Faculté de Droit,
d'Economie et de Finance

PhD-FDEF-1-2007

Thèse

Soutenue le 8 octobre 2007 à Luxembourg

En vue de l'obtention du grade académique conjoint de

Docteur de l'Université du Luxembourg
en Mathématiques Appliquées

et de

Docteur de la Faculté Polytechnique de Mons
en Sciences de l'Ingénieur

par

Patrick Meyer

Progressive Methods in Multiple Criteria Decision Analysis

Jury de thèse

Jean-Luc Marichal, *président du jury*
Université du Luxembourg

Marc Pirlot, *directeur de thèse*
Faculté Polytechnique de Mons

Vincent Mousseau
Université Paris Dauphine

Raymond Bisdorff, *directeur de thèse*
Université du Luxembourg

Philippe Fortemps
Faculté Polytechnique de Mons

Theodor Stewart
University of Cape Town

Thèse

Soutenue le 8 octobre 2007 à Luxembourg
En vue de l'obtention du grade académique conjoint de

Docteur de l'Université du Luxembourg
en Mathématiques Appliquées

et de

Docteur de la Faculté Polytechnique de Mons
en Sciences de l'Ingénieur

par

Patrick Meyer

Progressive Methods in Multiple Criteria Decision Analysis

August 2, 2007

This doctoral thesis is the fruit of research activities performed in the context of an assistant position at the University of Luxembourg in the Applied Mathematics Unit (SMA).

The printing of this thesis has been financially supported by the "Fonds National de la Recherche" of Luxembourg.

The page setting of this document has been carried out with LaTeX and the schemata have been drawn with Xfig, *both on Ubuntu GNU/Linux computers.*

Je dis des choses tellement intelligentes que le plus souvent je ne comprends pas ce que je dis.

Les Shadoks, Jacques Rouxel[1]

[1] I say things so smart that I generally do not understand what I say.

Acknowledgements

My life as a researcher started about nine years ago, when Marc Roubens offered me to work for him at the University of Liège. Thanks to him, I discovered the exciting and ever ongoing domain of Multiple Criteria Decision Analysis. I wish to express my gratitude to him, for having given me the opportunity, not to engage the professional path, which I would have taken otherwise.

In my unconventional research career, I have decided to start this PhD four years ago, after having met Raymond Bisdorff from the University of Luxembourg. During a discussion around a wonderful Scottish single malt whisky, I realised that it might be an interesting experience to start a doctoral thesis under his direction. He accepted to become my supervisor and since then we always had passionate discussions about various (research) topics. I would like to thank him for his permanent support, his everlasting kindness and the many pertinent remarks concerning my research.

My doctoral thesis was also co-directed by Marc Pirlot from the Faculty of Engineering of Mons. His very detailed reading of this work has greatly improved its quality. I would like to thank him for that and for his friendliness during our discussions on and around my doctoral research.

My doctoral activities would not have been possible without the supervisory committee which guided me during the last three years. It was composed of my two PhD supervisors, as well as Jean-Luc Marichal from the University of Luxembourg and Philippe Fortemps from the Faculty of Engineering of Mons. I wish to thank them all for their support and their interesting remarks on the preliminary versions of the manuscript.

I have published articles with various co-authors: Raymond Bisdorff, Fabien Boniver, Michel Grabisch, Ivan Kojadinovic, Stéphane Lallich, Philippe Lenca, Jean-Luc Marichal, Marc Roubens and Benoît Vaillant. Each of them has opened my eyes on new subjects and problems, and I would like to warmly thank them for that.

During the past few years I had the opportunity to meet a lot of different people with whom I had many passionate discussions on various more or less scientific subjects. As I would be very embarrassed if I left anybody out, I simply thank them all, no matter whether we shared the same opinion or had conflicting ideas. I am sure they will all recognise themselves.

I would also like to thank Lisa Ferreira for her permanent administrative support. Without her, I would probably have finished this PhD thesis much later, as she took over much of my administrative tasks at the University of Luxembourg.

My parents have supported me in any event during the 31 first years of my life. They have taught me to always be curious and to grow better in anything I undertake. I warmly thank them for that.

I especially thank Sophie Liénart for her patience and her love, and for having come along with me to Luxembourg. Without her, I am sure that I would have chosen different paths in life, and I would probably be a quite different person today.

Finally, I would like to apologise to all the people who had to endure my bad temper and irritability of the past months.

Contents

List of Figures

List of Tables

Preface

Helping to make the *right* decision, studying human *behaviours* when confronted to a decision problem, formalising their *preferences*, coping with a plurality of *points of view* in a decision process, ... All these topics, and many more, are the bread and butter of researchers active in the field of Multiple Criteria Decision Analysis (MCDA).

Roughly speaking, MCDA aims at helping a *decision maker* (DM), guided by an *analyst*, to prepare and make a decision where more than one point of view has to be considered. Research activities around MCDA have developed quite rapidly over the past years, and have resulted in various streams of thought and methodological formulations for the resolution of such decision problems.

Since the beginning of our research, we have always been interested in different MCDA problems, originating from the two major methodological trends, namely the *European* and the *American* schools. Quite regularly, we have been concerned with methods putting the DM in the centre of the decision process, aiming at determining his preferences in a *holistic* way, and providing him with results tending to make him *happy enough*.

Recently, we have been interested in quite specific decision processes, allowing to obtain the final *decision recommendation* via intermediate stages. Such methods, further called *progressive* MCDA methods, are iterative procedures which present *partial conclusions* to the DM, which can be refined at further steps of the analysis. This enables a DM, who is not completely satisfied with a recommendation, to further investigate the problem until a *satisfactory* solution can be found.

As we will show in this work, such methods allow to deal with multiple criteria decision problems involving impreciseness, missing information and limited economical resources. Indeed, progressiveness in MCDA permits to undertake a *prudent* construction of the output. As the ultimate recommendation does not necessarily have to be reached in one step, each partial conclusion exclusively exploits the information available at that moment. Consequently, issues linked to impreciseness or missing data can be treated later as the process goes on. Furthermore, in real-world problems, constrained by time and limited financial resources, the determination of certain evaluations of the alternatives can be postponed.

The subjects analysed in this work are coming from our quite heterogeneous research activities of the past few years. Under the direction of Marc Roubens, we started our research activities in the tradition of the American school, on methods and tools taking into account interactions between the points of view (via the so-called *Choquet integral*). Our work in this field is still continuing and has recently given birth, in collaboration with Ivan Kojadinovic and Michel Grabisch, to a software package, called Kappalab, which allows to put into practice some of our considerations in this field.

Later, Raymond Bisdorff gave us the opportunity to focus on procedures originating in the European stream of thought. Under his direction, we contributed to develop the Rubis method which allows to solve the problem of determining a single best alternative.

As a consequence, the document is structured in three parts as follows. Part **I** is intended to be a general introduction to the topics covered by this work. It is divided into two chapters. In the first one, we delimit the sphere of our activity and situate our discourse in a particular scientific context, namely *constructive* MCDA. Then, in Chapter 2, we define the concept of *progressiveness* and present its consequences on a decision process.

Then, Part **II** presents our methodological research on the *choice problematique* in the context of the European school for MCDA. It is divided into three chapters, where in the first one, we present the *bipolar-valued credibility calculus*, as well as the construction of a bipolar-valued outranking relation, which represent the foundations of the further developments of this part. Then, in Chapter 4, we present the Rubis method for the progressive determination of a single best alternative. Finally, in Chapter 5, we extend our considerations to the determination of k simultaneously best alternatives.

Finally, Part **III** gathers our research activities in the framework of the American school for MCDA. It is divided into four chapters. The first one focuses on Multiattribute Value Theory (MAVT) and the Choquet integral as an aggregation function. Then, in Chapter 7, we formalise the *capacity identification problem* and present different methods to determine the parameters of the Choquet integral via the DM's preferences. In the third chapter, we present how the different classes of problems of MCDA can be solved by means of our results. Finally, in Chapter 9, we present Kappalab, which is a package for the GNU R statistical system for capacity and integral manipulation on a finite setting and which can be used in the context of MAVT.

This structure unveils the two main branches of our research activities via Parts II and III, whereas Part I serves as a shell around the whole work. A reader familiar with MCDA techniques can straight off switch to Chapter 2 of the first part to be informed on progressiveness, before getting down to either one of the final two parts, which can be read in any order. However, an MCDA novice should start by reading the first introductory chapter in order to clearly locate the problems discussed in this work.

Note that Part II is inspired from our two articles [BMR07] and [MB07], whereas Part III is based on our four papers [MR05b, MR05a, MMR05] and [GKM07]. Nevertheless, our discussions contain some added value compared to the articles, as we have put our considerations in the light of progressiveness.

From a methodological point of view, our research has always focussed on practical aspects of MCDA processes, on the central role of the DM and on computational facets of MCDA methods. This quite *pragmatic* perception of decision analysis is the general guideline of our work, and quite regularly we will discuss practical implications of the underlying theoretical developments.

To help the reader to get through this text, important concepts are put in the *margin* of the text. They allow to have a synthetic vision of each section and to quickly go back to previously introduced notions. *margin notes*

Finally, note that the numbers after the bibliographical entries indicate the pages on which the articles have been cited.

Patrick Meyer
Luxembourg, June 2007

Main personal references

The second and the third parts of this work are based on the following six articles. They represent our contributions to the field of Multiple Criteria Decision Analysis. Note that all the papers are either published, or in press.

[BMR07] R. Bisdorff, P. Meyer, and M. Roubens. Ruby: a bipolar-valued outranking method for the best choice decision problem. *4OR, Quarterly Journal of the Belgian, French and Italian Operations Research Societies*, 2007. in press.

> *The main concern of this article is to present and motivate the* RUBIS *method for tackling the choice problem in the context of multiple criteria decision aiding. Its genuine purpose is to help a decision maker to determine a single best decision alternative. Methodologically we focus on pairwise comparisons of these alternatives which lead to the concept of bipolar-valued outranking digraph.*

[MB07] P. Meyer and R. Bisdorff. Exploitation of a bipolar-valued outranking relation for the choice of k best alternatives. In *FRANCORO V / ROADEF 2007: Conférence scientifique conjointe en Recherche Opérationnelle et Aide à la Décision*, pages 193–206, Grenoble, France, 20-23 February 2007.

> *This article presents the problem of the selection of k best alternatives in the context of multiple criteria decision aiding. We situate ourselves in the context of pairwise comparisons of alternatives and the underlying bipolar-valued outranking digraph. We present three formulations for the k-choice problem and detail how to solve two of them directly on the outranking digraph.*

[MMR05] J-L. Marichal, P. Meyer, and M. Roubens. Sorting multiattribute alternatives: The Tomaso method. *International Journal of Computers & Operations Research*, 32:861–877, 2005.

> *In this article we analyse the ordinal sorting procedure Tomaso for the assignment of alternatives to graded classes and we present a software based on this procedure. We illustrate it by two examples, and do some testing in order to show its usefulness.*

[MR05a] P. Meyer and M. Roubens. Choice, ranking and sorting in fuzzy multiple criteria decision aid. In J. Figuera, S. Greco, and M. Ehrgott, editors, *Multiple Criteria Decision Analysis: State of the Art Surveys*, pages 471–506. Springer, 2005.

> *In this chapter we survey several approaches to derive a recommendation from some preference models for multiple criteria decision aiding. We detail a sorting procedure for the assignment of alternatives to graded classes when the available information is given by interacting points of view and a subset of prototypic alternatives whose assignment is given beforehand. A software dedicated to that approach (Tomaso) is briefly presented. Finally we define the concepts of good and bad choices based on dominant and absorbant kernels in the valued digraph that corresponds to an ordinal valued outranking relation.*

[MR05b] P. Meyer and M. Roubens. On the use of the Choquet integral with fuzzy numbers in multiple criteria decision support. *Fuzzy Sets and Systems*, pages 927–938, 2005.

> *This paper presents a multiple criteria decision aiding approach in order to build a ranking and recommend a choice on a set of alternatives. The partial evaluations of the alternatives on the points of view can be fuzzy numbers. The aggregation is performed through the use of a fuzzy extension of the Choquet integral. We detail how to assess the coefficients of the aggregation operator by using alternatives which are well-known to the decision maker, and which originate from his domain of expertise.*

[GKM07] M. Grabisch, I. Kojadinovic, and P. Meyer. A review of capacity identification methods for Choquet integral based multi-attribute utility theory: Applications of the Kappalab R package. *European Journal of Operational Research*, 2007. in press.

The application of Multi-attribute Utility Theory whose aggregation process is based on the Choquet integral requires the prior identification of a capacity. The main approaches to capacity identification proposed in the literature are reviewed in this paper and their advantages and inconveniences are discussed. All the reviewed methods have been implemented within the Kappalab R package. Their application is illustrated on a detailed example.

Part I

Multiple Criteria Decision Analysis and progressiveness

Part I: Multiple Criteria Decision Analysis and progressiveness

S'il n'y a pas de solution, c'est qu'il n'y a pas de problème.

Les Shadoks, Jacques Rouxel[2]

Abstract

The first part of this work is dedicated to a general and intuitive introduction to Multiple Criteria Decision Analysis, and in particular to a concept that we call *progressiveness*. The latter is mainly a framework which guides the analysis in a particular way, which allows to extract intermediate conclusions requiring further interactions by the stakeholders of the process.

Consequently, the purpose of this part is to clearly situate and delimit the subject of this work and to prepare the reader for the following two parts by introducing fundamental concepts and the notation.

The notion of progressiveness can be found in various resolutions of real-world decision problems. In order to intuitively assess the meaning of this concept, let us present right away a short example which will position our discourse for the reader.

> **Example** *Imagine that a company searches for an appropriate candidate for a newly opened position. Such a recruitment process is generally performed in multiple steps, where at each stage, the set of applicants is more and more reduced. This possibly begins with a filtering process solely based on information from the candidates' résumés. Then, the remaining applicants are interviewed via phone and a second selection is made. Finally, a very small number of persons are asked to come to the company in order to have personal interviews with the head of the company.*

This example is quite explicit and shows how the final best applicant for the position is selected in a progressive process. At each intermediate stage, a certain number of candidates are eliminated on the basis of the information available at that moment. Note that this example will be further detailed later in this

[2] If no solution exists to a problem, there may be no such problem.

part.

This first part is divided into two chapters. The first one is a brief introduction to Multiple Criteria Decision Analysis and the approach to decision analysis which will guide us through this work. It is based on a succinct bibliographical review of literature on decision analysis. In Chapter 2, we introduce the concept of interactivity, define progressiveness and detail their implications on the decision analysis and the type of recommendation generated by such a process. These original reflections represent our personal contribution to the first part of this work.

Chapter 1

On Multiple Criteria Decision Analysis

Contents

The goal of this first chapter is to delimit the sphere of our activity and to situate our discourse in a particular scientific context, namely Multiple Criteria Decision Analysis (MCDA). Our main concern is to present the philosophical environment to which we adhere and which will represent the framework of this work.

In the first section we introduce MCDA and specify the terminology used in the sequel. Then, in Section 1.2 we define some conventions and concepts common to the three parts of this work. Finally, in Section 1.3 we present two distinct trends of thought in MCDA and briefly show how they can be brought together in a common framework.

1.1 A brief introduction

It is quite common among scientists to consider that mathematics should be used to serve human cognition in its broadest sense. It is therefore very tempting for a scientist to model real world problems by some strong logical principles in order to describe and explain them, or even to forecast future events. This way of thinking may be very applicable in *hard*[1] sciences, but it is less relevant in human sciences. Indeed, as soon as human behaviours or decisions have to be modelled, mathematical descriptions often run up against their irrationality.

In particular, when it comes to making decisions, human intervention often represents a major part of the difficulties which may be encountered. Therefore, in many practical situations, it is advisable to resolve these decision problems via a *scientific preparation* called *decision analysis* .

decision analysis

Such a decision analysis process requires in general at least two actors. On the one hand there is the so-called *decision maker* (DM) which is a person who will take the responsibility for the decision act. He furthermore bears certain values, priorities and *preferences* related to the particular decision problem. On the other hand there is an individual, who will facilitate the decision analysis process by investigating thoroughly the underlying problem. He is often called the *analyst* . In general, his task may be very vast and time consuming and may include different steps, as for example a clear formulation or a rational structuring of the problem. Furthermore, the analyst's work may not be accomplished without a central task, which consists of interactions and discussions with the stakeholders of the decision process. Note that in practise both actors could be groups of persons. In this study we nevertheless make the hypothesis that the DM can be represented by a single person. In particular this implies that the search for a preferential consensus among a set of DMs is not our concern. Finally, it may also happen that both stakeholders are represented by a single person who will play both the role of the DM and that of the analyst.

decision maker

analyst

Roughly speaking, in *classical* OR, a given decision problem is formulated in analytical terms and solved by means of an optimisation method. The output of such an analysis is (if it exists) a solution called the *optimum*. The DM intervenes in the delimitation of the problem and the validation of the solution. Note that the latter task is not trivial and is an essential part for the implementation of a solution in practice. In decision analysis, the DM is more involved in the possibly multiple stages of the resolution. In particular, the decision analysis methodology requires that the *preferences* of the DM are correctly modelled. This fundamental step involves that the consequences of each of the possible decision acts are thoroughly analysed in collaboration with all the stakeholders of the decision process. Therefore, in decision analysis, the DM has a central and paramount position.

preferences

One of the main challenges in real world decision problems is the multidimensional nature of the potential decision actions. Very often, even the

[1]The term *hard* is used here without any derogative meaning for non-hard sciences.

apparently simplest decisions that we have to make on a regular basis imply multiple preference dimensions. Consider for example the following classical problem.

> **Example 1** *On the one hand Sophie prefers restaurant a to b because its cook is more famous, but on the other hand she prefers restaurant b to a because the prices of its dishes are cheaper.*

Sophie's decision will depend on her current mood and financial situation and she might choose either a or b. We can see via this short, but very demonstrative example, that the two selected preferential dimensions are conflicting and that the preferences of Sophie (the DM) have a great influence on the final output. It has therefore been suggested to adopt a *multiple criteria* approach, based on a multidimensional description of the potential decision actions (see for example [Roy87a]). This multidimensional vision of the problem linked to a decision analysis approach defines the so-called *Multiple Criteria Decision Analysis* .

multiple criteria

MCDA

In short, MCDA's general aim is to help a DM to prepare and make a decision and to study decision problems where more than one point of view has to be considered. Its objective being not to force a decision at any cost, MCDA can range from a rational structuring of the decision problem to the elaboration of a recommendation.

From a historical point of view, the roots of MCDA go back at least to the 18th century, where the Marquis de Condorcet has been the first to systematically apply mathematics in social sciences. In 1785 he wrote the *Essay on the Application of Analysis to the Probability of Majority Decisions*, which deals with decision making in presence of multiple voters. The very foundations of MCDA have however been laid around the middle of the 20th century with Samuelson's theory of revealed preferences [Sam38], the gradual beginning of game theory [vM44], the emergence of social choice theory [Arr51] and the interest in psychological and mathematical aspects of decisions [LR57, Fis70].

