AI for addressing typical multicriteria issues. We can divide the relationships between AI and MCDM in three large sub-chapters, in the first ones we will focus on the new ideas and representations, including non-classical logic, coming from AI that have been introduced in MCDM and, in the third part, we will examine how AI and MCDM share the same concerns as regards interactivity, friendliness and user involvement.

Whereas we try to focus, in this chapter, on the use of AI in MCDM we must mention that many methods and tools, which were popularized by AI people, are now largely spread in multicriteria analysis as well as in other fields. Some of these tools are so widely used in multicriteria analysis that they deserve special chapters, among them genetic algorithms, heuristic and meta-heuristic methods (Chapters 6 and 16), interactivity and user-friendliness (Chapters 7, 9 and 10).

15.1 USING TRADITIONAL AI REPRESENTATIONS

15.1.1 Introduction

At the beginning of the AI venture were two main ideas attested in the seminal book of Newell and Simon [91], the first one is that computers are able to handle symbolic as well as numerical computing, the second is that heuristic reasoning is the general framework for human reasoning. We will return to the second idea in the third section of this chapter. Let us start with symbolic representations.

As a matter of fact, multicriteria decision analysis mainly nurtured in an OR environment more prone to numerical representations than to symbolic one. This is probably one of the reasons why, whereas there exist many works on aggregation starting from valued alternatives (we mean alternatives represented as numerical vectors), there are very few papers concerned with more qualitative forms of reasoning such that the design of the alternatives and/or of the criteria.

Among the representations issued from AI we will distinguish four main families: rule-based and object-based representations, neural networks, and finally, logic-based representations. This chapter is organized as follows: in the first subsection we will examine the MCDM use of AI representations. Then, because of the large number of applications of logic we will devote the whole subsection 15.2 to this topic and finally, we will show the contribution of AI to heuristic search and interactivity.

15.1.2 Rule-based representations

In multicriteria decision making there are several important steps to consider: defining the alternatives or, using Simon's vocabulary [133] designing the alternatives, defining the criteria (see chapter 1), evaluating the alternatives according to each criterion, then eventually the aggregation (the choice according to Simon's phases of decision process), not to mention the review phase. Among all these possibilities the choice phase generated most of the papers and few attempts were made to address the other issues. Nevertheless, we will try to survey all the proposals we are aware of. Let us begin by the most important, at least regarding the number of papers.

Aggregating or supporting the decision maker's choice by rules. Let us think of the multicriteria decision making process as a reasoning process. The decision maker choice relies on trade-offs between more or less contradictory criteria, *e.g.* earnings against risk. We can directly express these trade-offs by rules. Let us give an example, drawn from [15], concerning a car:

> If *SECURITY* is *high* and *COMFORT* is *very good*
> then *TECHNICAL-CHARACTERISTICS* are *very good.*

It is clear that the preceding rule aggregates the two criteria "*SECURITY*" and "*COMFORT*" under the name "*TECHNICAL-CHARACTERISTICS*". *Fundamentally a rule is an aggregating process.* When talking of a rule as an aggregating process we can, at least, imagine two situations. Either the rules are interpreted as a unstructured aggregation process depending on the alternatives and on the satisfaction levels. For example:

If *PRICE* (A) = *low* and *COLOR* (A) = *black* then *DESIRABILITY* (A) = *very high*

> If *MAX_SPEED*(B) $\leq$ 130 km/h and *PRICE*(B) = *medium*
> and *COMFORT*(B) = *medium* then *DESIRABILITY*(B) = *weak*

In this case the final evaluation of the cars depends on a variable unstructured collection of facts. We are in a classical expert system process in which each rule is a "local" mental aggregating process. By local, we mean, referring to the above example, that when the color is black (local value) then the price suffices for inferring the desirability, whereas for car B the color does not intervene but the speed and the comfort.

Different is the case where each criterion is seen as an aggregation of hierarchically lower criteria. For example the programming quality of a microcomputer may be divided into language and quality of user environment, and the user environment is further divided into hardware and operating system utilities [40]. Then, each rule gives the values of the "functional schema" [80], *e.g.* for the functional schema:

(LANGUAGE, USER ENVIRONMENT) $\rightarrow$ *QUALITY OF THE MICRO*

If *LANGUAGE* is *acceptable* and *USER ENVIRONMENT* is *good*
then the *QUALITY OF THE MICRO* is *good.*

If *LANGUAGE* is *poor* and *USER ENVIRONMENT* is *good*
then the *QUALITY OF THE MICRO* is *bad.*

This suggests that in many cases (diagnosis, evaluation), a knowledge base is nothing else than an aggregating machine of concepts.

Reasoning with schemata as suggested by Lévine-Pomerol [80] provides a very strong tool for knowledge structuring. As an example, let us give (figure 15.1 hereafter) the whole aggregating tree of the relevant concepts for defining the financial security in a reasoning process about the evaluation of a firm [80]. (In this example, the financial security is just one of the knowledge bases which cooperate in a multi-expert system in order to give a complete risk diagnosis).

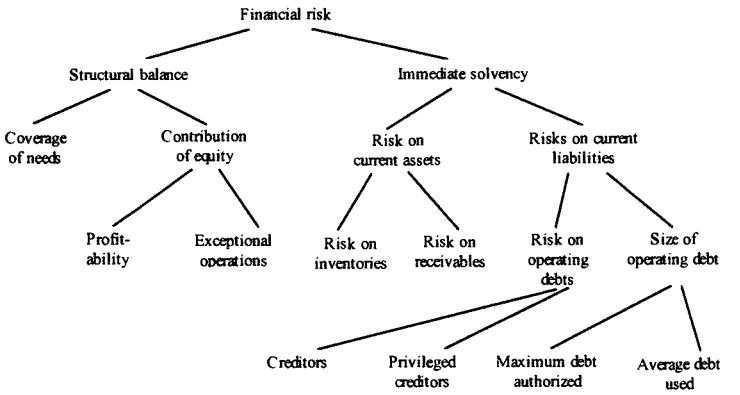


Figure 15.1: An aggregating tree (see [80])

As already noted in [80] a schema expresses a discrete functional relation which may be restated as a decision table. For example, the first above rule, about microcomputers, can be written as the following table (Tab. 15.1).

	Language	poor	poor	acceptable	acceptable
Premises	Environment	poor	acceptable	good	acceptable
Action	Qual. of micro	poor	acceptable	good	acceptable

Table 15.1: Decision table

This table defines a function from {poor, acceptable, good}×{poor, acceptable, good} into {poor, acceptable, good}. The problem is that being discrete, this function presents jumps which are regarded as drawbacks in many applications. By giving values in [0,1] and replacing the table by a continuous function (*e.g. QUALITY OF THE MICRO* = 0.4 *LANGUAGE* +0.6 *ENVIRONMENT*), Saaty's method [129, 130] avoids the gap effect inherent to the discrete world. Fuzzy rules

were introduced for the same reason (see Section 15.2). Saaty's method AHP (for Analytical Hierarchical Process) is based on hierarchical decomposition of the criteria followed by an additive aggregation. Roughly speaking aggregation by schemata of rules is a generalization of AHP to symbolic values.

Rather than replacing the rules by continuous functions, an alternative consists of smoothing the discrete function defined by the table 15.1, see [16, 17] and also [118] where various interpolation functions are used to pass from a discrete to a continuous function.

It is also noteworthy that a decision table, either complete or incomplete may be proceeded via an algorithm such that ID3 [117] in order to produce rules. It paves the way for building an aggregating tree from examples (see [118]). Similars ideas were developed in rough set theory (see Chapter 14).

Criticisms of the aggregation by rules. The aggregation process based on rules suffers from some drawbacks. Firstly, the construction of the rules is generally considered as being too empirical to be reliable and dependent on the knowledge engineer [157]. At least, one advantage is that the rules under their symbolic form are readable for the decision maker. However, it is true that if the rules are written in a messy file of multiple "local ideas", the readability disappears and the result of the aggregation depends on the alternatives. When the aggregation is organized along a semantic tree as explained above, the process is fully readable. As compared to numerical aggregation methods, aggregating by rules has the advantage that the appreciation of the alternatives depends on the value, because a table like table 15.1 is more informative than the function:

$$QUALITY\ OF\ THE\ MICRO = 0.4\ LANGUAGE + 0.6\ ENVIRONMENT$$

in which the parameters 0.4 and 0.6 do not depend on the value of *LANGUAGE* or *ENVIRONMENT*.

Another criticism pointed out in [66], [157] and emphasized in [67] is that the decision maker's preferences are not explicitly considered because the rules are generally not centered on the decision maker's objectives. This criticism is certainly true and important in the general setting of expert systems. As pointed out in [66], the explicit distinction between choice (or user preferences) and descriptive facts is fundamental both from decision theory and practice points of view. We will return on this question in the next section.

In the MCDM context, the situation is somewhat different, because the limit between facts and objectives heavily depends on the decision maker's interpretation. More precisely, it is the decision maker who draws the frontier between objectives and "factual" alternatives that cannot be changed. Nevertheless, we agree that it is very important to distinguish, as far as possible, the two concepts (see hereafter (Evaluating the alternatives) and [67]).

Supporting the decision maker's choice. Rules can also be used either to direct the exploration process in the alternative set. One can find such an example in [119]. In the same vein, rules are used in [56] to choose between basic multiattribute solutions

obtained by the simplex algorithm. Explaining the result of an aggregation procedure can be regarded as an undirect aid to decision makers. In [69] a rule-based system is designed, which explains the choice resulting from a additive aggregation procedure along a hierarchical tree of criteria similar to AHP [129].

Again with the purpose of helping the decision maker, we can mention the "recursive idea" of supporting him with a multicriteria method during the choice among multicriteria methods. This idea was considered as a desirable issue in [61, 62] and [3] in the context of multicriteria DSSs. The idea was implemented in [139]. Their method is a symbolic one relying on a criterion tree as described in the preceding section. [57] and [100] also address the same question by using rules. The same idea supporting the choice of the aggregation method is again taken up in [109] by using case-based reasoning (see Section 15.1.3). An expert system can also be used either to determine the weights according to the context [75] or modify them according to the "satisfaction levels" [89].

Evaluating the alternatives. It has been frequently reported by practitioners that decision makers are often upset by the evaluation of the alternatives according to their various criteria. The problem is equivalent to that of building an utility function (numerical or symbolic) starting from a given set of alternatives. Numerous papers have been devoted to the construction of numerical utility functions (see [43] and [113] for surveys).

To build a symbolic utility function (*i.e.* whose values belong to an ordered set of words) we can imagine to use an expert system. This idea appears in [94]. The simplest structure consists of an expert system processing various data to assess the value of the alternatives according to each criterion. The same idea was also implemented by [154] for customer evaluation. The main criticism against this technique is, as above, that it merges choice and descriptive items [66, 157].

To overcome this difficulty and to separate the choice (preference) from the descriptive (factual) part of the alternatives, it is proposed in [84] to define a structure in which each alternative is characterized by a fact base and each criterion by a rule base. The rule base is organized as a tree (see Figure 15.2). In this tree the concepts involved in one criterion are displayed according to their importance. The values are propagated along the tree (see Figure 15.2 for a part of the tree associated with the criterion "quality of the service" in the choice of a postal sorting machine).

According to these ideas, the architecture of the system is as follows (see figure 15.3). Applying the rules R_j attached to criterion j to fact base B_i attached to alternative i produces the value a_{ij} of alternative i according to criterion j.

This architecture has the advantage of separating the reasoning about the choice which is contained in the rule base associated to a criterion, from the evaluation which is made by the application of this rule base to the facts characterizing each alternative. The result of the system is a decision matrix which afterwards can be processed by any multicriteria method.

Figure 15.2: Definition of the criterion "quality of the service"

Rule 1
IF RISK OF UNINTENTIONAL DROPS=NO
IF NUMBER OF MANUAL TRANSFERS=LOW
THEN RISK OF DAMAGE TO PACKAGES =LOW

Rule 2
IF RISK OF UNINTENTIONAL DROPS=YES
IF NUMBER OF MANUAL TRANSFERS=HIGH
THEN RISK OF DAMAGE TO PACKAGES
=IMPORTANT

Rule 12
IF EMERGENCY STOPPING TIME=LONG
IF MISDIRECTED PACKAGES=YES
IF SAFETY FEATURE=NO
THEN RISK OF MISDIRECTING=HIGH

Rule 5
IF EMERGENCY STOPPING TIME=SHORT
IF MISDIRECTED PACKAGES=NO
IF SAFETY FEATURE=YES
THEN RISK OF MISDIRECTING=LOW

Table 15. 2 : Some rules associated with the tree of the figure 15 .2.

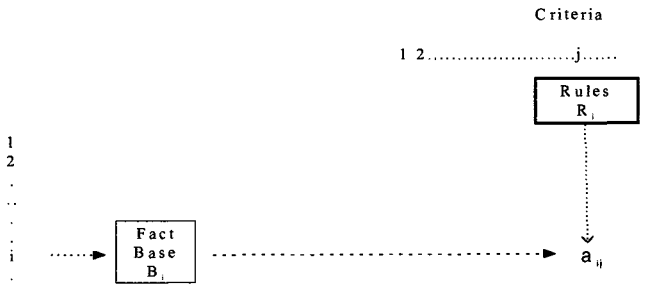


Figure 15. 3: Architecture of a multiattribute decision system [84]

One must also notice that, on the one hand, the difference between concepts used in a criterion and those considered as facts describing some characteristics of the alternatives is purely interpretative. For example, the "emergency stopping time" can be interpreted as a part of the criterion "quality of the service" (Figure 15.2) or it could be seen as a characteristic of the sorting machine. On the other hand, one has

also to decide among the concepts used to define a criterion. Which one deserves to be the root of the tree and consequently becomes one of the criterion used to control the choice process. For instance, is the "*risk of misdirecting*" a criterion for the choice of the machine or simply a given element of the "*quality of the service*"? There is no immediate answer, it has to be discussed with the decision maker.

It is often difficult to distinguish between the design of the alternatives and their evaluation. This is particularly the case when rule-based methods are used. A typical example of a rule-based system used to evaluate the alternatives according to several "high level attributes" is provided by [50]. Then, the choice is made by ELECTRE I method. The system also displays explanations about the evaluation.

Let us also mention the using of cognitive maps and rules to develop a multicriteria knowledge-based system [147]. It is noticeable that, in this system, the cognitive map appears as a generalization of a AHP-like hierarchical criterion tree.

Designing the alternatives. In many real situations, the first step is not even to assess the alternatives but to understand what the alternatives are. This is generally the case for economical decisions where the alternatives actually are scenarios. For example, you do not decide between increase or decrease the prices of the product but between two or generally much more scenarios, increase the price and begin an advertising campaign and differentiate your product (alternative A) vs. decrease the price and increase production capacity (alternative B) and so on. The alternatives contain different temporal actions and/or consequences. The first difficulty is thus to have a clear idea of what an alternative is and how to characterize it before assessing it! In our example, assume that the criteria are the net profit, the gross sales and the market share. It is not obvious to evaluate alternative A vs. B, it needs very serious studies to characterize the results of A and B according to the chosen criteria.

Let us give another very specific example. The problem is to assess the robustness of railway timetables [115]. Given a railway network (RN) and the machines (M) we can define theoretical timetables (TT). Now, when an incident occurs, the train schedules are perturbed and it is evident that the importance of the perturbation depends on RN, M, and TT. With an unlimited number of railroads, an incident on a train would have no consequence on the other trains, but it is not the case. An incident generally delays many trains, but the importance of the total delay depends on (RN, M, TT). Assume that you want to assess an RN investment devoted to improve the situation in case of incident. It means that we have to assess alternatives of the form (RN, M, TT) regarding the robustness in case of incident. This is a typical case where human reasoning is unable to construct, without help, what we have called a *fully expanded alternative* (FEA). The fully expanded alternative associated to an alternative (RN, M, TT) is the real timetable incorporating all the delays caused by an incident. To "propagate" the incidents the decision maker needs support. We have designed such a multicriteria DSS [115]. In this system, an expert system plays the role of the dispatchers, *i.e.* it makes the decisions, resulting from an incident and its consequences. This decision results into actions on the trains perturbed by the incident (typical actions are delay, cancel or deviate trains). When the perturbed

timetable resulting from the preceding decisions is known, and only when this is done, we can assess the importance of the delays, the number of concerned commuters and three other criteria. In this framework, the expert system or more generally the DSS is "simply" used to bridge the gap between the alternatives and the "fully expanded alternatives". We believe that this type of problem will become more and more frequent as people addresses more complex problems. In fact, a scenario is a sequence of intertwined events and actions but, by using robust actions designed to be good enough whatever the realization of the event is, decision makers tend to reason on a sequel of actions. This is the reason why we have coined the term FEA for a sequence of robust (sub)actions [112]. The point is that computer support is necessary to deal with scenarios and FEA complexity.

Another example of the use of an expert system in order to build an alternative is presented in [33] where a frame-based system determines the possible pathology of a patient. The plausibility of these pathologies are then evaluated by a multicriteria method (PROMETHEE). This example can be interpreted as a first step toward the introduction of case-based representation in MCDM.

15.1.3 Case-based and object representations

Up to now, as far as we know, very few attempts exist to merge MCDM and *case-based reasoning* (CBR). However, it seems that analogical reasoning is one of the usual mode of decision making. One can suppose that the same is true for multicriteria decision making. The principle of case-based reasoning is known: an action is undertaken when the current state is more or less similar to one of the recorded cases. The similarity is measured by a similarity function defined on any couple of cases. The difficulties are to define a good representation of the cases encompassing all the requisite context and to define an adequate similarity function. These two difficulties seem to have prevented MCDM designers to put forward many case-based systems. One of the first attempt goes back to 1992 with a system designed by Angehrn and Dutta [2]. In this system the CBR module is intended to act as an adviser to "recognize whether a multicriteria approach was used successfully in other similar situations before, and provide information about the types of alternatives, criteria and preference structures used" [2]. The second role of the system is as a "story teller". In this case the system allows the user to replay a previous case. The main component of the system is a case library which contains structured multicriteria problems with their solutions. This system remains one of the most completed examples of CBR structuration for multicriteria problems.

Another CBR system was developed in [109] to solve the classical problem of the choice of the best aggregation procedure. In this system a case is defined by the type of the problem (choice, ranking, sorting,...), the number of alternatives, the number of criteria, the weights, etc... Then, when a new case is introduced, according to a kind of hierarchical similarity function, the most relevant multicriteria method is chosen.

15.1.4 Neural networks

It is argued in [148] that "*Decision making is a natural candidate for connectionist modeling because it is a complex activity that is generally performed intuitively and that can benefit from the computational advantage of the neural parallel processing*". Several recent works focused on the use of neural networks to represent and simulate individual choice behavior. Among the studies related to multicriteria decision making and neural networks, one can roughly distinguish two main activities concerning two different decision tasks: multiattribute categorisation tasks and multiattribute choice tasks (see *e.g.* [148, 63, 77, 64]). In this subsection, we propose a brief survey of some representative studies in this domain.

Neural networks for decisions associated with multiattribute categorization tasks. Assigning multiattribute objects to categories is a standard problem in Artificial Intelligence and many studies involving neural networks for learning and simulating categorization rules have been carried out independently of any perspective related to decision making. However, in many MCDM problems, decisions concerning the alternatives depend on their membership in one or several categories. In such problems, the assignment of an alternative to a category is based on its multiattribute profile and assignment rules can be automatically extracted by a neural network. The basic idea consists of training a neural network on a reference set of alternatives classified by the decision maker. Then new alternatives are introduced in the system and classified (see [88] and [19]).

Several dynamic networks methodologies have also been proposed for multiattribute categorization tasks (see [63, 77, 64]). They are based on Adaptive Resonance Theory (ART) networks proposed in [87]. This kind of networks consists of various interconnected modules encoding attributes of alternatives, categories and goals of the decision makers (prototypes of categories). The nodes associated to categories and attributes are connected to each other with connections modifiable by associative learning, leading to interlevel resonant feedback (for more details see [77]). The assignment of an alternative to a category is possible only if the alternative matches the prototype of the category to a certain threshold. When an alternative does not match with the prototypes, a new category is created and the unassigned alternative becomes its prototype. An interesting variation of the ART network called CATEG_ART has been proposed and tested [63, 64]. This network is designed to implement a cognitive model based on the Moving Basis Heuristics (MBH, see Section 15.3.3) and to automatically produce categorization rules synthesized by polynomials. For example, the polynomial:

$$P(C) = X_1^2 X_4^3 + X_2^1 X_5^4 + X_3^5$$

is composed of a disjunction of monoms meaning that, if an alternative has been placed by the decision maker into category C, it is because it verifies ($X_1 = 2$ and $X_4 = 3$) or ($X_2 = 1$ and $X_5 = 4$) or $X_3 = 5$. Once the neural network has been trained on the reference set of alternatives, the assignment rules created allow the classification of

new alternatives whom categories are unknown. The category of a new alternative will be the one of the monomial selected by resonant competition.

Neural networks for decision making in multiattribute choice tasks. We briefly present here a dynamic neural network methodology proposed in [148] for multiattribute decision making. Unlike some standard networks that fit reference data by choosing optimal weights or coefficients within a predetermined mathematical model, the network is regarded here as a descriptive model aiming at approximating the decision maker behavior when facing a multicriteria problem. Consider a set of alternatives $A = \{a_1, ..., a_m\}$ where each alternative is described by n attributes $X_1, ..., X_n$. Suppose that the attractiveness of each alternative a_i with respect to each attribute X_j has been evaluated by the decision maker and quantified by positive coefficients w_{ij} for $i = 1, ..., m, j = 1, ..., n$. Then, as proposed in [148], we can consider the following decision network's architecture constructed from the prior knowledge of the decision situation (see figure 15.4, adapted from [148]).

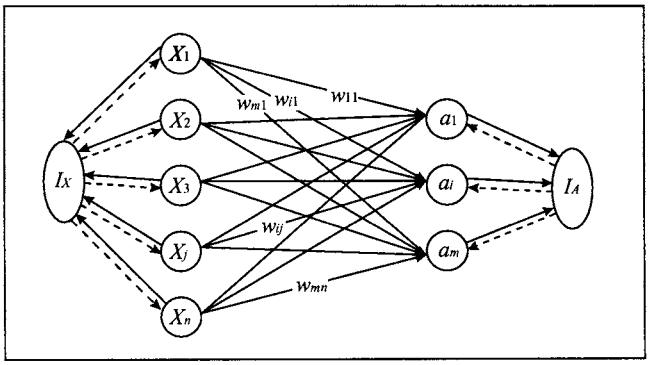


Figure 15.4: A network's architecture for multiattribute decision making

This dynamic neural network consists of two connected subnetworks; the left one is composed of competing attributes assemblies representing "the current state of decision maker's mind" (what are the current focused attributes); the right one corresponds to competing alternatives assemblies representing the current potential decision. Full lines represent excitatory and dashed lines represent inhibitory connections. I_X and I_A are inhibitory assemblies that mediate competition in left and right subnetworks respectively. For example, when inhibitory neurons of I_X get exitation from all the assemblies in the subnetwork, they return inhibition provoking the activity decay of some attributes. Arrows of type (X_j, a_i) represent excitatory connections weighted by the coefficient w_{ij} (if w_{ij} exceeds a certain threshold, the a_i assemblies is activated each time attribute X_j is scanned).

After an initial activation, the network is autonomous and evolves dynamically. Most of the time, it converges towards a stable state. However, when dynamic variations of thresholds are allowed, the network may also adopt a periodic or chaotic behavior with some stable periods. When stability is reached, the activated alternatives make up the selection proposed by the network.