In the late 50s, an important step towards pragmatic foundations of decision analysis is made by Simon's *bounded rationality* theory [Sim57]. It states that in real-life decision problems, different factors limit the extent to which a DM can make a fully rational decision. Therefore he only owns bounded rationality and he will choose an option taking into account the limitations of both knowledge and cognitive capacity. Strictly mathematically, this decision might not be optimal, but it will tend to make the DM *happy enough*. This vision underlies the whole discourse presented in this work.

bounded rationality

By the end of the 60s, first methods to solve multidimensional decision problems start to appear. In 1968, Roy inaugurates the branch of outranking methods [Roy68] whereas in 1976, Keeney and Raiffa broadened value theory to the multidimensional case [KR76]. These two distinct trends of thought yield two different methodological conceptions: the so-called European and American schools.

From a philosophical point of view, the European and the American school are often said to be conflictual in the way they conceive the output of the decision analysis. Indeed, it is quite widely accepted that the former's goal is to give a *recommendation* (*decision aiding*) , whereas the latter seeks to approach an optimal solution (*decision making*) . However, as we will see later, the practice of methods issued from both schools of thought shows that this mutually exclusive vision is too restrictive (see also [SL03]).

decision aiding
decision making

From a methodological point of view, there are important differences in the tools generated from both trends. On the one hand, the European school developed around discrete methods via outranking relations, where the recommendation is built upon pairwise comparisons of the different options [Roy68]. On the other hand, the American school grew around utility and value methods to obtain a total comparability of the options [KR76].

At this point, it is useful to differentiate between MCDA *approaches*, *methods* and *models*. An MCDA approach is a general framework which guides the plan of attack to a given decision problem. It underlies a certain number of logical, pragmatical and philosophical principles which endow it with a conceptual coherence. An MCDA method is situated within one or more particular approaches and it is a regular and systematic way of dealing with a decision problem. Finally, an MCDA model is a mathematical representation of a decision problem. In short, an MCDA method implements an approach and leads to a model. We call a *recommendation* the output of an MCDA method.

approach
method
model
recommendation

In the literature on decision theory it is common to differentiate between four types of approaches (see for example [Roy93] and [BRT88]): normative, descriptive, prescriptive and constructive. The differences lie in the signification of the model which is built, in the way the model is obtained, and in the interpretation of the results which are presented to the DM [DT04, BMP+06].

The objective of *normative* decision analysis approaches is to derive models from norms and standards which are set up beforehand and which are commonly accepted. Such models aim at being universal in the sense that they apply to any DM who wants to act in accordance with the underlying rationality (see for example [vM44], [LR57] or [Wak89]). The goal of *descriptive* approaches is to infer models from the observation of how DMs behave when confronted to a certain and precise decisional problem. These models are then applicable to any DM who has to face a similar situation (see for example [Sch88]). *Prescriptive* approaches try to unveil models for a given DM based on his system of values and on preference-related answers. Such models are not intended

normative
descriptive
prescriptive

to be general and are only applicable in the context from which they originate (see for example [Roy85] or [BS01]). Finally, *constructive* approaches build models based on the preferences of a particular DM in a precise decision problem. In this case, the interaction between the analyst and the DM helps to construct and uncover his preferences and has therefore a significant influence on the final output (see for example [LPB83] or [Ros89]).

constructive

In short, models built in a normative context are based on generally accepted norms whereas in descriptive approaches, they are established on empirical observations. The purpose of models induced in prescriptive approaches is to discover a system of values of a DM which exists prior to the decision analysis whereas in constructive approaches these preferences tend to be constructed simultaneously with the model.

Note here that a decision analysis method may in practice belong to more than one approach. It might indeed be advisable in certain situations to adopt different aspects of different approaches in the resolution of a decision problem. As stated in [DT04], it can be interesting, for example, when adopting a prescriptive construction of the model, to impose some rationality principles issued from a normative approach in order to facilitate the dialogue with the DM and to allow to draw strong conclusions.

In this work we focus on what we call *progressive interactive* decision analysis methods. As we will show, these methods fall within the framework of constructive approaches. Their goal is to obtain a recommendation by proceeding in steps and steadily by increments via recurrent interactions with the DM.

Roy [Roy85] has stated that the objective of an MCDA is to solve one of the following four typologies of problems (or *problematiques*): determine one alternative considered as the best one (*choice*), assign each decision option to a clearly defined ordered category (*sorting*), rank the alternatives from the best to the worst one (*ranking*), describe the options and their impact in a formalised way (*description*). In this work, we present our contributions to the first three formulations. Besides we detail a further proposal that we call the k-*choice* decision problematique (also called *portfolio* problematique in [BS03]). It is an extension of the standard choice problem to the determination of k best alternatives ($k > 1$).

problematique
choice
sorting
ranking
description
k-*choice*

1.2 Common definitions and concepts

In this section we present some of the conventions that we use throughout this work. Note that we only introduce concepts common to the following two parts

Restaurant	Cook	Price (€)	Cuisine
a	very famous	100	Italian
b	not at all famous	80	Italian
c	little famous	50	Luxembourgish
d	little famous	100	French

Table 1.1: Sophie's selection of restaurants

and that further notions are defined as we go along.

alternative

The starting point of our discussion is a finite set X of $p > 1$ potential decision objects (also called *alternatives*) . They represent the possible options on which the DM has to make his decision. As our discourse is situated in a multidimensional framework, these alternatives are evaluated on a finite set $N = \{1, \ldots, n\}$ of $n > 1$ *attributes*. Let $g_i : X \to X_i$ be a descriptor which allows to assess the alternatives on attribute i of N, where X_i is the set of levels of the associated scale. It is now possible to represent an alternative x of X by its corresponding *evaluation profile* $(g_1(x), \ldots, g_n(x))$. To illustrate these different concepts, let us return to the short example of restaurant selection.

attribute

> **Example 1** (*continued*) *Recall that Sophie has to select a restaurant. Let us imagine that she makes a first selection for which she retains four restaurants (= X). She has decided to evaluate each of them on three attributes (= N), namely the reputation of the cook, the average price of a meal and the type of cuisine. Table 1.1 summarises how she evaluated the four restaurants a, b, c and d on the three attributes. For attribute "type of cuisine" ($i = 3$), the set X_3 is equal to $\{French, Italian, Luxembourgish\}$. Restaurant c can for example be represented by its evaluation profile $(little\ famous, 50, Luxembourgish)$.*

Roy [Roy85] underlines that the set X has in a first step to be clearly identified and validated by the DM and that the attributes represent all the dimensions that have consequences on the objective of the decision analysis. As we will show later, in particular, two alternatives having the same evaluations on all the selected attributes should be considered as indifferent.

criterion

A *criterion* is the combination of an attribute with supplementary information derived from the DM's preferences. For short, it is a numerical function which represents the attribute together with some of the DM's preferences, as, for example, an order of the different evaluation levels. Nevertheless, concerning this point, both previously cited methodological schools diverge in the way these preferences are put into practical effect.

Let us present on the restaurant selection problem the concept of preferential information.

> **Example 1** (*continued*) *Imagine that Sophie expresses the following preferences concerning the restaurant selection problem: her preference goes for very famous cooks, she would like to pay as little as possible and she prefers French cuisine to Italian one, which in turn she prefers to Luxembourgish meals.*
>
> *Considering these observations, one can easily check that there exists no* optimal *restaurant in X, one that would* dominate *all the other ones. As a consequence, an MCDA on this problem should reveal a* compromise *alternative, satisfactory for Sophie.*

1.3 Two methodological philosophies

In this section, our objective is to present how both methodological philosophies, the European and American schools, can be brought together in the framework of constructive approaches. We start by presenting synthetically the methodological grounds of the two trends of thought and show in Section 1.3.3 how methods issued from both schools can be considered in constructive approaches.

Note that, as classically done, the asymmetric part of a binary relation $\succeq$ will be denoted by $\succ$ and its symmetric part by $\sim$.

1.3.1 Building and exploiting an outranking relation

The objective of outranking methods is to build a relation on the alternatives, called the *outranking relation*, and to exploit it in order to solve one of the MCDA problematiques defined in Section 1.1. This relation then represents the preferences of the DM based on pairwise comparisons of the elements of X and is not necessarily transitive or complete.

One of the particularities of outranking methods is that the relation built on the set X permits three types of comparisons of alternatives, namely preference, indifference and incomparability. According to Roy [Roy90], they allow to represent hesitations of the DM which may result from phenomena like uncertainty, conflicts or contradictions.

Bouyssou [Bou90] defines a *criterion* as a real valued function on the set X of alternatives, such that it appears meaningful to compare two alternatives x and y according to a particular point of view on the sole basis of their two evaluations.

As mentionned in Section 1.2, we denote by $g_i(x)$ the performance of alternative x on criterion[2] i of N. Such a *performance function* g_i ($i \in N$) can be regarded, without any loss of generality, as a real-valued function s.t. (see [Roy90]):

$$\forall x, y \in X,\ g_i(x) \geq g_i(y) \Rightarrow x \text{ is at least as good as } y \text{ on attribute } i.$$

thresholds

Furthermore, to each preference dimension are associated variable preference, indifference and veto *thresholds*. For a quantitative attribute, the set of values taken by the criterion function can be identical to X_i. For qualitative criteria however, the values taken by the criterion function have to be chosen carefully. In both cases, these values can only be interpreted when linked to the different thresholds in the pairwise comparison phase.

criterion

For short, in the context of outranking methods, we call a *criterion* the association of an attribute i of N with the corresponding criterion function g_i and the different preferential thresholds.

consistent family of criteria

In this framework, the set N of criteria is supposed to be *consistent* . Consistency is defined via three properties: exhaustiveness, coherence and irredundancy. The family of criteria is *exhaustive*, if all the consequences which allow the preferential comparison of any two alternatives have been taken into account. To illustrate this, let x and y be two alternatives of X. If $g_i(x) = g_i(y)$ for each i of F, then necessarily x has to be considered as *indifferent* from y. If not, certain points of view have not been considered in the family of criteria. Second there must be a *coherence* between local preferences modelled at the level of the individual criteria and overall preferences modelled over the whole family N. If $g_i(x) = g_i(y)$ for each i of $F \setminus \{k\}$ and $g_k(x) \geq g_k(y)$, then necessarily x has to be considered as at least as good as y. Finally the criteria should be *irredundant* in the sense that the family of criteria is considered as minimal with respect to the preceding two conditions (see [Roy85, Bou90, Bis02]). This implies in particular that there should not be more criteria in N than strictly necessary. From a technical point of view, Roy and Bouyssou have described a set of operational tests which allows to check the consistency of a family of criteria (see [RB93]).

outranking

Considering two alternatives x and y of X, an *outranking* S between x and y holds (xSy) if it is reasonable to accept, from the DM's point of view, that x is at least as good as y. From this definition it is easy to derive that x and y are considered as *indifferent* if simultaneously xSy and ySx, that an incomparability situation originates from the complete absence of outranking between x and

[2] In certain situations it may be necessary to join different attributes with the same preferential semantics into one criterion. This implies that there might not be a one-to-one correspondence between the attributes and the criteria. Nevertheless, as the construction of the family of criteria is not the topic of this work, and to avoid confusions, we denote the criteria by the same labels as the attributes.

y (neither xSy nor ySx), and that x is strictly preferred to y if xSy and not ySx.

The construction of the outranking relation is done via pairwise comparisons of the alternatives on each of the criteria. They are based on differences of evaluations which are then compared to preference, indifference and veto thresholds (fixed in accordance with the DM's preferences) in view of elaborating the outranking relation. An additive aggregation of such local relations is then performed via a weighted sum. This requires that to each criterion is associated its importance coefficient (or *weight*). Finally, this calculation produces a valued outranking relation on the set X which can then be seen as a valued digraph, called the *outranking digraph* .

weight

outranking digraph

Different ways of constructing the outranking relations have been proposed in the literature on MCDA methods. Among the most famous ones, one can find the ELECTRE-like methods (see for example [KR76, RB93] for their detailed description) or the PROMETHEE-like methods (see for example [BM02] for an extensive presentation).

The second step of an outranking method is to exploit the outranking digraph in order to solve one of the MCDA problematiques mentionned in Section 1.1. As the outranking relation is not necessarily complete or transitive, this task is in general quite difficult and requires a clear understanding of the semantics linked to the outranking relation.

In Part II we present the construction and exploitation of a particular outranking relation (called the *bipolar-valued* outranking relation) in order to solve the choice and the k-choice problematique. On the basis of the corresponding digraph, we show and pragmatically justify how to determine in a progressive manner the potential candidates for a choice recommendation.

1.3.2 Building and exploiting an overall value function

The goal of Multiattribute Value Theory (MAVT) [KR76] is to build a numerical representation of the preferences of the DM on X.

In other words, MAVT seeks at modelling the preferences of the DM, supposed to be a weak order, represented by the binary relation $\succeq$ on X, by means of an *overall value function* $U : X \to \mathbb{R}$ such that,

overall value function

$$x \succeq y \iff U(x) \geq U(y), \qquad \forall\, x, y \in X.$$

Note that the preference relation induced by such an overall value function is necessarily a complete weak order.

The overall value function U can be determined via many different methods, presented for example in [vE86, Chapter 8] in the context of an additive value function model. Ideally, such methods should consist in a discussion with the DM in the *language* of his expertise, and avoid technical questions linked to the model which is used.

Concerning the overall value function, a commonly used model is the *additive value function model*. In such a case,

$$U(x) := \sum_{i=1}^{n} w_i u_i(x_i), \qquad \forall\, x = (x_1, \dots, x_n) \in X,$$

marginal value function

where we write x_i for $g_i(x)$, and where the functions $u_i : X_i \to \mathbb{R}$ are called the *marginal value functions* and w_i is the *weight* associated to criterion i, $\forall i \in N$. As far as the marginal value functions are considered and depending on the selected MAVT model, for any $x \in X$, the quantity $u_i(x_i)$ is sometimes interpreted as a measure of the satisfaction of the value x_i for the DM.

Another model is the *weighted sum model*, which can be written as

$$U(x) := \sum_{i=1}^{n} w_i g_i(x), \qquad \forall\, x \in X.$$

criterion

From now on, in the context of MAVT, the term *criterion* is used to designate the association of an attribute $i \in N$ with the corresponding marginal value function u_i.

Choquet integral

As we will see in Part III, the average value function model is only applicable if *mutual preferential independence* (see e.g. [Vin92]) among the criteria can be assumed. This independence may however be hardly verified in many real-world applications. It has therefore been suggested to consider more complex models, as for example the *Choquet integral* , which can be considered as a natural extension of the weighted sum model (see for example [Gra92, Mar00a, LG03]).

trade-off

The marginal value functions and the parameters of the overall value function are often determined together. In the case of the weighted sum, this amounts to determining the importance of the n criteria (as *trade-offs* that a DM would be willing to make). In such an additive situation, it might be realistic to ask the DM to provide such a weight vector. In more complex models, however, the number of parameters can become huge and their meaning be unclear for a DM.

preference elicitation

In this latter case it might be advisable to determine these parameters directly from the DM by a proper questioning, called *preference elicitation*. Hence the DM can for example provide preferential information on a (small) subset of the set X of alternatives or on interactions between the criteria. The precise form of these prior preferences are discussed in Chapter 7.

Once the overall value function has been determined, each element of X is associated with a real number. As a consequence, the alternatives become comparable, and, as we will show in Chapter 8, it is possible to solve quite conveniently the different problematiques presented earlier.

1.3.3 MAVT and outranking methods in a constructive approach

As a researcher or practitioner of MCDA, quite frequently one may be confronted to the thesis that there exists a clear dichotomy between the *decision aiding* characteristic of the methods issued from the outranking school and the *decision making* feature of MAVT methods. In this section we show that this separation can be overcome and that in practice, both trends of thought can be brought together in the framework of constructive approaches.

The thesis that the methods issued from the European MCDA school fit quite well in the context of constructive approaches has been thoroughly studied and motivated by Bouyssou and Roy (see for example [RB86, RB93]). It originates from the assumption that a DM's preferences are in general poorly formulated and can be variable over time and context (see also the concept of bounded rationality in [Sim79] and [Mar78]). The goal of decision aiding methods is therefore to build a new and continuously evolving model of reality. This *context-dependent representation* of the problem [Vin92] is consequently the result of a constructive elicitation process. It is important to note that the validation of such a model is in general hard to achieve, but is in practice done via its acceptance by the DM (see also [LMO83] for a discussion on model validation).

From the literature on problems solved by MAVT it appears that the practice of these methods also belongs to a constructive approach. In this context, Bouyssou and Roy show that the adoption of a constructive approach does not at all signify the rejection of MAVT methods [RB93, p. 595]. Similarly, the authors of [vE86] state that it is a mistake to think that a DM has *numbers in his head* which wait to be elicited. This clearly supports the idea that in an MAVT method, the preferences have to be determined in a constructive manner.

Practical applications of MAVT methods are presented in [Bel99] in the framework of *decision conferencing*. The latter is a collaborative way to support shared decision making problems where all the involved parties are gathered together to thoroughly discuss and analyse the problem. MAVT is used to systematically model the different views of the participants in order to enhance the understanding of the problem. From their implementation in real world problems, it appears that such methods fit in the category of constructive approaches. The elicitation of the preferences is indeed done interactively and in small steps, and it is not supposed that the preferences exist beforehand in the participants' minds.

In [SL03], the authors argue that in practice, due to their axiomatic foundations, certain MAVT-based models may sometimes not be well adapted to describe the preferences of DMs. Nevertheless, they state that it is still possible to consider those models as a *first order approximation* for a decision situation. The axioms can indeed quite easily be explained to a DM and should be used as guidelines for the elicitation of his preferences.

In practical applications, these underlying axioms could be viewed as restrictive locks which don't allow the DM to express his preferences freely. However, the thesis of Stewart and Losa [SL03] is that these theoretical foundations should provide a guidance in the analysis in order to avoid certain biases, or even to explicit the latter ones for further consideration.

Both trends of thoughts can also be reconciled in other fields than their usage in a particular decision analysis approach. Stewart and Losa [SL03] also compare their partially compensatory behaviours and the way they deal with incomparabilities. Nevertheless, these considerations would lead us too far away from one of the main objectives of this first part of our work, namely to discuss the interactive and progressive aspects of outranking and MAVT methods, which is the purpose of the following chapter.

Chapter 2

Progressive interactive Multiple Criteria Decision Analysis

Contents

The goal of this chapter is to give a clear definition of progressiveness in interactive MCDA methods and to explain and delimit the way we use it throughout this work.

In Section 2.1 we first focus on the interaction between the DM and the analyst. Then, in the second section, we introduce the concept of progressiveness, and outline its use in both methodological schools of thought.

Note that the subject of progressiveness is rediscussed later in Parts II and III after we have introduced the MCDA methods and tools issued from our research.

2.1 Interactivity

A dictionary definition of *interactive* is *capable of acting on or influencing each other*[1]. Furthermore, in the context of computer science, it is defined as *interacting with a human user, often in a conversational way, to obtain data or commands and to give immediate results or updated information*[2]. As we will show, these two definitions describe quite well the spirit of interactivity that we consider in this work.