Notice that an inhibition parameter controls the average number of simultaneously activated attributes in the left networks. It may be used for

representing more or less requiring dominance rules to compare the alternatives (when only one attribute is activated at each time, the decision process can be compared to the "elimination by aspect" process, see [145, 146]). Another inhibition parameter attached to the I_A assemblies controls the level of competition among the alternatives and therefore the average number of alternatives that can be activated at each time. Other parameters (noise factors) control the competition and inertia among the alternative. Depending on the values attached to these parameters, the network represented on figure 15.4 can simulate various decision behaviors. Among them let us mention the elimination by aspects, the dominance rule, conjunctive and disjunctive models and lexicographic decision rules (for more details see [148]).

15.1.5 Using classical logic

Besides traditional representations by rules, frames and objects, many AI researchers are fond of logic-based languages, such as PROLOG. Some of these researches received attention in MCDM. First of all, an application of PROLOG to multicriteria decision is proposed in [120]; it is observed that using a PROLOG rule such that:

$$COMFORT(X) \wedge TECH.\ CHARACT.\ (X) \rightarrow QUALITY\ (X)$$

amounts to a lexicographic search on criteria COMFORT and TECH. CHARACT in this order of importance. A second application concerns preferences. The idea consists of translating the preferences of the decision maker about the criteria into a preorder on the criteria endowed with desirable properties. Such a system has been developed in [101]. Namely, let J be the set of criteria and 2^J the set of parts of J. We can define on 2^J a relation $>_J$ reflecting the importance of criteria coalitions using the following rules:

$\forall$ I, K, $L \in 2^J$, $I >_J K$ and $K >_J L \Rightarrow I >_J L$ (transitivity)
$\forall$ I, K, $L \in 2^J$, $I >_J K \Rightarrow I \cup L >_J K$ (amplification)
$\forall$ I, K $\in 2^J$, $I >_J K \Rightarrow$ not$(K >_J I)$ (asymmetry)
$\forall I, K, L \in 2^J$, $I >_J K$ and $I \cap L = \varnothing \Rightarrow I \cup L >_J K \cup L$ (additivity)

PROLOG programming is particularly well suited to make the heavy calculations introduced by the above rules which complete $>_J$ from the partial preorder given by the decision maker. When the relation is completed, it can be used to replace the weights in outranking methods such that ELECTRE III for example [134]. It suffices to state: a is strictly preferred to b *iff* $\{j / a P_j b\} >_J \{j / b P_j a\}$

The EXTRA system proposed in [101] is entirely developed from these PROLOG-based ideas. In an additive framework, this method about comparison of sets of criteria amounts to compare different sums of weights. It is reminiscent of Churchman and Ackoff revisited [70].

Many other examples showing the interest and relevance of logic for preference modeling should be considered. For example, multivalued logic, and most of non-standard logics may be used advantageously to construct sophisticated models of preferences. Since this domain has steadily developed in the 10 past years the next section is devoted to this topic.

15.2 MODELS BASED ON NON-CLASSICAL LOGIC

Most of decision support activities require the use of mathematical models representing the available knowledge about the practical decision problem to deal with. In real decision situations, knowledge representation is complicated by the presence of conflicting opinions, and the uncertainty or imprecision about the consequences of potential acts or decisions. In the face of such complexity, decision support systems involve specific tools for preference modeling and reasoning with imperfect information. In this respect, logic languages are useful because they allow the compact representation of the various objectives of the agents and their preferences. This section illustrates the interest of non-classical logics for preference modeling and preference aggregation in the context of imperfect information.

15.2.1 The use of [0, 1]-valued logic

Preference modeling with valued relations. Preference modeling in a decision problem with a finite number of alternatives aims at providing, for all possible pairs (a, b) of alternatives, answers to questions of type: "is a comparable to b", "is a indifferent to b", "is a preferable to b"? However, in most applied problems, experts are reluctant to make categorical assertions about alternatives and the available preference information is rarely sufficient to provide yes/no type answers to these questions. If nevertheless, the expert or the decision maker is constrained to produce crisp judgments when comparing alternatives, he/she will have to sacrifice a large part of his expertise and the resultant mathematical model may lose much of its adequacy to the real situation. Thus, modeling preferences by valued (or fuzzy) relations provides a more flexible language, allowing the expression of intermediate degrees of preferences, between the categorical "yes" or "no". It does not prevent the analyst or the decision maker to draw clear-cut conclusions in a later stage of the decision process.

The use of fuzzy sets for preference modeling has been deeply investigated since the seminal works of Belman and Zadeh [8] and Orlovski [96] and the language of fuzzy relations has shown to be useful for preference modeling and MCDM (see *e.g.* [14, 18, 20, 27, 47, 48, 49, 73, 74, 97, 98, 99, 103, 104, 105, 122, 152, 153]).

Basically, valued relations are used in multicriteria methods to represent either ill-defined or ill-known preferences. The second case mainly occurs when a preference relation must be constructed from uncertain criterion values. This point will be illustrated in Section 15.2.4. In the first case, valued relations are mainly used to represent a continuum of situations from total indifference to strict preference, even if criterion values are precisely known. Consider for instance a multicriteria problem with a finite discrete set of alternatives A and a set of criterion function g_j, $j = 1, ..., n$, to be maximized. Following [125, 22, 104, 47, 105], a valued strict preference relation P_j can be constructed on each dimension j from criterion values by setting:

$$P_j(a, b) = f_j(g_j(a), g_j(b)) \tag{15.1}$$

where f_j is a real-valued function, non-decreasing of the first argument, non-increasing of the second argument and such that $f(x, x) = 0$ for all possible x. In a

similar way, a valued indifference relation I_j can be constructed on each dimension j from criterion values, using the following definition:

(15.2) $$I_j(a, b) = 1 - \max\{f_j(g_j(a), g_j(b)), f_j(g_j(b), g_j(a))\}$$

Examples of such relations are depicted on figure 15.5

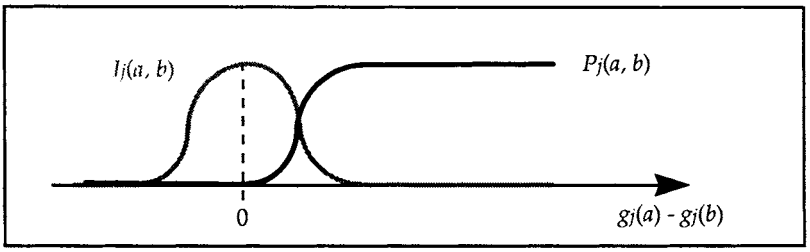


Figure 15.5: Valued preference relations defined from criterion values

In this section, we will only consider models in which preferences between alternatives are described by fuzzy relations whose valuation lies between 0 and 1. From the mathematical viewpoint, a fuzzy relation R defined on a set A is a fuzzy subset of the Cartesian product $A \times A$. It is characterized, by a membership function v defining, for each pair (a, b), the value $v(R(a, b)) \in [0, 1]$ representing the confidence we have in the statement "the relation R holds for the pair (a, b)". Such relations both extend additive preference models used in *difference measurement* [72] and in *threshold measurement* [137]. From a logical viewpoint, any binary relation R can be seen as a binary predicate $R(x, y)$ interpreted in a multivalued logic whose semantic domain is [0, 1] (any truth value within the unit interval is admissible; 0 denotes the "false" and 1 denotes the "true"). Hence, $v(R(a, b))$ represents the truth value of the closed formula $R(a, b)$. In order to represent more complex propositions about preferences, this language must be extended to all well-formed formulae of the standard predicate calculus. Denoting *WFF* the set of well-formed formulae, the interpretation of any complex formula F in *WFF* is defined from the interpretation of its atomic components using the following rules:

(15.3) $$v(\neg A) = N(v(A))$$
(15.4) $$v(A \wedge B) = T(v(A), v(B))$$
(15.5) $$v(A \vee B) = V(v(A), v(B))$$
(15.6) $$v(A \supset B) = I_T(v(A), v(B))$$
(15.7) $$v(A \equiv B) = E_T(v(A), v(B))$$
(15.8) $$v(\exists x\, A(x)) = sup_x(v(A(x)))$$
(15.9) $$v(\forall x\, A(x)) = inf_x(v(A(x)))$$

where A and B are well-formed formulae, $\neg$, $\wedge$, $\vee$, $\supset$, and $\equiv$ are usual logical connectives representing negation, conjunction, disjunction, implication and equivalence respectively, N is a function defined from [0, 1] onto itself, T, V, I_T, E_T are functions defined from $[0, 1]^2$ into [0, 1] as follows:

- $N(x) = \phi^{-1}(1 - \phi(x))$ for some automorphism ϕ on [0, 1],

- $T(x, y)$ is a symmetric, associative and monotonic function with 1 as neutral element (t-norm),
- $V(x, y)$ is a symmetric, associative and monotonic function with 0 as neutral element (t-conorm). Usually, $V(x, y) = N(T(N(x), N(y)))$ so as to preserve the De Morgan law,
- I_T is the quasi-inverse of T defined by: $I_T(x, y) = \sup\{\varepsilon \in [0, 1] : T(x, \varepsilon) \leq y\}$
- E_T is defined as a double implication: $E_T(x, y) = T(I_T(x, y), I_T(y, x))$

In this context, a formula F in *WFF* is said to be a *tautology* when $v(F) = 1$ for all interpretations v and a *contradiction* when $v(F) = 0$ for all interpretations v.

Using this language, any logical property or concept expressed using a classical preference relation can be extended in the context of fuzzy preferences [102]. As a first example, we show how the semi-transitivity property attached to semi-orders is extended when preferences are fuzzy. For crisp binary relations, semi-transitivity is defined as follows [27, 108]:

$$\forall x\ \forall y\ \forall z\ \forall w\ [(R(x, y) \wedge R(y, z)) \supset (R(x, w) \vee R(w, z))]$$

In the multivalued logic considered above, this logical formula is true if and only if the following condition holds for any t-uple (x, y, z, w):

$$v\ [(R(x, y) \wedge R(y, z)) \supset (R(x, w) \vee R(w, z))] = 1$$

Using equations (15.3)-(15.9), the above equation can be rewritten as follows:

$$I_T(v(R(x, y) \wedge R(y, z)), v(R(x, w) \vee R(w, z))) = 1$$
$$v(R(x, y) \wedge R(y, z)) \leq v(R(x, w) \vee R(w, z))$$
$$T(R(x, y), R(y, z)) \leq V(R(x, w), R(w, z))$$

Hence, choosing $T(x, y) = \min(x, y)$ leads to a familiar condition considered by many authors as the natural extension of semi-transitivity in the context of fuzzy preferences. It is often used in the definition of fuzzy semi-orders (see [47, 108]):

$$\forall x\ \forall y\ \forall z\ \forall w,\quad \min(R(x, y), R(y, z)) \leq \max(R(x, w), R(w, z))$$

It has been shown in [104] that this property hold for any relation P_j defined from criterion values as presented before (see equation (15.1)).

As a second example, let us show how the concept of non-dominated alternatives is translated when preferences are valued. Consider a (crisp) weak preference relation defined on a finite set X of alternatives. Within X, an element x is said to be non-dominated if and only if no other element in X is strictly preferred to x. This property can be formalized as follows :

$$ND(x) \equiv (\forall\ y\ (R(y, x) \supset R(x, y))) \tag{15.10}$$

with the following predicates:

Symbol	*Arity*	*Domain*	*Interpretation*
ND	1	X	$ND(x)$: "x is non-dominated"
R	2	$X \times X$	$R(x, y)$: "x is weakly preferred to y"

Assuming equation (15.10) is a tautology, we get:

$$v[ND(x) \equiv (\forall y\,(R(y, x) \supset R(x, y)))] = 1$$

and therefore: $E_T(v[ND(x)], v[\forall y\,(R(y, x) \supset R(x, y))]) = 1.$

Hence
$$v(ND(x)) = v[\forall y\,(R(y, x) \supset R(x, y))]$$
$$= \inf_{y \in X} v[R(y, x) \supset R(x, y)] = \inf_{y \in X} I_T[R(y, x), R(x, y)]$$

Moreover, whatever the t-norm T we consider, we get:

$$v(ND(x)) = 1 \Leftrightarrow \inf_{y \in X} I_T[R(y, x), R(x, y)] = 1$$
$$\Leftrightarrow \forall y \in X,\ I_T[R(y, x), R(x, y)] = 1$$
$$\Leftrightarrow \forall y \in X,\ R(y, x) \leq R(x, y)$$

For instance, if T is the Lukasiewicz t-norm defined by $T(\alpha, \beta) = \max(\alpha + \beta - 1, 0)$, then $I_T(\alpha, \beta) = 1 - \max(\alpha - \beta, 0)$ and we get:

$$v(ND(x)) = 1 - \sup_{y \in X} \max(R(y, x) - R(x, y), 0)$$

Hence, in the context of fuzzy preference relations, the subset X^{ND} of non-dominated elements becomes a fuzzy subset of X characterized by membership values $v(ND(x))$ for all x in A. From this fuzzy subset, we can derive a crisp subset UND defined as the core of ND (*e.g.* the crisp subset of elements in A such that $v(ND(x)) = 1$). Hence we obtain:

$$X^{ND} = \{(x,\ 1 - \sup_{y \in X} \max(R(y, x) - R(x, y), 0)),\ x \in X\}$$
$$X^{UND} = \{x \in X : \forall y \in A,\ R(y, x) \leq R(x, y)\}$$

Notice that we came up to the definitions of non-dominated sets X^{ND} and X^{UND} exactly as they have been originally proposed in [96].

[0, 1]-valued logic for preference aggregation. Suppose now that the decision maker is able to formulate logical principles specifying what should be a good alternative or a well-stated preference. [0, 1]-valued logic can be used to support the definition of an aggregation function reflecting these principles through a purely numerical process. This can be illustrated by the following logical sentences:

LS1: "An alternative is *relevant iff* there exists at least one important criterion according to which it is good"

LS2: "An alternative is *good iff* it is good according to all the important criteria"

LS3: "An alternative a is *opposable* to b *iff* there exists at least one important criterion according to which a is better to b"

LS4: "An alternative a is *preferred* to b *iff* a is better than b according to all the important criteria"

Consider the following predicates:

Symbol	*Arity*	*Domain*	*Interpretation*
relev	1	A	*relev*(x): "x is relevant"
good	1	A	*good*(x, y): "x is good"
imp	1	N	*imp*(j): "j is important"
pref	2	$A \times A$	*pref*(x, y): "x is preferred to y"
good1	2	$A \times N$	*good1*(x, j): "x is good according to criterion j"
oppo	2	$A \times A$	*oppo*(x, y): "x is opposable to y"
pref1	3	$A \times A \times N$	*pref1*(x, y, j): "x is preferred to y according to j"

where A is the set of alternatives and $N = \{1, ..., n\}$ is the set of criteria. Using these symbols, logical sentences LS1, ..., LS4 can be represented by the following equations:

$$\forall x\ [\exists j\ \ (imp(j) \wedge good1(x,j)) \equiv relev(x)]$$
$$\forall x\ [\forall j\ \ (imp(j) \supset good1(x,j)) \equiv good(x)]$$
$$\forall x \forall y\ [\exists j\ \ (imp(j) \wedge pref1(x,y,j)) \equiv oppo(x,y)]$$
$$\forall x \forall y\ [\forall j\ \ (imp(j) \supset pref1(x,y,j)) \equiv pref(x,y)]$$

If we try to combine these equations with equations (15.3)-(15.9), denoting $r(x) = v(relev(x))$, $g(x) = v(good(x))$, $o(x, y) = v(oppo(x, y))$, $p(x, y) = v(P(x, y))$, $x_j = v(good1(x, j))$ the degree to which x satisfies the j^{th} objective, $w_j = v(imp(j))$ the degree to which criterion j is important, $p_j(x, y) = v(pref1(x, y, j)))$ we obtain the following results:

$$r(x) = V(T(w_1, x_1), ..., T(w_n, x_n))$$
$$g(x) = T(I_T(w_1, x_1), ..., I_T(w_n, x_n))$$
$$o(x, y) = V(T(w_1, p_1(x, y)), ..., T(w_n, p_n(x, y)))$$
$$p(x, y) = T(I_T(w_1, p_1(x, y)), ..., I_T(w_n, p_n(x, y)))$$

For instance, choosing $T(x, y) = \min(x, y)$ and $V(x, y) = \max(x, y)$ we get:

$$r(x) = \max(\min(w_1, x_1), ..., \min(w_n, x_n))$$
$$g(x) = \begin{cases} 1 \ \text{ if } \forall j \in N,\ \ x_j \geq w_j \\ \min\{x_j :\ x_j < w_j\} \ \text{ otherwise} \end{cases}$$

$$o(x, y) = \max(\min(w_1, p_1(x, y)), ..., \min(w_n, p_n(x, y)))$$

$$p(x, y) = \begin{cases} 1 \ \text{ if } \forall j \in N,\ \ p_j(x,y) \geq w_j \\ \min\{p_j(x,y) :\ p_j(x,y) < w_j\} \ \text{ otherwise} \end{cases}$$

Usually, a coefficient $w_k = 1$ is attached to the most important criterion and to all other criteria of same importance. In this case, $r(x)$ and $g(x)$ are compromise aggregators taking their values in the interval $[\min\{x_j : j \in N\}, \max\{x_j : j \in N\}]$. $r(x)$ is a weighted maximum of partial satisfaction indices (see [36]). In the function $g(x)$, coefficients w_j can be understood as aspiration levels since $g(x)$ is a minimum operator restricted to criterion values that do not completely satisfy the aspirations of the decision maker. In the same way, $o(x, y)$ and $p(x, y)$ define fuzzy relations

reflecting a compromise between one-dimensional preferences. Many other sophisticated operators have been studied for aggregating fuzzy sets or preference relations. The interested reader should consult [29, 160, 38, 47, 54, 46, 12]).

These illustrative examples show that multivalued logics associated with fuzzy sets provide powerful algebraic tools for the systematic and careful extension of conventional preference models. In the framework of MCDM, the main interest of constructing and processing fuzzy preferences is the increased expressiveness of models and their ability to preserve continuous dependence between inputs (criterion values) and outputs (recommendations) in decision support systems (see [104]). An alternative approach to extend the conventional language of preference is to use another non-classical logic allowing the explicit representation of partial inconsistencies and conflicting information. The following section provides some recent works relying on Belnap's logic and its extensions.

15.2.2 The use of Belnap's logic and its extensions

The basic assumption in Belnap's logic [9, 10] is that the extension of the formula $\neg A$ is not necessarily coincident with the complement of the extension of A. Such an assumption avoids the systematic inconsistency of the sentence $A \wedge \neg A$ (inconsistency can be recovered with $A \wedge \sim A$ where $\sim$ denotes the complement of A, *i.e.* an interpretation belongs to $\sim A$ if and only if it does not belong to A). In the same way, $A \vee \neg A$ is no longer a tautology. Such a logic is well fitted to the representation of partial ignorance and contradiction. Using such a logic is of particular interest in the context of MCDM where preference modeling requires specific tools to better handle the possibly incomplete or conflicting information.

In this context, preference modeling implies to decide whether an alternative a is preferred to an alternative b or not. As mentioned above, finding decisive arguments is not always possible, even after a careful investigation of the possible criterion values of alternatives a and b. Basically, four typical situations can be distinguished:

- s1: there exist strong arguments supporting the preference of a over b and no significant argument conflicting with this preference,
- s2: there exist weak arguments supporting the preference of a over b and sufficiently strong arguments are conflicting with this preference,
- s3: there exists little argument supporting the preference of a over b but there exists no sufficient argument for deciding that a is not better than b,
- s4: there exists strong arguments supporting the preference of a over b but there exists also strong arguments conflicting with this preference.

The first and second situations are usual and can be well represented with the classical logic. In the first case (s1), the preference of a over b is well justified and $P(a, b)$ must be interpreted as *true*; in the second case (s2), the preference of a over b is not justified and $P(a, b)$ must be interpreted as *false*. The two other situations are less conventional. In situation s3, interpreting $P(a, b)$ as *true* seems irrelevant because of the lack of positive reasons justifying this preference, but interpreting

$P(a, b)$ as *false* is not better justified. Actually, the most natural interpretation for preference is *unknown*, since we lack of decisive arguments. In situation s4, it is also impossible to decide whether $P(a, b)$ is *true* or *false*. However, the lack of decisiveness is not due to the absence of significant information but to the strong conflict between arguments supporting preference and non-preference. In order to distinguish this situation from the previous one, interpreting $P(a, b)$ as *contradictory* seems natural. Thus four truth values t, f, u, k representing the *true*, *false*, *unknown* and *contradictory* are considered so as to distinguish the decision situations s1, ..., s4. More precisely, denoting $\Delta P(a, b)$ (resp. $\Delta\neg P(a, b)$) the presence of significant arguments supporting $P(a, b)$ (resp. $\neg P(a, b)$), the possible interpretations of the atomic formula P are defined as follows:

$$v(P) = t \Leftrightarrow \Delta P \wedge \neg\Delta\neg P \quad (15.11)$$
$$v(P) = f \Leftrightarrow \neg\Delta P \wedge \Delta\neg P \quad (15.12)$$
$$v(P) = u \Leftrightarrow \neg\Delta P \wedge \neg\Delta\neg P \quad (15.13)$$
$$v(P) = k \Leftrightarrow \Delta P \wedge \Delta\neg P \quad (15.14)$$

The interpretations of non-atomic formulae can be derived using the following truth tables (see [28]):

$v(P)$	t	k	u	f
$v(\neg P)$	f	k	u	t

$$v(A \wedge B) = f_\wedge(v(A), v(B))$$
$$v(A \vee B) = f_\vee(v(A), v(B))$$
$$v(A \supset B) = f_\supset(v(A), v(B))$$
$$v(A \equiv B) = f_\wedge(f_\supset(v(A), v(B)), f_\supset(v(B), v(A)))$$

with the following truth tables:

$f_\wedge$	t	k	u	f
t	t	k	u	f
k	k	k	f	f
u	u	f	u	f
f	f	f	f	f

$f_\vee$	t	k	u	f
t	t	t	t	t
k	t	k	t	k
u	t	t	u	u
f	t	k	u	f

$f_\supset$	t	k	u	f
t	t	k	u	f
k	t	t	u	u
u	t	k	t	k
f	t	t	t	t

Moreover, for any predicate $A(x)$ with the instance domain $\{x_1, ..., x_n\}$, the interpretation of quantified formulae is given by:

$$v(\exists x\, A(x)) = f_\vee(v(A(x_1)), ..., v(A(x_n)))$$
$$v(\forall x\, A(x)) = f_\wedge(v(A(x_1)), ..., v(A(x_n)))$$

The interest of a four-valued logic to express ignorance and partial inconsistency is well known in the Artificial Intelligence community since the important work of [9, 10] and this idea was even present earlier in the literature (see [32]). However, surprisingly enough, the idea of adapting such a logic to knowledge representation

and preference modeling in the context of MCDM is relatively recent [28, 141, 142, 143, 144]. We present now an illustrative example of the use of this logic in MCDM.

Four-valued logic for preference aggregation. Multicriteria aggregation aims at synthesizing possibly conflicting preferences. At this step, the four-valued logic can be used to obtain a compact representation of overall preferences, allowing the typical situations (preference, absence of preference, conflict, ...) to be clearly distinguished [144]. For example, coming back to the multicriteria problem with n criterion functions $g_j, j = 1, ..., n$, defining n preference relation (*e.g.* semi-orders) on the set of alternatives (see equation (15.1)). By definition, a set function m measuring the importance of each coalition of criteria fulfills the following conditions:

$$m(2^N) = 1, \qquad m(\varnothing) = 0$$
$$\forall A, B \in 2^N, \quad A \subset B \;\Rightarrow\; m(A) \leq m(B)$$

where $N = \{1, ..., n\}$ is the set of criteria indices.