In decision aiding, the term *interactive* is associated with a particular type of methods which are quite well known and documented in the literature (see for example [VV89] for an extensive overview). Each of them is based on various principles or philosophical assumptions.

interviewer
protocol

The general structure of such an interactive method is a dialogue between a DM and an *interviewer* (the latter can be the analyst or a software for example). This discussion follows in general a precise *protocol*, which ensures a coherent construction of the output and which depends on the method which is used. The answers of the DM generate stepwisely the desired model for the decision analysis. This process stops either if the DM is satisfied with the current response, or if the interviewer considers that a dead end has been reached and that further questionings will not improve the quality of the solution.

interactivity

In this work we will regard interactivity from a less specific point of view. In the sequel we consider *interactivity* in the process of the elicitation of the DM's preferences, in the sense of the previously mentionned dictionary definitions.

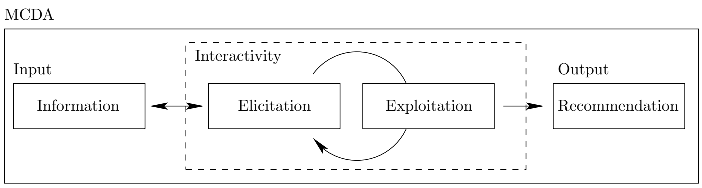

Figure 2.1: General scheme of an MCDA process

Figure 2.1 represents a standard decision analysis process (inspired from [GM98]), in which the interactive step is highlighted around the loop which includes a preference elicitation and an exploitation phase. As already mentionned in Chapter 1, any MCDA process applied to real world problems requires a strong interaction between the analyst and the DM. Therefore this

[1]The American Heritage Dictionary of the English Language, Fourth Edition.
[2]Dictionary.com Unabridged (v 1.1).

discussion phase is necessary for a proper determination of the DM's preferences as well as for an adequate construction of the recommendation.

The interactive loop can be regarded as an alternation of discussion and calculation phases. The DM is asked to express some information about his preferences, which are then incorporated in the chosen mathematical model. The output of this calculation step is then confronted to the DM and his preferences, which restarts the loop. Such a procedure clearly aims at supporting a *self-learning* of the DM's preferences and the interaction plays an active part in their development and evolution [Roy87b]. *self-learning*

Besides the resolution of a given decision problem, such an interactive procedure inside the process aims at various goals. First, in a constructive context, it is clearly meant to enlighten the DM, in the sense that the questioning should provide him with intellectual insight on a decision problem. This means that the ultimate goal is not solely to propose a recommendation, but also to let him actively participate in its construction by defining and eliciting his own preferences. As a direct consequence, the DM's confidence in the output of the process will be strengthened and he will more easily tend to validate it.

Second, in such interactive constructions, the preference elicitation algorithm stops either when the DM is satisfied with the output, or when the interviewer decides that it is not useful to continue the iterations. The latter case occurs when enough information has been gathered to provide a *good* recommendation, or when a dead end situation has been reached (further questionings would lead to contradictions, for example).

Third, such interactive questionings are generally performed in a way which avoids that too technical questions are asked to the DM. Ideally he should only be questionned via the *language* of his *domain of expertise*. In particular, a qualitative enquiry is preferable to a quantitative one (recall the statement from [vE86] that a DM *does not have numbers in his head*). *domain of expertise*

We suppose in this work that the *input information* (see left part of Figure 2.1) is given beforehand. This implies that the construction of the different alternatives, as well as the determination of the criteria which have to be considered in the decision, are performed in an earlier stage. These tasks are far from being trivial, require a lot of time and effort, but they are out of the scope of this work. Note that interactivity as considered here allows that this initial information is modified during the questioning process, in case the DM considers that it is necessary. *input information*

Let us now describe the progress of the interactive construction. Its starting point is an initialisation phase which requires that the DM expresses some *initial preferences* concerning the decision problem. They depend on the MCDA *initial preferences* method which is used and may be of different types. In this work, in the context of outranking methods, we require from the DM that he indicates the weights of the criteria and some indifference, preference and veto thresholds on the values taken by the criteria functions. In the context of MAVT methods, we ask the DM to provide a partial weak order over a subset of X (the *reference set*), and, if possible, some intuitions on the interaction between some of the criteria or their importance.

In case the DM has difficulties to express some of the required initial information, it is the analyst's role to help him to identify some of his preferences via a discussion. For example, in MAVT-based methods, it is possible to propose a certain number of alternatives (possibly fictitious) to the DM so that he can express some preferential statements. A quite convenient procedure is to present very similar evaluation vectors to the DM and to ask him to compare them. Ideally, these alternatives only differ on two dimensions, which allows to estimate a tradeoff between the concerned criteria and to build the marginal value functions (see for example [vE86, Chapter 8] for interactive elicitation procedures).

Once the initial preferences are implemented, a first run of the chosen MCDA method is performed. This allows to determine a first recommendation which is in accordance with the preferences expressed by the DM. In the context of MAVT, it may already happen at this stage that the initial preferences expressed by the DM are not compatible with the selected model. In such a case, different options can be considered: a revision of the DM's preferences guided by the axioms underlying the model, the choice of a more flexible (and thus more complex) model or the selection of a model which gives a satisfactory solution by violating some of the DM's expressed preferences.

The DM is then confronted to the output of the MCDA method, in order to continue the interactive construction of the recommendation and his preferences. If the output is not in accordance with his expertise, he can correct some of his statements or let his preferences evolve. The role of the analyst is then to adapt the parameters of the method in order to satisfy the DM. In outranking methods, this can be done by fine-tuning the weights or the thresholds to better fit to the DM's preferences. In MAVT methods, the analyst can, for example, ask the DM to add some further alternatives to the reference set which will enrich the numerical model. All in all, during these discussions, the goal of the analyst is to determine whether the DM is satisfied with the current recommendation or whether it needs to be adjusted.

The discussion with the DM and his confrontation to the output of a method allow him to obtain a synthesised view of the preferences which he expressed. Such a process therefore clearly fits in the context of constructive approaches. Two options appear at that moment: either he is satisfied with the result or he detects certain incompatibilities with his expertise. In the latter case, he may have identified new preferential information which he can inject in the method. By proceeding this way he can stepwisely identify and construct his preferences.

It is important to mention here that at any step of this discussion, previous affirmations may be *revoked*. Indeed, if the DM detects that his current perception of his preferences is not compatible with earlier statements, he can revoke them and replace them at any time with corrected information. In particular, in Figure 2.1, one can see that it is always possible to leave the interactive loop to the left, in order to modify or update the input information. In practice this is done if a new option is identified during the interactive questioning, if the evaluations of an alternative require readjustments or if an existing decision action becomes obsolete. *revocability*

As already explained, in the context of MAVT methods, a non negligible issue which may occur at any of the phases, is an incompatibility of the preferences expressed by the DM with the chosen MAVT model. Nevertheless, this should not be considered as a restrictive blockage, but rather as a guidance for the constructive determination of the DM's preferences [SL03]. It is clear that if these violations cannot be overcome, the chosen MAVT model is too restrictive. In such a case, it should either be enriched or the DM must accept approximate solutions which violate some of the preferences that he expressed. If neither of these two solutions is acceptable for him, it might be advisable to use another MCDA method.

The interactive process is in general stopped if the DM is satisfied with the output of the method. At that stage, a major benefit for the DM is that he has probably gained further insight on his preferences in the context of the considered decision problem. The output of the method is then a recommendation which is in accordance with the chosen decision analysis problematique.

2.2 Progressiveness

In this section we introduce the concept of progressiveness in MCDA. Without detailing a particular protocol, we indicate the major guidelines which should underlie a progressive method. We also discuss the consequences and properties of such processes.

Similarly as for interactivity, it is a general framework which guides a decision analysis in a particular way. To start, let us once again cite a dictionary definition which underlines well the meaning of progressiveness which we would like to point out. The adjective *progressive* can be defined by *moving forward;*

advancing; proceeding in steps; continuing steadily by increments[3]. Thus, one can already foresee that progressive MCDA methods will fit well in the context of constructive approaches.

progressiveness

A *progressive MCDA method* is an iterative procedure which presents intermediate recommendations to the DM which have to be refined at a further step of the MCDA. The concept of progressiveness therefore intervenes in practice in the determination of the recommendation, rather than in the elicitation of a DM's preferences.

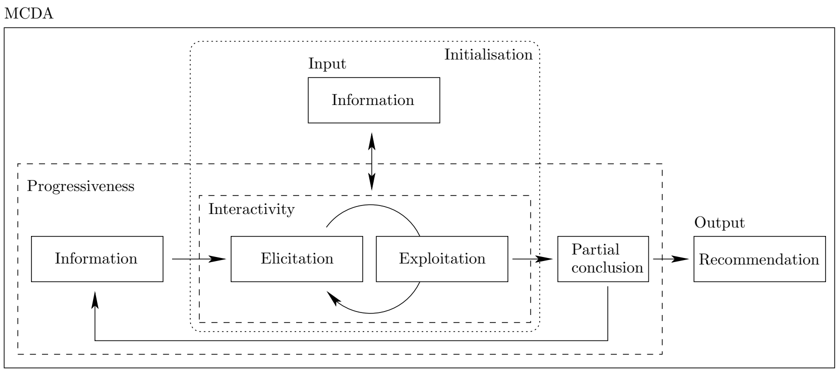

Figure 2.2: General scheme of a progressive MCDA process

partial conclusions

Figure 2.2 represents progressiveness in an MCDA method. As one can see, it is a framework around interactivity, which controls the construction of a final recommendation via intermediate *partial conclusions*. The entry point to a progressive method is an initialisation phase which generates a first recommendation in accordance with the DM preferences, via an interactive questioning (see Section 2.1). If the DM is not completely satisfied with this output, a progressive process can then be initiated to further refine the partial conclusion.

robustness

Note here that progressiveness should not be confused with *robustness*. In the literature, robustness may have several meanings, such as flexibility, prudence or stability (see [BMP+06, Section 7.5] for several definitions). Intuitively, robustness leads to consider several different sets of values of the parameters of an MCDA method and to look for recommendations which are *good* for almost all sets of values [BMP+06].

Before coming to a detailed discussion on progressiveness and its consequences on an MCDA, let us start by presenting a short example which should

[3] The American Heritage Dictionary of the English Language, Fourth Edition.

illustrate the necessity of such methods in real-world applications. The example is situated in the choice problematique.

> **Example 2** *Consider the problem of the selection of a candidate for a position in a company. The human resources department of the company collects a huge amount of résumés and other publicly available information on potential candidates. The latter are evaluated on a certain number of criteria according to the general objectives of the company. After this first evaluation step, a filtering is performed to reject the largest possible amount of individuals and to retain only those among which the most interesting one is situated.*
>
> *These few applicants are then called, and a telephonic interview is performed by the head of the human resources department. Each of the individuals is then reanalysed, some of his evaluations are refined and possibly he is evaluated on some other supplemental criteria. Afterwards, a filtering is again performed.*
>
> *Finally, a very limited number of candidates is invited to the company for an interview with the head of the department proposing the open position. This final analysis, followed by a filtering step, is intended to unveil the best candidate among the remaining applicants.*

Algorithm 1, presented hereafter, summarises the general structure of a progressive MCDA method. As we will see, the algorithm shows that the final recommendation is constructed via *intermediate partial conclusions* which are analysed by refining the currently available information. We call X^i (resp. N^i) the set of alternatives (resp. criteria) at stage i of the progressive resolution. Note that in order to simplify our discourse, N^i also contains the supplementary information linked to the criteria, like importance weights, interactions or thresholds. We write R^i for the output of the MCDA at stage i. As long as the final recommendation is not determined, the R^is are called *partial conclusions*.

As already discussed for Figure 2.1, the starting point is a first interactive loop which allows to generate an initial recommendation R^0. During this interactive loop, the initial sets X and N may be modified by the DM in order to allow him to better express his preferences. Therefore the output of this first stage is given by the triple (X^0, N^0, R^0). If the DM is satisfied with the recommendation R^0, the MCDA process can stop here and there is no necessity for a progressive resolution.

If the DM is not satisfied with this first output, then the current recommendation is reanalysed in view of refining it further to satisfy the DM. The precise form of these complementary analyses may depend on the type of problematique and the mathematical model which are chosen. Nevertheless, very generally, these improvements may take the following form:

- Enrich the currently available data by further information;

Algorithm 1 Progressive MCDA method

Input: (X, N)

1. *Initialisation:*
   - $i \leftarrow 0$
   - Interactive elicitation of the DM's preferences and determination of a recommendation R^0
   - *Output:* (X^0, N^0, R^0)
2. *Progressive Loop:*

   While the DM is not satisfied with R^i, do

   1. *Interactive elicitation*
      - Enrich (X^i, N^i) by exploiting currently available information via an interactive questioning of the DM
      - $i \leftarrow i + 1$
   2. *Partial conclusion*
      - Based on (X^{i+1}, N^{i+1}) generate a partial conclusion R^{i+1}
   3. *Output:* $(X^{i+1}, N^{i+1}, R^{i+1})$

Output: The final recommendation R^k.

- Solve issues related to missing data;
- Focus on a subset of alternatives to refine the recommendation.

Such supplementary investigations are then combined with a new interactive loop which finally produces a new partial conclusion. The output of such a progressive loop is then the triple $(X^{i+1}, N^{i+1}, R^{i+1})$. If the DM is again not satisfied with the current recommendation R^{i+1}, the progressive loop is restarted.

This iteration is continued until the DM accepts the recommendation R^k as the final one. In practice this means that he is satisfied with R^k or that he can get along with it on his own to elaborate his final decision.

Note that the different elements of this process can easily be identified in Example 2 which uses an initialisation phase (selection via the résumés) and two times the loop of the algorithm.

prudence

The use of a progressive decision analysis method can be motivated by (at least) three reasons. First, it can be justified by *prudence*. As the ultimate recommendation does not necessarily have to be reached in one step, each partial conclusion can focus exclusively on the information available at that moment. This means that at each intermediate stage, only strongly motivated affirmations are made. In the context of outranking relations, this could for example mean that the comparability of the alternatives is not forced at any cost.

Second, progressiveness is also motivated by *economical constraints*. Indeed, at a given moment, only limited financial or temporal resources may be available. In Example 2 this is clearly demonstrated by the progressive approach to the candidate selection. It would indeed not be possible to have personal interviews in the company with all the initial applicants as such consultations are in general very costly and time demanding.

economical constraints

Third, as the DM's preferences, as well as the final recommendation, are actively constructed via small steps, such methods are motivated by a constructive approach to the problem.

Note also that progressiveness is a methodological context which is well adapted to deal with *missing values*. This does not mean that progressiveness allows solving problems involving missing values by default. It is rather a framework which allows to postpone the issue of incomplete information to later stages. In such a case, it is obviously necessary that the algorithm underlying the MCDA is able to deal with such issues. In a progressive process, such incomplete information can then be completed, if necessary, as the decision analysis process goes along. This prudent way of dealing with these difficulties allows to make no hypotheses on the missing data (approximations, default values) and to consider them as such until they can be determined.

missing values

Finally, it is important to note that if progressiveness is adopted for an MCDA problem, its resolution will be guided by a *no return* policy. We call this characteristic of progressiveness *irrevocability* of previous partial conclusions. Example 2 clearly underlines this feature, as at each step, a certain number of candidates are put aside, for good reasons, and their rejection is never reappraised at later stages of the analysis.

irrevocability

Irrevocability is a consequence of prudence and the limited economical resources pointed out above. To illustrate this, consider again the choice problematique. As we will detail hereafter, the goal is be to narrow at each step the current set of potential alternatives for the choice. If revocability was allowed, nothing would guarantee to converge to the desired solution. Furthermore, in a situation involving limited time or money resources, such reappraisals of previous conclusions are certainly not appropriate.

Nevertheless, in practical situations, this irrevocability condition might be too strong. One could indeed imagine a situation, where during the progressive process, new *potentially good* decision actions appear, which the DM would like to include in the current set of alternatives. Such a manipulation does not strictly fit in a progressive context as described above. Therefore, it should be performed with much care, and the analyst must control that it does not generate contradictions with affirmations of previous steps of progressiveness.

Let us now turn to a more detailed description of progressiveness in the two methodological trends for MCDA.

2.2.1 Progressiveness in outranking methods

Recall that in outranking methods, the relation built on the set X permits three types of comparisons of alternatives, namely preference, indifference and incomparability. The latter is a strong argument for the use of progressiveness in such methods. As already mentioned, progressiveness finds a very natural context of application in the choice problematique. Hereafter we describe what the three stages of Algorithm 1 could be in this context. Other formulations could of course be found. Note that an extended study of a progressive method for the choice problematique is the subject of the second part of this work.

The first stage of the algorithm consists of the analysis of the initial set of potential candidates for the choice. During this phase, the preferences of the DM have to be determined and implemented in the choice algorithm. A first recommendation R^0 is then generated by removing the alternatives which can obviously be put aside, considering the currently available information. This has to be performed via clearly defined rules which justify this cutting off. In that case, R^0 should contain hardly comparable alternatives.

If the DM is not satisfied with R^0, the progressive analysis can be started. The alternatives of R^0 (or X^0) are then reanalysed and the input data for the next loop is adjusted or completed with additional information. One goal of this step is to try to revoke the hard comparabilities between the potential choice candidates.

During this step, the data is enriched with new information from the DM and finally a partial conclusion R^{i+1} is generated. If it satisfies the DM, the process can be stopped. Else the progressive loop is reinitiated and the current recommendation R^{i+1} is refined.

monotone reduction

Note that due to the property of irrevocability discussed earlier, the initial set X of alternatives is *monotonically reduced* after each loop of the progressive method and that ideally, the final recommendation may consist of a single alternative.

As already said, at a given stage, the partial conclusion R^{i+1} should contain alternatives which are hardly comparable. The reasons for such difficult comparabilities can be of various types, as, for example, the existence of missing values in the evaluation vectors of the alternatives. In such a case, the objective of next phase is to try to solve this issue by investigating further the problem in order to obtain this previously unavailable information.

In such a choice context, we clearly have $R^{i+1} \subset R^i$ and $X^{i+1} \subset X^i$.

Progressiveness may also be applicable in the framework of the ranking problematique. In general, outranking methods generate a ranking of the alternatives containing incomparabilities and indifferences (a partial weak order). Similarly as in the choice problematique, the DM may not be satisfied with such a ranking, and in a further step of the analysis, he may be interested in resolving the incomparabilities via further investigations to create a *less partial* order. This issue is nevertheless out of the scope of this work. Note that in [Lam07]. the author develops a progressive process to reach a compromise ranking from multiple rankings originating from a group of DMs. The goal of such a method, based on *prudent orders*, is to eliminate incomparability situations in the compromise ranking in a progressive manner through interactions with the group of DMs.

Finally note that progressiveness could also be used to discriminate between alternatives which are considered as indifferent, at a given stage of the process. In such a situation, a deeper analysis may add additional information on the pairwise comparisons which will create a *less weak* order.

2.2.2 Progressiveness in MAVT methods

The objective of MAVT methods is to build an overall value function which is a representation of a weak order over the alternatives of X. Hence in this case, the objective of progressiveness cannot be to overcome incomparabilities between alternatives. Nevertheless, the use of progressiveness in MAVT has a similar goal as in outranking methods, namely to make the recommendation more accurate.