Considering the following coalitions:

$$[a\ P\ b] = \{j \in \{1, ..., n\}: g_j(a) - g_j(b) > q_j\}$$
$$[b\ P\ a] = \{j \in \{1, ..., n\}: g_j(b) - g_j(a) > q_j\}$$

we define the following conditions:

(15.15) $$\Delta P(a, b) \Leftrightarrow m([a\ P\ b]) > \gamma$$

(15.16) $$\Delta \neg P(a, b) \Leftrightarrow m([b\ P\ a]) > \delta$$

Hence equations (15.11)-(15.14) lead to a four-valued relation synthesizing the main possible situations in the aggregation. Figure 15.6 illustrates the four basic situations when $m(J)$ represents the proportion of voters in the coalition J, for any $J \subset 2^n$. The *true* reflects a significant consensus for the preference $P(a, b)$ without any significant discordant criterion, the *unknown* represents a situation where the coalitions supporting the preference are not significant and the *contradictory* reveals the presence of strongly conflicting coalitions of criteria.

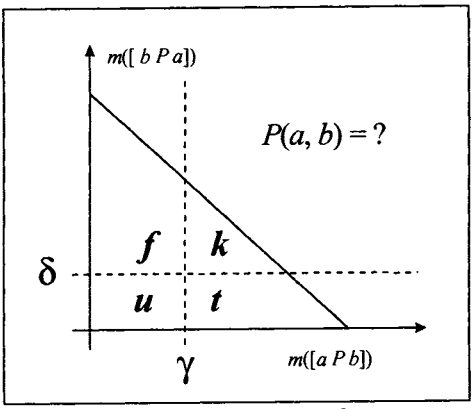


Figure 15.6: Four-valued logic for preference aggregation

Continuous extension of the four-valued logic for preference modeling. The previous construction has been extended in [106] for the aggregation of fuzzy preference relations. In [106], a continuous extension of the four-valued logic has been introduced. It is based on a continuum of truth values characterized by the set of 2×2 matrices whose elements M_{ij} verify the following conditions:

$$\forall i,j \in \{1,2\},\ M_{ij} \geq 0$$
$$M_{11} + M_{12} + M_{21} + M_{22} = 1$$
$$\min\{M_{12}, M_{21}\} = 0$$

The interpretation of such matrices is very intuitive since any proposition P of the predicate calculus is represented by the matrix:

$$v(P) = \begin{pmatrix} f(P) & k(P) \\ u(P) & t(P) \end{pmatrix}$$

where $f(P)$, $k(P)$, $u(P)$, $t(P)$ denote the degree to which the proposition P is *true*, *contradictory*, *unknown* and *true*. The functions f, k, u and t are defined on the set of well-formed formulae and valued in the unit interval. Following the approach adopted in the four valued logic, it is possible to define functions f, k, u and t by balancing positive and negative arguments to justify any proposition. Denoting $c(P)$ and $c(\neg P)$ the confidence we have in propositions ΔP and $\Delta\neg P$ respectively, a possible translation of equations (15.11)-(15.14) using multivalued logic is:

(15.17) $$t(P) = \min(c(P),\ 1 - c(\neg P))$$

(15.18) $$k(P) = \max(c(P) + c(\neg P) - 1,\ 0))$$

(15.19) $$u(P) = \max(1 - c(P) - c(\neg P),\ 0))$$

(15.20) $$f(P) = \min(1 - c(P),\ c(\neg P))$$

Logical operations like negation, conjunction, disjunction and implication have been defined in this framework but there is no space in this chapter for presenting them in details. For more details see [106]. We present now the possible use of such a continuous extension for preference aggregation.

Continuous extension of four-valued logic for preference aggregation. Another feature of this logic is the possibility to refine four-valued preference models resulting from the aggregation of multiple criteria. The basic idea is to allow the continuous transition from $\Delta P(a, b)$ to $\neg\Delta P(a, b)$, and from $\Delta\neg P(a, b)$ to $\neg\Delta\neg P(a, b)$. Equations (15.15)-(15.16) do not allow this continuity since $\Delta P(a, b)$ and $\Delta\neg P(a, b)$ are 0-1 predicates defined by using cutting thresholds γ and δ. However, these predicates could be interpreted within the framework of multivalued logic. For example, consider the following definitions:

(15.21) $$c(\Delta P(a,b)) = \min\left(1,\ \max\left(0, \frac{m([a\ P\ b]) - \gamma^-)}{\gamma^+ - \gamma^-}\right)\right)$$

$$c(\Delta P(a,b)) = \min\left(1, \max\left(0, \frac{m([b\,P\,a]) - \delta^-)}{\delta^+ - \delta^-}\right)\right) \tag{15.22}$$

where γ^- is the threshold below which the preference is not defensible, γ^+ is the threshold beyond which the preference is well supported, δ^- is the threshold below which the preference $P(a, b)$ remains possible, δ^+ is a veto threshold beyond which the preference $P(a, b)$ is no longer defensible ($\gamma^+ \geq \gamma^- \geq \delta^-$ and $\gamma^+ \geq \delta^+ \geq \delta^-$). The resulting truth value of $P(a, b)$ is characterized by $t(P(a, b))$, $k(P(a, b))$, $u(P(a, b))$, $f(P(a, b))$ defined by equations (15.17)-(15.20). The possible values of these indices are depicted on figure 15.7 (to be compared to figure 15.6), for $\gamma^- = 0.3$, $\gamma^+ = 0.5$, $\delta^- = 0.1$, $\delta^+ = 0.3$.

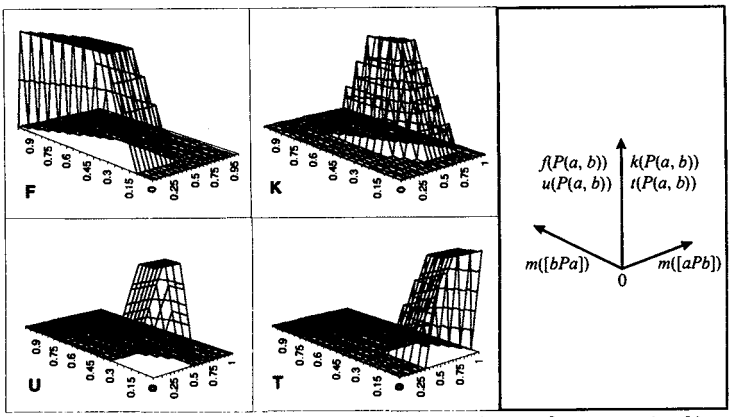


Figure 15.7: Possible truth values for the preference $P(a, b)$

15.2.3 Preference expression as a reasoning process

The previous examples on Belnap's logic illustrate how preferences can emerge from a reasoning process. This reasoning process depends on context and information about future. The reasoning must be non-monotonic because the knowledge about future may change. Thus, we enter into the realm of non-classical logics and various proposals have been made to accommodate these problems. A first attempt using first order logic with "Truth Maintenance Systems" [30] can be found in [140]. Further works inspired by the default logic [138] distinguish between beliefs and preferences, the latter depending on states of the world [31]. A recent work [90] aims at merging case-based reasoning and default reasoning (or negligibility). In fact, the idea of distinguishing between beliefs (information about future) and preferences (utility of the decision maker about results) goes back to Savage [131]. In [131] a probabilistic information is deduced from the preferences on the actions. However, in practice, a probabilistic a priori information is required to derive usual decision criteria, this is the "Bayesian" point of view. Unfortunately, the available information is generally poorer than requested by the model in terms of accuracy and quantity. Several suggestions have been made to cope with partial and rough information about future and relax the assumption of a probalilistic world. One can use non-

additive measures such as belief functions [132] and then, apply a so-called "pingnistic" transformation to deduce manageable probabilities and decision criteria [135]. One can also decide to replace probabilities by possibilities. This domain of "qualitative decision theory" is nowadays very active. A complete reconstruction of Decision Theory under Uncertainty [92, 131] within a possibilistic framework has been developed in [39, 34], while [76, 42] also parallel Savage's framework for a qualitative decision theory.

Actually, it is worth noticing that there is a formal analogy between decision under uncertainty and multicriteria decision making. In Savage [131], each elementary event belonging to the set S of states of nature entails a different result for an action a. The result (or consequence) is a function $r(a, s)$. Assume that s is a criterion and a an alternative, the evaluation of the alternative a according to criterion s is also a function $r(a, s)$. Thus, subsets of S may be regarded as a coalition of criteria and the relative likelihood of states or events as the relative importance of criteria or coalitions. This striking analogy is well illustrated in [34, 35, 42]. These papers emphasize the relationships between multicriteria decision making and various models for decision making under uncertainty proposed by AI people. Let us give, in the next section, a flavour of the possiblilistic approach of decision under uncertainty.

15.2.4 Possibility theory and qualitative decision theory under uncertainty

One important source of complexity in multicriteria decision problems is the imperfect knowledge of the consequences of alternatives. In this context, preference modeling requires the use of specific tools and concepts allowing the representation of partial ignorance and the comparison of uncertainty distributions. As mentioned above, the classical approaches to decision making under uncertainty are based on probabilities. Even before Savage's important contribution, the initial axiomatic approach proposed by von Neumann and Morgenstern [92] advocates the use of the *expected utility* criterion, based on the use of numerical probabilities and utilities about the consequences of alternatives. However, when the information about the relative likelihood of outcomes is not sufficient to define precise probabilities the standard expected utility criterion becomes useless. Several ways to overcome this difficulty have been investigated, based on the use of belief functions [132, 58, 59], or capacitities [25] or possibilities [39].

Since only poor information is available in many situations, it is also difficult to construct a cardinal utility function for the comparison of the consequences of alternatives. In many situation a more qualitative decision theory based on ordinal information is necessary. This shows the specific interest of considering the concepts and tools introduced in the field of Artificial intelligence. The classical Artifical Intelligence community is traditionally attached to purely symbolic approaches for knowledge representation and some of the models may be used advantageously in

the context of decision making. For example, possibility theory is one of the most popular alternatives to classical probability theory for the representation of uncertainty in preference models (see [37, 105]). In order to show how uncertainty and preference can be jointly handled for a *qualitative decision theory*, we briefly recall the possibilistic counterpart of the von Neumann and Morgenstern expected utility, as it has been introduced in [39].

Consider a decision problem where the consequences of alternatives on some dimension j are uncertain and characterized by possibility distributions defined on a finite set of consequences X. Formally, each alternative is represented by a lottery π, that is a function defined from X into [0, 1] and such that $\pi(x)$ represent the possibility of consequence x with $\sup_{x \in X}\{\pi(x)\} = 1$. Note that $\pi(x) = 0$ means that the consequence x is impossible for the alternative represented by π, $\pi(x) = 1$ means that x is a really plausible (unsurprising) consequence, and $\pi(x) \leq \pi(y)$ means that consequence y is at least as likely as consequence x. It is shown in [39] that, under a set of reasonable conditions (transitive and complete comparability of lotteries, certainty equivalence, risk aversion, independence, reduction of lotteries, continuity), the decision maker's preferences on dimension j can be represented by the criterion function g_j^- to be maximized. This function is defined on the set of lotteries as follows:

$$g_j^-(\pi) = \min_{x \in X} \max(N(\pi(x)), g_j(x)) \tag{15.23}$$

where N is a negation defined on the unit interval (*e.g.* $N(\alpha) = 1 - \alpha$).

Criterion function g_j^- reflects a rather pessimistic point of view since it can be interpreted as the extension of the Wald maximin criterion with possibilistic weights. An alternative view, more optimistic, can be adopted within the framework of qualitative decision theory under uncertainty, by substituting "risk aversion" by the "risk prone" axiom. In this case, decision maker's preferences can be represented by another criterion function $g_j^+(\pi)$ (to be maximized) defined as follows [39]:

$$g_j^+(\pi) = \max_{x \in X} \min(\pi(x), g_j(x)) \tag{15.24}$$

Since by definition there exists at least one $x_0 \in X$ such that $\pi(x_0) = 1$, we have:

$$\min_{x \in S(\pi)}\{g_j(x)\} = g_j^-(\pi) \leq g_j^+(\pi) = \max_{x \in S(\pi)}\{g_j(x)\}$$

where $S(\pi) = \{x \in X : \pi(x) > 0\}$. Therefore, $g_j^-(\pi)$ and $g_j^+(\pi)$ are compromise operators that can be understood as natural counterparts of expected value expressed in the max-min algebra. Interpreting $g_j(x)$ as the degree to which consequence x satisfies the jth objective of the decision maker, $g_j^+(\pi)$ (resp. $g_j^-(\pi)$) measures the possibility (resp. the certainty) that the outcome of π is good on dimension j. Intervals of type $[g_j^-(\pi), g_j^+(\pi)]$ provide a compact representation of lotteries that allow an interval order structure to be defined as follows:

$$\pi \, P_j \, \pi' \quad \Leftrightarrow \quad g_j^-(\pi) > g_j^+(\pi')$$
$$\pi \, I_j \, \pi' \quad \Leftrightarrow \quad [g_j^+(\pi), g_j^-(\pi)] \cap [g_j^-(\pi'), g_j^+(\pi')] \neq \varnothing$$

Sharper preference structures are given by complete preorders of type $R^- = P^- \cup I$ or $R^+ = P^+ \cup I$ where P^-, P^+ and I are defined as follows:

(15.25) $\pi P_j^- \pi' \Leftrightarrow [g_j^-(\pi) > g_j^-(\pi')$ or $(g_j^-(\pi) = g_j^-(\pi')$ and $g_j^+(\pi) > g_j^+(\pi'))]$
(15.26) $\pi P_j^+ \pi' \Leftrightarrow [g_j^+(\pi) > g_j^+(\pi')$ or $(g_j^+(\pi) = g_j^+(\pi')$ and $g_j^-(\pi) > g_j^-(\pi'))]$
(15.27) $\pi I_j \pi' \Leftrightarrow [g_j^-(\pi) = g_j^-(\pi')$ and $g_j^+(\pi) = g_j^+(\pi')]$

As pointed out by [105], this approach may also be used to generalize the construction of preference relations from criterion values when alternatives are represented by possibilistic lotteries. Suppose for instance that a preference relation P_j has been defined from g_j using a function f_j as in equation (15.1). From relation P_j we can derive two relations P_j^- and P_j^+ extending P_j for the comparison of lotteries:

(15.28) $P_j^-(\pi, \pi') = \inf_{(x, y) \in X \times X} \max(1 - \pi(x), 1 - \pi'(y), f_j(g_j(x), g_j(y)))$
(15.29) $P_j^+(\pi, \pi') = \sup_{(x, y) \in X \times X} \min(\pi(x), \pi'(y), f_j(g_j(x), g_j(y)))$

Example 15.1 Consider $X = \{x_1, ..., x_5\}$ a set of consequence on dimension j such that $g_j(x_1) > g_j(x_2) > g_j(x_3) > g_j(x_4) > g_j(x_5)$ and consider 3 alternatives (acts) represented by 3 lotteries π, π' and π'' such that $\pi(x_2) = \pi''(x_4)$, $\pi(x_5) = \pi''(x_1)$ and $0 < \pi(x_5) < \pi(x_2)$ (see figure 15.8).

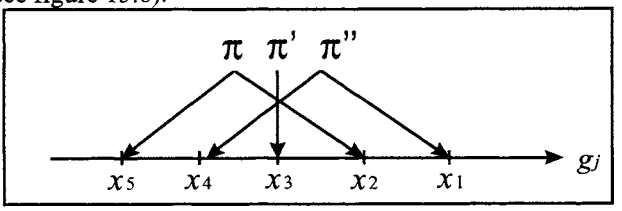


Figure 15.8: Comparison of lotteries π, π' and π''

In this case, the possible states of the world are all the possible triplets of outcomes for π, π', π''. Thus we get the following four states: $s_1 = (x_5, x_3, x_4)$, $s_2 = (x_5, x_3, x_1)$, $s_3 = (x_2, x_3, x_4)$, $s_4 = (x_2, x_3, x_1)$ and the following "multicriteria decision table" (see table 15.3):

	s_1	s_2	s_3	s_4
π	x_5	x_5	x_2	x_2
π'	x_3	x_3	x_3	x_3
π''	x_4	x_1	x_4	x_1

Table 15.3: Outcomes of alternatives π, π' and π''

Consistently with the definition of possibility distributions we set $\pi(x_2) = \pi''(x_4) = 1$ and $\pi(x_5) = \pi''(x_1) = \alpha \in (0, 1)$. Hence s_3 is the more plausible state and s_1, s_2, s_4 have equal possibilities. Suppose that the preference relation P_j is defined on X by:

$$f_j(x,y) = \begin{cases} 1 \ \textit{if} \ g_j(x) > g_j(y) \\ 0 \ \textit{otherwise} \end{cases}$$

We get $x_1 P_j x_2 P_j x_3 P_j x_4 P_j x_5$. The associated relations P_j^- and P_j^+ obtained by equations (15.28)-(15.29) are depicted on figure 15.9.

Notice that it is easy to derive transitive preference relations from these valued graphs (*e.g.* $\pi\ P_j^+\ \pi' \Leftrightarrow (P_j^+(\pi, \pi') > P_j^+(\pi', \pi))$). This would not be the case with a probabilistic representation of uncertainty. Suppose for instance that the same problem has been represented with 3 probabilistic lotteries p, p' and p'' instead of π, π' and π''. In order to represent the relative likelihood of consequences we have to set $p(x_2) = p''(x_4) = \beta$, $p(x_5) = p''(x_1) = 1 - \beta$ and $0.5 < \beta < 1$. Hence s_3 is the more probable state and the probability of the preference of type $p\ P_j p'$ is given by:

$$P_j(p, p') = \sum_{(x,y) \in X \times X} p(x)\, p'(y)\, f_j(x, y)$$

This leads to the preference graph depicted on the right of figure 15.9. Unfortunately this graph does not provide a clear information since the preference relation defined by: $p\ P_j p' \Leftrightarrow (P_j^+(p, p') > P_j^+(p', p)))$ is cyclic for $\beta \in [1/2, 1/\sqrt{2}\ ]$.

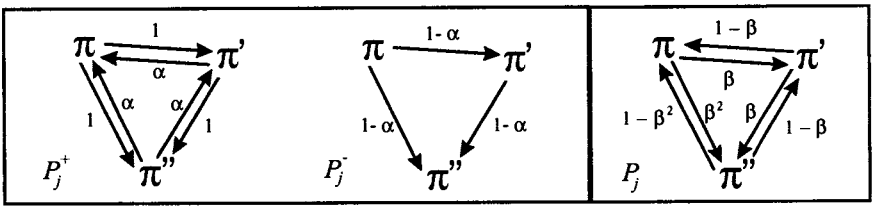


Figure 15.9: Possibility, necessity and probability of preference

Thus we observed that an additive (probabilistic) representation of uncertainty does not allow to export the transitivity from preferences on crisp consequences to preferences on lotteries. This example highlights a typical advantage of possibility theory for handling uncertainty in multicriteria decision support systems. For a deeper analysis of such problems the interested reader should consult [34, 42].

Suppose now that criterion values $g_j(x)$ and possibilities $\pi(x)$, $\pi'(x)$ and $\pi''(x)$ are known for all x in X and expressed on the same scale [0, 1]. For example, $\pi(x_2) = \pi''(x_4) = \pi'(x_3) = 1$, $\pi(x_5) = \pi''(x_1) = 0.6$, $g_j(x_1) = 0.9$, $g_j(x_2) = 0.7$, $g_j(x_3) = 0.5$, $g_j(x_4) = 0.3$, $g_j(x_5) = 0.1$. Using equations (15.23, 15.24) we obtain the following criterion values: $g_j^-(\pi) = 0.4$, $g_j^+(\pi) = 0.7$, $g_j^-(\pi') = 0.5$, $g_j^+(\pi') = 0.5$, $g_j^-(\pi'') = 0.3$, $g_j^+(\pi'') = 0.6$. Hence by equations (15.25)-(15.27) we get: $\pi\ P^+ \pi''\ P^+ \pi'$ and $\pi'\ P^- \pi\ P^- \pi''$.

Notice that, in both orderings, π is preferred to π'' but the position of π' remains unclear. It depends on the attitude of the decision maker towards risk. Notice also that comparing lotteries through their "expected" utilities always leads to transitive preferences. This is the main interest of making comparable the uncertainty scale and the preference scale. ■

The various concepts developed in this section illustrate either qualitative or quantitative approaches developed in the artificial intelligence community and show their potential interest for knowledge representation and reasoning in MCDM. Some other important contributions of AI to decision analysis concern the learning of preference models and decision rules. These aspects are partly evoked in the section devoted to the neural approach but other models based either on rough sets or

evolutionary algorithms could be considered as well (see Chapters 14 and 16 of this book).

15.3 Heuristic search and interactivity

15.3.1 Introduction

Perhaps the most striking impact of AI on decision aid, generally speaking, or more specially, in MCDM is the spreading of heuristic search methods. At the beginning is the idea put forward in [91] that human problem solving results into an exploration among a state space, each state containing some information about the solution. This exploration obeys the very simple rules of heuristic search [91]. Heuristic search is a systematic trial and error process controlled by the evaluation of each state. This evaluation is generally carried out by an evaluation function which formalizes the heuristic knowledge about the search. However, in many cases it is difficult to formalize this heuristic knowledge and the evaluation of each state is very contextual and idiosyncratic. This is the reason why decision making requests human intervention [110, 83].

In a heuristic search the purpose is to find a sequence of well-defined operators leading from the initial state to the target one (Figure 15.10). In any attained current state the explorer has to decide, according to his evaluation, whether :

- he chooses one of its son states as the new current state (advance)
- he still continues to develop new sons (continue)
- he backtracks to an already developed and recorded state.

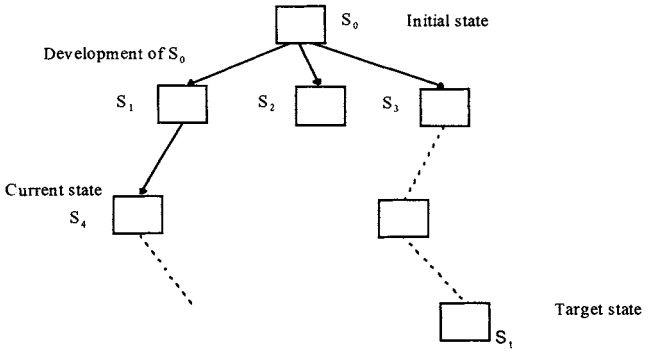


Figure 15.10: Heuristic search

This three-forked decision is the basis for any interactive decision process (see Section 15.3.4). However, this general framework is now appropriated by many researchers in optimization and decision aid. Thus, we will restrict ourselves to applications to MCDM. Among these applications the most direct consists of applying heuristic search process to choose among the alternatives into the set of all available alternatives (*choice set*).

15.3.2 Heuristic search in the choice set

The set of all the alternatives is considered as a state space. Thus, after defining the operators, the decision maker can perform a heuristic search in this space. This idea was exploited in PRIAM [78, 79]. A resulting software, named MULTIDECISION, was developed involving PRIAM algorithm (see [113], ch.9). In PRIAM, each state essentially contains the "satisfaction levels" attained by the decision maker at this step of the search. By "satisfaction level" we mean a vector valued on each attribute which is not necessarily an alternative, this vector indicates the level of satisfaction that the decision maker would like to get for each criterion. The system also indicates the number of possible alternatives still above the satisfaction level. This is the reason why the method is somewhat reminiscent of elimination by aspects [145, 146]. At the view of the alternative proposed by the system, the decision maker evaluates whether he is satisfied or how promising is the continuation of the exploration (see figure 15.11, below). Relying on this evaluation the decision maker triggers or not an operator to get new alternatives proposed by the system. He can also decide to backtrack to an already met alternative. The whole process is very similar to the A* algorithm [93].