The DM may consider that the ranking generated by such a method might be too rough and that too many alternatives can be considered as equivalent. Such a situation can originate from two non-exclusive observations. First the overall value function generated equivalence classes by assigning the same overall values to many alternatives (real indifferences). Second, if the chosen model allows to give a meaning to the difference of overall evaluations, the alternatives having very close values may be considered as indifferent. In such a case, it might be interesting to focus on those problematic elements of X in order to try to discriminate them further, if necessary. In practice this amounts to establishing that their values are ordered in a definite way.

In case the input data contains nonstatistical uncertainty, the output of an MAVT method can be an interval order where the overall evaluations of the alternatives also suffer from impreciseness. In such a case, progressiveness may be used to focus on the evaluation of a certain number of alternatives in order to make them more precise. As a consequence, such a progressive refinement would allow to further discriminate between the selected alternatives in the ranking.

At this point it is again possible to clearly observe the distinction between robustness and progressiveness. Indeed, robustness does not aim at searching further information on the problem to determine a more accurate output. Its goal is rather to obtain a recommendation that could be justified by any possible sets of input values.

In Part III we will represent impreciseness in the evaluations of the alternatives by so-called *fuzzy numbers* and we will detail how progressiveness can be applied in such a framework.

In the particular case of the choice problematique, progressiveness allows again to focus on a subset of alternatives (namely the first positions of the ranking) to determine the single best one.

Note that in Section 8.3 of Chapter 8, after having discussed the resolution of different problematiques, we will briefly return to progressiveness in MAVT methods.

Part II

Progressive choice methods in an outranking framework

Part II: Progressive choice methods in an outranking framework

On n'est jamais aussi bien battu que par soi même.

Les Shadoks, Jacques Rouxel[4]

Abstract

The second part of this work is situated in a framework of outranking methods, requiring a pairwise comparison of the alternatives. More precisely we focus on the choice problematique and its extension to the determination of $k > 1$ best alternatives. Note beforehand that our developments are based on the pioneering works of Roy and Bouyssou on these subjects, but quite quickly we will move away from this path by reformulating the foundations of the resolution of the choice problematique.

This part is divided into three chapters, each of them presenting some personal contributions. The first chapter introduces some preliminary considerations on what we call the *bipolar-valued credibility calculus*. It represents the roots of the developments concerning the choice problematique.

Then, in Chapter 4 we revisit the choice problematique and detail the Rubis method for the determination of a choice recommendation. The considerations of these two first chapters are based on our article [BMR07].

Finally, in the third chapter, we discuss the k-choice problematique and its various formulations. Two of them are solved by a modification of the original outranking relation and by means of the Rubis algorithm. Note that this chapter is inspired from our article [MB07].

[4] You are never so well beaten than when you are beaten by yourself.

Chapter 3

Preliminary considerations

Contents

The objective of this first chapter of Part II is to introduce fundamental concepts which we will use later in the description of our proposal (the Rubis method) to solve the choice problematique. As we are situated in a context where we perform pairwise comparisons of alternatives, we detail in this chapter the construction of an outranking relation.

The chapter is structured as follows. In Section 3.1 we introduce the so-called *bipolar-valued credibility calculus* which allows us to perform logical operations on propositions which are considered more or less credible by a DM. Then, in Section 3.2 we detail the construction of the outranking relation on the set of alternatives, based on the bipolar-valued credibility calculus. Finally, in the last section, we recall results from [BPR06] on the determination of kernels in bipolar-valued outranking digraphs.

Note that the developments of this chapter are inspired from our article [BMR07].

3.1 Bipolar-valued credibility calculus

Later in this work, we will detail the Rubis method for constructing *choice recommendations* and its extension to the k-choice problematique. In order to easily understand the considerations of the following chapters, we start in this section by establishing the backbone of the RUBIS method, namely the so-called bipolar-valued credibility scale, modelling the credibility of the validation of preferential statements.

Let ξ be a propositional statement like "*alternative x is a choice recommendation*" or "*alternative x is at least as good as alternative y*". In a decision process, a DM may either accept or reject these statements following his belief in their validation [Bis00]. Such a degree of credibility (or *credibility* for short) may be represented via a credibility scale $\mathcal{L} = [-1, 1]$ supporting the following semantics.

credibility

Let ξ and ψ be two propositional statements with which are associated credibilities r and $s \in \mathcal{L}$:

validation
non-validation
indetermined

(1) If $r = +1$ (resp. $r = -1$) then it is assumed that ξ is *clearly validated* (resp. *clearly non-validated*). If $0 < r < +1$ (resp. $-1 < r < 0$) then it is assumed that ξ is *more validated than non-validated* (resp. *more non-validated than validated*). If $r = 0$ then ξ could either be validated or non-validated, a situation we call *indetermined*.

(2) If $r > s$ then it is assumed that the validation of ξ is more credible than that of ψ (or that the non-validation of ψ is more credible than that of ξ).

disjunction
conjunction

(3) The credibility of the *logical disjunction* $\xi \vee \psi$ (resp. the *logical conjunction* $\xi \wedge \psi$) of these statements equals the credibility of the statement that is the most (resp. the less) credible of both, i.e. $\max(r, s)$ (resp. $\min(r, s)$).

negation

(4) The credibility of the non-validation of ξ equals $-r \in \mathcal{L}$, which also denotes the credibility of the validation of the *logical negation* of ξ (written $\neg\xi$).

Definition 3.1.1. *The credibility associated with the validation of a propositional statement ξ, defined on a credibility domain $\mathcal{L}$ and verifying properties (1) to (4) is called* a bipolar-valued characterisation *of ξ.*

It follows from Property (3.1) that the graduation of credibility degrees concerns both the *affirmation* and the *negation* of a propositional statement (see, e.g., [Win84]). Starting from +1 (certainly validated) and −1 (certainly non-validated), one can approach a central position 0 by a gradual weakening of the absolute values of the credibility degrees. This particular point in $\mathcal{L}$ represents an *indetermined* situation concerning the validation or non-validation of a given propositional statement [Bis00, Bis02].

Definition 3.1.2. *The* degree of determination *of the validation (for short* determinateness*) $D(\xi)$ of a propositional statement ξ is given by the absolute value of its bipolar-valued characterisation: $D(\xi) = |r|$.*

For both a clearly validated and a clearly non-validated statement, the determinateness equals 1. On the opposite, for an indetermined statement, this determinateness equals 0.

This establishes the central degree 0 as an important *neutral* value in the bipolar-valued credibility calculus. Propositions characterised with this degree 0 may be either seen as *suspended* or as *missing* statements [Bis02]. The credibility degree 0 represents a temporary delay in characterising the validation or non-validation of a propositional statement. In the framework of progressive decision aiding, this feature allows us to easily cope with currently indetermined preferential situations that may eventually become determined to a certain degree, either as validated or non-validated, in a later stage of the decision analysis process.

The following section introduces the concept of bipolar-valued outranking digraph which is the preferential support for the RUBIS choice decision analysis method.

3.2 Bipolar-valued outranking relation

Recall that $X = \{x, y, z, \ldots\}$ is a finite set of p alternatives evaluated on a finite, coherent family $N = \{1, \ldots, n\}$ of n criteria.

To each criterion j of N we associate its *relative significance weight* represented by a rational number w_j from the open interval $]0, 1[$ such that *weight*

$$\sum_{j=1}^{p} w_j = 1.$$

Besides, to each criterion j of N is attached a criterion function g_j, with values in $[0, 1]$, and which allows to compare the performances of the decision objects on the corresponding preference dimension (see Section 1.3.1 for further details).

Let $g_j(x)$ and $g_j(y)$ be the performances of two alternatives x and y of X on criterion j. Let $\Delta_j(x, y)$ be the difference of the performances $g_j(x) - g_j(y)$.

With each criterion j of N is associated a certain number of thresholds, which allow to represent a DM's preferences more accurately when comparing two alternatives. In this work we consider four such thresholds:

- an indifference threshold $q_j(g_j(x)) \in [0, 1[$;
- a preference threshold $p_j(g_j(x)) \in [q_j(g_j(x)), 1[$;
- a weak veto threshold $wv_j(g_j(x)) \in [p_j(g_j(x)), 1] \cup \{2\}$;

- and, a strong veto threshold $v_j(g_j(x)) \in [wv_j(g_j(x)), 1] \cup \{2\}$.

veto

The complete absence of *veto* is modelled here via the value 2. All these threshold functions are supposed to verify the standard non-decreasing monotonicity condition (see [RB93, page 56]).

Note that the thresholds are not absolute and may depend on the value taken by alternative x on criterion j.

outranking

Let S be a binary relation on X. Classically, an *outranking* situation xSy between two alternatives x and y of X is assumed to hold if there is a sufficient majority of criteria which supports an "*at least as good as*" preferential statement and there is no criterion which raises a veto against it [Roy85]. The validation of such an outranking situation may quite naturally be expressed in the bipolar credibility calculus defined in Section 3.1. Our formulation is based on the classical Electre definition of the outranking index. Nevertheless the reader should notice some slight but important differences, due to the semantics of the underlying bipolar valuation.

local concordance

Indeed, in order to characterise a local "*at least as good as*" situation between two alternatives x and y of X for each criterion j of N (called *local* or *partial concordance*) we use the following function $C_j : X \times X \;\rightarrow\; \{-1, 0, 1\}$ such that:

$$C_j(x,y) = \begin{cases} 1 & \text{if} \quad \Delta_j(x,y) > -q_j(g_j(x))\,; \\ -1 & \text{if} \quad \Delta_j(x,y) \leqslant -p_j(g_j(x))\,; \\ 0 & \text{otherwise}\,. \end{cases}$$

Credibility 0 is assigned to $C_j(x,y)$ in case it cannot be determined whether alternative x is at least as good as alternative y or not (see Section 3.1).

local veto

Similarly, the *local* or *partial veto* situation for each criterion j of N is characterised via a veto-function $V_j : X \times X \;\rightarrow\; \{-1, 0, 1\}$ where:

$$V_j(x,y) = \begin{cases} 1 & \text{if} \quad \Delta_j(x,y) \leqslant -v_j(g_j(x))\,; \\ -1 & \text{if} \quad \Delta_j(x,y) > -wv_j(g_j(x))\,; \\ 0 & \text{otherwise}\,. \end{cases}$$

Again, according to the semantics of the bipolar-valued characterisation, the veto function V_j renders an indetermined response when the difference of performances is between the weak and the strong veto thresholds wv_j and v_j.

Figure 3.1 represents both functions for a fixed $g_j(x)$.

The overall outranking index $\widetilde{S}$, defined for all pairs of alternatives $(x,y) \in X \times X$, conjunctively combines an overall concordance index, aggregating all

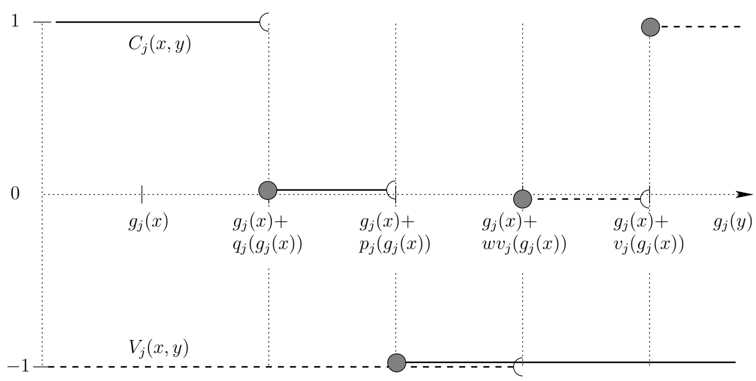

Figure 3.1: Local concordance (—) and veto (- -) functions, with $g_j(x)$ fixed

local "*at least as good as*" statements, and the absence of veto on each of the criteria. For any two alternatives x and y of X we have:

$$\widetilde{S}(x,y) = \min\{\widetilde{C}(x,y), -V_1(x,y), \ldots, -V_n(x,y)\}, \tag{3.1}$$

where the overall concordance index $\widetilde{C}(x,y)$ is defined as follows:

$$\widetilde{C}(x,y) = \sum_{j \in N} w_j \cdot C_j(x,y). \tag{3.2}$$

The min operator in Formula 3.1 translates the conjunction between the overall concordance index $\widetilde{C}(x,y)$ and the *negated* partial veto indexes $-V_j(x,y)$ ($\forall j \in N$). In case $V_j = -1$ for all $j \in N$ (absence of partial veto on all criteria), the resulting outranking index $\widetilde{S}$ equals the overall concordance index $\widetilde{C}$.

Following Formulae (3.1) and (3.2), $\widetilde{S}$ is a function from $X \times X$ to $\mathcal{L}$ representing the credibility of the validation or non-validation of an outranking situation observed between each pair of alternatives. $\widetilde{S}$ is called the bipolar-valued characterisation of the outranking relation S, or for short, the *bipolar-valued outranking relation*. *bipolar-valued outranking relation*

The maximum value $+1$ of the valuation is reached in the case of *unanimous concordance*, whereas the minimum value -1 is obtained either in the case of *unanimous discordance*, or if there exists a strong veto situation on at least one criterion.

The median situation 0 represents a case of *indeterminateness*: either the arguments in favour of an outranking are compensated by those against it or, a positive concordance in favour of the outranking is outbalanced by a potential (weak) veto situation.

β-cut

Let us now show how this indetermination degree can be extended to a larger range of values. To do so, we define the concept of *β-cut* of the bipolar-valued credibility scale. Let $\beta \in]0,1]$. The β-cut relation $\widetilde{S}_\beta$ of $\widetilde{S}$ is defined as follows, for each $(x,y) \in X \times X$:

$$\widetilde{S}_\beta(x,y) = \begin{cases} 0 & \text{if } |\widetilde{S}(x,y)| < \beta; \\ \widetilde{S}(x,y) & \text{else.} \end{cases}$$

This modification of the original bipolar-valued outranking relation $\widetilde{S}$ allows to collapse symmetrically a given range of values around 0 on this indetermination point. As a consequence, it is possible to consider a larger interval of values as indetermined in order to take into account *majority-related impreciseness*. A DM could decide to require such a modification of the outranking relation, if he considers that a simple majority is not sufficient to consider that an outranking situation is validated. Figure 3.2 schematically represents the effect of a β-cut on the bipolar-valued credibility scale.

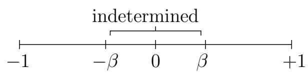

Figure 3.2: The values in $]-\beta,\beta[$ are collapsed on 0

It is easy to recover the semantics linked to this bipolar-valued characterisation from our earlier considerations (see Section 3.1). For any two alternatives x and y of X,

- $\widetilde{S}(x,y) = +1$ means that assertion "xSy" is *clearly validated*.
- $\widetilde{S}(x,y) > 0$ means that assertion "xSy" is *more validated than non-validated*.
- $\widetilde{S}(x,y) = 0$ means that assertion "xSy" is *indetermined*.
- $\widetilde{S}(x,y) < 0$ means that assertion "xSy" is *more non-validated than validated*.
- $\widetilde{S}(x,y) = -1$ means that assertion "xSy" is *clearly non-validated*.

outranking digraph

Definition 3.2.1. *The set X associated to a bipolar-valued characterisation $\widetilde{S}$ of the outranking relation $S \in X \times X$ is called a* bipolar-valued outranking digraph, *denoted $\widetilde{G}(X,\widetilde{S})$.*

crisp outranking relation

The *crisp outranking relation* S can be constructed via its bipolar-valued characterisation. S is the set of pairs (x,y) of $X \times X$ such that $\widetilde{S}(x,y) > 0$. We write $G(X, S)$ the corresponding so-called *crisp outranking digraph* associated to $\widetilde{G}(X, \widetilde{S})$.

Let us present these concepts via a short example.

Example 3 *Consider the set* $X = \{a, b, c, d, e\}$ *of alternatives evaluated on a coherent family* $N = \{1, \ldots, 5\}$ *of criteria of equal weights (see left part of Table 3.1). To each criterion is associated a preference scale in* $[0, 1]$ *and an indifference threshold of* 0.1*, a preference threshold of* 0.2*, a weak veto threshold of* 0.6*, and a strong veto threshold of* 0.8*.*

Based on the performances of the five alternatives on the criteria, we compute the bipolar-valued outranking relation $\widetilde{S}$ *shown in the right part of Table 3.1. The crisp outranking digraph* $G(X, S)$ *associated to the bipolar-valued outranking digraph* $\widetilde{G}(X, \widetilde{S})$ *is shown in Figure 3.3.*

	coherent family of criteria					$\widetilde{S}$				
alternatives	1	2	3	4	5	a	b	c	d	e
a	0.52	0.82	0.07	1.00	0.04	1.0	-0.2	-1.0	0.6	0.4
b	0.96	0.27	0.43	0.83	0.32	0.4	1.0	0.2	0.2	0.4
c	0.85	0.31	0.61	0.41	0.98	0.2	0.4	1.0	0.4	0.6
d	0.30	0.60	0.74	0.02	0.02	-1.0	-1.0	-1.0	1.0	-1.0
e	0.18	0.11	0.23	0.94	0.63	0.2	0.2	-0.4	0.0	1.0

Table 3.1: Example 3: performance table and bipolar-valued outranking relation

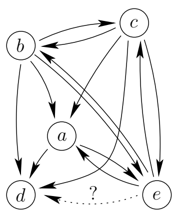

Figure 3.3: Example 3: associated crisp digraph and indetermined arc

It is worthwile noting in the previous example the dotted arc from alternative e to d, representing an indetermined outranking (see Figure 3.3). This situation is not expressible in a standard Boolean-valued characterisation of the outranking. Consequently, the negation of the *strictly positive* part of $\widetilde{S}$ is in general not identical to the complement of S in $X \times X$. This therefore requires the use of the *dotted* arcs to better represent $\widetilde{S}$ via its associated crisp digraph.

The reader, familiar with the ELECTRE methods, may have noticed much ressemblance between the bipolar-valued characterisation $\widetilde{S}$ and the classical ELECTRE-type valuations of an outranking relation. It is important to notice, however, that the latter do not necessarily respect the semantics of the bipolar credibility calculus.

In particular, the 1/2 value does in general not have the meaning of indetermined validation which is given here to the 0 credibility degree. The bipolar valuation of the outranking relation presented here is solely based on sums and differences of weights of individual criteria. This is less obvious in some ELECTRE-type valuations, where the local concordance index can take values in the real unit interval (see for example ELECTRE III). As a consequence, the meaning conveyed to some of these valuations is not clear.

order

size

Let us now introduce some further concepts which are used in this article. The *order* n of the digraph $\widetilde{G}(X, \widetilde{S})$ is given by the cardinality of X, whereas the *size* p of $\widetilde{G}$ is given by the cardinality of S.

A *path of order* $m \leqslant n$ in $\widetilde{G}(X, \widetilde{S})$ is a sequence $(x_i)_{i=1}^m$ of alternatives of X such that $\widetilde{S}(x_i, x_{i+1}) \geq 0$, $\forall i \in \{1, \ldots, m-1\}$. A *circuit of order* $m \leqslant n$ is a path of order m such that $\widetilde{S}(x_m, x_1) \geq 0$.

odd chordless circuit

Definition 3.2.2. *An* odd chordless circuit $(x_i)_{i=1}^m$ *is a circuit of odd order* m *such that* $\widetilde{S}(x_i, x_{i+1}) \geq 0$, $\forall i \in \{1, \ldots, m-1\}$, $\widetilde{S}(x_m, x_1) \geq 0$ *and* $\widetilde{S}(x_i, x_j) < 0$ *otherwise.*

Following a result by [Bou06] which extends the results of [Bou96] to the bipolar-valued case, it appears that, apart from certainly being reflexive, the bipolar-valued outranking digraphs do not necessarily have any particular relational properties such as transitivity or total comparability. Indeed he shows that, with a sufficient number of criteria, it is always possible to define a performance table such that the associated crisp outranking digraph renders any given reflexive binary relation. This result bears a negative algorithmic consequence. Indeed, as we will show in Chapter 4, solving the choice problematique based on a bipolar-valued outranking relation is a non-trivial algorithmic problem in case of non-transitive and partial outrankings.

It is important to underline here that the starting point of this study is deliberately a *given* performance table, a set of threshold and veto functions as well as significance weights which are all clearly defined and have been acknowledged by the DM. Consequently, tackling impreciseness issues in these data is out of the scope of this work. For first attempts to cope with this topic in a bipolar-valued credibility calculus framework, see [Bis04].

Historically, in the context of outranking relations, the progressive choice problematique has been solved by using the independent outranking set, i.e., the *kernel* of a digraph [Roy68, Roy85]. Let us define in the next section this concept in a bipolar-valued outranking digraph and show how it can be determined.

3.3 Kernels in the bipolar-valued digraph

Definition 3.3.1. *Let Y be a non-empty subset of X.*

1. *Y is said to be* outranking *(resp.* outranked *) in $\widetilde{G}(X, \widetilde{S})$ if and only if $x \notin Y \;\Rightarrow\; \exists y \in Y : \widetilde{S}(y, x) > 0$ (resp. $\widetilde{S}(x, y) > 0$).* *outranking* *outranked*

2. *Y is said to be* independent *(resp.* strictly independent*) in $\widetilde{G}(X, \widetilde{S})$ if and only if for all $x \neq y$ in Y we have $\widetilde{S}(x, y) \leqslant 0$ (resp. $\widetilde{S}(x, y) < 0$).* *(strictly) independent*

3. *Y is called an* outranking *(resp.* outranked*)* kernel *if and only if it is an outranking (resp. outranked) and independent set.* *kernel*

4. *Y is called a* determined *outranking (resp. outranked)* kernel *if and only if it is an outranking (resp. outranked) and strictly independent set.*

It follows from these definitions that, if $\widetilde{S}$ only takes negative values ($S = \emptyset$), X is an outranking and an outranked kernel.

Let us illustrate these concepts on Example 3.

> **Example 3** (*continued*) *In the crisp digraph G (see Figure 3.3) we can observe two determined outranking kernels, namely the singletons $\{b\}$ and $\{c\}$. The digraph also contains one outranked kernel, namely the pair $\{d, e\}$. Note that alternatives d and e are independent (but not strictly independent) from each other.*

A set Y can be characterised via bipolar-valued membership assertions $\widetilde{Y} : X \to \mathcal{L}$, expressing the credibility of the fact that $x \in Y$ or not, for all $x \in X$. $\widetilde{Y}$ is called a *bipolar-valued characterisation* of Y, or for short a *bipolar-valued set* in $\widetilde{G}(X, \widetilde{S})$. The semantics linked to this characterisation can again be derived from the properties of the bipolar-valued scale $\mathcal{L}$ (also see Section 3.1): *bipolar-valued set*

– $\widetilde{Y}(x) = +1$ means that assertion "$x \in Y$" is *clearly validated*;

– $\widetilde{Y}(x) > 0$ means that assertion "$x \in Y$" is *more validated than non-validated*;

– $\widetilde{Y}(x) = 0$ means that assertion "$x \in Y$" is *indetermined*;

– $\widetilde{Y}(x) < 0$ means that assertion "$x \in Y$" is *more non-validated than validated*;

– $\widetilde{Y}(x) = -1$ means that assertion "$x \in Y$" is *clearly non-validated*. Equivalently, one can say that assertion $x \notin Y$ is *clearly validated*.

In the following paragraphs, we recall useful results from [BPR06]. They allow us to establish a link between the classical graph theoretic and algebraic representations of kernels (via their bipolar-valued characterisations).

Proposition 3.3.1. *The outranking (resp. outranked) kernels of $\widetilde{G}(X, \widetilde{S})$ are among the bipolar-valued sets $\widetilde{Y}$ satisfying the respective following bipolar-valued kernel equation systems:*

$$\max_{y \neq x}[\min(\widetilde{Y}(y), \widetilde{S}(y, x))] = -\widetilde{Y}(x), \quad \textit{for all } x \in X; \tag{3.3}$$

$$\max_{y \neq x}[\min(\widetilde{S}(x, y), \widetilde{Y}(y))] = -\widetilde{Y}(x), \quad \textit{for all } x \in X. \tag{3.4}$$

sharpness

Let $\mathcal{Y}^+$ and $\mathcal{Y}^-$ denote the set of bipolar-valued sets verifying respectively kernel equation systems (3.3) and (3.4) above. Let $\widetilde{Y}_1$ and $\widetilde{Y}_2$ be two elements of $\mathcal{Y}^+$ (or $\mathcal{Y}^-$). $\widetilde{Y}_1$ is said to be *at least as sharp as* $\widetilde{Y}_2$ (denoted $\widetilde{Y}_2 \preceq \widetilde{Y}_1$) if and only if for all x in X either $\widetilde{Y}_1(x) \leqslant \widetilde{Y}_2(x) \leqslant 0$ or $0 \leqslant \widetilde{Y}_2(x) \leqslant \widetilde{Y}_1(x)$. The $\preceq$ relation defines a partial order (antisymmetrical and transitive) [Bis97]. If $\widetilde{Y}(x) \neq 0$ for each x in X, $\widetilde{Y}$ is called a *determined* bipolar-valued set.

determined bipolar-valued set

Theorem 3.3.1 (Bisdorff, Pirlot, Roubens, 2006)**.**

1. *There exists a one-to-one correspondence between the maximal sharp determined sets in $\mathcal{Y}^+$ (resp. $\mathcal{Y}^-$) and the determined outranking (resp. outranked) kernels in $\widetilde{G}$.*

2. *Each maximal sharp set in $\mathcal{Y}^+$ (resp. $\mathcal{Y}^-$) characterises an outranking (resp. outranked) kernel in $\widetilde{G}$.*

Proof. The first result, specialised to determined sets, is proved in [BPR06, Theorem 1]. The second one results directly from the kernel equation systems of Proposition 3.3.1. □

The maximal sharp sets in $\mathcal{Y}^+$ (resp. $\mathcal{Y}^-$) deliver thus *outranking* (resp. *outranked*) kernel characterisations. Let us illustrate this result on Example 3.

> **Example 3** (*continued*) *Recall that the crisp outranking digraph G contains two outranking kernels and one outranked kernel. The bipolar-valued characterisations of these kernels are shown in Table 3.2.*
>
> *The outranking kernel $\{c\}$ is more determined than $\{b\}$ and is therefore the more credible instance. Indeed, one can easily verify that*

$\widetilde{Y}$	a	b	c	d	e
$\{b\}$	-0.2	**0.2**	-0.2	-0.2	-0.2
$\{\mathbf{c}\}$	-0.2	-0.4	**0.4**	-0.4	-0.4
$\{d, e\}$	-0.6	-0.2	-0.4	**1.0**	**0.0**

Table 3.2: Example 3: bipolar-valued characterisations of the kernels

the degrees of logical determination of the membership assertions for $\{c\}$ *are higher than those for* $\{b\}$ *(see Definition 3.1.2). Concerning the outranked kernel* $\{d, e\}$*, it is worthwhile noting that alternative* d *belongs to it with certainty, whereas the belonging of alternative* e *to this kernel depends on the indetermined situation* dSe*.*

In the context of a progressive method, if the latter outranking becomes more true than false at a later stage, then e *can be dropped from the kernel without any regret. On the opposite, if the outranking becomes more non-validated than validated, then* e *remains part of the then determined kernel* $\{d, e\}$*.*

In the past, Bisdorff and Roubens [BR03] have promoted the *most determined* outranking kernel in a bipolar-valued outranking digraph $\widetilde{G}$ as a convenient choice recommendation in a progressive resolution of the choice problematique.

Example 3 (*continued*) *The reader can indeed easily verify in the performance table of his example (see Table 3.1) that alternative* c *is performing better than alternative* b*. Alternative* a *has very contrasted performances and* d *indisputably presents the worst performances.*

However, well founded criticisms against the capacity of the outranking kernel concept to generate, in general outranking digraphs, a satisfactory and convincing choice recommendation led us to propose a new method. The following chapter deals with the choice problematique and details the RUBIS method for determining an adequate choice recommendation.

Chapter 4

The choice problematique

Contents

This chapter represents the core of the second part of this work. It contains our developments on the choice problematique in the framework of a bipolar-valued outranking relation. It is mainly motivated by the unsatisfactory way the classical Electre methods deal with the choice problematique. We therefore present in this chapter a new proposal for computing provisional choice recommendations from a bipolar-valued outranking digraph called the Rubis method. Our approach is based on new pragmatic and logical foundations of the progressive choice problematique in the tradition of the pioneering work of Roy and Bouyssou [RB93].

In the first section, we introduce the choice problematique, present the existing proposals and introduce the foundations of the Rubis method. Then, in Section 4.2 we determine the graph theory-related object which will represent the choice recommendation. Finally, in Section 4.3 we present the Rubis algorithm and some of the properties of the Rubis choice recommendation.

Note that the considerations of this chapter are mainly based on our article [BMR07].

4.1 Foundations of the Rubis choice decision aiding methodology

Apart from the European multiple criteria decision aiding community [Roy85, RV96], the progressive resolution of the choice problematique has attracted quite little attention in the Operational Research (OR) field. Seminal work on it goes back to the first article of Roy on the Electre I methods [Roy68]. After [Kit93], interest in solving the choice problematique differently from the classical optimisation paradigm has reappeared. An early work of [BR96] on valued kernels has resulted in new attempts to tackle the progressive choice problematique directly on the valued outranking digraph. After first positive results [Bis00], methodological difficulties appeared when facing highly non-transitive and partial outranking relations.

In this section, we first revisit the choice problematique in order to identify the type of pragmatic decision aiding we are interested in. A brief comparison with the classical Electre method will underline similarities and differences with the Rubis method. Finally, we present new foundations for the choice decision aiding methodology.

4.1.1 The choice problematique

From a *classical* OR point of view, the choice problematique is the search for one best or optimal alternative. From a *decision aiding* point of view, however, the assistance we may offer the DM depends on the nature of the decision aiding process we support.

4.1.1.1 Choice and elimination

Following the tradition [Roy85, RB93], we call *choice problematique* the category of decision problems consisting of the search for a *single best* alternative. Symmetrically to this, we define the *elimination problematique* as the category of decision problems, whose objective is to search for a *single worst* alternative.

choice and elimination problematique

The interest of considering both opposite problematiques will appear later in Sections 4.1.3 and 4.2.3, where we show that, due to the intransitivity of the outranking relation, certain sets of alternatives can be considered a potential choice, as well as, a potential elimination recommendation (which makes the recommendation ambiguous in both problematiques).

Following the symmetric design of the bipolar credibility calculus, both the choice, and the elimination problematique can be tackled similarly. As the first one is much more common, we will in the sequel exclusively focus on the choice problematique.

4.1.1.2 Type of decision aiding process

Recall that a decision aiding *method* is a regular and systematic way of dealing with a given decision problem. A *choice recommendation* is the output of such a decision aiding method in the particular context of the choice problematique.

choice recommendation

Very generally, one may distinguish between two general kinds of choice problematiques, depending on the nature of the underlying decision problem. On the one hand, choice problems which require the single best alternative to be uncovered in a single decision aiding step, and, on the other hand, choice problems which allow to progressively uncover the single best alternative through the implementation of an iterative, progressive multiple step decision aiding process.

In the first case, a choice recommendation must always propose a single best alternative, whereas in the second case, the choice recommendation is a provisional advice that should, given the current available information, propose all plausible candidates for a final solution. It is in fact a set of potentially best alternatives which has to be refined via further interactions with the DM. It is important to clearly distinguish between a *current* and the *eventual* choice recommendation consisting ideally of the single best alternative. If not, this last recommendation requires to be further analysed by the DM himself, in view of determining his ultimate choice.

eventual choice recommendation

As already pointed out, our interest lies in this latter category of problems, where the ultimate recommendation can be determined progressively. We therefore focus in the sequel on the resolution of this progressive decision aiding problem, in the tradition of the classical ELECTRE methods.

4.1.2 The ELECTRE choice decision aiding method

The progressive choice problematique is extensively discussed and promoted in the context of multiple criteria decision aiding in [RB93], where the authors explain that it is important that the non-retained alternatives for the current choice recommendation are left out for well-founded reasons, acknowledged and approved by the DM. Instead of forcing the decision aiding procedure to elicit a single best alternative at any cost, it is indeed preferable to obtain a set Y of potential candidates for the choice, as long as it can be plainly justified on the basis of the currently available preferential information.

Starting from this methodological position, Roy defines two principles for the construction of a choice recommendation. A subset Y of X is a choice recommendation if:

1. Each alternative which is not selected in Y is outranked by at least one alternative of Y;

2. The number of retained alternatives in the set Y is as small as possible.

The first principle counterbalances the second one. Indeed, it tends to keep the cardinality of the choice recommendation high enough to guarantee that no potentially best alternative is missed out. The second principle tends to keep its cardinality as small as possible in order to focus on the single best choice.

outranking kernel

In the context of the ELECTRE methods, Roy [Roy68, Roy85] proposes to use as a provisional choice recommendation the concept of *outranking kernel*. One can indeed easily check that this recommendation verifies both principles.

ELECTRE I

According to Roy, a choice recommendation has furthermore to be unique. The existence of a unique outranking kernel is, however, only guaranteed when the digraph does not contain any circuits at all [Ber70]. To avoid a possible emptiness or multiplicity of outranking kernels, Roy [Roy68] initially proposed in the ELECTRE I method to consider the alternatives belonging to maximal circuits as ties. These circuits are then collapsed on single nodes, which results in an outranking digraph which always admits a unique outranking kernel.

ELECTRE IS

The alternatives gathered in such a maximal circuit might, however, not be all equivalent and behave differently when compared to alternatives exterior to the circuit. Furthermore, the validation of the arcs of such a circuit may be problematic due to imprecision in the data or the preferential information provided by the DM. All in all, these difficulties in the clear interpretation of those circuits led to the development of the ELECTRE IS method [RB93]. There, robustness considerations allow to remove certain arcs of the outranking digraph leading to a circuit-free graph containing a unique outranking kernel.

Note finally that in both methods, the outranking relation is not viewed on a valued credibility scale. The double requirement of sufficient concordance and absence of vetoes is used instead for a crisp validation of pairwise outranking situations.

In this work we do not follow the same approach, even if the bipolar-valued framework would allow it. We prefer to investigate the option of avoiding perturbations and modifications of the original outranking digraph to find a solution of the choice problem. We therefore revisit the very foundations of a progressive choice decision aiding methodology in order to discover how the bipolar-valued concept of outranking kernel may deliver a satisfactory choice recommendation without having to express doubts about a given bipolar-valued characterisation of the outranking relation.

4.1.3 New foundations for a progressive choice decision aiding methodology

Let us now introduce five principles (two from the previous discussion and three new ones) that the construction of a choice recommendation in a progressive decision aiding method should follow.

$\mathcal{P}_1$: **Non-retainment for well motivated reasons**
Each non-retained alternative must be put aside for well motivated reasons in order to avoid to miss any potentially best alternative.

A similar formulation is that each non-retained alternative must be considered as worse as at least one alternative of the choice recommendation.

$\mathcal{P}_2$: **Minimal size**
The number of alternatives retained in a choice recommendation should be as small as possible.

This requirement is obvious when recalling that the goal of the choice problematique is to find a single best alternative and that ultimately, a choice recommendation containing a single element concludes the progressive decision aiding process.

$\mathcal{P}_3$: **Efficient and informative refinement**
Each step of the progressive decision aiding must deliver an *efficient and informative refinement* of the previous recommendation.

The currently delivered recommendation should focus on new and previously unknown preference statements, such that the progressive decision aiding process can converge to a single choice recommendation as quickly and efficiently as possible.

Note that a progressive decision aiding process is not required to go on until a single best alternative can be recommended. As already mentioned, it may be up to the DM to determine the ultimate choice from the eventual recommendation of the decision aiding.

Principle $\mathcal{P}_\mathbf{3}$ is quite similar to the previous principle and appears to make it redundant. In the following section, however, when implementing the RUBIS method, their strategic difference will become apparent.

> $\mathcal{P}_\mathbf{4}$: **Effective recommendation**
> The recommendation should not correspond simultaneously to a *choice* and an *elimination* recommendation.

This principle avoids the formulation of ambiguous recommendations, i.e. both outranking and outranked sets of alternatives, which could appear in intransitive and partial outranking relations.

It is worthwhile noting that in a situation where all decision alternatives are either considered to be pairwisely equivalent or incomparable, no effective choice recommendation can be made.

> $\mathcal{P}_\mathbf{5}$: **Maximal credibility**
> The choice recommendation must be as credible as possible with respect to the preferential knowledge available in the current step of the decision aiding process.

As the credibility degrees in the bipolar-valued outranking digraph represent the more or less overall concordance or consensus of the criteria to support an outranking situation, it seems quite natural that in the case of several potential choice recommendations, we recommend the one(s) with the highest determinateness of the membership assertions.

As mentioned before, the first two principles are identical to those proposed by Roy (see Section 4.1.2). However, alone they are not sufficient to generate satisfactory choice recommendations. The three additional principles $\mathcal{P}_\mathbf{3}$, $\mathcal{P}_\mathbf{4}$, and $\mathcal{P}_\mathbf{5}$, not satisfied by neither ELECTRE I nor ELECTRE IS, will show their operational value when translated in Section 4.2 into properties in the bipolar-valued outranking digraph. Let us finish by defining the concept of RUBIS choice recommendation.

RUBIS *choice recommendation*

We call a RUBIS *choice recommendation (RCR)*, a choice recommendation which verifies the five pragmatic principles above.

Our goal in the following section is to determine which graph theory-related object these properties characterise as a convincing choice recommendation.