The quality of the exploration performed by the decision maker depends on his shrewdness to understand that he cannot increase all the criteria together. In other words, when he gains on some criteria he has to diminish his wishes on some others. This is a kind of learning process about trade-offs However, the software must provide good operators that help the decision maker to make a wise exploration and, more important, to reach efficient points (Pareto boundary). In fact, this kind of exploration is inherently interactive and heuristic because, at each step, the decision maker is faced to an alternative proposed by the software. This raises two questions: Does the process leads to an efficient point and has the decision maker sufficient information to control his exploration and to evaluate each state attained by the system? To answer the first question, the idea in PRIAM was to provide the decision maker with an efficient point, dominating the current point, by prolonging the preceding move up to the Pareto boundary.

To answer the second question about the evaluation of the current state, various ideas can be put forward. The system can display some alternatives in the "neighborhood" of the current point to show to the decision maker some other possibilities. It is also possible to give an aggregated real value of each alternative displayed by the system. This is a kind of scoring system. In MULTIDECISION, some of the alternatives already met by the decision maker and considered as "satisficing" ones are displayed, so that the decision maker can easily backtrack and explore different trade-offs. With this facility and the possibility to know the distance to the Pareto boundary, and thus to get efficient points, the exploration with MULTIDECISION is very interactive and attractive. Consequently, the idea has been exploited again in some other packages [44, 45]. The same idea of interactive search is also used in AIM [86, 87] at one of the steps of the method. PRIAM was also adapted to continuous multiobjective programming [116].

15.3.3 The moving base heuristic

The moving base heuristic (MBH) is an interactive choice procedure proposed in [4, 5, 6] for multicriteria decision problems. As introduced in [6], this heuristic can be seen as a sophistication of the elimination by aspects heuristic (EBA) [145, 146] since it is based on an iterative algorithm consisting of progressive refinement of the initial set of alternatives. MBH relies on the following assumptions:

- the process cannot be reduced to elimination and may include selection steps
- at each level of the selection/elimination process the decision maker can consider several attributes at the same time and even several coalitions of attributes
- the decision maker only reacts to differences of a given amplitude (this implies discrimination thresholds)

Basically, MBH heuristic can be used to compare two alternatives. In this case, the procedure proposed in [4] consists of the following steps:

1. consider a set of attributes and thresholds on these attributes
2. compare two alternatives with respect to these attributes
3. if the observed differences are greater than the fixed thresholds
   then chose the dominating alternative
   else if there exists other relevant sets of attributes
   then select a set of attributes and thresholds and go to step 2
   else make no decision

Any session of this procedure can be summarized by a polynomial $M_1 + ... + M_k$. where M_j characterizes a set of relevant attributes at steps j with the associated thresholds. For example, consider a problem with 5 integer-valued criteria $\{g_1, g_2, g_3, g_4, g_5\}$ with thresholds values 2, 1, 5, 3, 1. In such a case,

$X_1^2 X_4^3$ means $a\,P\,b \Leftrightarrow (g_1(a)\text{-}g_1(b)) > 2 \wedge (g_4(a)\text{-}g_4(b)) > 3$

$X_2^1 X_3^5 X_5$ means $a\,P\,b \Leftrightarrow (g_2(a)\text{-}g_2(b)) > 1 \wedge (g_3(a)\text{-}g_3(b)) > 5 \wedge (g_5(a)\text{-}g_5(b)) > 1$

Suppose that the first monomial $X_1^2 X_4^3$ corresponding to the base $\{g_1, g_4\}$ does not enable to discriminate between a and b at the first stage of the *MBH*. Then, the second monomial corresponding to the base $\{g_2, g_3, g_5\}$ is used at the second stage, leading to the decision a is preferred to b. This is a combination of lexicographic orders and semi-orders with thresholds. A compact representation of this entire session is given by the polynomial: $P = X_1^2 X_4^3 + X_2^1 X_3^5 X_5^1$.

Thus, the MBH can be used to compare each pair of alternatives. Since there is little chance for a single polynomial to be compatible with all sessions used to compare all pair of alternatives, the resulting strict preference relation is not necessarily transitive. The interested reader should consult [4] to find a necessary

and sufficient condition for a compatible choice polynomial to exist, and a procedure to identify minimal choice bases.

In order to understand to what extend the *MBH* is an extension of the *EBA*, assume that a_* is a given satisfaction level. Then, MBH can be used to compare each feasible alternative to a_*. Thus, a MBH polynomial acts as a sequence of elimination by groups of aspects.

15.3.4 Interactivity

The overall framework for interactivity in MCDM is drawn in [126] who distinguishes between elicitation of the decision maker's preferences and progressive articulation of the preferences. We think that in the context of MCDM analysis, the second situation is prevalent. This is the reason why we consider interactive methods as progressive articulation methods intended to obtain more and more knowledge from the decision maker in order to jointly build the decision maker's preferences. By joint construction we mean that the decision maker and the system cooperate to the progressive elicitation of the preferences.

Although progressive elicitation methods in MCDM go back to the very beginning (*e.g.* [123, 11, 124]), and were independently developed from AI, one must stress the numerous similarities between progressive articulation method in MCDM and AI.

Actually, progressive articulation methods in MCDM always rely on the same process : the system displays to the decision maker an alternative (current alternative) and then according to the decision maker's reaction try to find another "better" alternative. This process is nothing else than a heuristic search in which the exploration is interpreted as a learning procedure. This procedure is twofold: the decision maker becomes progressively conscious of his own preferences and trade-offs and the system learns the preferences of the decision maker. The principles of progressive articulation methods are displayed on Figure 15.11, see also [149, 150] for a comprehensive view about interactive methods in a pure multicriteria framework. Among heuristic search methods, [53] distinguishes between so-called search-oriented and learning-oriented, depending on the role of the decision maker. The two concepts refer to heuristic search as presented here, but in a "search-oriented" process, the decision maker is supposed to have a given preference structure to unveil, whereas in "learning oriented" procedures, the decision maker elicits his preferences via the interaction. We do not adopt this distinction because we believe that: 1) the first case does not really exist in multicriteria decision making (*i.e.* the decision maker never has a given, a priori, preference structure in mind); 2) the observer has no means to distinguish between search and learning oriented behaviors.

On Figure 15.11, the current state is generally an alternative or a satisfaction level which is displayed to the decision maker by the system. Once the reaction of the decision maker is known the system determines another alternatives (which

becomes) the new current state. The words "continue", "advance" and "backtrack" refer to heuristic search (Section 15.3.2). When the information provided by the decision maker is not a feasible satisfaction level, for instance, an ideal point, the system generally calculates the alternative(s) that is (are) the closest of this satisfaction level by using a scalarizing function [159] or a provisional aggregation function [151]. One can find an overview of these methods in this book (Chapter 9).

In some sense, the preceding interpretation of progressive articulation methods links research of a compromise in a MCDM setting and problem solving. While it is possible to study interactivity in MCDM without any reference to AI (see *e.g.* [150]), we think that heuristic methods offer a good framework to think about interactivity and man-system cooperation. First of all, heuristic search emphasizes the role of evaluation. This is exactly what the decision maker has to do when a new current alternative is displayed by the system. Secondly, an important point introduced by Simon's work is that this heuristic search leads to a "satisficing" issue. Whereas many works relative to interactive MCDM seek to find some kind of efficient alternative by using a real-valued aggregation function, it is clear that this is not the problem for decision makers who only need a satisficing compromise between more or less contradictory criteria.

Thus, AI views eventually pave the way to more realistic and simpler procedures. See for example TRIPLE C [1] and AIM [86, 87] for interactive methods closer to AI than traditional interactive methods. This simplicity is often appreciated by decision makers [156, 95].

Finally, we do not develop in this sub-section the problems raised by user-friendliness of the interfaces and man-machine communication in the framework of MCDM. However, a lot of ideas in these domain come from AI people. We refer to [113] for an introduction about visualization methods in MCDM.

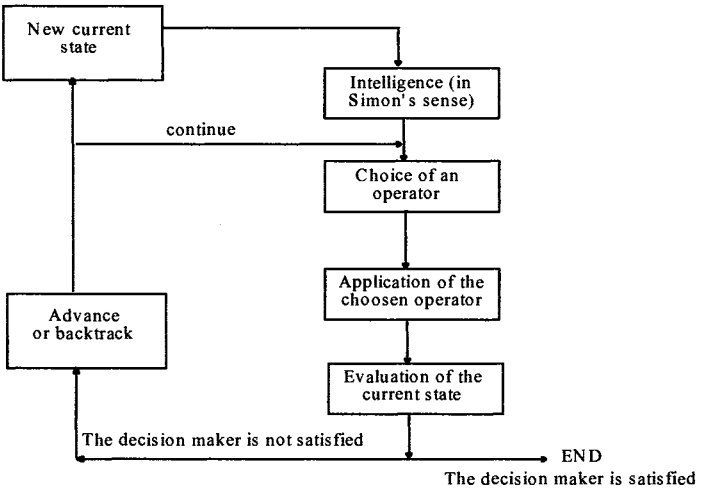


Figure 15.11 : Principle of progressive articulation methods

15.3.5 Multicriteria DSSs

A presentation of multicriteria DSSs (MCDSSs) would be difficult without any reference to AI because many of the most recent MCDSSs involve knowledge-based systems and various representations issued from AI. However, we do not intend to address the problem of MCDSS design and development in its full generality, referring to [111] for a critical survey. We just want to point out some possibilities opened by AI technique for the development of MCDSSs. To structure our presentation, we will refer to Simon's analysis of the decision processes [133]. Thus we will distinguish the four, more or less recursive, decision phases: intelligence, design, choice and review. Of these four phases, there are no theoretical reasons explaining why MCDSS designers generally address only the third. However, there are probably practical and cultural reasons. We will not review here the issues related to choice because this is the main subject of multicriteria analysis and therefore of this book. Let us just survey the contributions of AI addressing the three other phases.

Intelligence. "Intelligence" refers to the information process essential in any decision support system because, as Simon said, information constrains the decision. Information is related to the data base and the data base management system contained in any DSS. The functionalities related to the data base (recording, retrieval, structuration, etc.) are not specific to MCDSSs. Consequently we will not devote a special subsection to this theme but will refer to the general literature on DSSs or Executive Information System (EIS). However, let us add that we are not aware of a single MCDSS noticeable from this point of view. This is probably one weakness which unfortunately, as we will see, prevents designers from addressing the fourth phase (review). These questions have recently received a renewed attention in the framework of datarwarehousing, since datawarehousing generally point to multidimensional decision tools and multidimensional spread-sheets.

Design. There are two specific targets for MCDSS designers: the alternatives and the criteria. Let us begin with the criteria. The first task consists of defining a family of criteria. As explained in [21] (see also [128, 113]) it is often difficult to decide what the criteria must be. For example, how to measure the damage caused to the inhabitants living near an airport. As far as we know, the process of specifying the criteria remains a human one, mainly relying on the discussion between people interested in the problem.

The second question is that of the consistence of the family of criteria [128]. Although it seems possible to bring some support to the decision maker for this particular task, the notion is probably too recent to be already taken into account in MCDSSs.

Once, everybody agrees about the family of criteria, assuming that the alternatives are known, it remains to complete the decision matrix, *i.e.* to evaluate each alternative according to the criteria. This evaluation theoretically depends on the posterior aggregation procedure, but this fact is generally ignored by the designers so that the assessment is generally independent on the aggregation.

We are not aware of any significant AI application to the design of criteria which tackles the preceding problems. One can say the same thing about the weights, although there exist several packages bringing some support for weight evaluation (see [113], ch. 4).

The second point concerns the design of the alternatives. This is a prominent point because many implicit assumptions relative to the alternatives are not realistic. The main implicit assumption is that the set of alternatives is well-defined and fixed (B. Roy, in [26]). This assumption, which is generally not satisfied, suffices in many cases to prevent decision makers from using multicriteria analysis. The second assumption being that the alternatives be mutually exclusive. For instance, it is not allowed to mix them to get a new synthetic alternative.

We have already addressed the question of the alternative design in Section 15.1.2. Knowledge based systems are generally necessary to expand simple alternatives into robust scenarios (or FEA, see Section 15.1.2) that can be evaluated. In our previous example of timetable robustness [115] an expert system was introduced to make some intermediary decisions. In a MCDSS for the choice among the many possible variants of a future privation law [81], expert systems were designed to simulate actors reactions (companies, unions, stockholders). Many multicriteria negotiation systems use AI methods to draw up the proposals of the negotiators (progressive building of the alternatives). Among these systems, let us mention those using rule representations: MEDIATOR [60], DECISION MAKER [51], and NEGO [68], see also [82] and [65] for surveys. A system relying on multi-agent technique has also been designed in [41].

In many MCDSSs, the user has to introduce the alternatives, but in some cases the system can help to generate them (see 15.1.2). Sometimes the system provides some help for assessing the feasibility of the alternatives and generating some new ones. For example, in [107] the alternatives are introduced via a spread-sheet which allows the user to test the feasibility and the quality of the alternatives he has in mind.

More generally many MCDSSs tends to become multi-structural, in the sense that they are based on many types of models [81]. In many cases, theses multi-structural MCDSSs are in fact "intelligent" DSSs, this means that they involve knowledge-based modules [7, 13, 23, 33, 52, 75, 115, 136]. *Without any doubt, the main effort to introduce MCDSSs in organizations concerns the design of fully expanded alternatives using the resources of various models and AI techniques within multi-structural MCDSSs.*

Review and learning. Whereas it is not clearly stated in [133], we think that one of the most important functions of "review" is learning and we believe that the main support that could be provided to organizations should especially concern "learning". In many cases, we have observed that the decision is thought as a one shot game whereas most decisions are more or less repetitive. Moreover, human memory has some known biases and, for that reason, cannot accurately analyse the decision ex-post. Thus, it is an illusion to think that one can reach any learning capabilities for

complex decisions without the help of a computer and a DSS. Nevertheless, up to now, very little seems to have been done in this domain.

There are many possibilities related to learning, review and ex-post analysis. First, in some senses, a decision maker can learn the effect of the values he has given to the weights. Similarly, in outranking methods, the decision maker learns to modify the concordance and discordance factors, [134] and [155].

In some interactive procedures based on aspiration levels, whose prototypes are given by DIDASS in continuous programming [55] and PRIAM in a discrete setting [79], the decision maker certainly learns to behave with conflicting goals and he gains a certain ability for trade-offs. The same idea is also present in certain graphic methods (*e.g.* TRIPLE C [1], or the Pareto Race [71]).

The invitation to enlarge the heuristic search of the decision maker by showing alternatives not yet studied can also be regarded as a kind of learning, since this may lead to new compromises different from those which were envisaged. MULTIDECISION [113] and AIM [87] have such a feature.

To learn about decision making, it is most important to record the exact conditions prevailing when a decision was made. The main items are: what were our ideas about the possible alternatives (why?), what was our evaluation of the alternatives and, what were our probabilities, our expectations, our perception of the environment? It is clear that the previous questions cannot be answered without an appropriate MCDSS which, together, supports the decision maker and records the data describing the context of a decision. Designing such a MCDSS is the price to pay for learning and making serious comparisons to discover possible flaws in the past decisions. This is also the only way to record reliable data to shed some light on future decisions. To our knowledge, some EISs have hardly attacked some of these questions but no MCDSS. This kind of learning still seems to be an empty field open to MCDSSs.

15.4 Conclusions

We have tried to provide a panorama of AI representations and tools used in MCDM. Many of these tools (rules, objects, non-classical logic, neural networks, fuzzy sets, rough sets, possibilities, etc) are now commonly used and the users no longer refer to AI. Similarly, whereas heuristics methods were for a long time absent from operational research culture and, consequently, from MCDM, they are now widely spread thanks to interactive systems. Hence, one can wonder what the future impact of AI on MCDM is? Right now, the most active interfaces between AI and MCDM seem to be case-based reasoning as regards symbolic representations and decision under uncertainty for numerical issues. This latter topic is potentially very important for MCDM because, up to now, very few MCDM models deal with the uncertainty about the alternatives. Although there is a formal analogy between multicriteria analysis and decision making under uncertainty (see Section 15.2.3), the subject has not been thoroughly studied. This is a prominent issue because, in most

real situations, alternatives and results are uncertain. Another weakness of MCDSSs is the oversimplification of the alternatives they are able to cope with [114]. In order to meet user requirements, we urgently need to design better MCDM tools supporting the decision maker in look-ahead and scenario management. This is the reason why new models of uncertainty promoted by AI researchers seem of a particular interest.

Finally, we restrict ourselves to the exchange from AI to MCDM whereas it would be possible to follow the invert current. Up to now, AI researchers have largely ignored MCDM issues. However, in heuristic search, (Section 15.3.1) a multicriteria evaluation of the current state would be more informative than a scalar-valued function focusing on a single aspect. In fact, a scalar-valued evaluation function may be regarded as an aggregation function implicitly synthesizing multiple aspects. It would be wise to make the aggregation process explicit by introducing a vector valued evaluation function (see [85, 136] and [158], for first steps in this direction).

Other AI fields could also take advantage of MCDM. On the one hand, in planning, AI systems generally stumble on task partial orderings. Therefore, various way of generating a total ranking from these partial orderings should be worth studying. On the other hand, autonomous agents in distributed AI often make decision by means of very simple one-dimensional decision functions. Using MCDM procedures would produce more sophisticated behaviors.

In less than twenty years, AI and decision making (particularly MCDM) have woven many relationships. These links are rather asymmetric, because whereas MCDM has borrowed many ideas and tools issued from AI, relatively few exchanges followed the opposite direction. MCDM people still have to continue their work to popularize their domain...

REFERENCES

[1] Angehrn A. "TRIPLE C, Overview and how to use it", User's Guide, INSEAD, 1989.

[2] Angehrn A., Dutta S., 1992, "Integrating case-based reasoning in multi-criteria decision support systems". In: T. Jelassi, M. R. Klein and W.M. Mayon-White (eds.): *Decision Support Systems Experiences and Expectations*, North-Holland, pp.133-150.

[3] Balestra G., Tsoukiàs A., 1990, "Multicriteria analysis represented by artificial intelligence techniques", Journal of the Operational Research Society 41, 5, pp.419-430.

[4] Barthélemy J.P., Mullet E.: "Choice basis: a model for multi-attribute preference", British Journal of Mathematical and Statistical Psychology, 39, 1986, pp.106-124.

[5] Barthélemy, J.P., Mullet E.: "Choice basis: a model for multi-attribute preferences: some further evidence". In: E.E. Roskam (ed): *Mathematical Psychology in Progress*, Springer, 1989, pp.179-196.

[6] Barthélemy J.P., Mullet E.: "A model of selection by aspects", Acta Psychologica, 79, 1992, pp.1-19.

[7] Bayad R., Pomerol J.-Ch., 1992, "An intelligent DSS for the reinforcement of urban electrical power networks", in Decision Support Systems : Experiences and Expectations, T. Jelassi, M. R. Klein and W.M. Mayon-White, Eds., Noth-Holland, pp.153-165.

[8] Belman R.E., Zadeh L.A., Decision Making in a Fuzzy-Environment, Management Science, 17, 1970, B141-B164.

[9] Belnap N.D.: "How a computer should think", proceedings of the Oxford International Symposium on Contemporary Aspects of Philosophy, Oxford, England, 1976, pp. 30-56.

[10] Belnap N.D.: "A useful four-valued logic". In.: G. Epstein, J. Dunn (eds): *Modern uses of multiple valued logics*, D. Reidel, Dordrecht, 1977, pp. 8-37.

[11] Benayoun R., de Mongolfier J., Tergny J., Larichev O.I.:, "Linear programming with multiple objective functions: Step method STEM", Mathematical Programming 1, 1971, pp.366-375.

[12] Benferhat S., Dubois D., Prade H.: "From Semantic to Syntactic Approaches to Information Combination in Possibilistic Logic". In: B. Bouchon-Meunier (ed): *Aggregation and Fusion of imperfect Information*, Physica-Verlag, 1998, pp. 141-161.

[13] Beurrier A., Demota Ph., Lévine P. and Pomerol J.-Ch.: "Distribution of empty mailcars by an expert system: a case study with comparison with O.R. approaches", Journal of the Operation al Research Society, 4, 1990, pp. 527-537.

[14] Bezdek J., Spilman B., Spilman R.: "A fuzzy relation space for group decision theory", Fuzzy Sets and Systems, 2, 1978, pp. 197-212.

[15] Bohanec M., Bratko I., Rajkovic V.: "An expert system for decision making". In: H.G. Sol, (ed): *Processes and Tools for Decision Support, North Holland*, 1983, pp.235-248.

[16] Bohanec M., Urh B., Rajkovic V.:, "Evaluating options by combined qualitative and quantitative methods", 13th Conference on Subjective Probability, Utility and Decision Making, 1991, Fribourg.

[17] Bohanec M., Rajkovic V.: "DEX : an expert system shell for decision support", Sistemical n° 1, 1990, pp.145-158.

[18] Bondareva O.N.: "Revealed fuzzy preferences". In: J. Kacprzyk and M. Fedrizzi, (eds): *Multiperson Decision Making Using Fuzzy Sets and Possibility Theory*, Kluwer Academic Publishers., 1990, pp. 71-79.

[19] Boscarino M., Gianoglio G. and Ostanello. A.: "A Neural Network Design to Support Decision Making with a Complex Multicriteria Problem", Journal of Decision systems, 2, 1993, pp.149-171.

[20] Bouchon B.: "Preference defined from fuzzy questions". In: J. Kacprzyk and S.A. Orlovski (eds): *Optimization Models Using Fuzzy Sets and Possibility Theory*, D. Reidel Publ. Co., Dordrecht, Boston , Lancaster, Tokyo, 1987, pp.110-122.

[21] Bouyssou D.: "Problèmes de construction de critères", Cahier du LAMSADE n°89, Université de Paris-Dauphine, 1989, Paris.

[22] Brans J.P., Vincke Ph.: "A preference ranking organization method", Management Science, 31, 6, 1984, pp. 647-656.

[23] Burrell Ph., Duan Y. and Boskovic A.: "The need for hybrid decision systems in complex marketing domains", Journal of Decision Systems, 4, 4, pp.319-337.

[24] Carpenter G.A, Crossberg S. (1987): "A massively parallel architecture for a self-organizing neural pattern recognition machine", Computer Vision, Graphics, and Image Processing, 37, pp.54-115.

[25] Chateauneuf, A., Jaffray J.Y., "Some characterizations of lower Probabilities and other Monotone capacities through the use of Möbius inversion", Mathematical social sciences, 1989, pp. 263-283.

[26] Courbon J.C., Dubois D., Roy B., "Autour de l'aide à la décision et de l'intelligence artificielle", Rapport LAFORIA/IBP 94/01, 1994.

[27] Doignon J.P., Monjardet B., Roubens M., Vincke Ph.: "Biorders families, valued relations and preference modelling", Journal of Mathematical Psychology, 30, 4, 1986, pp. 435-480.

[28] Doherty P., Driankov D., Tsoukiàs A.: "Partial Logics and Partial Preferences", Proceedings of the CEMIT 92 international conference, Tokyo, 1992, pp.525-528.