$\widetilde{S}$	a	b	c	d	e
a	1.0	0.2	-1.0	-0.7	-0.8
b	-0.6	1.0	0.8	1.0	0.0
c	-1.0	-1.0	1.0	0.2	0.8
d	0.6	-0.6	-1.0	1.0	-0.4
e	-1.0	-0.8	-0.4	-0.6	1.0

Table 4.1: Example 4: the bipolar-valued outranking relation

4.2 Tackling the choice problematique

Let us note beforehand that obvious RUBIS choice recommendations exist in case the outranking relation is transitive, namely the set of all maximal alternatives. However, as already mentioned earlier, the crisp outranking digraphs that we obtain from the bipolar-valued characterisation of an outranking relation are in general not transitive. This clearly motivates the necessity to find a procedure which computes a choice recommendation verifying the five principles for any possible reflexive binary relation.

Throughout this section, we illustrate our discourse via the following didactic example[1].

Example 4 *Let $\widetilde{G}(X, \widetilde{S})$ be a bipolar-valued outranking digraph, where $X = \{a, b, c, d, e\}$ and $\widetilde{S}$ is given in table 4.1 and the associated crisp digraph $G(X, S)$ is represented in figure 4.1.*

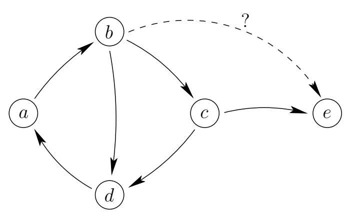

Figure 4.1: Example 4: associated crisp digraph and an indetermined arc

Let us now analyse the previously mentioned principles one by one and present their translations in terms of the concepts presented in Chapter 3. Note that all the directed concepts linked to an outranking property can symmetrically be reused in an elimination problematique via the corresponding outranked properties.

[1] B. Roy, 2005, private communication.

4.2.1 Non-retainment for well motivated reasons (principle: $\mathcal{P}_1$)

In terms of the bipolar-valued outranking relation, principle $\mathcal{P}_1$ amounts to saying that each non-retained alternative should be outranked by at least one alternative of the choice recommendation.

> $\mathcal{R}_1$: **Outranking**
> An RCR is an outranking set in $\widetilde{G}(X, \widetilde{S})$.

> **Example 4** (*continued*) *The sets* $\{a, b, e\}, \{b, c, d\}$*, as well as* $\{a, b, c\}$ *for instance, are all outranking sets.*

4.2.2 Minimal size and efficient and informative refinement (principles: $\mathcal{P}_2$ and $\mathcal{P}_3$)

In this Section we show that these two principles are closely linked. To rewrite principle $\mathcal{P}_2$ of *minimal size* in the present context, we first need to define some concepts related to graph theory.

Definition 4.2.1.

1. *The* outranking neighbourhood $\Gamma^+(x)$ *of a node (or equivalently an alternative)* x *of* X *is the union of* x *and the set of alternatives which are outranked by* x*.*

2. *The* outranking neighbourhood $\Gamma^+(Y)$ *of a set* Y *is the union of the outranking neighbourhoods of the alternatives of* Y*.*

3. *The* private outranking neighbourhood $\Gamma^+_Y(x)$ *of an alternative* x *in a set* Y *is the set* $\Gamma^+(x) \setminus \Gamma^+(Y \setminus \{x\})$*.*

For a given alternative x of a set Y, the set $\Gamma^+_Y(x)$ represents the *individual* contribution of x to the outranking quality of Y. If the private outranking neighbourhood of x in Y is empty, this means that, when x is dropped from this set, Y still remains an outranking set. From this observation one can derive the following definition.

irredundancy

Definition 4.2.2. *An outranking (resp. outranked) set* Y *is said to be* irredundant *if all the alternatives of* Y *have non-empty private outranking (resp. outranked) neighbourhoods.*

In view of these considerations, we transcibe the *minimal size* principle into its formal counterpart, which is that of *irredundancy* of the set.

> **Example 4** (*continued*) $\{a, b, e\}$*,* $\{b, c, d\}$*,* $\{b, e, d\}$*, and* $\{a, c\}$ *are irredundant outranking sets.* $\{a, b, c\}$*, listed in the context of principle* $\mathcal{R}_1$*, is not irredundant outranking because alternative* b *has an empty private neighbourhood in this set.*

Let us now switch to principle $\mathcal{P}_3$ (*efficient and informative refinement*), whose primary objective is to avoid that, in the case of a provisional choice recommendation, the DM may notice a best sub-choice without any further analyses.

We require therefore that a choice recommendation Y should be such that the digraph restricted to the nodes of Y does not contain any obvious sub-choice recommendation. Consequently, at each stage of the decision aiding process, the provisional choice recommendation must focus on new and previously indetermined or unknown preference statements. Let us illustrate this with a short example.

Example 5 *Consider the problem shown on the crisp digraph represented in figure 4.2.*

Both highlighted sets $Y_1 = \{a, b\}$ *and* $Y_2 = \{a, d, e, f, \ldots, z\}$ *verify the principles* $\mathcal{P}_1$ *and* $\mathcal{P}_2$ *as outranking irredundant sets. A DM could be tempted to prefer* Y_1 *to* Y_2 *because of its lower cardinality. Nevertheless,* Y_1 *contains information which is already confirmed at this stage of the progressive search, namely that the statement "a outranks b" is validated. In the case of the choice* Y_2*, the next step of the search will focus on alternatives which presently are incomparable.*

If a further analysis step would focus on the set Y_1*, then it is quite difficult to imagine that the DM will be able to forget about the already confirmed validation of the statement "a outranks b". He will most certainly consider a as the choice, which might however not be the best decision alternative, as a is not outranking any of the alternatives of* $\{d, e, f, \ldots, z\}$*.*

According to principle $\mathcal{P}_3$ *we therefore recommend* Y_2 *as a choice recommendation.*

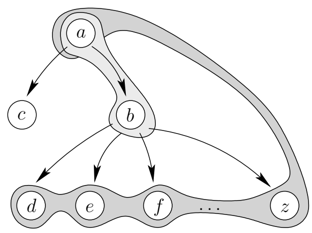

Figure 4.2: Example 5: an unstable ($\{a, b\}$) and a stable ($\{a, d, e, f, \ldots, z\}$) set

In view of the previous considerations and the output generated by principles $\mathcal{P}_1$ and $\mathcal{P}_2$, it is quite natural to define the concept of stability as follows:

stability

Definition 4.2.3. *An outranking (resp. outranked) set Y in $\widetilde{G}(X, \widetilde{S})$ is said to be* stable[2] *if and only if the induced subgraph $\widetilde{G}_Y(Y, \widetilde{S}|_Y)$ does not contain any irredundant outranking (resp. outranked) subset.*

The outranking (resp. outranked) kernels (see Definition 3.3.1) of an outranking digraph verify this property of stability. Nevertheless, as already mentioned and as it is shown in the following property, the existence of an outranking (resp. outranked) kernel is not guaranteed in an outranking digraph.

Property 4.2.1. *If a digraph $\widetilde{G}(X, \widetilde{S})$ has no outranking (resp. outranked) kernel, it contains a chordless circuit of odd order.*

Proof. This property represents the contraposition of Richardson's general result: If a digraph contains no chordless circuit of odd order, then it has an outranking (resp. outranked) kernel [Ric53]. □

The outranking kernel gives indeed a potential choice recommendation in case the outranking digraph does not contain any chordless circuit of odd order. Consider now the case where a potential choice recommendation, resulting from principles $\mathcal{P}_1$ and $\mathcal{P}_2$, consists of a chordless circuit $Y = \{a, b, c\}$ of order 3 such that aSb, bSc and cSa. Such a choice recommendation is clearly neither a kernel nor is it a stable recommendation. Nevertheless, it may be an interesting provisional recommendation because it presents three alternatives to the DM which do not contain obvious information on the possible single choice at this step of the progressive search. In fact, a, b and c can be considered as equivalent potential candidates for the choice in the current stage of the decision process.

These considerations show that neither the concepts of stability and irredundancy nor that of outranking kernel are in fact sufficient for guiding the search for a choice recommendation in a general outranking digraph. In the first case, potentially interesting choice recommendations are left out and in the latter case, nothing guarantees the existence of a kernel in an outranking digraph. In order to overcome these difficulties, we introduce the concept of *hyperindependence*, an extension of the independence property discussed in Section 3.3.

(strict) hyperindependence

Definition 4.2.4. *A set Y is said to be (strictly)* hyperindependent *in $\widetilde{G}$ if it consists of chordless circuits of odd order $p \geq 1$ which are (strictly) independent of each other.*

Note that in Definition 4.2.4 above, singletons are assimilated to chordless circuits of (odd) order 1. Principles $\mathcal{P}_2$ and $\mathcal{P}_3$ can now be translated into the following formal property:

$\mathcal{R}_2$: **Hyperindependence**
An RCR is a hyperindependent set in $\widetilde{G}(X, \widetilde{S})$.

[2]Note that in graph theory, the term *stability* is often used to designate what we call *independence*. Here we use the term in a different sense.

As a direct consequence, we can define the concept of *hyperkernel*.

Definition 4.2.5. *A hyperindependent (resp. strictly hyperindependent) outranking (resp. outranked) set is called an outranking (resp. outranked)* hyperkernel *(resp.* determined hyperkernel*).* *hyperkernel*

Note that the outranking and outranked hyperkernels obviously verify the property of stability (see Definition 4.2.3).

> **Example 4** (*continued*) *Set* $\{a, b, d, e\}$ *(see Figure 4.1) is an outranking hyperkernel. The indetermined outranking relation between* b *and* e *implies that the set is not strictly hyperindependent. Note here that this obvious potential choice recommendation would have been left out if the search was restricted to outranking kernels.*

In case the outranking digraph does not contain any chordless circuits of odd order 3 and more, the outranking kernels of the digraph deliver potential choice recommendations verifying the first two Rubis principles.

In the general case however, the RCR will consist of at least one outranking hyperkernel of the digraph.

4.2.3 Effective and maximally credible recommendation (principles $\mathcal{P}_4$ and $\mathcal{P}_5$)

In order to translate principle $\mathcal{P}_4$ (*effective recommendation*), we introduce the concept of *strict outranking* set.

Recall that one can associate an outranking (resp. outranked) set Y with a bipolar-valued characterisation $\widetilde{Y}^+$ (resp. $\widetilde{Y}^-$). It may happen that both kernel characterisations are solutions of the respective kernel equation systems of Proposition 3.3.1. In order to determine in this case whether Y is in fact an outranking or an outranked set, it is necessary to specify which of its bipolar-valued characterisations is the more determined.

We extend therefore the concept of determinateness of propositional statements (see Definition 3.1.2) to bipolar-valued characterisations of sets.

Definition 4.2.6. *The* determinateness $D(\widetilde{Y})$ *of the bipolar-valued characterisation* $\widetilde{Y}$ *of a set* Y *is given by the average value of the determinateness degrees* $D(\widetilde{Y}(x))$ *for all* x *in* X*.* *determinateness*

In view of the bipolar definition of the global outranking and concordance indexes (Formulae 3.1 and 3.2), which solely balance rational significance weights, we define here the overall determinateness of a bipolar-valued set characterisation as the mean of all the individual membership determinatenesses. Nevertheless other aggregation operators could be used as well.

We can now define the concepts of strict and null set as follows:

Definition 4.2.7.

null set

1. *A set Y which is outranking and outranked with the same determinateness, i.e., $D(\widetilde{Y}^+) = D(\widetilde{Y}^-)$ is called a* null set.

strictness

2. *A set Y for which $D(\widetilde{Y}^+) > D(\widetilde{Y}^-)$ (resp. $D(\widetilde{Y}^-) > D(\widetilde{Y}^+)$) is called a* strict *outranking set (resp. outranked set).*

One can now translate the principle of effectiveness $\mathcal{P}_4$ into the following formal property:

$\mathcal{R}_3$: **Strict outranking**
An RCR is a strict outranking set in $\widetilde{G}(X, \widetilde{S})$.

This concept allows to solve the problem raised by the following example.

Example 6 *Consider the crisp outranking digraph represented on figure 4.3* [3] *(for the sake of simplicity we suppose that all the arcs which are drawn (resp. not drawn) represent a credibility of the outranking of* 1 *(resp.* -1*)).*

$\{a\}$ and $\{c\}$ are both irredundant outranking sets with the same maximal determinateness 1*. However, one can easily see that alternative a compares differently with b than c does. Set $\{c\}$ is clearly a null set. If we now require the three properties $\mathcal{R}_1$, $\mathcal{R}_2$ and $\mathcal{R}_3$ to be verified, only the set $\{a\}$ can be retained as a potential choice recommendation.*

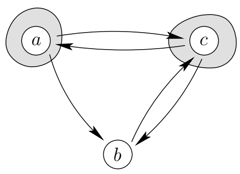

Figure 4.3: Example 6: illustration of the necessity of Property $\mathcal{R}_3$

An immediate consequence of the effectiveness principle is that a bipolar-valued outranking digraph, which is completely symmetrical, i.e., with equal credibility degrees for all xSy and ySx, does not admit any RCR. Every outranking set will automatically be a null set. Indeed, without any asymmetrical preferential statements, it is impossible to derive any preferential discriminations that would support a convincing choice recommendation.

[3] Inspired from [RB93].

$\widetilde{Y}$	a	b	**c**	d	e	$D(\widetilde{Y})$	
{**c**}	-0.2	-0.4	**0.4**	-0.4	-0.4	**0.36**	RCR
$\{b\}$	-0.2	**0.2**	-0.2	-0.2	-0.2	**0.20**	-

Table 4.2: Example 3: illustration of the maximal credibility principle

Finally, principle $\mathcal{P}_5$ (*maximal credibility*) involves again the idea of determinateness of bipolar-valued sets (see Definition 4.2.6). In the case of multiple potential choice recommendations, we recommend the most determined one, i.e., the one with the highest determinateness. Let $\widetilde{\mathcal{Y}}$ be the set of sets verifying $\mathcal{R}_1$, $\mathcal{R}_2$ and $\mathcal{R}_3$ in $\widetilde{G}(Y, \widetilde{S})$.

$\mathcal{R}_4$: **Maximal determinateness**
An RCR is a choice in $\widetilde{G}(X, \widetilde{S})$ that belongs to the set

$$\widetilde{\mathcal{Y}}^* = \big\{\widetilde{Y}' \in \widetilde{\mathcal{Y}} | D(\widetilde{Y}') = \max_{\widetilde{Y} \in \widetilde{\mathcal{Y}}} D(\widetilde{Y})\big\}. \tag{4.1}$$

Example 3 (*continued*) *Recall that in this example (see Section 3.2), we determined two outranking kernels which were potential choice recommendations (see Table 4.2). The determinateness of the kernel* $\{c\}$ (0.36) *is significantly higher than that of kernel* $\{b\}$ (0.20). *Following property* $\mathcal{R}_4$*, we recommend in this case the first solution, namely kernel* $\{c\}$.

In this section, we have presented the translation of the five Rubis principles into properties of sets of alternatives defined in the bipolar-valued outranking digraph. Detailed motivations for these principles have been given. They lead quite naturally to the new concept of outranking hyperkernel of an outranking digraph.

Remember that, in Section 4.1.3, we called a Rubis choice recommendation, a recommendation verifying the five pragmatic principles. The maximally determined strict outranking hyperkernel being a consequence of the translation of these principles into formal properties of a bipolar-valued outranking digraph, it consequently gives an adequate Rubis *choice recommendation.* RCR

The following section focuses on the construction of the hyperkernels and proposes a general algorithm for computing the Rubis choice recommendations in a given (non-symmetrical) bipolar-valued outranking digraph.

4.3 Computing the RUBIS choice recommendation

We start by presenting an algorithm which allows to determine the hyperkernels of an outranking digraph, before presenting some of their properties.

4.3.1 Determination of the hyperkernels

If $\widetilde{G}(X,\widetilde{S})$ contains chordless circuits of odd order ($\geqslant 3$), the original outranking digraph is modified into a digraph that we will call the *chordless-odd-circuits-augmented* (COCA) outranking digraph $\widetilde{G}^{\mathcal{C}}(X^{\mathcal{C}},\widetilde{S}^{\mathcal{C}})$.

Intuitively, the main idea is to "hide" the problematic circuits behind new nodes which are added to the digraph in a particular way. This may appear to be a problematic perturbation of the original information. Nevertheless, as we will see later, such a transformation does not affect the original problem but only helps to find further solutions.

The procedure to obtain the COCA digraph $\widetilde{G}^{\mathcal{C}}$ is iterative. The initial digraph is written $\widetilde{G}_0(X_0,\widetilde{S}_0)$, and is equal to $\widetilde{G}(X,\widetilde{S})$. At step i, the set of nodes becomes $X_i = X_{i-1} \cup \mathcal{C}_i$, where $\mathcal{C}_i$ is a set of nodes representing the chordless circuits of odd order of $\widetilde{G}_{i-1}(X_{i-1},\widetilde{S}_{i-1})$. These nodes are called *hypernodes*.

The outranking relation $\widetilde{S}_{i-1}$ is augmented by links between the nodes from X_{i-1} and those from $\mathcal{C}_i$ in the following way (the resulting relation is written $\widetilde{S}_i$)[4]:

$$\forall C_k \in \mathcal{C}_i \begin{cases} \widetilde{S}_i(C_k,x) = \bigvee_{y\in C_k} \widetilde{S}_{i-1}(y,x) & \forall x \in X_{i-1}\setminus C_k\ , \\ \widetilde{S}_i(C_k,x) = +1 & \forall x \in C_k\ , \end{cases} \tag{4.2}$$

$$\forall x \in X_{i-1}, C_k \in \mathcal{C}_i \begin{cases} \widetilde{S}_i(x,C_k) = \bigvee_{y\in C_k} \widetilde{S}_{i-1}(x,y) & \text{if } x \notin C_k\ , \\ \widetilde{S}_i(x,C_k) = +1 & \text{if } x \in C_k\ . \end{cases} \tag{4.3}$$

The iteration is stopped at step r for which $|X_r| = |X_{r+1}|$.

This constructions permits the hypernodes to inherit the outranking, outranked and independence properties from their corresponding odd chordless circuits. As we will see later, this inheritance property allows to construct the outranking and outranked hyperkernels.

[4]For the sake of simplicity, an element C_k of $\mathcal{C}_i$ will represent a node of X_i as well as a the set of nodes of X_{i-1} representing the circuit C_k.

We then define $\widetilde{G}^{\mathcal{C}}(X^{\mathcal{C}}, \widetilde{S}^{\mathcal{C}})$ as the digraph $\widetilde{G}_r(X_r, \widetilde{S}_r)$. As the order of the original digraph $\widetilde{G}$ is finite, the number of circuits it may contain is also finite. Therefore, the iteration is a finite process. Note that this iterative approach is necessary because of the fact that new chordless circuits of odd order may appear when new hypernodes are added to the digraph.

Figure 4.4 presents such a case. First the chordless circuit $\{a, b, c\}$ of order 3 is detected. A new node labelled $\{a, b, c\}$ is added. Then the chordless circuit $\{\{a, b, c\}, d, e\}$ of order 3 is detected. Again, a new node labelled $\{\{a, b, c\}, d, e\}$ has to be added. No further odd chordless circuit can then be detected.

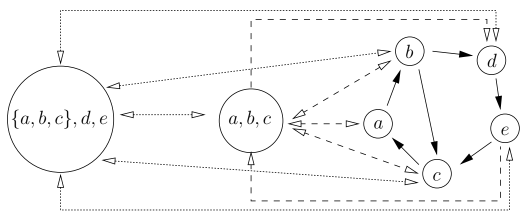

Figure 4.4: An iteration generates a new odd circuit

The outranking (resp. outranked) hyperkernels of $\widetilde{G}(X, \widetilde{S})$ are then determined by searching the classical outranking (resp. outranked) kernels of $\widetilde{G}^{\mathcal{C}}(X^{\mathcal{C}}, \widetilde{S}^{\mathcal{C}})$ [Bis97, Bis06a].

4.3.2 Properties of the COCA outranking digraph

Let us now focus on the properties of this extension of the outranking digraph.

Property 4.3.1. *The outranking (resp. outranked) kernels of* $\widetilde{G}(X, \widetilde{S})$ *are also outranking (resp. outranked) kernels of* $\widetilde{G}^{\mathcal{C}}(X^{\mathcal{C}}, \widetilde{S}^{\mathcal{C}})$.

Proof. If $\widetilde{G}$ does not contain an odd chordless circuit, $\widetilde{G}^{\mathcal{C}} = \widetilde{G}$ and the property is trivial.

Let us suppose that $\widetilde{G}$ contains at least one odd chordless circuit. Let Y be an outranking kernel of $\widetilde{G}$ (the case of the outranked kernels can be treated similarly). We must prove that Y is also an outranking kernel of $\widetilde{G}^{\mathcal{C}}$.

First, the elements of Y are independent in $\widetilde{G}$ and $\widetilde{G}^{\mathcal{C}}$ because no relation is added between elements of X in $X^{\mathcal{C}}$. Second, as Y is an outranking set in $\widetilde{G}$, each element of $X \setminus Y$ is outranked by at least one element of Y. In particular, if C_k is an odd chordless circuit of X, each node of C_k is also outranked by at least one element of Y (in X). Due to the special way $\widetilde{S}^{\mathcal{C}}$ is built, the node representing C_k in $X^{\mathcal{C}}$ is also also outranked by at least one element of Y. This remains true if at least one node of C_k belongs to Y. □

Following from the construction principle of the COCA digraph (see Equations 4.2 and 4.3), a hypernode *inherits* the outranking, outranked and independence characteristics of its corresponding odd chordless circuit. Furthermore, the individual nodes of each odd chordless circuit are outranked by and are outranking the hypernode with a credibility of $+1$ (which could be called *indifference*).

Property 4.3.2. *The digraph* $\widetilde{G}^{\mathcal{C}}(X^{\mathcal{C}}, \widetilde{S}^{\mathcal{C}})$ *contains at least one outranking (resp. outranked) kernel.*

Proof. If $\widetilde{G}$ contains an outranking (resp. outranked) kernel, then via Property 4.3.1, this remains valid for $\widetilde{G}^{\mathcal{C}}$.

Let us suppose that $\widetilde{G}$ contains no outranking kernel (a similar proof can be done for outranked kernels). Via Property 4.2.1, this implies that $\widetilde{G}$ contains at least one odd chordless circuit.

Consequently, there exists at least one irredundant outranking set Y in X which contains at least one arc of at least one odd chordless circuit. This results from the fact that there exists no bipartition of an odd chordless circuit into an outranking and an outranked kernel.

Let us suppose that the irredundant outranking set Y contains a single such arc, belonging to the odd chordless circuit C_k, which we denote by xSy (the case of multiple such arcs can be treated similarly). By construction, in $\widetilde{G}$, there exists a hypernode which inherits the outranking and outranked properties of C_k. In particular, it inherits the fact that x and y are independent of the remaining nodes of Y.

Consequently, the set $Y \setminus \{x, y\} \cup C_k$ is irredundant, outranking and independent in $\widetilde{G}^{\mathcal{C}}$. Consequently it is a kernel of $\widetilde{G}^{\mathcal{C}}$. □

An important consequence of the previous properties is given in the following theorem.

Theorem 4.3.1. *Any outranking digraph* $\widetilde{G}(X, \widetilde{S})$ *contains at least one outranking (resp. outranked) hyperkernel.*

Proof. This theorem follows directly from Property 4.3.2: if any COCA outranking digraph contains at least one outranking (resp. outranked) kernel, then consequently, any outranking digraph contains at least one outranking (resp. outranked) hyperkernel. □

Let us now present and discuss the computing of a Rubis choice recommendation in the following section.

Algorithm 2 The RUBIS algorithm

Input: $\widetilde{G}(X, \widetilde{S})$,

1. Construct the associated COCA digraph $\widetilde{G}^{\mathcal{C}}(X^{\mathcal{C}}, \widetilde{S}^{\mathcal{C}})$,
2. Extract the sets $\widetilde{K}^{+}$ and $\widetilde{K}^{-}$ of all outranking and outranked hyperkernels from $\widetilde{G}^{\mathcal{C}}$,
3. Eliminate the null kernels from $\widetilde{K}^{+}$,
4. Rank the elements of $\widetilde{K}^{+}$ by decreasing logical determinateness,

Output: The first ranked element(s) in $\widetilde{K}^{+}$.

4.3.3 The RCR algorithm

The first step of the RCR algorithm is by far the most difficult to achieve, as the number of odd chordless circuits in a bipolar-valued outranking digraph can be huge.

To study this operational difficulty, we have compiled a sample of 1000 bipolar-valued outranking digraphs. One possibility would have been to generate them as such, without considering underlying performance tables. This would have produced a set of digraphs which would not be very representative for *real* MCDA problems. Instead, we chose to generate them from performances of 20 alternatives evaluated randomly on 7 to 20 criteria with random weights distributions and random thresholds.

In nearly 98% of the sample, the time to compute the COCA digraph on a standard desktop computer is less than a second. In one case, we observe an execution time of around 30 seconds (due to a high number of odd chordless circuits in the digraph).

In Table 4.3, we note that nearly 75% of the sample digraphs do not admit any odd chordless circuit at all. In 100% of the observations less than 10 hypernodes are added to the original outranking digraph.

The second step of the RCR algorithm concerns the extraction of hyperkernels from the COCA digraph. From a theoretical point of view, this step is well-known to be computationally difficult [Chv73]. However, this difficulty is directly linked to the arc-density, i.e., the relative size of the digraph. Indeed, only very sparse digraphs, showing an arc-density lower than 10% in the range of digraph orders which are relevant for the choice decision aiding problematique (10-30 alternatives), may present difficulties for the search of kernels. Figure 4.5 shows the histogram of the distribution of the arc-density on the

number of odd chordless circuits	#	rel. freq. (%)	cum. freq. (%)
0	735	73.5	73.5
1	116	11.6	85.1
2	65	6.5	91.6
3	25	2.5	94.1
⋮	⋮	⋮	⋮
9	1	0.1	100.0

Table 4.3: Number of odd chordless circuits in random bipolar-valued outranking digraphs of order 20 (1000 observations).

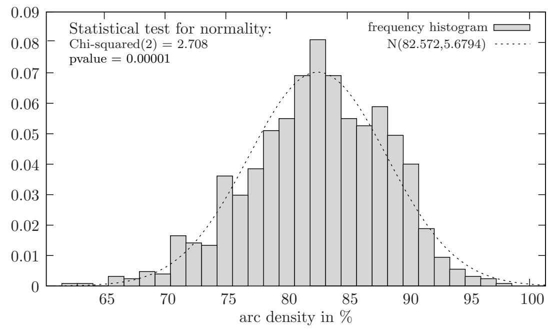

Figure 4.5: Histogram of the arc density of a sample of 1000 COCA digraphs of order 20

sample of 1000 random outranking digraphs of order 20. For this test sample, we observe a very high mean density of 82.6% with a standard deviation of 5.7%. Consequently, determining hyperkernels is in general a task which is feasible in a very reasonable time. Indeed, the mean execution time with its standard deviation for this step of the algorithm are around a thousandth of a second on a standard desktop computer.

Finally, eliminating the null hyperkernels and sorting the strict outranking hyperkernels in decreasing order of determinateness is linear in the order of the digraph and involves no computational difficulty at all.

Let us mention a few more results related to this simulation, which allow to better understand the structures of the underlying problems. The average number of outranking hyperkernels is 6.629 per digraph (standard deviation of 3.053). Among these, the average number of strict outranking hyperkernels is

6.000 (standard deviation of 2.565) and the average number of null sets is 0.208 (standard deviation of 0.526). Among the strict outranking hyperkernels, on average 0.112 were containing hypernodes.

Let us illustrate the RCR algorithm on Example 4 (see Section 4.2).

Example 4 (*continued*) *The bipolar-valued outranking digraph of this example (see Figure 4.1) contains a chordless circuit of order 3, namely* $\{a, b, d\}$. *The original digraph* $\widetilde{G}$ *is extended to the digraph* $\widetilde{G}^{\mathcal{C}}$ *which contains a hyper-node representing* $\{a, b, d\}$.

The corresponding outranking digraph admits an outranking kernel $\{a, c\}$ *and a hyperkernel* $\{\{a, b, d\}, e\}$ *which is both outranking and outranked, but not with the same degree of determinateness (see Table 4.4). The first one is significantly more determined than the second one. Consequently, the* RUBIS "choice recommendation" *is* $\{\{a, b, d\}, e\}$, *where alternative* e *is in an indetermined situation.*

$\widetilde{S}^{\mathcal{C}}$	a	b	c	d	e	$\{a,b,d\}$	D
a	0.1	0.2	-1.0	-0.7	-0.8	1.0	
b	-0.6	1.0	0.8	1.0	0.0	1.0	
c	-1.0	-1.0	1.0	0.2	0.8	0.2	
d	0.6	-0.6	-1.0	1.0	-0.4	1.0	
e	-1.0	-8	-0.4	-0.6	1.0	-0.6	
$\{a,b,d\}$	1.0	1.0	0.8	1.0	0.0	1.0	
$\{\{a,b,d\},e\}^+$	-0.6	-0.6	-0.6	-0.6	**0.0**	**0.6**	**0.5**
$\{a,c\}$	**0.2**	-0.2	**0.2**	-0.2	-0.2	-0.2	0.2
$\{\{a,b,d\},e\}^-$	0.0	0.0	0.0	-0.6	0.0	0.6	0.2

Table 4.4: Example 4: the associated COCA digraph with the bipolar-valued characterisations of its outranking (+) and outranked (-) hyperkernels.

Before finishing this chapter, let us return to a concept which was introduced earlier in Section 3.2, namely the β-cut of a bipolar-valued outranking relation.

4.3.4 On the β-cut of the bipolar-valued outranking relation

As already said, in practice, a β-cut applied on the RUBIS choice recommendations is used to extend the indetermination to a larger range of values. This allows the DM to change the *majority threshold* and to focus on *strong* conclusions.

Let us illustrate its use in practice, and return to Example 3.

Example 3 (*continued*) *Imagine that the DM considers that for an outranking statement to be validated (resp. non-validated), it requires that a* weighted *majority of more than* 65% *of the criteria supports it.*

This means that all propositions which have credibility degrees below 0.3 and above -0.3 should be considered as indetermined. The effect of this β-cut is presented in Table 4.5. It shows in particular that at this level of β it is necessary to consider $\{a, c\}$ as a RCR, where a is in an indetermined situation ($\widetilde{S}(c, a) = 0$ after the β-cut; see Table 3.1 and Figure 4.6).

$\widetilde{Y}$	a	b	$\mathbf{c}$	d	e	$D(\widetilde{Y})$	
$\{\mathbf{a}, \mathbf{c}\}$	**0.0**	-0.4	**0.4**	-0.4	-0.4	**0.32**	RCR
$\{b\}$	0.0	0.0	0.0	0.0	0.0	0.0	-

Table 4.5: Example 3: illustration of the effect of a β-cut

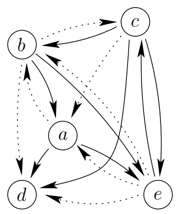

Figure 4.6: Example 3: the β-cut digraph

Let us finish this section by indicating that all the examples of this chapter have been computed with the free Python module *digraphs* [Bis06b] which allows to manipulate bipolar-valued digraphs and to determine the RCR from a given performance table.

4.4 Further remarks on progressiveness in the Rubis method

It is obvious that the Rubis method fits in the framework of progressive methods. In Section 2.2.1 we outlined the general structure of a progressive resolution of the choice problematique. The Rubis method is totally in line with this presentation.

In practice, it may happen that the RUBIS algorithm produces no RUBIS choice recommendation because all the outranking hyperkernels are *null* sets. In such a situation it is recommendable to reanalyse the whole problem in order to construct a less symmetrical outranking relation. Another solution would be to use alternative definitions of the degree of determinateness D or the strictness of a set.

It may also happen that the RUBIS algorithm generates multiple RUBIS choice recommendations with the same degree of determinateness. In such a situation, two options appear: either the DM selects one of the recommendations and continues the process on this set, or he considers the union of the recommendations for the next step of the progressive analysis. The second option is obviously more *prudent*, as any potential best alternative is reanalysed at a later stage.

In less problematic cases, at each step, the DM is asked to focus on progressively smaller sets, containing the potential candidates for the choice. The hard comparability of the alternatives of those subsets originates from three possible situations:

- Odd chordless circuits;
- Strict independence;
- Indetermination.

In the first case, the alternatives belonging to an odd chordless circuit can be seen as more or less equivalent. It is therefore useful to continue the progressive analysis to try to discriminate them further in order to be able to eliminate some of them from the choice recommendation.

In the second situation, alternatives from the choice are considered as strictly independent. A further step of the progressive analysis is then meant to determine new or more precise information on those alternatives, in order to make them preferentially comparable. Again, the goal is to reduce the choice set by eliminating supplementary alternatives.

Finally, in indetermination situations, the pairwise comparisons of such alternatives should be enriched with more precise or supplementary information in view of generating a higher discrimination. As discussed earlier, indeterminations might also be the result of a β-cut of the outranking relation. In such a case, it could be advisable to lower the value of β in order to resolve certain indetermined situations.

All in all, in any of the possible states, the objective of a progressive analysis of the problem is to reduce the choice set in view of uncovering the single best alternative.

Chapter 5

The k-choice problematique

Contents

In this chapter we discuss the problem of choosing k best alternatives in the context of outranking-based models. We call this typology of problems the *k-choice problematique*. Our study is grounded on the methodological studies of Chapter 4, where we detailed the construction of a recommendation for the choice problematique. As we will show, this problem can be seen as a particular case of the k-choice problematique, where $k = 1$. Furthermore we will underline that k-choice decision analyses can have different formulations depending on the operational objectives of the underlying required decision problem.

k-choice problematique

The chapter is organised as follows. In Section 5.1, we present three formulations of the k-choice problematique. The second section deals with the resolution of two of these problems and in Section 5.3, we present a small example which shows the differences between the two latter formulations.

Note that the considerations of this chapter are mainly based on our article [MB07]. They represent our latest research in outranking methods, and are consequently only first results. Future developments on the k-choice problematique will be our priority in our future research activities.

5.1 Three formulations of the k-choice problematique

Solving the choice problematique amounts to helping the DM to determine a single alternative which can be considered as the best one. It is however less obvious to give a single definition of the k-choice problematique. Indeed, different formulations to the selection of k best alternatives can be given:

$\mathcal{K}_1$ Search for the first k best alternatives (k first-ranked);

$\mathcal{K}_2$ Search for a set of k alternatives better than any other coalition of k alternatives (best k-team);

$\mathcal{K}_3$ Search for a set of k alternatives better than all the other alternatives (best k-committee).

Other formulations of the k-choice problematique might possibly exist, but in this work we will focus on these three definitions.

Let us now detail these formulations one by one.

5.1.1 $\mathcal{K}_1$: considerations on the k first-ranked problem

This first formulation $\mathcal{K}_1$ corresponds probably to what people have commonly in mind when they think about "*selecting the best k alternatives*" among a set of decision objects:

> Consider the k objects ranked in the first k positions of a complete order or a weak order (a ranking).

As we already mentionned, in the context of MCDA, such a ranking can hardly be obtained in the framework of pairwise comparisons of alternatives. Indeed, the outranking relation which results from such pairwise comparisons is in general neither transitive nor complete (some alternatives may be incomparable in terms of the outranking relation). Furthermore, in case the outranking relation is a partial order (or a partial weak order), it is difficult to conceive what the k first positions of the ranking could be.

These observations show that the outranking relation can difficultly be used directly to solve the problem of the k first-ranked alternatives. To overcome this problem, the outranking relation must first be exploited in order to build a total order or at least a weak order (see for example [Bis99] or the Electre II, III and IV methods [RB93] or the PROMETHEE I and II method [BM94, BM05]).

If such an exploitation is adopted, it might then possible to rank (possibly with ties or incomparabilities) the alternatives from the *best* to the *worst* one and to determine the k first ones. In case of indifferences in a weak order, selecting the k first ones might not be possible and it will be necessary to select $k' > k$ alternatives and proceed via a supplementary decision aiding step to determine the k first ranked alternatives. A similar difficulty arises in case of a partial order.

Note here that another possibility to achieve total comparability of the alternatives in MCDA is via Multiattribute Value Theory, where a weak order on the alternatives is generated by means of an overall utility function. This subject is the purpose of Part III of this work.

From the previous considerations, it is possible to derive a quite obvious, but important and very general property concerning the k first-ranked problem.

Property 5.1.1. *Let Y_i be the set of i first-ranked alternatives. $\forall k \leq n$, if Y_{k-1} and Y_k exist, we have*

$$Y_{k-1} \subset Y_k.$$

This property is simply a translation of a quite natural intuition: the $k-1$ first-ranked alternatives also belong to the set of k first-ranked alternatives. The possible non-existence of Y_{k-1} or Y_k is simply due to the difficulty which arises in case of ties in weak orders or of incomparabilities in partial orders.

It is also obvious that in the case where $k = 1$, the k first-ranked problem amounts to selecting the first (and therefore best) alternative in the ranking.

Due to the necessity of exploiting the outranking relation in order to obtain a ranking, we will not explore this option further here. We will rather focus on the remaining two formulations $\mathcal{K}_2$ and $\mathcal{K}_3$, which are solely based on the outranking digraph, and which produce different results in general.

5.1.2 $\mathcal{K}_2$: introducing the best k-team problem

The second formulation of the k-choice problem in a set X of alternatives can be summarised by the following intuitive procedure:

> Search for a subset Y of X of cardinality k which *is better than* any other set of cardinality k.

The main difficulty lies in the formal definition of the "*is better than*"-relation for this particular case. Nevertheless, before dealing with this problem, let us first present practical situations in which the determination of the best k-team is applicable.

A first potential practical context is given by any situation where teams of k persons have to compete against each other and where each person has been evaluated individually on the family N of criteria. In our context, a pairwise comparison of all possible teams of k persons is then performed on the basis of the individual outrankings.

A second kind of situations is given by facility location problems, where k locations have to be selected simultaneously. Again, all the eligible combinations of k locations are pairwisely compared to determine which one is the most appropriate.