[29] Dombi J.: "Basic concepts for a theory of evaluation : the aggregative operator", European Journal of Operational Research, 10, 1982, pp. 282-293.

[30] Doyle J.: "A truth maintenance system", Artificial Intelligence 12, 1979, pp. 231-272.

[31] Doyle J. and Wellman M.P.: "Representing preferences as ceteris paribus comparatives", In: S. Hanks, S. Russell and M.P. Wellman (eds), *AAAI Spring Symposium notes*, 1994, pp.69-75.

[32] Dubarle D.: "Essai sur la généralisaton naturelle de la logique usuelle", manuscript (1963) published posthumously in *Mathématique, informatique et sciences humaines*, 107, 1989, pp.17-73.

[33] Du Bois Ph., Brans J.P., Cantraine F., Mareschal B.: "Medicis : an expert system for computer-aided diagnosis using the PROMETHEE multicriteria method", European Journal of Operational Research 39, 1989, pp.284-292.

[34] Dubois D., Fargier H., Prade H.: "Decision Making under Ordinal Preferences and Comparative Uncertainty". In: D. Geiger, P. Shenoy (eds), *Uncertainty in Artificial Intelligence*, Morgan & Kaufmann, San Francisco, 1997, pp. 157-164.

[35] Dubois, D., Grabisch, M., Modave, F. and Prade H.: "Un parallèle entre décision dans l'incerain et décision multicritère dans une perspective IA", in RFIA-98 proceedings, vol II, 1998, pp. 225-234.

[36] Dubois D., Prade H.: "Weighted minimum and maximum operations in fuzzy set theory", Information Science, 39, 1986, pp. 205-210.

[37] Dubois D., Prade H.: Possibility Theory - An Approach to the Computer Processing Uncertainty, Plenum Press, New-york, 1988.

[38] Dubois D., Prade H.: "Aggregation of possibility measures", In: J. Kacprzyk and M. Fedrizzi, (eds): *Multiperson Decision Making Using Fuzzy Sets and Possibility Theory*, Kluwer Academic Publishers, Dordrecht, 1990, pp. 55-63.

[39] Dubois D., Prade H.: "Possibility theory as a Basis for qualitative Decision Theory", In: Proceedings of IJCAI-95, 1995, pp.1924-1930.

[40] Efstathiou J., Rajkovic V., Bohanec M., "Expert systems and rules based decision support systems". In G. Mitra (ed): Computer Assisted Decision Making, Elsevier Science Publishers, 1986, pp.165-174.

[41] Espinasse B., Picolet G., Chouraqui E., "Negotiation support systems: a multicriteria and multi-agent approach", European Journal of Operational Research 103, 1997, pp.389-409.

[42] Fargier H., Perny P.: "Axioms for Qualitative Decision Theory". In: proceedings of the ECAI'98 Workshop "Decision Theory meets artificial Intelligence" , 1998, pp.77-93.

[43] Farquhar P.H., "Utility assessment methods", Management Science 30 n° 11, 1984, pp.1283-1300.

[44] Fiala P., "Problem solving methods in multicriteria analysis", Diskussionbeitrag Fernuniversität Hagen n° 181, 1991.

[45] Fiala P., "Artificial intelligence methods in multicriteria analysis". In Cerny M., Glückaufova D., Loula D. Eds.,Multicriteria Decision Making, Methods, Algorithms, Applications, Czechoslovak Academy of Sciences, Prague, 1992, pp.58-65.

[46] Fodor J., Calvo T.: "Aggregation functions defined by t-norms and t-conorms", In: B. Bouchon-Meunier (ed): *Aggregation and Fusion of imperfect Information*, Physica-Verlag, 1998, pp. 36-48.

[47] Fodor J., Roubens M.: *Fuzzy Preference Modelling and Multicriteria Decision support*, Kluwer Academic Publishers, 1994.

[48] Fodor J., Orlovski S., Perny P., Roubens M.: "The use of fuzzy preference models in multiple criteria choice, ranking and sorting". In. R. Slowinski (ed): *Handbook of Fuzzy Sets in Decision Analysis, Operations Research and Statistics*, Kluwer Academic Publishers, 1998, pp. 69-101.

[49] Fodor J., Perny P., Roubens M.: "Decision Making and Optimization", In: E. Ruspini, P. Bonissone and W. Pedrycz (eds): *Handbook of Fuzzy Computations*, Chapter 5, to appear.

[50] Främling K., Graillot D., Bucha J.: "Uncertainty handling in a waste placement decision support system", in 4th International Conference on Database and Expert Systems Applications, Prague, 1993.

[51] Fraser N.M., Hipel K.W., "Conflict analysis as negotiation support systems", in B.R. Munier and M.F. Shakun Eds., Compromise Negotiation and Group Decision, Reidel, 1988, pp.225-243.

[52] Gandibleux X., Rosenthal-Sabroux C. and Libert G., "Multicriteria approach for intelligent decision support in supervisory control", Annals of Operations Research 51, 1994, pp.327-348.

[53] Gardiner L.R. and Vanderpooten D.: "Interactive Multiple Criteria Procedures: Some Reflections". In: Joao Climaco (ed): *Multicriteria Analysis, Proceedings of the XIth conference on MCDM*, 1997, pp.290-301.

[54] Grabish M., Nguyen H.T., Walker E.A.: "Fundamentals of uncertainty calculi, with application to fuzzy inference", Kluwer Academic, 1995.

[55] Grauer M., Lewandowski A. and Wierzbicki A.: "DIDASS, Theory, Implementation and Experiences", In: M. Grauer and A. Wierzbicki (eds): *in Interactive Decision Analysis - Lecture note in Economics and Mathematical Systems* 229, Springer, 1984, pp. 22-30.

[56] Honda N., Mimaki T., "Multiobjective decision method using heuristic rules", in Artificial Intelligence in Economics and Management, L.F. Pau Ed., 1986, pp. 157-165.

[57] Hong H.B., Vogel D.R.: "Data and Model Management in a generalized MCDM-DSS", Decision Sciences, 22, 1991, pp.1-25.

[58] Jaffray J.Y.: "Linear Utility theory for Belief functions", Op. Res. Letters, 8, 1989, pp.107-112.

[59] Jaffray J.Y.: "Dynamic Decision Making with belief functions". In: R.R Yager, J. Kacprzyk , M. Fedrizzi (eds*): Advances in the Demster-Shafer theory of Evidence*, John Wiley & Sons, 1994.

[60] Jarke M., Jelassi M.T., Shakun M.F.: "MEDIATOR: toward a negotiation support system", European Journal of Operational Research 31, 1987, pp.314-334.

[61] Jelassi T., Jarke M., Stohr E.A.: "Designing a generalized multiple criteria decision support system", Journal of Management Information Systems 1/4, 1985, pp.24-43.

[62] Jelassi T.: "MCDM: from "stand-alone" methods to integrated and intelligent DSS". In: Y. Sawaragi, K. Inone and H. Nakayama (eds): Toward Interactive and Intelligent Decision Support Systems, Lecture Notes in Economics and Mathematical Systems 286, 1986, pp.90-99.

[63] Kant J.D.: CATEG_ART: "A neural network for automatic extraction of human categorization rules", in proceedings of the International Conference on Artificial Neural Networks, October 9-13, 1995, Paris, pp.479-484.

[64] Kant J.D., Levine D.S.: "A Neural Network for Decision Making Under the Influence of reinforcement", proceedings of the International Conference on Artificial Neural Networks, 1997, Houston.

[65] Karacapilidis N.I., Pappis C.P.: "A framework for group decision support systems: combining AI tools and OR techniques", European Journal of Operational Research 103, 1997, pp.373-388.

[66] Keeney R.L.: "Value-driven expert systems for decision support", Decision Support Systems 4 n° 4, 1988, pp.405-412.

[67] Keeney R.L.: Value-focused thinking, Harvard University Press, Cambridge, MA, 1992.

[68] Kersten G.E., Matwin S., Michalowski W., Szpakowicz S.: "Representating the negotiation process with a rule-based formalism", Theory and Decision 25, 1988, pp.225-257.

[69] Klein D.A. Shortliffe E.H.: "A framework for explaining decision-theoretic advice", Artificial Intelligence 67, 1994, pp.201-243.

[70] Knoll A.L., Engelberg A.: "Weighting multiple objectives. The Churchman-Ackoff technique revisited", Computer and Operation Research vol. 5, 1978, pp.165-177.

[71] Korhonen P. and Wallenius J.: "A Pareto race", Naval Research Logisitics, 35, 1988, pp.165-175.

[72] Krantz D.H., Luce R.D., Suppes P., Tversky A.: Foundation of measurement, Academic, New York, vol. I, 1971.

[73] Kuzmin V.B., Ovchinnikov S.V.: "Group decision I in arbitrary spaces of fuzzy binary relation", Fuzzy Sets and Systems, 4, 1980a, pp. 53-62.

[74] Kuzmin V.B., Ovchinnikov S.V.: "Design of Group decision II in arbitrary spaces of partial ordered fuzzy relations", Fuzzy Sets and Systems, 4, 1980b, pp. 153-165.

[75] Labat J.M., Futtersack M.: "In SIAD multicritère pour la planification des aides aux chômeurs de longue durée", Decision Systems Journal 1 n° 1, 1992, pp.79-92.

[76] Lehmann D.: "Generalized Qualitative Probability: Savage revisited", conference of uncertainty in AI, 1996, pp. 381-388.

[77] Leven S.J., Levine D.S.: "Multiattribute Decision Making in Context: A Dynamic Neural Network Methodology", Cognitive Science, 20, 1996, 271-299.

[78] Lévine P., Pomerol J.-Ch.: "An interactive program for the discrete multiobjective problem", Methods of Operation Research 45, 1983, pp.479-480.

[79] Lévine P., Pomerol J.-Ch.: "PRIAM, an interactive program for choosing among multiple attribute alternatives", European Journal of Operational Research, Vol. 25, n°2, 1986, pp.272-280.

[80] Lévine P., Pomerol J.-Ch.: "Knowledge representation by schemata in financial expert systems", Theory and Decision 27 n°1/2, 1989a, pp.147-161.

[81] Lévine P., Pomerol J.-Ch.: Systèmes interactifs d'aide à la décision et systèmes experts, 1989b, Hermès, Paris.

[82] Lévine P. and Pomerol J.-Ch. "Negotiation support systems: an overview and some knowledge-based examples", In: R. Avenhaus, H. Karkar and M. Rudnianski (eds): *Defense Decision Making*, Springer, 1991, pp.241-256.

[83] Lévine P., Pomerol J.-Ch., "The role of the decision maker in DSSs and representation levels", Proceedings of the 28th Hawaï International Conference on System Sciences, J.F. Nunamaker and R.H. Sprague Eds., IEEE, vol. 3, 1995, pp.42-51.

[84] Lévine P., Pomerol M.-J. and Saneh R.: "Rules integrate data in a multicriteria decision system", IEEE Trans. on Systems Man and Cybernetics 20 n° 3, pp.678-686, 1990.

[85] Lirov Y.:"Algorithmic multi-objective heuristics construction in the A* search", Decision Support System 7, 1991, pp.159-167.

[86] Lotfi V., Zionts S.: "AIM, aspiration-level interactive method for multiple criteria decision making", User's Guide, State University of New York at Buffalo, 1990.

[87] Lotfi V., Stewart T.J., Zionts S.: "An aspiration-level interactive model for multiple criteria decision making", Computer and Operations Research, 19, 7, 1992, pp.671-681.

[88] Malakooti B., Zhou Y.Q., "Feedforward artificial neural networks for solving discrete MCDM problems", Management Science 40, 1992, pp.1542-1561.

[89] Mandic N.J., Mamdani E.H.: "A multiple attribute decision making model with fuzzy rule-based modification of priorities". In: E.J. Zimmerman, L.A. Zadeh and B.R. Gaines (eds): *Fuzzy Sets and Decision Analysis*, 1984, pp.285-306.

[90] Monnet J.-M., Lagrange J.-Ph., Pomerol J.-Ch.: "A formal framework for decision making in case-based reasonning", IPMU, Paris , Juillet 1998.

[91] Newell A., Simon H.A.: Human problem solving, Prentice Hall, Englewood Cliffs, 1972.

[92] von Neumann J., Morgenstern O: *Theory of games and economic behaviour*, Princeton University Press, Princeton, New-Jersey, 1944.

[93] Nilsson N.: Principles of Artificial Intelligence, 1982, Springer Verlag.

[94] O'Leary D.E.: "Multiple criteria decision making in accounting expert systems", Comptes rendus des 6ièmes journées internationales sur les systèmes experts et leurs applications, Avignon, 1986, pp.1017-1035.

[95] Olson D.L., Mechitov A.I. and Moshkovich H.M. : "Cognition Effort and Learning Features of Decision Aids : Review of Experiments", Journal of Decision Systems, 7, n°2, 1998, in Press.

[96] Orlovski S.: "Decision making with a fuzzy preference relation", Fuzzy Sets and Systems, 1, 1978, pp. 155-167.

[97] Ovchinnikov S.: "Preference and choice in a fuzzy environment". In: J; Kacprzyk an S.A. Orlovski (eds): *Optimization Models Using Fuzzy Sets and Possibility Theory*, D. Reidel Publ. Co, Dordrecht, Boston, Lancaster, Tokyo, 1987, pp. 91-98.

[98] Ovchinnikov S., Ozernoy V.: "Using fuzzy binary relations for identifying noninferior decision alternatives", Fuzzy sets and systems, 25, 1988, pp.21-32.

[99] Ovchinnikov S., Roubens M. : "On fuzzy strict preference, indifference and incomparability relations", Fuzzy Sets and Systems, 47, 1992, pp. 313-318.

[100] Ozernoy V.M.: "Using expert system technology for multiple criteria decision making : a survey", Proceedings of the tenth conference on Multicriteria Decision Making, 1992, pp.271-280.

[101] Pasche C.: "EXTRA : an expert system for multicriteria decision making", European Journal of Operational Research 52, 1991, pp.224-234.

[102] Perny P.: "Defining fuzzy covering relations for decision aid". In: B. Bouchon, R. Yager, L. Zadeh (eds): *Fuzzy Logic and soft Computing, Advances in Fuzzy Systems - Applications and Theory*, 4, 1995, pp. 109-118.

[103] Perny P.: "Multicriteria filtering methods based on concordance and non-discordance principles", Annals of Operations Research, 80, 1998, pp. 137-165.

[104] Perny P., Roy B.: "The use of fuzzy Outranking Relations in Preference Modelling", Fuzzy sets and Systems, 49, 1992, 33-53.

[105] Perny P., Roubens M.: "Preference modelling". In: R. Slowinski (ed): *Handbook of Fuzzy Sets in Decision Analysis, Operations Research and Statistics*, Kluwer Academic Publishers, 1998, pp. 3-30.

[106] Perny P., Tsoukiàs A.: "On the continuous extension of four valued logic for preference modelling", Proceedings of the 7th IPMU conference, Paris, 1998, pp. 302-309.

[107] Perny P., Vanderpooten D.: "An interactive multiobjective procedure for selecting medium-term countermeasures after nuclear accidents", Journal of Multicriteria Decision Analysis 7, 1998, pp.48-60.

[108] Pirlot M., Vincke Ph.: *Semiorders: properties, representations, applications*, Kluwer Academic Publishers, Dordrecht, Boston, London, 1997.

[109] Petrovic S.: "Case-based reasoning in the DSS for multicriteria analysis", Preprint presented at EURO Conference on DSS, Bruges, 1996.

[110] Pomerol J.-Ch.: "Editorial", Decision Systems Journal 1 n° 1, 1992, pp.11-13.

[111] Pomerol J.-Ch.: "Multicriteria DSSs: State of the art and problems", Central European Journal for Operations Research and Economics, 3, 1993, pp. 197-211.

[112] Pomerol J.-Ch.: " Scenario development and practical decision making under uncertainty: robustness and "risk-control"", Proceedings CESA-98, IEEE, 2, 1998, pp.238-242.

[113] Pomerol J.-Ch., Barba-Romero S.: Choix multicritère dans l'entreprise, Hermès, 1993, Paris. English translation to appear, Kluwer, 1999.

[114] Pomerol J.-Ch and Brézillon P.: "Organisational experiences with multicriteria decision support systems: problems and issues", proceedings of the 30th Hawaii international conference on system sciences, IEEE pub., vol.3, 1997.

[115] Pomerol J.-Ch., Roy B., Rosenthal-Sabroux C., Saad A.: "An "Intelligent" DSS for the muticriteria evaluation of railway timetables", Foundations of Computing and Decision Sciences 20, 1995, pp.219-238.

[116] Pomerol J.-Ch., Trabelsi T.: "An adaptation of PRIAM to continuous multiobjective linear programming", European Journal of Operational Research 31, 1987, pp.335-341.

[117] Quinlan J.R.: "Discovering rules by induction from large collections of examples". In: Michie (ed): *Expert systems in the microelectronic age*, Edinburgh University Press, 1979.

[118] Rajkovic V., Bohanec M., Batagelj V.: "Knowledge engineering techniques for utility identification", Acta Psychologica 68, 1988, pp.271-286.

[119] Reyes J., Barba-Romero S.: "Expert systems and multicriteria decision making", presented at EURO VIII, 1986, Lisbon.

[120] Rommel Y.: "Une procédure interactive reposant sur les techniques d'apprentissage symbolique : bibliographie sur l'apprentissage symbolique par induction à partir d'exemples et présentation de la procédure interactive PLEXIGRAS", Cahier LAMSADE n°67, Université de Paris Dauphine, Paris, 1991.

[121] Roubens M., Vincke Ph.: "Fuzzy preferences in an optimization perspective". In: J; Kacprzyk an S.A. Orlovski (eds): *Optimization Models Using Fuzzy Sets and*

Possibility Theory, D. Reidel Publ. Co, Dordrecht, Boston, Lancaster, Tokyo, 1987, pp.77-90.

[122] Roubens M., Vincke Ph.: "Fuzzy possibility graphs and their application to ranking fuzzy numbers". In: "Non-conventional Preference Relations in Decision Making, J. Kacprzyck and M. Roubens (eds), 1988, pp. 119-128.

[123] Roy B.: "Classement et choix en présence de points de vue multiples, la méthode ELECTRE", RIRO vol.2 n°8, 1968, pp.57-75.

[124] Roy B.: "Management scientifique et aide à la décision", Direction Scientifique de la SEMA rapport n°86, Paris, 1974.

[125] Roy B.: "ELECTRE III: un algorithme de classement fondé sur une représentation floue des préférences en présence de critères multiples", Cahiers du CERO, 20, 1, 1978, pp.3-24.

[126] Roy B.: "Meaming and validity of interactive procedures as tools for decision making", European Journal of Operational Research 31, pp.297-303, 1987.

[127] Roy B.: *Multicriteria Methodology for Decision Aiding*, Kluwer Academic Publishers, 1996.

[128] Roy B., Bouyssou D.: Aide multicritère à la décision: méthodes et cas, 1993, Economica, Paris.

[129] Saaty T.L.: The analytic hierarchy process, 1980, New-York, Mc Graw-Hill.

[130] Saaty T.L.: "How to make a decision: the analytic hierarchy process", Interfaces 24, 1994, pp.19-43.

[131] Savage L.J.: "The foundations of Statistics", Dover, 2nd edition, 1972.

[132] Shafer G.: *A Mathematical Theory of Evidence*, Princeton University Press, 1976.

[133] Simon H.A.: The new science of management decision, revised edition, 1977, Prentice-Hall, Englewood Cliffs.

[134] Skalka J.M., Bouyssou D., Bernabeu Y.A.: "ELECTRE III et IV, aspects méthodologiques et guide d'utilisation", Cahier du LAMSADE n°25, Université de Paris Dauphine, 1983, Paris.

[135] Smets Ph.: "Constructing the pingnistic probability function in a context of uncertainty", In: M. Henrion, R. Shachter, L.N. Kanal and J.F. Lemmer (eds): Uncertainty in AI, North Holland, 1990, pp.29-39.

[136] Stewart B.S. and White C.C.: "Multiobjective A*", Journal of the Association for computinq Machinery, 38, 1991, pp.775-814.

[137] Suppes P., Krantz D.H., Luce R.D., Tversky A., Foundations of measurement, Academic, San Diego, 1989, vol. II.

[138] Tan S.W. and Pearl J.: "Qualitative Decision Theory", in Proceedings of AAAI-94, 1994, pp.928-933.

[139] Teghem J., Delhaye C., Kunsh P.L.: "An interactive decision support system (IDSS) for MCDA", Math. Comput. Model., vol. 12, n°10/11, 1989, pp.1311-1320.

[140] Tsoukiàs A.: "Preference modeling as a reasoning process: a new way to face uncertainty in Multiple Criteria Decision Support Systems", EJOR, 55, 1991, pp.309-318.

[141] Tsoukiàs A.: "A first Order, four valued, weakly paraconsistent logic", Cahier du Lamsade No 139, Paris Dauphine University, 1996.

[142] Tsoukiàs A., Vincke Ph.: "A new axiomatic foundation of partial comparability theory", Theory and Decision, 39, 1995, pp. 79-114.

[143] Tsoukiàs A., Vincke Ph.: "Double threshold orders: a new axiomatization", Report SMG IS 3/96, Institut de statistiques, Université de Bruxelles, 1996.

[144] Tsoukiàs A., Vincke Ph.: "Extended preference structures in MCDA". In: J. Climaco (ed): Multicriteria Analysis, Springer Verlag, Berlin, 1997, pp. 37-50.

[145] Tversky A.: "Elimination by aspects", Psychological Review 76, 1972a, pp.31-48.
[146] Tversky A.: "Choice by elimination", Journal of Math. Psycology 4, 1972b, pp.175-201.
[147] Ulengin F., Topçu I.: "Cognitive map: KBDSS integration in transportation planning", Journal of the Operational Research Society 48, 1997, pp.1065-1075.
[148] Usher M., Zakay D.: "A Neural Network Model for Attribute-Based Decision Processes", Cognitive Sciences, 17, 1993, pp. 349-396.
[149] Vanderpooten D.: "The interactive approach in MCDA: a technical framework and some basic conceptions", Mathematical Computing Modelling 12, 1989, pp.1213-1220.
[150] Vanderpooten D.: "The construction of prescriptions in outranking methods". In: C.A. Bana e Costa (ed): *Reading in MCDM*, Springer, 1990a, pp.184-215.
[151] Vanderpooten D.: "L'approche interactive dans l'aide multicritère à la décision, Thèse de l'Université de Paris-Dauphine, 1990b, Paris.
[152] Valverde L.: "On the structure of F-indistinguishability operators, Fuzzy sets and systems, 17, 1985, pp. 313-318.
[153] Van de Walle B., De Baets B., Kerre E.: "Characterizable fuzzy preference structures", Annals of Operations Research, to appear.
[154] Vasko F. J., Gegger, M.L., Lutsko L.G., Newhart D.D. and Renn D.J.: "A Multiple-Attribute Based Expert System for Customer/Order Evaluation", Journal of Operational Research Society, 41, 11, 1990, pp.1077-1080.
[155] Vetschera R.: "Sensitivity analysis for the ELECTRE multicriteria method", Zeitschrift, Operations Research, 30, 1986, pp. B99-B117.
[156] Wallenius J.: "Comparative evaluation of some interactive approaches to multicriterion optimization", Management Science 21, 1978, pp.1387-1396.
[157] White III C.C.: "A survey on the integration of decision analysis and expert systems for decision support", IEEE Transactions on System Man and Cybernetics Vol. SMC 20, 1990, pp.358-364.
[158] White III C.C., Stewart B.S. Carraway R.L.: "Multiobjective, preference-based search in acyclic OR-graphs", European Journal of Operational Research 56, 1992, pp.357-363.
[159] Wierzbicki A.: "Interactive decision analysis and interpretative computer intelligence". In: M. Grauer and A. Wierzbicki (eds): *Interactive Decision Analysis.* Lecture Notes in Economic and Mathematical Systems 229, Springer, 1984, pp. 2-19.
[160] Yager R.: "On Ordered Weighted Averaging Aggregation Operators in Multicriteria Decision Making", IEEE Trans. on Syst., Man and Cybernetics, 18, 1988, pp.183-190.

16 EVOLUTIONARY ALGORITHMS AND SIMULATED ANNEALING FOR MCDM

Andrew J. Chipperfield, James F. Whidborne
and Peter J. Fleming

CONTENTS

T. Gal et al. (eds.), *Multicriteria Decision Making*

Abstract: This chapter describes two stochastic search and optimization techniques, evolutionary algorithms and simulated annealing, both inspired by models of natural processes (evolution and thermodynamics) and considers their role and application in multiple criteria decision making and analysis. The basic single criteria algorithms are first presented in each case and it is then demonstrated with an example problem how these may be modified and set up to deal with multiple design criteria. Whilst the example employed considers the design of a robust control system for a high speed maglev vehicle, the approaches and techniques have a far wider range of application.

16.1 INTRODUCTION

Real-world problems will usually involve the satisfaction of multiple performance measures, or objectives, which should be solved simultaneously. In some situations, the objective functions may be solved in isolation from one another and an insight concerning the best that may be obtained in each performance domain obtained. However, when considering the overall problem, suitable solutions will seldom be found in this way. An optimal performance in one objective domain will often imply an unacceptably low performance in one or more of the remaining objectives necessitating the need for some sort of compromise solution to be reached. Suitable solutions to problems posing such conflicts should offer an "acceptable", though possibly sub-optimal in a single objective sense, performance across all the objectives. In this case "acceptable" is then a problem-dependent and subjective notion.

The simultaneous solution of multiple, possibly competing, objective functions is unlikely to yield a single utopian solution. Instead, the solution of a multiobjective optimization (MO) problem is a set of Pareto-optimal solutions which in most practical situations is likely to be very large. Subsequently there is a difficulty in representing the set of Pareto-optimal solutions and in choosing a suitable solution from this set when there is no information regarding the relative performance of each objective. The size of the solution set can, however, be reduced by including a set of objective function goals which must also be satisfied.

Various non-linear programming methods have been developed to solve the MO problem (see, for example, [3]). However, this is not a trivial task as practical problems are generally non-convex, multimodal and frequently non-smooth or exhibit discontinuities. These traditional approaches use deterministic transition rules, generally to implement a form of hill climbing, and as such can only be expected to work well if the problem is small and has few local minima, i.e. distinct regions in decision variable space that yield Pareto-optimal solutions. Additionally, they will require a good estimate of the solution if they are not to converge to some local, sub-optimal solution. For larger, more realistic prob-

lems or ones that may have many local minima, algorithms with probabilistic transition rules offer greater potential for success.

This chapter introduces two such approaches, evolutionary algorithms and simulated annealing. After an introduction to the basic components of an evolutionary algorithm, the notion of a decision maker is presented and used as the basis for constructing a multiobjective genetic algorithm (MOGA). An example from robust control system design, the electro-magnetic suspension of a high speed maglev vehicle, is then used to illustrate the approach. An alternative stochastic approach to solving multiple criteria optimization problems, simulated annealing, is then described and demonstrated on a similar controller design problem.

16.2 EVOLUTIONARY ALGORITHMS

Evolutionary algorithms are based on computational models of fundamental evolutionary processes such as selection, reproduction and mutation, as shown in Fig. 16.1. Individuals, or current approximations, are encoded as strings composed over some alphabet(s), e.g. binary, integer, real-valued etc., and an initial population of chromosomes, `Chrom` in Fig. 16.1, is produced by randomly sampling these strings. Once a population has been produced it may be evaluated using an objective function or functions that characterize an individual's performance in the problem domain. Where the encoding of chromosomes uses a mapping from the decision variables to some other alphabet, e.g. real-values encoded as binary strings, it will be necessary to decode the chromosomes before the objective function may be evaluated and a cost vector, `Cost`, assigned to the population. The objective function(s) is also used as the basis for selection and determines how well an individual performs in its environment. A fitness value is then derived from the raw performance measure given by the objective function(s) and is used to bias the selection process towards promising areas of the search space. Highly fit individuals will be assigned a higher probability of being selected for reproduction than individuals with a lower fitness value. Therefore, the average performance of individuals can be expected to increase as the fitter individuals are more likely to be selected for reproduction and the lower fitness individuals get discarded. Note that individuals may be selected more than once at any generation (iteration) of the EA and that the temporary vector of selected individuals, `Sel`, may therefore contain more than one copy of any individual in the original population.

Selected individuals are then reproduced, usually in pairs, through the application of genetic operators and these new individuals may then overwrite their parents in the vector `Sel`. These operators are applied to pairs of individuals with a given probability and result in new offspring that contain material

exchanged from their parents. The offspring from reproduction are then further perturbed by mutation. These new individuals then make up the next generation, **Chrom**. These processes of selection, reproduction and evaluation are then repeated until some termination criteria are satisfied, e.g. a certain number of generations, **Gen**, completed, a mean deviation in the performance of individuals in the population or when a particular point in the search space is reached.

```
procedure EA {
    initialize(Chrom);
    while not finished do {
        Cost = objv_fun(decode(Chrom));
        Sel = select(Chrom, Cost);
        Sel = reproduce(Sel);
        Chrom = mutate(Sel);
        Gen = Gen + 1;
     }
}
```

Figure 16.1 An evolutionary algorithm.

Although similar at the highest level, many variations exist in EAs. A comprehensive discussion of the differences between the various EAs can be found in [22].

16.3 MULTIOBJECTIVE OPTIMIZATION AND DECISION MAKING

The use of multiobjective optimization (MO) recognizes that most practical problems require a number of design criteria to be satisfied simultaneously, viz:

$$\min_{\mathbf{p}\in\Omega} \mathbf{F}(\mathbf{p}) \tag{16.1}$$

where $\mathbf{p} = [p_1, p_2, \ldots, p_q]$ and Ω define the set of free variables, $\mathbf{p}$, subject to any constraints and $\mathbf{F}(\mathbf{p}) = [f_1(\mathbf{p}), f_2(\mathbf{p}), \ldots, f_n(\mathbf{p})]$ are the design objectives to be minimized.

Clearly, for this set of functions, $\mathbf{F}(\mathbf{p})$, it can be seen that there is no one ideal 'optimal' solution, rather a set of Pareto-optimal solutions for which an improvement in one of the design objectives will lead to a degradation in one or more of the remaining objectives. Such solutions are also known as non-inferior or non-dominated solutions to the multiobjective optimization problem.

Generally, members of the Pareto-optimal solution set are sought through solution of an appropriately formulated nonlinear programming problem. A number of approaches are currently employed including the ϵ-constraint, weighted sum and goal attainment methods [12]. However, such approaches require precise expression of a, usually not well understood, set of weights and goals. If the trade-off surface between the design objectives is to be better understood, repeated application of such methods will be necessary. In addition, nonlinear programming methods cannot handle multimodality and discontinuities in function space well and can thus only be expected to produce local solutions.

Evolutionary algorithms, on the other hand, do not require derivative information or a formal initial estimate of the solution region. Because of the stochastic nature of the search mechanism, genetic algorithms (GAs) are capable of searching the entire solution space with more likelihood of finding the global optimum than conventional optimization methods. Indeed, conventional methods usually require the objective function to be well behaved, whereas the generational nature of GAs can tolerate noisy, discontinuous and time-varying function evaluations [7]. Moreover, EAs allow the use of mixed decision variables (binary, n-ary and real-values) permitting a parameterization that matches the nature of the design problem more closely. Single objective GAs, however, do still require some combination of the design objectives although the relative importance of individual objectives may be changed during the course of the search process.

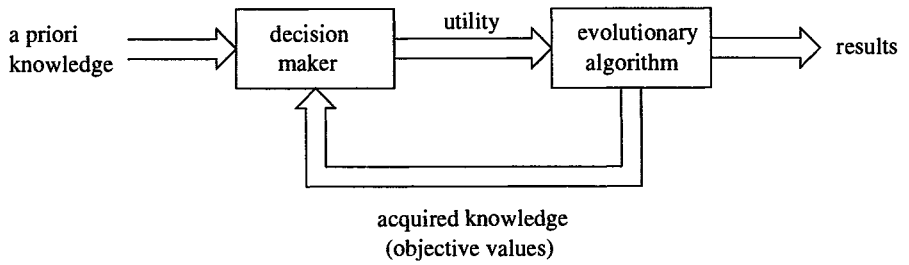


Figure 16.2 A general multiobjective evolutionary optimizer

A general view of multiobjective evolutionary optimization has been proposed by Fonseca and Fleming [6] and is illustrated in Fig. 16.2. The decision maker block represents a utility assignment strategy, which may be anything from a straight-forward weighted sum approach to an intelligent decision maker or human operator. The EA is employed to generate a set of candidate solu-

tions according to the utility level assigned by the decision maker to the current set of solution estimates. The decision maker thus influences the production of new solution estimates and as these are evaluated they provide new trade-off information which can be used by the decision maker to refine its current goals and preferences. The effect of any changes in the decision process, perhaps arising from taking recently acquired information into account, is seen by the EA as a change in environment. In the next section, a multiobjective decision making process, based on a Pareto-ranking approach, is described and a multiobjective evolutionary algorithm developed.

16.4 MULTIOBJECTIVE GENETIC ALGORITHMS

The notion of fitness of an individual solution estimate and the associated objective function value are closely related in the single objective GA described earlier. Indeed, the objective value is often reffered to as fitness although they are not, in fact, the same. The objective function characterizes the problem domain and cannot therefore be changed at will. Fitness, however, is an assigned measure of an individual's ability to reproduce and, as such, may be treated as an integral part of the GA search strategy.

As Fonseca and Fleming describe [6], this distinction becomes important when performance is measured as a vector of objective function values as the fitness must necessarily remain scalar. In such cases, the scalarization of the objective vector may be treated as a multicriterion decision making process over a finite number of candidates - the individuals in a population at a given generation. Individuals are therefore assigned a measure of utility depending on whether they perform better, worse or similar to others in the population and, possibly, by how much. The remainder of this section describes the main differences between the simple EA outlined earlier and MOGAs.

16.4.1 Decision Strategies

In the absence of any information regarding the relative importance of design objectives, Pareto-dominance is the only method of determining the relative performance of solution estimates. Non-dominated individuals are all therefore considered to be 'best' performers and are thus assigned the same fitness [7], e.g. zero. However, determining a fitness value for dominated individuals is a more subjective matter. The approach adopted here is to assign a cost proportional to how many individuals in a population dominate a given individual, Fig. 16.3. In this case, non-dominated individuals are all treated as desirable.

If goal and/or priority information is available for the design objectives then it may be possible to differentiate between some non-dominated solutions. For example, if degradations in individual objectives still allow those goals to be

satisfied whilst also allowing improvements in other objectives that do not already satisfy their design goals, then these degradations should be accepted. In cases where different levels of priority may be assigned to the objectives then, in general, it is only important to improve the high priority objectives, such as hard constraints, until the corresponding design goals are met, after which improvements may be sought in the lower priority objectives.

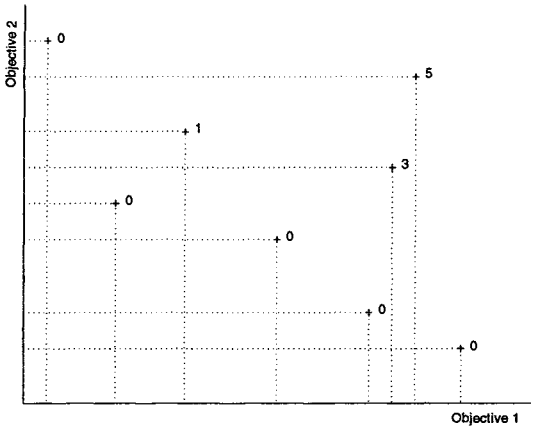


Figure 16.3 Pareto ranking

These considerations have been formalized in terms of a transitive relational operator, *preferability*, based on Pareto–dominance, which selectively excludes objectives according to their priority and whether or not the corresponding goals are met [5]. For simplicity only one level of priority is considered here. Consider two objective vectors $\mathbf{u}$ and $\mathbf{v}$ and the corresponding set of design goals, $\mathbf{g}$. Let the smile, $\overset{\mathbf{u}}{\smile}$ denote the components of $\mathbf{u}$ that meet their goals and the frown $\overset{\mathbf{u}}{\frown}$ those that do not. Assuming minimization, one may then write

$$\mathbf{u}^{\overset{\mathbf{u}}{\smile}} \leq \mathbf{g}^{\overset{\mathbf{u}}{\smile}} \quad \wedge \quad \mathbf{u}^{\overset{\mathbf{u}}{\frown}} > \mathbf{g}^{\overset{\mathbf{u}}{\frown}}, \tag{16.2}$$

where the inequalities apply component wise. This is equivalent to

$$\forall i \in \overset{\mathbf{u}}{\smile},\ u_i \leq g_i \quad \wedge \quad \forall i \in \overset{\mathbf{u}}{\frown},\ u_i > g_i, \tag{16.3}$$

where u_i and g_i represent the components of $\mathbf{u}$ and $\mathbf{g}$, respectively. Then, $\mathbf{u}$ is said to be preferable to $\mathbf{v}$ given $\mathbf{g}$ if and only if

$$(\mathbf{u}^{\overset{\mathbf{u}}{\frown}} \text{ p}< \mathbf{v}^{\overset{\mathbf{u}}{\frown}}) \vee \left\{(\mathbf{u}^{\overset{\mathbf{u}}{\frown}} = \mathbf{v}^{\overset{\mathbf{u}}{\frown}}) \wedge \left[(\mathbf{v}^{\overset{\mathbf{u}}{\smile}} \not\leq \mathbf{g}^{\overset{\mathbf{u}}{\smile}}) \vee (\mathbf{u}^{\overset{\mathbf{u}}{\smile}} \text{ p}< \mathbf{v}^{\overset{\mathbf{u}}{\smile}})\right]\right\}, \tag{16.4}$$

where $\text{p}<$ is a dominance operator such that $\mathbf{u}\ \text{p}< \mathbf{v}$ denotes that $\mathbf{u}$ dominates $\mathbf{v}$, i.e.

$$\forall i \in \{1,\ldots,n\}, u_i \leq v_i \wedge \exists i \in \{1,\ldots,n\} : u_i < v_i. \tag{16.5}$$

Hence $\mathbf{u}$ will be preferable to $\mathbf{v}$ if and only if one of the following is true:

1. The violating components of $\mathbf{u}$ dominate the corresponding components of $\mathbf{v}$.
2. The violating components of $\mathbf{u}$ are the same as the corresponding components in $\mathbf{v}$, but $\mathbf{v}$ violates at least one other goal.
3. The violating components of $\mathbf{u}$ are equal to the corresponding components of $\mathbf{v}$, but $\mathbf{u}$ dominates $\mathbf{v}$ as a whole.

16.4.2 Fitness Mapping and Selection

After a cost has been assigned to each individual, selection can take place in the usual way. Suitable schemes include rank-based cost to fitness mapping [1] followed by stochastic universal sampling [2] or tournament selection, also based on cost, as described by Ritzel et al. [18].

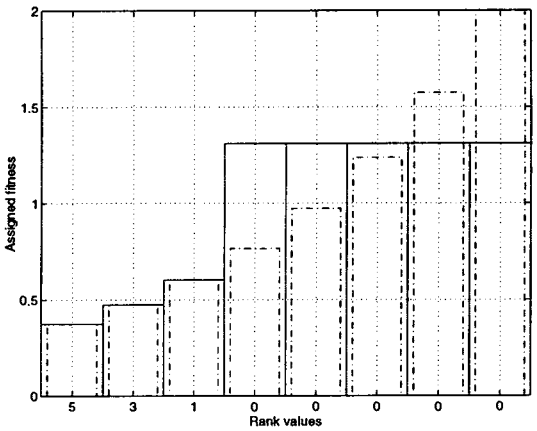


Figure 16.4 Rank-based fitness assignment

Exponential rank-based fitness assignment is illustrated in Fig. 16.4. Here, individuals are sorted by their cost - in this case the values from Fig. 16.3 - and assigned fitness values according to an exponential rule (determined by

the particular ranking method employed) in the first instance, shown by the narrow bars in Fig. 16.4. A single fitness value is then derived for each group of individuals sharing the same cost, through averaging, and is shown in the figure by the wider bars.

16.4.3 Fitness Sharing

Even though all preferred individuals in the population are assigned the same level fitness, the number of offspring that they will produce, which must obviously be integer, may differ due to stochastic nature of EAs. Over generations, these imbalances may accumulate resulting in the population focusing on an arbitrary area of the trade-off surface, known as *genetic drift* [9]. Additionally, recombination and mutation may be less likely to produce individuals at certain areas of the trade-off surface, e.g. the extremes, giving only a partial coverage of the trade-off surface.

Originally introduced as an approach to sampling multiple fitness peaks, fitness sharing [8] helps counteract the effects of genetic drift by penalizing individuals according to the number of other individuals in their neighbourhood. Each individual is assigned a *niche count*, initially set to zero, which is incremented by a certain amount for every individual in the population, including itself. A *sharing function* determines the contribution of other individuals to the niche count as a function of their mutual distance in genotypic, phenotypic or objective space. Raw fitness values are then weighted by the inverse of the niche count and normalized by the sum of the weights prior to selection. The total fitness in the population is re-distributed, and thus shared, by the population. However, a problem with the use of fitness sharing is the difficulty in determining the niche size, σ_{share}, i.e. how close together individuals may be before degradation occurs.

An alternative, but analogous, approach to niche count computations are kernel density estimation methods [19] as used by statisticians. Instead of a niche size, a smoothing parameter, h, whose value is also ultimately subjective, is used. However, guidelines for the selection of suitable values for h have been developed for certain kernels, such as the standard normal probability density function and Epanechnikov kernels. The Epanechnikov kernel may be written as [19]

$$K_e(d/h) = \begin{cases} \frac{1}{2}c_n^{-1}(n+2)[1-(d/h)^2] & \text{if } d/h < 1 \\ 0 & \text{otherwise} \end{cases}, \tag{16.6}$$

where n is the number of decision variables, c_n is the volume of the unit n-dimensional sphere and d/h is the normalized Euclidean distance between individuals. Apart from the constant factor, $\frac{1}{2}c_n^{-1}(n+2)$, this kernel is a particular

case of the family of power law sharing functions proposed by Goldberg and Richardson [8].

Silverman [19] gives a smoothing factor that is approximately optimal in the least mean integrated squared error sense when the population follows a multivariate normal distribution for the Epanechnikov kernel $K_e(d)$ as

$$h = \left[8c_n^{-1}(n+4)(2\sqrt{\pi})^n/N\right]^{1/(n+4)} \tag{16.7}$$

for a population with N individuals and identity covariance matrix. Where populations have an arbitrary sample covariance matrix, S, this may simply be 'sphered', or normalized, by multiplying each individual by a matrix R such that $RR^T = S^{-1}$. This means that the niche size, which depends on S and h, may be automatically and constantly updated, regardless of the cost function, to suit the population at each generation and used directly to perform sharing in Euclidean decision variable spaces.

16.4.4 Mating Restriction

Mating restrictions are employed to bias the way in which individuals are paired for reproduction [4]. Recombining arbitrary individuals from along the trade-off surface may lead to the production of a large number of unfit offspring, called lethals, that could adversely affect the performance of the search. To alleviate this potential problem, mating can be restricted, where feasible, to individuals form within a given distance of each other, σ_{mate}. A common practise is to set $\sigma_{\text{mate}} = \sigma_{\text{share}}$ so that individuals are allowed to mate with one another only if they lie within a distance h from each other in the 'sphered' space used for sharing [6].

16.4.5 Progressive Preference Articulation

As the population of the MOGA evolves, trade-off information will be acquired. In response to the optimization so far, the operator may wish to investigate a smaller region of the search space or even move on to a totally new region. This can be achieved by resetting the goals supplied to the MOGA which, in turn, affects the ranking of the population and modifies the fitness landscape concentrating the population on a different area of the search space. The priority of design objectives may also be changed interactively using this scheme.

The introduction of a small number of random individuals at each generation, say 10–20%, has been shown to make the EA more responsive to sudden changes in the fitness landscape as occurs when the optimization is changed interactively [11]. This technique may also be employed by a MOGA and is used in the example presented in the next section.

16.4.6 Algorithm Description

A pseudo-code outline of the multiobjective genetic algorithm is shown in Fig. 16.5. The population is initialized and the chromosomes are decoded, if necessary, and then evaluated according to the multiple objective functions. Preference-based ranking, `pref_rank` in Fig. 16.5, assigns a non-unique cost to each individual dependent on its dominance in the population such that all non-dominated individuals are ranked zero, as described in section 16.4.1. As well as the vector of performance goals, `GoalV` an additional vector of objective priority levels, `PriV` may also be specified although this is not used in the example here.

The niche counts, `Share`, are calculated using a kernel estimator based on the Epanechnikov kernel. The decoded decision variables, `DVar` are passed to the function twice as they are both the sample data and the points where the population density should be estimated. The default smoothing parameter `Sigma` (h) and a matrix R, such that `DVar * R` has identity covariance matrix, are also returned by the estimation function for use later during mating restriction.

The function `ranking` uses `Share` to perform fitness sharing between individuals of equal cost as part of the fitness assignment procedure. Individuals can now be selected for reproduction, in this case by stochastic universal sampling, and allowance should be made at this point if random chromosomes are to be inserted into the population after mutation so that only the required number of individuals are selected. Mating restriction is implemented by reordering the selected individuals in `Sel` so that consecutive pairs correspond to individuals within the required distance `Sigma` of one another within normalized decision variable space wherever possible (`restrict` in Fig. 16.5).

Recombination of individuals may now proceed as normal and the resulting population mutated. If random chromosomes are to be appended to the population then this should occur after mutation so that they will have to survive selection before they can reproduce with the main population. This is most likely to occur when the fitness landscape changes, as a result of changes in `GoalV` or `PriV`, and the population is no longer well adapted to it.

16.5 EMS CONTROL SYSTEM DESIGN EXAMPLE

In this section we present the example we will use to demonstrate the application of the multiobjective evolutionary and simulated annealing algorithms. The problem considered is the design of a control system for high-speed magnetically levitated (maglev) vehicle and was chosen as it is sufficiently complex as to be realistic yet un-amenable to efficient direct solution using conventional techniques. The remainder of this section describes the controller design procedure, the system under consideration and the design objectives. However, as

```
procedure MOGA {
    initialize(Chrom);
    while not finished do {
        DVar = decode(Chrom);
        ObjV = multi_obj_fun(DVar);
        Cost = pref_rank(ObjV, GoalV);
        [Share, Sigma, R] = epanechnikov(DVar, DVar);
        Fitn = ranking(Cost, Share);
        Sel = select(Chrom, Fitn);
        Sel = reproduce(restrict(decode(Sel)*R, Sigma));
        Chrom = mutate(Sel);
    }
}
```

Figure 16.5 A multiobjective genetic algorithm.

many readers will be unfamiliar with the intricacies of control system design, the characteristics of the design objectives and parameters will be described in less formal terms in sections 16.6 and 16.8 and the remainder of this section may be referred to as required.

16.5.1 The Loop-Shaping Design Procedure

For effective control system design using analytical robust control optimization methods, such as McFarlane and Glover's H_∞ Loop-Shaping Design Procedure (LSDP) [14], suitable weighting functions are required. The H_∞-norm of a transfer function matrix $F(s)$ is the peak over all frequncies of the maximum singular value, $\bar{\sigma}(F(j\omega))$, at each frency, ω, i.e.

$$\|F\|_\infty \triangleq \sup_\omega \bar{\sigma}\left(F(j\omega)\right). \tag{16.8}$$

In controller design, this norm can be used as a measure of the robustness of the stability of the control system (see, for example, [21]). The selection of good weighting functions to achieve both a small H_∞-norm and explicit closed loop performance can involve considerable effort from the designer. To help this process, multiobjective optimization techniques can be used to search for weighting functions which give a design that satisfies a set of closed loop performance and stability robustness goals [27].

For the LSDP, the plant model $G = \tilde{M}^{-1}\tilde{N}$, is a normalized left coprime factorization (NLCF) of G. A perturbed model G_P is defined as $G_P = (\tilde{M} +$

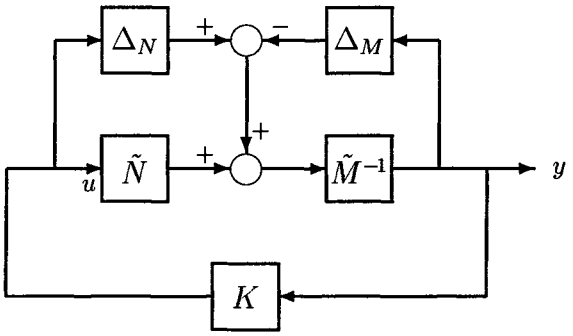


Figure 16.6 Robust stabilization with respect to coprime factor uncertainty

$\Delta_M)^{-1}(\tilde{N} + \Delta_N)$ where $\Delta_M, \Delta_N \in RH_\infty$. A controller K_0 which maximizes this class of perturbed models such that the system remains stable can be synthesized from the solution of 2 algebraic Ricatti equations. The optimal controller obtains

$$\gamma_0 = \inf_{\text{stabilizing } K} \left\| \begin{bmatrix} W_1^{-1}K \\ W_2 \end{bmatrix} (I - GK)^{-1} \begin{bmatrix} W_2^{-1} & GW_1 \end{bmatrix} \right\|_\infty . \tag{16.9}$$

To meet closed-loop performance requirements, the nominal plant G is augmented with pre- and post-compensators W_1 and W_2 respectively, so that the augmented plant G_s is $G_s = W_2 G W_1$. An optimum feedback controller K_0 is synthesized which robustly stabilizes the NLCF of G_s. The final feedback controller K is then constructed by simply combining K_0 with the weights to give $K = W_1 K_0 W_2$. The minimized norm γ_0 is a design indicator of the success of the loop-shaping as well as a measure of the robustness of the stability property.

If closed-loop performance objective functions, f_i, and goals, g_i, are defined, then the weighting functions may be parameterized by $\mathbf{p}$ and the problem formulated as a multiobjective optimization problem.

Problem: The problem is to find W_1, W_2 and hence K such that

$$\begin{aligned} &\gamma_0(W_1, W_2) \leq g_1, \\ &\text{and } f_i(W_1, W_2) \leq g_i \text{ for } i = 2 \ldots n. \end{aligned} \tag{16.10}$$

16.5.2 EMS Design Model

The control design problem is for a 140 m/s maglev vehicle consisting of a chassis supporting a passenger cabin by means of a secondary suspension of

airsprings and hydraulic shock absorbers. The chassis is levitated by means of dc electromagnets under active control providing an attractive force to the guideway. The aim of the control is to provide stability to an inherently unstable system, to maintain the airgap, and to ensure the quality of the ride for the passengers.

The design model [13, 20] is for a single electro-magnet considering the vertical movement of the chassis and passenger cabin. The secondary suspension is modeled as a linear spring damper system, and the primary electro-magnetic suspension is described by a nonlinear first order differential equation for the vertical force being derived from the magnet force law and the current/voltage relation. The model configuration is shown in Fig. 16.7.

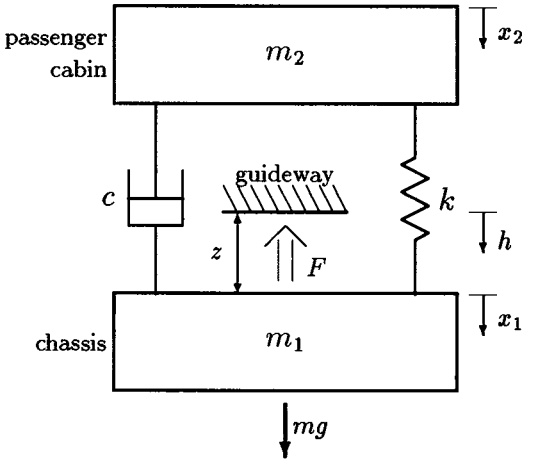


Figure 16.7 Maglev vehicle model

The force, F, exerted by the magnet is

$$F(i,z,t) = \frac{K_m}{2}\left[\frac{i(t)}{z(t)}\right]^2, \tag{16.11}$$

where, from steady state considerations, $K_m = 2mg(z_0/i_0)^2$; and where i is the current (nominal value i_0), z is the gap between magnet and guideway (nominal value z_0), m is the total mass of the vehicle and g is the gravitational constant. If R is the total resistance of the circuit (including the amplifier output resistance and the magnet winding resistance), then for an instantaneous voltage $v(t)$ across the magnet winding, the excitation current $i(t)$ is controlled by

$$v(t) = Ri(t) + K_m \frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{i(t)}{z(t)}\right]. \tag{16.12}$$

The chassis has mass m_1 and the passenger cabin has mass m_2. The secondary suspension has a spring constant k and damping constant c. The relationship between the two is assumed linear and satisfies Newton's law:

$$m_1\ddot{x}_1 + c(\dot{x}_1 - \dot{x}_2) + k(x_1 - x_2) = mg - F, \tag{16.13}$$

$$m_2\ddot{x}_2 + c(\dot{x}_2 - \dot{x}_1) + k(x_2 - x_1) = 0, \tag{16.14}$$

where x_1 is the absolute position of the chassis, and x_2 is the absolute position of the passenger cabin.

The air gap is related to the absolute chassis position by

$$z(t) = x_1(t) - h_g(t), \tag{16.15}$$

where $h_g(t)$ is the disturbance resulting from variations in the guideway profile.

The measurement of the air gap, z, is always used for feedback, and the following measurements are also available: the secondary gap, $x_1 - x_2$, the chassis acceleration, $\ddot{x}_1$, the passenger cabin acceleration, $\ddot{x}_2$, and the magnet current, i.

16.5.3 Performance for Control System Design

The performance measures used must reflect the objective of the control, namely the maximum airgap and the quality of the ride. In addition, there is a constraint on the amount of control voltage that can be applied. Hence, performance indices based on the maximum variation in the airgap, z, the maximum variation in control voltage v and the maximum acceleration experienced by the passengers $\ddot{x}_2$ are proposed.

The major disturbance to the system is from variations in the guideway height, and the following bound on the guideway variations has been suggested for a 140 m/s vehicle [17] :

$$D = \sup\left\{\left|\dot{h}_g(t)\right| : t \geq 0\right\} = 30 \text{ mm/s}. \tag{16.16}$$

Now, from [28, 26], for all possible $h_g(t)$ such that $\sup\{\left|\dot{h}_g(t)\right| : t \geq 0\} \leq D$,

$$\sup\{|y_i(t, h_g)| : t \geq 0\} = D\int_0^\infty |y_i(\tau, 1)|\,\mathrm{d}\tau, \tag{16.17}$$

where $y_i(\tau, 1)$ is the unit step response of the ith output of the linear closed-loop system. Thus, using (16.16) and (16.17), nominal performance functions for the linear closed-loop system can be defined on the airgap, passenger acceleration and control voltage.

To measure the performance of the closed-loop non-linear system, the responses of the non-linear system outputs $y_i'(t, h_{test})$ to a test input, $h_{test}(t)$, can be calculated, and the maximum values evaluated. The test input, $h_{test}(t)$, (shown at the top of Fig. 16.8) was chosen to correspond to a severe guideway disturbance [Equation (16.16)] and represents a level guideway encountering a constant gradient of 30 mm/s immediately followed by a negative gradient of 30 mm/s.

A maximum power spectral density (p.s.d.) $\Phi_{max}(\omega)$ of the passenger cabin acceleration has been recommended by the US Department of Transportation as a minimum ride quality standard [20]. A performance functional can be defined based on this recommendation. The p.s.d. of the track variations can be modeled as $\Phi_{h_g h_g}(\omega) = Av/\omega^2$, where A depends on the track quality and v is the speed. The p.s.d. of the passenger cabin is thus

$$\Phi_{\ddot{x}_2\ddot{x}_2}(\omega) = \left|T_{\ddot{x}_2 h_g}(\omega)\right|^2 \frac{Av}{\omega^2}, \tag{16.18}$$

where $T_{\ddot{x}_2 h_g}$ represents the transfer function between h_g and $\ddot{x}_2$.

16.6 DESIGN USING EVOLUTIONARY ALGORITHMS

The maglev vehicle is inherently unstable and thus the primary goal of the controller is to provide stability. An electromagnet excited by a constant voltage will either clamp to the rail or fall as the attractive forces decrease with an increasing airgap. However, as well as ensuring the stability of the system, there are three other important considerations that are required when assessing the effectiveness of a control scheme.

1. In order to avoid undesirable contact between the guideway and skids the allowed air gap should be maintained between 7 and 17 mm with a nominal value of 12 mm. Contact leads to vibration, noise, friction and possible damage to the vehicle and/or guideway and thus the control, or value of this output, is critical as the air gap has a definite bound. The nominal air gap should not be large as the lifting capacity decreases with a growing air gap and feasible power consumption and magnet weights would be un-realizable. The minimum air gap of 7 mm assures a sufficient safety margin to accommodate the failure of a single magnet [13] and therefore the maximum error between the actual air gap and nominal air gap is 5 mm.

2. Vertical acceleration experienced by passengers may be used as a measure of ride comfort. Typically, this should not exceed 0.5 m/s^2 in either direction [13]. However, this is not a rigid requirement and it may be allowed to increase to increase to as much as 1.0 m/s^2.

3. Additionally, the control effort required should be within feasible limits of the electromagnets. Thus, the required control voltage should be within $\pm$ 600V [13].

The objective functions for the design of the maglev EMS control system using the evolutionary algorithm are thus defined as

$$
\begin{aligned}
f_1 &= \gamma_0, \\
f_2 &= D\int_0^\infty |z(\tau,1)|\mathrm{d}\tau, \\
f_3 &= D\int_0^\infty |\ddot{x}_2(\tau,1)|\mathrm{d}\tau, \\
f_4 &= D\int_0^\infty |v(\tau,1)|\mathrm{d}\tau, \\
f_5 &= \max_t\{|z'(t,h_{test})|\}, \\
f_6 &= \max_t\{|\ddot{x}_2'(t,h_{test})|\}, \\
f_7 &= \max_t\{|v'(t,h_{test})|\}.
\end{aligned}
\tag{16.19}
$$

The first objective is the H_∞-norm, equation (16.9), and is employed as a measure of the overall success of the design in terms of its stability robustness. Objectives $f_{2,3,4}$ are obtained from the step response of a linear model of the system and controller, defined by the free design parameters, and correspond to airgap, passenger acceleration and control voltage respectively. To measure the performance of the controlled non-linear system, the airgap, passenger acceleration and control voltage are determined for a severe disturbance in the guideway profile (shown at the top of Fig. 16.8). A satisfactory controller should ensure that the bounds are not exceeded when this disturbance is encountered and are measured by objectives $f_{5,6,7}$ respectively. From [27, 24, 13], suitable goals for the objective functions are $g_1 = 5$, $g_2, g_5 =$ 5mm, $g_3, g_6 =$ 500mm/s and $g_4, g_7 =$ 600 V.

Three possible candidate pre-plant weighting function configurations are considered, $W_1^{(1)} = p_1$, $W_1^{(2)} = p_1(s+p_2)/(s+p_3)$, or $W_1^{(3)} = p_1(s^2+p_2 s+p_3)/(s^2+p_4 s+p_5)$. The post-plant weighting function W_2 configuration options are

$$W_2 = \mathrm{diag}\,(W_{2,1}, W_{2,2}, \ldots, W_{2,q}),$$

where q measurements are used. The possible configurations for each diagonal element $W_{2,i}, i = 1,\ldots,q$, of W_2 are either $W_{2,i}^{(1)} = p_n$ or $W_{2,i}^{(2)} = p_n(s+p_{n+1})/(s+p_{n+2})$. A measurement which is not used is defined as $W_{2,i}^{(0)}, i =$

$2, \ldots, 5$, however because the measurement on z is always used, the weighting function on the air gap, $W_{2,1}^{(0)} = 1$, is a candidate.

In the other words, the weighting function configuration set consists of vectors, $\mathbf{X}$, with coordinates X_j, such that:

$$X_j \in \begin{cases} \{1,2,3\} & \text{if } j = 1 \\ \{0,1,2\} & \text{if } j = 2,3,\ldots,6. \end{cases} \tag{16.20}$$

This configuration set is associated with the parameter set, Π, consisting of the admissible design parameters, p_k, $k = 1,2,\ldots,20$. The design vectors, $\mathbf{P}$, are defined with the help of its subsets, i.e., $\mathbf{P} \subset \Pi$.

The problem for H_∞ design of the EMS control system is, therefore, converted to the determination of admissible pairs $(\mathbf{X}, \mathbf{P})$, satisfying the inequalities:

$$f_i \leq g_i, \quad (i = 1,2,\ldots,7). \tag{16.21}$$

It can be also regarded as a multicriteria optimization problem for simultaneous minimization of the functionals, f_i, under the corresponding constraints. The application of standard optimization techniques to this problem, however, needs solution of $3^6 = 729$ optimization subproblems in order to evaluate all possible configurations and define the feasible ones with respect to the performance criteria.

The MOGA was implemented as described in section 16.3 with a population of 70 real-valued individuals. The intermediate recombination operator [16] was applied with probability 0.7 during reproduction and breeder GA mutation [16] was then employed with probability 0.1. The use of adaptive mutation rates may have been more appropriate for this example and representation, although the (seemingly) high mutation rate is consistent with the use of real-valued operators and the average number of active parameters. No fine-tuning of the operator rates was attempted.

A typical solution found during the optimization by the MOGA was a configuration vector,

$$\mathbf{X} = (\ 3,\ 0,\ 1,\ 0,\ 0,\ 0\ ), \tag{16.22}$$

and design parameter vector,

$$\begin{aligned} \mathbf{P} = (\ & 299.8786,\ 2.1687,\ 25.1303,\ 173.6903, \\ & 36.7175,\ 0.7365,\ 0.1473,\ 0.5089, \\ & 0.6537,\ 0.4082,\ 0.2274,\ 0.3738, \\ & 0.6708,\ 0.3344,\ 0.4342,\ 0.6547, \\ & 0.1696,\ 0.9729,\ 0.3389,\ 0.6848\ ), \end{aligned} \tag{16.23}$$

corresponding to the performance vector,

$$\mathbf{F} = (3.1773, 4.9986, 350.8752, 194.4865, 4.4619, 204.2840, 105.2997), \tag{16.24}$$

which clearly satisfies (16.21).

Thus, only measurements of the air gap, z, and the secondary airgap, $x_1 - x_2$, were required with the weighting functions

$$W_1(s) = 300\frac{(s^2 + 2.17s + 25.1)}{(s^2 + 173.7s + 36.7)} \tag{16.25}$$

and W_2 =diag(1, 0.654). The response of the maglev vehicle to the input signal $h_{test}(t)$ are shown in Fig 16.8 where it can be seen that all of the outputs remain within the critical bounds of the system specification and that satisfactory control has been achieved.

However, for such responses it is difficult to determine the relative merits of one controller against another over the entire set of potential solutions. This is particularly true if on-line preference articulation is to be used to guide the search during an optimization. An alternative is the trade-off graph of the type shown in Fig. 16.9.

Here, each line represents a non-dominated solution found by the MOGA illustrating typical trade-offs between controller designs for the maglev vehicle. The x-axis shows the design objectives, the y-axis the performance of controllers in each objective domain and the cross-marks in the figure show the design goals. Crossing lines between adjacent objectives indicates that there is a trade-off between those two objectives while parallel lines show that there is no conflict in the current population of solution estimates. The order of the objectives along the x-axis may, of course, be varied in order to assess trade-off between other design objectives.

In Fig. 16.9, only the preferred individuals, those that satisfy the design goals, are shown. When no individuals satisfy all the design goals, the non-dominated or Pareto optimal solutions are displayed. Trade-offs between adjacent objectives result in the crossing of the lines between them whereas concurrent lines indicate that the objectives do not compete with one another. For example, in Fig. 16.9, airgap and passenger cabin acceleration (objectives 2 and 3) appear to compete quite heavily whilst passenger cabin acceleration and control voltage (objective 4) do not exhibit the same level of competition. Note, however, that all of the potential solutions shown satisfy all of the design criteria and are thus all equally valid.

Examination of the controllers found by the MOGA reveals that none of the preferred candidate solutions have pre-plant weighting functions $W_1^{(1)}$ or $W_1^{(2)}$. However, Pareto-optimal solutions with these weighting functions were found that violated the robustness criteria (objective 1) whilst providing better responses to the test signal for the time-domain performance measures.

Examination of the preferred controllers post-plant weighting functions reveals that the simplest satisfactory control schemes may be realized using mea-

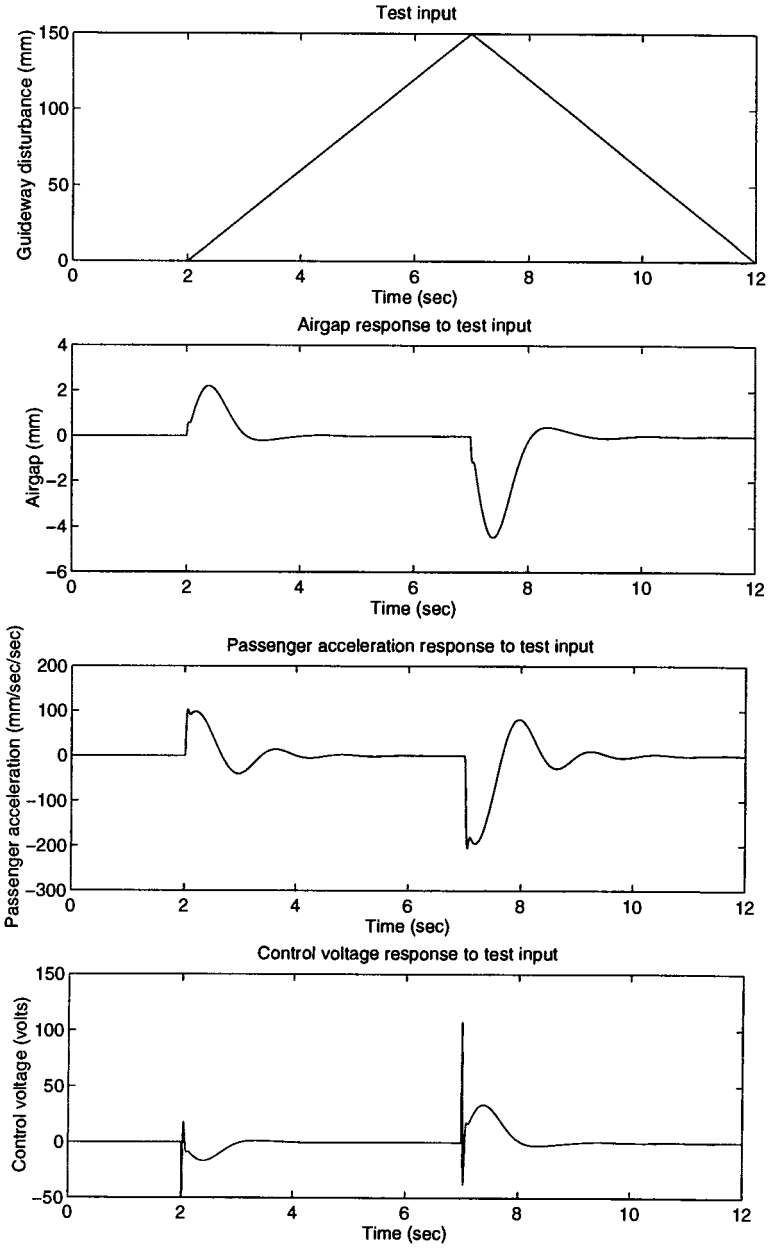


Figure 16.8 Nonlinear system responses to test input h_{test}

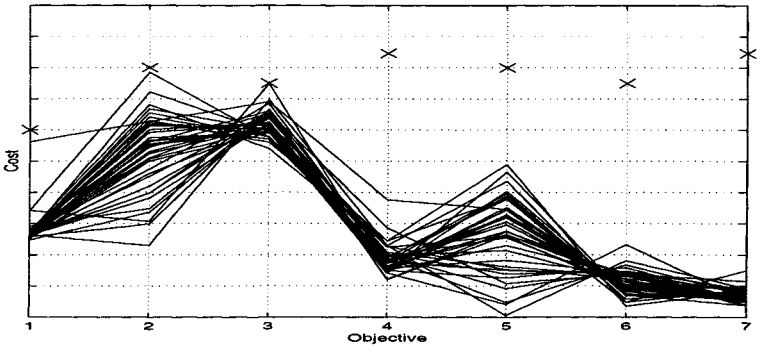


Figure 16.9 Design objective trade-offs

surements for the airgap, z, and only one of either secondary airgap, $x_1 - x_2$, or passenger cabin acceleration, $\ddot{x}_2$, weighted with a pure gain, p_n. A number of other satisfactory control strategies were also found although they required more measurements or greater complexity in the weighting functions.

Having satisfied the original design goals, the control engineer is now free to enhance the performance of the controller. That is, the relative degree of under- or over-attainment of the design goals is clearly visible in Fig. 16.9 and the designer may take advantage of this information if the initial optimization requires the setting of new design goals. For example, the passenger cabin acceleration criteria (objectives 3 and 6) may be relaxed in order to try to improve the airgap error response, control voltage or stability measure. The final choice of the controller to be employed can thus be made from a number of different satisfactory solutions, possibly derived from different performance goals or even design methodologies. For example, the more complex controllers in this example appear to offer a greater level of robustness while the simpler ones would be less expensive to implement.

16.7 SIMULATED ANNEALING

Another probabilistic-based approach to solving multiobjective optimization problems is simulated annealing. Like evolutionary algorithms, simulated annealing approaches do not require (or deduce) any functional derivative information, and are thus unaffected by discontinuities and non-linearities. Simulated annealing is a search procedure which uses local hill climbing, but in a modified manner. The algorithm makes small steps to search the local topography. If the step results in an improved solution, the new solution is accepted;

otherwise it is accepted with a probability which is initially close to unity, but as the algorithm progresses is reduced in stages until it is close to zero. By accepting some solutions which are worse, the algorithm has the chance of climbing out of local minima and finding a global optimum.

The algorithm presented in this chapter is based on a simulated annealing scheme developed by Vanderbilt and Louie [23], but has been extended to multiobjective optimization. The algorithm searches for an $\mathbf{p} \in \Omega$, such that the following inequalities

$$f_i(\mathbf{p}) \leq g_i \; \forall \; i = 1 \ldots n \tag{16.26}$$

are satisfied. The algorithm could be easily modified to solve (16.1).

16.7.1 The Metropolis algorithm

The simulated annealing algorithm has its origins in the statistical mechanical annealing of solids. A simple algorithm was proposed by Metropolis et al. [15] which can be used to simulate the annealing process. The temperature is reduced in stages. At each temperature, the system is perturbed, and the change in energy calculated. The perturbed state is accepted as the new state if the energy has decreased. If the energy has increased, the new state is accepted with a probability, $P(\Delta E)$, given by $\exp(-\Delta E/(k_b T))$

$$P(\Delta E) = \exp(-\Delta E/(k_b T)), \tag{16.27}$$

where $-\Delta E$ is the change in energy, T is the temperature and k_b is the Boltzmann constant. This acceptance rule is referred to as the Metropolis criterion. The basic Metropolis algorithm is shown in Fig. 16.10.

16.7.2 Multiobjective Optimization by Simulated Annealing

The Metropolis algorithm can be used in a simulated annealing scheme to solve optimization problems. This may be done by replacing the energy with the objective function, using the temperature as a control parameter, and by assuming that the role of the states of a solid is taken by a description of a system configuration. Many practical simulated annealing techniques have been developed for combinatorial optimization problems, however, Vanderbilt and Louie [23] first proposed that the basic scheme be used for searching over a continuous variable space. The main difference with combinatorial methods is the perturbation mechanism for searching the parameter space.

There are 5 main features of the Metropolis algorithm which must be considered in developing a simulated annealing scheme: (i) the description of the system, (ii) the perturbation mechanism, (iii) the energy function, (iv) the acceptance criterion and (v) the cooling scheme.

```
Set T = initial temperature
Set M = a number of steps
Set k_b = cooling rate constant
Until terminating condition is true do{
   for i = 1 to M do{
      Make small perturbation on system
      Calculate change in energy ΔE
      If ΔE ≤ 0
         accept perturbation
      else
         accept perturbation with probability
                                        P(ΔE) = exp(−ΔE/(k_b T))
   }
   Reduce Temperature
}
```

Figure 16.10 The basic Metropolis algorithm.

The description of the system. The system configuration must be represented in some way which describes the space of possible designs. For continuous parameter problems, this is generally just by a parameter vector $\mathbf{p} \in \mathbb{R}^q$ with a parameter set $\Omega \subset \mathbb{R}^q$, thus $\{\mathbf{p} : \mathbf{p} \in \Omega\}$ describes the set of possible designs.

The perturbation mechanism. A vector $u \in \mathbb{R}^q$ of independent random numbers $(u_1, u_2, \ldots, u_q)$ is generated where each u_i is chosen independently from the interval $[-\sqrt{3}, \sqrt{3}]$, so they have zero mean and unit variance. Thus the vector $u = (u_1, u_2, \ldots, u_q)$ occurs with a constant probability density inside a hypercube of volume $(2\sqrt{3})^q$ and zero outside. A step Δp to a trial point $\tilde{\mathbf{p}}$ is taken $\tilde{\mathbf{p}} = \mathbf{p} + \Delta p$ where

$$\Delta p = Qu, \tag{16.28}$$

where the matrix $Q \in \mathbb{R}^{q \times q}$ controls the step distribution. Random steps with a desired covariance matrix $s \in \mathbb{R}^{q \times q}$ can be generated from (16.28) and by solving for Q,

$$s = QQ^T. \tag{16.29}$$

A procedure for choosing s so that it adapts to the topography of the objective function has been proposed by Vanderbilt and Louie [23]. The excursions

of the random walk are used as the measure of the local topology, so that the search adapts itself to the local topology. Firstly, from [23], the available phase space $\mathcal{E}(T)$ is vaguely defined as a function of the temperature T to be

$$\mathcal{E}(T) = \{\mathbf{p} : E(\mathbf{p}) - E_{\min}(\mathbf{p}) \lesssim T\}, \tag{16.30}$$

where $E(\cdot)$ is the energy function which is defined in the next section. If the axes of s are poorly aligned with the topology of $E(T)$, much time will be wasted exploring fruitless directions of search. So, at the end of the lth set of M steps, the first and second moments of the walk segment are calculated, where

$$A_i^{(l)} = 1/M \sum_{m=1}^{M} p_i^{(m;l)} \tag{16.31}$$

and

$$S_{ij}^{(l)} = 1/M \sum_{m=1}^{M} \left[p_i^{(m;l)} - A_i^{(l)}\right] \left[p_j^{(m;l)} - A_j^{(l)}\right] \tag{16.32}$$

where $\mathbf{p}^{(m;l)}$ is the value of $\mathbf{p}$ on the mth step of the lth set. Thus S describes the actual shape of the walk over the lth set. To choose s for the next iteration set, $l+1$,

$$s^{(l+1)} = \frac{\chi_s}{\beta M} S^{(l)}, \tag{16.33}$$

where $\chi_s > 1$ is a growth factor, and β is based on a geometric average over the random variables Δp. Typically, $\chi_s = 3$ and $\beta = 0.11$ [23].

The idea behind this scheme is that the steps are initially small, but grow as the annealing progresses until the walk can cover the phase space $\mathcal{E}(T)$. Steps outside the phase space will then begin to be rejected, and the size of the walk will reduce as the phase space $\mathcal{E}(T)$ gets smaller as the temperature is reduced. The size and shape of S and hence s will thus adapt to the topology of $E(\mathbf{p})$. In this way, the whole possible phase space is covered in a relatively efficient manner.

The energy function. The energy function E is used to convert the multi-objective problem to a single objective minimax problem by

$$E(\mathbf{p}) = \max\left\{\max\left\{\frac{f_i(\mathbf{p}) - g_i}{w_i}, 0\right\} : i = 1, \ldots, n\right\}, \tag{16.34}$$

where $w_i, i = 1, \ldots, n$ are positive weightings chosen *a priori* by the designer to reflect the relative importance of each objective function f_i.

The acceptance criterion. The algorithm uses the Metropolis acceptance criterion. Thus, if the energy change is $\Delta E = E(\tilde{\mathbf{p}}) - E(\mathbf{p})$, where $\tilde{\mathbf{p}}$ is a trial point, $\mathbf{p}$ is the current point and E is as for (16.34), then if $\Delta E \leq 0$, the trial point is accepted; otherwise, the trial point is accepted with a probability given by $\exp(-\Delta E/(k_b T))$.

The cooling scheme. After every M steps of the algorithm at the temperature T, the new temperature T^* is set to $T^* = \chi_T T$, where $0 < \chi_T < 1$. The choice of χ_T greatly affects the efficiency of the annealing and its ability to climb out of local minima; if χ_T is too small, the temperature will reduce too quickly, and the process will get stuck in a local minimum. If χ_T is too large, the process will become very inefficient.

In addition, the efficiency of the algorithm and its ability to climb out of local minima is dependent on the initial temperature. Vanderbilt and Louie [23] suggest that the initial temperature is chosen on the basis of the variance of the energy $E(\mathbf{p})$ of a random sample of points $\mathbf{p}$, and χ_T is chosen by trial and error. From the tests by Vanderbilt and Louie [23] and in [25], a slow anneal, with $\chi_T = 0.99$, seems to result in a solution, but it is at the expense of a large number of iterations.

16.7.3 The Multiobjective Simulated Annealing Algorithm

The complete multiobjective simulated annealing algorithm is shown in Fig. 16.11. The algorithm iterates until a solution which satisfies (16.26) is found. If a solution does not exist, the algorithm iterates until cool i.e. $T^{(l)} \leq T_{min}$. Note that if a solution is not found, there is no guarantee that a solution does not exist. In order that a solution to (16.1) is obtained, the termination criterion and the energy function can be changed.

A test should be made to check that $\tilde{\mathbf{p}} \in \Omega$, and if not, trials are made until a trial $\tilde{\mathbf{p}} \in \Omega$ is obtained. An alternative is to characterize Ω by a set of inequality constraints, with a penalty function if $\tilde{\mathbf{p}} \notin \Omega$.

```
Define f_i, g_i and w_i for i = 1,...,n
Set M, χ_T, χ_s, β, and k_b
Set initial and minimum temperatures, T^(0) and T_min
Set initial point p^(0) = (p_1^(0), p_2^(0), ..., p_q^(0))
Calculate objective function f(p^(0))
Calculate initial energy E(p^(0)) by (16.34)
Set initial covariance matrix s^(0)
Calculate Q^(0) by solving (16.29)
Set iteration number k = 0
Set cooling stage number l = 0
Set cooling stage counter m = 0
until f_i(p^(k)) ≤ g_i ∀i = 1,...,n or T^(l) ≤ T_min do {
   Generate random vector u = (u_1, u_2, ..., u_q) where
                                          u_i = random[-√3, √3]
   Generate trial point p̃ = p^(k) + Q^(k)u
   Calculate f(p̃) and E(p̃)
   Calculate energy change ΔE = E(p̃) - E(p^(k))
   if
      ΔE ≤ 0, accept move
   else
      if exp(-ΔE/(k_b T^(l))) > random[0,1), accept move
   if move accepted {
      Set new point p^(k+1) = p̃
      Increment k = k + 1 and m = m + 1
      if m = M {
         Calculate A_i^(l) and S_ij^(l) from (16.31) and (16.32)
         Calculate s^(l+1) from (16.33) and Q^(l+1) by solving (16.29)
         Reduce temperature T^(l+1) = χ_T T^(l)
         Set m = 0 and increment l = l + 1
      }
   }
}
```

Figure 16.11 Multiobjective simulated annealing algorithm

16.8 DESIGN USING SIMULATED ANNEALING

For the design of the maglev EMS control system using the simulated annealing algorithm, the objective functions are defined as

$$
\begin{aligned}
f_1 &= \gamma_0, \\
f_2 &= D\int_0^\infty |z(\tau,1)|\mathrm{d}\tau, \\
f_3 &= D\int_0^\infty |\ddot{x}_2(\tau,1)|\mathrm{d}\tau, \\
f_4 &= D\int_0^\infty |v(\tau,1)|\mathrm{d}\tau, \\
f_5 &= \max_t\{|z'(t,h_{test})|\}, \\
f_6 &= \max_t\{|\ddot{x}_2'(t,h_{test})|\}, \\
f_7 &= \max_t\{|v'(t,h_{test})|\}, \\
f_8 &= \max_\omega\{\Phi_{\ddot{x}_2\ddot{x}_2}(\omega) - \Phi_{max}(\omega)\}.
\end{aligned}
\tag{16.35}
$$

Here, objectives $f_{1,\ldots,7}$ are the same as those specified in (16.19). An additional objective, f_8, described in section 16.5.3 defines the power spectral density and is used as a measure of ride quality standard. From [27, 24, 13], suitable goals for the objective functions are $g_1 = 5$, $g_2, g_5 = 5$mm, $g_3, g_6 = 500$mm/s, $g_4, g_7 = 600$ V and $g_8 = 0$.

Measurements of the air gap, z, and the passenger cabin acceleration, $\ddot{x}_2$, are used for feedback. The weighting function configurations are

$$
W_1 = p_1\frac{(s^2+p_2s+p_3)}{(s^2+p_4s+p_5)}, \quad W_2 = \mathrm{diag}\left(p_6\frac{(s+p_7)}{(s+p_8)}, p_9\right). \tag{16.36}
$$

The secondary suspension stiffness and damping factors c and k are included as design parameters p_{10} and p_{11}.

The weighting vector for the energy function (16.34) is defined as $\mathbf{w} = (0.1, 0.5, 10, 10, 0.5, 10, 10, 0.4)$; this reflects the relative importance of the objective functions. The initial temperature was set to $T = 5$.

The multiobjective simulated annealing algorithm was run with the LSDP and a design vector found which met all the design goals except the performance functional based on the power spectral density specification, which was only marginally exceeded. The performance objective functions of the design are

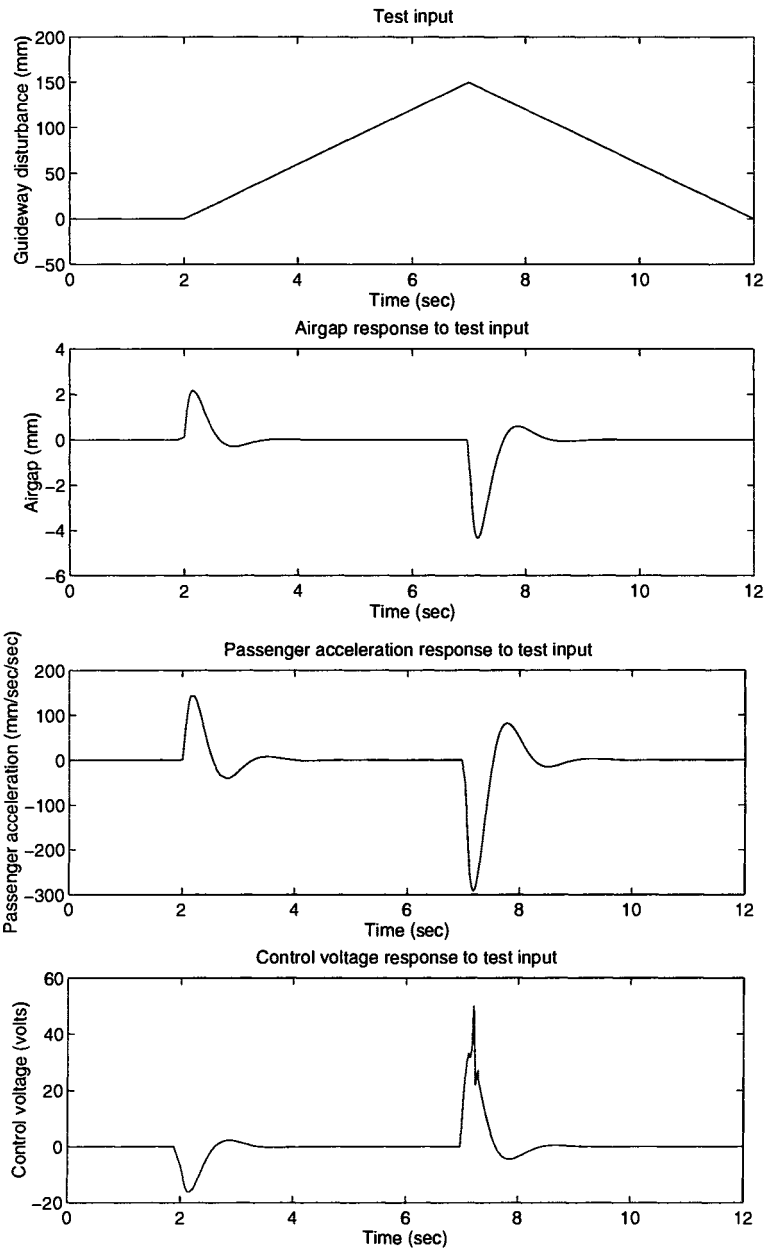


Figure 16.12 Nonlinear system responses to test input h_{test}

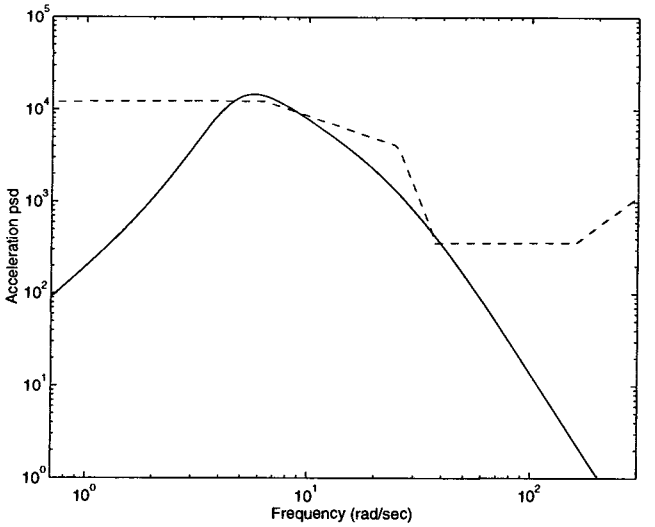


Figure 16.13 Passenger cabin psd (——) and ride quality standard psd (- - -)

$$
\begin{aligned}
f_1 &= 2.64, \\
f_2 &= 5.06 \text{ mm}, \\
f_3 &= 391.4 \text{ mm/s}^2, \\
f_4 &= 201.9 \text{ V}, \\
f_5 &= 4.38 \text{ mm}, \\
f_6 &= 291.1 \text{ mm/s}^2, \\
f_7 &= 33.6 \text{ V}, \\
f_8 &= 0.075.
\end{aligned}
\tag{16.37}
$$

The designed weighting functions are

$$
W_1 = 903.2 \frac{(s^2 + 169.6s + 485.4)}{(s^2 + 69.5s + 488.3)}, \tag{16.38}
$$

$$
W_2 = \text{diag}\left(426.9 \frac{(s + 256.2)}{(s + 379.2)}, 3.23\right), \tag{16.39}
$$

and the designed secondary suspension stiffness and damping factors are $c = 90.3$ and $d = 20.0$. The non-linear simulation response and power spectral density of the final design are shown in Figs. 16.12 and 16.13.

16.9 DISCUSSION AND CONCLUSIONS

This chapter has presented two approaches to solving multiple criteria optimization problems, multiobjective genetic algorithms and simulated annealing, and demonstrated their application on a realistic industrial problem of robust control system design. Both approaches were shown to yield acceptable solutions to the design problem, satisfying a set of competing design criteria.

The MOGA approach, utilizing evolutionary algorithms that are known to perform well on broad classes of of ill-behaved problems, offers a number of properties desirable in the solution of multiple criteria problems. The ability to simultaneously handle many candidate solutions is well suited to most multiple criteria problems as are the mechanisms to promote diversity in the population. Indeed, this diversity and corresponding richness in trade-off information may be used to refine the initial preferences or goals until a suitable compromise is obtained, possibly through the use of on-line preference articulation. A drawback of the MOGA approach may be the length of execution time due to the need to evaluate whole populations of candidate solutions.

The simulated annealing algorithm, offering a simple parameterization of the search space, is shown to be an efficient approach when the choice of initial design parameters is difficult or where hill-climbing methods fail or are unable to locate solutions other than local ones. In such circumstances, the possibly larger number of iterations over the conventional approach may be justified by a better quality solution. A drawback of the simulated annealing algorithm presented here compared with the MOGA is the need to define weightings for the energy function, although these could be calculated by some method automatically.

References

[1] Baker, J.E.: "Adaptive selection methods for genetic algorithms". In: John J. Grefenstette (ed): *Genetic Algorithms and Their Applications: Proc. of the First International Conference on Genetic Algorithms*, Lawrence Erlbaum, 1985, pp. 101–111

[2] Baker, J.E.: "Reducing bias and inefficiency in the selection algorithm". In: Grefenstette [10], pp. 14–21

[3] Becker, R.G., A.J. Heunis, D.Q. Mayne: Computer-aided design of control systems via optimization. IEE Proc. D 126(1979)573–578

[4] Deb, K., D.E. Goldberg: "An investigation of niche and species formation in genetic function optimization". In: J. David Schaffer (ed): *Proc. of the Third International Conference on Genetic Algorithms*, Morgan Kaufmann, San Mateo, CA, 1989, pp. 42–50

[5] Fonseca, C.M.: *Multiobjective Genetic Algorithms with Application to Control Engineering Problems.* PhD thesis, University of Sheffield, 1995

[6] Fonseca, C.M., P.J. Fleming: An overview of evolutionary algorithms in multiobjective optimization. Evolutionary Computation 3(1995)1–16

[7] Goldberg, D.E.: *Genetic Algorithms in Search, Optimization and Machine Learning.* Addison-Wesley, Reading, Massachusetts, 1989

[8] Goldberg, D.E., J. Richardson: "Genetic algorithms with sharing for multimodal function optimization". In: Grefenstette [10], pp. 41–49

[9] Goldberg, D.E., P. Segrest: "Finite markov chain analysis of genetic algorithms". In: Grefenstette [10], pp. 1–8

[10] Grefenstette, J. J., (ed): *Genetic Algorithms and Their Applications: Proceedings of the Second International Conference on Genetic Algorithms.* Lawrence Erlbaum, 1987

[11] Grefenstette, J.J.: "Genetic algorithms for changing environments". In: R. Männer and B. Manderick (eds): *Parallel Problem Solving From Nature 2,* North–Holland, 1992, pp. 137–144

[12] Hwang, C.-L., A.S.M. Masud: *Multiple Objective Decision Making – Methods and Applications,* Vol. 164 of *Lecture Notes in Economics and Mathematical Systems.* Springer-Verlag, Berlin, 1979

[13] Kortüm, W., A. Utzt: Control law design and dynamic evaluations for a maglev vehicle with a combined lift and guidance suspension system. ASME J. Dyn. Syst. Meas. & Control 106(1984)286–292

[14] McFarlane, D.C., K. Glover: *Robust Controller Design Using Normalized Coprime Factor Plant Descriptions,* Vol. 138 of *Lect. Notes Control & Inf. Sci.* Springer-Verlag, Berlin, 1990

[15] Metropolis, N., A. Rosenbluth, M. Rosenbluth, A. Teller, E. Teller: Equation of state calculations by fast computing machines. J. Chem. Phys. 21(1953)1087–1092

[16] Mühlenbein, H., D. Schlierkamp-Voosen: Predictive models for the breeder genetic algorithm: I. continuous parameter optimization. Evolutionary Computation, 1(1993)25–49

[17] Müller, P.C.: "Design of optimal state-observers and its application to maglev vehicle suspension control". In: *Proc. 4th IFAC Symp. Multivariable Technological Systems,* Fredericton, Canada, 1977, pp. 175–182

[18] Ritzel, B.J., J.W. Eheart, S. Ranjithan: Using genetic algorithms to solve a multiple objective groundwater pollution containment problem. Water Resources Research 30(1994)1589–1603

[19] Silverman, B. W.: *Density Estimation for Statistics and Data Analysis,* Vol. 26 of *Monographs on Statistics and Applied Probability.* Chapman and Hall, London, 1986

[20] Sinha, P.K.: *Electromagnetic Suspension: Dynamics and Control.* Peter Peregrinus Ltd., London, 1987

[21] Skogestad, S., I. Postlethwaite: *Multivariable Feedback Control: Analysis and Design.* John Wiley and Sons Ltd., Chichester, England, 1996

[22] Spears, W.M., K.A. De Jong, T. Bäck, D.B. Fogel, H. De Garis: "An overview of evolutionary computation". In: *Machine Learning: ECML-93 European Conference on Machine Learning, Lecture notes in Artificial Intelligence*, Vol. 667 of *Lecture notes in Artificial Intelligence*, Springer-Verlag, 1993, pp. 442–459

[23] Vanderbilt, D., S.G. Louie: A Monte Carlo simulated annealing approach to optimization over continuous variables. J. Comp. Physics 56(1984)259–271

[24] Whidborne, J.F.: EMS control system design for a maglev vehicle - A critical system. Automatica 29(1993)1345–1349

[25] Whidborne, J.F., D.-W. Gu, I. Postlethwaite: Simulated annealing for multi-objective control system design. IEE Proc. D 144(1996)582–588

[26] Whidborne, J.F., G.P. Liu: *Critical Control Systems: Theory, Design and Applications.* Research Studies Press, Taunton, U.K., 1993

[27] Whidborne, J.F., I. Postlethwaite, D.-W. Gu: Robust controller design using H_∞ loop-shaping and the method of inequalities. IEEE Trans. on Contr. Syst. Technology 2(1994)455–461

[28] Zakian, V.: A performance criterion. Int. J. Control 43(1986)921–931

INDEX