Very generally, the best k-team problem is applicable in any decision situation where sets of alternatives have to be compared in view of choosing one of them. Let us now turn to a more formal definition of the best k-team problem.

k-set
k-team

We call a *k-set* a set of k alternatives. In the context of $\mathcal{K}_2$, we call a *k-team* such a k-set. Recall that the available information is an outranking relation on the single alternatives. Our objective is here to build an outranking relation on the k-teams. In this framework, it is quite natural to require that the following conditions are verified by a k-team:

$\mathcal{T}_1$ **Inheritance**
A k-team *inherits* the outranking and outranked properties of its members;

$\mathcal{T}_2$ **Intra-team indiscernibility**
A k-team is considered as a entity from the *outside*;

$\mathcal{T}_3$ **Exclusive inter-team comparisons**
Two k-teams are exclusively compared on basis of *inter-team* information.

The first property originates from the following observation. If an alternative $y \in X$ certainly outranks an alternative $y' \in X$ and if y and y' respectively belong to k-teams Y and Y', then this positive information for Y and negative information for Y' should be reflected in the way the two k-sets are compared.

The objective of the second property is to make a k-team act as a *coherent entity*. In such a situation, the elements of a k-team should act together as a coalition. Consequently, when compared to the other members of the k-team, a given alternative's weakness or strength should not be regarded.

The third property is useful in the case where two teams which have a non-empty intersection are compared. It is a consequence of the second property on the comparison of k-teams. As the k-sets should be considered as an entity, using *intra-team* information for their comparisons is not appropriate. Consequently, we require that two k-teams are only compared on basis of information which is not linked to their intersection.

Let us present a short example which allows to better understand the consequences of the three principles $\mathcal{T}_1$ to $\mathcal{T}_3$.

Example 7 *Consider a set of 4 alternatives $X = \{a, b, c, d\}$ and an outranking relation built from pairwise comparisons of these decision objects, $S = \{(a, d), (b, c), (c, d)\}$. Recall that S can be recovered from its bipolar-valued characterisation $\widetilde{S}$ (see Table 5.1 and Figure 5.1). Let us analyse how the three 2-teams $\{a, b\}$, $\{a, c\}$ and $\{b, c\}$ should be pairwisely compared.*

Sets $\{a, b\}$ and $\{b, c\}$ have alternative b in common, which outranks alternative c. The arc between b and c is internal *to the set $\{b, c\}$. In accordance with principle $\mathcal{T}_3$, this information should not be taken into account when comparing $\{a, b\}$ and $\{b, c\}$. Therefore with the available outranking information, these two sets are incomparable.*

Sets $\{a, b\}$ and $\{a, c\}$ have alternative a in common. In that case, the arc between b and c is clearly inter-team *information and the set $\{a, b\}$ should outrank the set $\{a, c\}$.*

$\widetilde{S}_2$	a	b	c	d
a	1	< 0	< 0	> 0
b	< 0	1	> 0	< 0
c	< 0	< 0	1	> 0
d	< 0	< 0	< 0	1

Table 5.1: Example 7: generic table representing $\widetilde{S}$

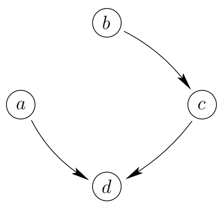

Figure 5.1: Example 7: crisp outranking digraph

This example intuitively explains how the comparison of the sets of alternatives in the k-team problem should be performed on the basis of the outranking relation built on pairs of alternatives. Note that a more detailed analysis of this example is given in Section 5.3.

The three properties above lead quite naturally to the following literal definition of the outranking relation on the set of k-teams. The following definition outlines how the k-team should inherit the outranking and outranked characteristics of its elements, in accordance with principle $\mathcal{T}_1$.

k-team outranking relation

Definition 5.1.1. *Let $Y, Y' \subset X$ be two k-teams. Y outranks (resp. is outranked by) Y' if $\exists(y, y') \in (Y \times Y') \setminus (Y \cap Y')^2$ s.t. ySy' (resp. $y'Sy$).*

We will detail the construction of the outranking relation between pairs of sets of alternatives in Section 5.2. Note already at this point that due to principle $\mathcal{T}_2$, no condition is imposed on the incomparability (or independence) of the alternatives in a k-team.

In this definition, the outranking relation between k-teams is build *disjunctively*. Note here that other definitions could of course be deduced from the principles. For example, one could require that a k-team Y outranks another k-team Y' if all the alternatives of Y outrank those of Y'. A weaker formulation would be to consider the outranking validated if a majority of elements of Y outranks those of Y'.

In practice certain combinations of alternatives may be meaningless or unachievable. This can be illustrated, e.g., in a facility location problem, where it is important that the locations are geographically well spread. In such a situation, some sites may have very good individual evaluations, but it would not make sense to combine them with some other locations. It might therefore be recommendable to have a prior analysis in order to determine the eligible combinations of alternatives.

In Section 5.2.1 we will show how to solve the k-team problem based on the previous considerations, Definition 5.1.1 and the Rubis choice method. Let us finish by situating the k-team problem in the context of progressive decision aiding methods.

In practice the ultimate objective of this problematique is to determine a unique k-team which is considered as the *best* set of k alternatives. Nevertheless, in a progressive method, it may be necessary to go through a few intermediate steps, where at each step, some k-teams are rejected for well motivated reasons (similarly to what is done in the Rubis choice method).

5.1.3 $\mathcal{K}_3$: introducing the k-committee problem

In this subsection we start by giving a third intuitive definition of what the selection of k best alternatives could be:

> Search for a subset Y of X of cardinality k which is in its integrality *better than* all the other alternatives.

Again, the problem here is to understand what the "*is better than*"-relation signifies in this particular case. Similarly as for the previous formulation, let us start by presenting a type of problem that the search for the best k-committee could address.

A potential practical context is given by any situation where in a set X of persons, a subset Y of k of them has to direct, pilot or command the remaining ones (for example a committee). In that case, each non-retained person of $X \setminus Y$ has to be considered as "*less preferred*" than Y in its collectiveness.

The main difference with the previous formulation $\mathcal{K}_2$ is that here, sets of alternatives have to be compared to single alternatives. In this context, a k-set is called a *k-committee*. Again, the initial information is given by an outranking relation on the set of alternatives. The objective here is to build an outranking relation which allows to compare k-sets to single alternatives. We therefore require in this context that the following principles are verified for any k-committee. *k-committee*

$\mathcal{C}_1$ **Inheritance**
A k-committee *inherits* the outranking and outranked properties of its members;

$\mathcal{C}_2$ **Intra-committee indiscernibility**
A k-committee should be considered as an *entity* from the *outside*;

$\mathcal{C}_3$ **Inter-committee comparisons**
k-committees are pairwisely compared via the alternatives they outrank.

Principles $\mathcal{C}_1$ and $\mathcal{C}_2$ are identical to $\mathcal{T}_1$ and $\mathcal{T}_2$ for the best k-team problem.

Principle $\mathcal{C}_3$ clearly shows the main difference between the best k-team and the best k-committee problem. Committees are compared via the single alternatives they outrank, whereas teams are compared to other teams.

Let us again analyse on basis of Example 7 how committees behave in an outranking digraph.

> **Example 7** (*continued*) *In this example, the following three 2-committees are considered: $\{a, b\}$, $\{a, c\}$ and $\{b, c\}$.*
>
> *Set $\{a, b\}$ has to be compared to alternatives c and d. c is outranked by b and d is outranked by a. Therefore, the set $\{a, b\}$ should outrank both b and d (inheritance principle $\mathcal{C}_1$). Set $\{b, c\}$ has to be compared to alternatives a and d. A similar reasoning as before leads to the fact that the set $\{b, c\}$ should outrank d but be incomparable to a. Finally, the set $\{a, c\}$ should be outranked by alternative b and outrank alternative d.*

Note again that a detailed analysis of this example in the case of the best k-committee problem is presented in Section 5.3.

The k-committee problem can now be defined as the search for a set Y of k alternatives which is better than all the alternatives which are not in

Y. Combined to the three principles, this leads to a literal definition of the outranking relation in the case of the k-committee problem.

k-committee outranking relation

Definition 5.1.2. *Let $Y \subset X$ be a k-committee and $x \in X$. Y outranks x if $\exists y \in Y$ s.t. ySx.*

Note again that other definitions could be deduced from the pragmatic principles and that no condition is given on the incomparability (or independence) of the alternatives in a k-committee. Furthermore, similarly as for the k-team problem, it may be necessary to determine first the feasible k-committees by a filtering process.

5.2 Solving $\mathcal{K}_2$ and $\mathcal{K}_3$

In this section we present how the latter two formulations of the k-choice problematique can be solved. In both cases the bipolar-valued outranking digraph, built on the set of alternatives, needs to be modified to obtain the desired choice recommendation. As we will show, the RUBIS method presented in Chapter 4 will be the backbone of the resolution algorithms. The main motivation to use this method is because it is based on pragmatic foundations which are developed in Chapter 4.

5.2.1 $\mathcal{K}_2$: best k-team

In view of the discussions of Section 5.1.2 one can easily see that the best k-team problem can be solved in an outranking digraph $\widetilde{G}^t(X^t, \widetilde{S}^t)$ where the nodes represent all possible sets of k alternatives.

If the outranking relation $\widetilde{S}^t$ is appropriately defined (see Definition 5.1.1), then the progressive search for the best k-team amounts to the progressive search for the RUBIS choice recommendation in that new digraph.

Let us now detail the construction of $\widetilde{G}^t$.

The nodes X^t of $\widetilde{G}^t$ represent all admissible subsets of k alternatives of X (recall that certain subsets could possibly be considered as obsolete). The cardinality of X^t is therefore at most $\binom{n}{k}$, which might be quite large for certain combinations of n and k. We label the nodes of X^t by capital letters in the sequel[1] as they represent subsets of X.

The outranking relation $\widetilde{S}^t$ is built as follows (based on Definition 5.1.1):

$$\begin{aligned}&\forall (V, W) \in X^t \times X^t : \\ &\quad \widetilde{S}^t(V, W) = \max\{\widetilde{S}(v, w) : (v, w) \in (V \times W) \setminus (V \cap W)^2\}.\end{aligned} \tag{5.1}$$

[1]For the sake of simplicity, an element Z of X^t will represent a node of $\widetilde{G}^t(X^t, \widetilde{S}^t)$ as well as a subset of X.

The crisp outranking relation S^t associated to $\widetilde{S}^t$ can be recovered as the set of pairs $(V, W) \in X^t \times X^t$ such that $\widetilde{S}^t(V, W) > 0$.

As mentionned in Section 3.1, the max operator models the credibility of the disjunction of logical statements. In fact, for a k-team Y to outrank another k-team Y', it is sufficient that one alternative of Y (positively) outranks another alternative of Y'. Furthermore, this aggregation of the outrankings allows in some way to model *complementarity* among the different alternatives of a k-team. All in all, the construction of $\widetilde{S}^t$ as detailed in formula 5.2.1 clearly satisfies principles $\mathcal{T}_1$, $\mathcal{T}_2$ and $\mathcal{T}_3$.

It is now obvious that all the concepts introduced in Chapter 4 can be used in $\widetilde{G}^t(X^t, \widetilde{S}^t)$ and have a signification in $\widetilde{G}(X, \widetilde{S})$ in terms of subsets of alternatives. For example, a hyperindependent choice in $\widetilde{G}^t(X^t, \widetilde{S}^t)$ is a choice in $\widetilde{G}(X, \widetilde{S})$ which is composed of independent odd chordless circuits of subsets of X.

The objective of $\mathcal{K}_2$ can now be reinterpreted in $\widetilde{G}^t(X^t, \widetilde{S}^t)$. The goal of $\mathcal{K}_2$ in $\widetilde{G}^t(X^t, \widetilde{S}^t)$ is to select one unique k-set (or node) which is considered as the best one. This definition is very comparable to the search for one best alternative in an outranking digraph (see Section 4.1.1).

Consequently, in the context of a progressive method for the determination of the best k-team, the solution is to apply the Rubis method to the digraph $\widetilde{G}^t(X^t, \widetilde{S}^t)$. As already mentioned, it will exploit the bipolar-valued outranking relation in order to extract at least one maximally determined strict outranking hyperkernel (the choice recommendation). The elements of this hyperkernel are subsets of k elements of X which are incomparable, in an indetermined situation, or, considered as equivalent in an odd chordless circuit.

In the case where the choice recommendation is unique and only contains one element V of $\widetilde{G}^t(X^t, \widetilde{S}^t)$, then the problem is solved and V is a subset of k alternatives of X which can be considered as the best k-team.

If the choice recommendation contains more than one element of $\widetilde{G}^t(X^t, \widetilde{S}^t)$, then these k-teams should not be considered as the *best* ones, but merely as a collection of hardly comparable subsets of alternatives, among which the best k-team can be found. Similarly, in the case of multiple choice recommendations of equal determinateness, it is recommendable to continue the progressive search with the union of the elements of the choice recommendations. Indeed, the only certain information is that some k-teams could be set aside for well motivated reasons. In the next step of the progressive method the decision maker can restrict his analysis to these potential k-teams and refine their evaluations or evaluate their members on further criteria.

Note that one can easily verify that the search for one best alternative as defined in Chapter 4 is a particular case of the best k-team problem for which $k = 1$.

5.2.2 $\mathcal{K}_3$: best k-committee

Recall that the goal of the search for the best k-committee is to determine a set Y of cardinality k which is as a whole better than all the other alternatives. The problem will again be solved in a modified bipolar-valued outranking digraph $\widetilde{G}^c(X^c, \widetilde{S}^c)$, but this time its construction is less obvious than for $\mathcal{K}_2$. We will nevertheless show that after a proper construction of $\widetilde{G}^c$, the Rubis method can again be applied in that new digraph to solve the best k-committee problem.

This time we focus on comparisons between sets of nodes and single alternatives. Therefore, the set X^c is defined as the union of X and a set of supplemental nodes which represent all possible subsets of k nodes of X. We use the same conventions as for $\mathcal{K}_2$ and consequently label the supplemental nodes (called *k-nodes*) in X^c by capital letters. Note that the nodes of X in X^c are still labelled with lower case letters.

k-node

The construction of $\widetilde{S}^c$ is somewhat trickier. The original relation $\widetilde{S}$ is included into $\widetilde{S}^c$, which is then built as follows:

$$\begin{aligned}
\forall (V,W) \in X^c \times X^c: &\quad \widetilde{S}^c(V,W) = 1; && (1)\\
\forall (V,w) \in X^c \times X^c \text{ s.t. } w \in V: &\quad \widetilde{S}^c(V,w) = 1 \text{ and } \widetilde{S}^c(w,V) = 1; && (2)\\
\forall (V,w) \in X^c \times X^c \text{ s.t. } w \notin V: &\quad \widetilde{S}^c(V,w) = \max\{\widetilde{S}(v,w) : v \in V\}; && (3)\\
\forall (v,W) \in X^c \times X^c \text{ s.t. } v \notin W: &\quad \widetilde{S}^c(v,W) = \max\{\widetilde{S}(v,w) : w \in W\}; && (4)\\
\forall (v,w) \in X^c \times X^c: &\quad \widetilde{S}^c(v,w) = \widetilde{S}(v,w). && (5)
\end{aligned}$$

Let us explain this construction in further details.

Formula (1) puts any two k-committees in a situation of equivalence. This is merely technical and will become clear in the presentation of Algorithm 3. Formula (5) simply expresses that the original outranking relation is included in $\widetilde{S}^c$.

In Formula (2) k-committees are quite naturally considered as equivalent to their members. Formulae (3) and (4) allow the comparison of k-committees to the remaining alternatives, pursuant to Definition 5.1.2.

Let us now turn to the determination of the best k-committee, at a given step of the progressive search. In $\widetilde{G}^c$, the search for the set of k alternatives,

Algorithm 3 The best k-committee algorithm

Input: $\widetilde{G}^c(X^c, \widetilde{S}^c)$

1. Search for the set $\mathcal{I}^c$ of irredundant outranking choices of $\widetilde{G}^c(X^c, \widetilde{S}^c)$ containing exclusively k-nodes;
2. $\forall\ Y^c \in \mathcal{I}^c$:
   - Remove any k-nodes from X^c which are not in $\mathcal{I}^c$ $(:= X^c_*)$;
   - if $|Y^c| = 1$ then determine the Rubis best choice in $\widetilde{G}^c_*(X^c_*, \widetilde{S}^c)$, containing exclusively k-nodes;
   - else :
     i. modify $\widetilde{S}^c$ as follows into $\widetilde{S}^c_*$:

$$\begin{aligned} &\widetilde{S}^c_*(V, W) = -1 && \forall (V, W) \in Y^c \times Y^c; \\ &\widetilde{S}^c_*(x, y) = \widetilde{S}^c(x, y) && \text{else.} \end{aligned}$$

     ii. Determine the Rubis best choice in $\widetilde{G}^c_*(X^c_*, \widetilde{S}^c_*)$, containing exclusive k-nodes;
3. Select the most determined bipolar-valued Rubis best choice(s) among all those determined at step 2;

Output: a single (resp. a set of) Rubis choice recommendation(s).

---

which is in its entirety better than any other alternative, amounts to determining at least one k-node V which outranks all the alternatives $x \in X$.

Algorithm 3 presents the general resolution scheme for the k-committee problem. Due to the particular way we construct $\widetilde{G}^c$ (and in particular $\widetilde{S}^c$), the output of the first step is one or more irredundant outranking choices containing exclusively k-nodes (the potential candidates for the best k-committee). This clearly shows that the first stage of the algorithm is used for filtering purposes.

In the second step, the Rubis best choice algorithm is applied to a modified graph for each irredundant outranking choice determined in the first step. Each digraph $\widetilde{G}^c_*$ is composed of the original alternatives and an irredundant outranking choice. The modification of the outranking relation consists in removing the "equivalence" arcs that link the potential k-committees in the outranking choice. This allows the Rubis algorithm to determine the desired strict choice recommendation.

Similarly as earlier for the k-team problem, if the output is not a single k-node, then the progressive search must be reapplied to the set of potential best k-committees (and the set of alternatives which compose the k-committees).

5.3 Illustrative example

In this section we develop a detailed description of the Example 7 presented in Section 5.1. In order to simplify the notations, we will label the alternatives of X^t by concatenations of the labels of the alternatives of X. For example, the node of X^t representing the subset $\{a, b, c\}$ of X will be labelled abc. We will furthermore suppose that any possible k-team and any k-committee is feasible.

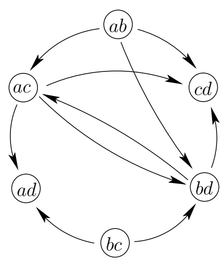

Figure 5.2: Crisp outranking digraph for the best 2-team problem

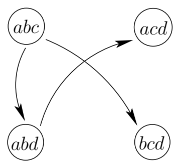

Figure 5.3: Crisp outranking digraph for the best 3-team problem

Two possible best k-team searches can be performed on this example (namely for $k = 2$ and $k = 3$). Both situations are represented on Figures 5.2 and 5.3.

The Rubis choice recommendation for the 2-team problem is given by the set $\big\{\{a, b\}, \{b, c\}\big\}$. These two potential candidates as a best 2-team are incomparable and are therefore selected for a further analysis. This signifies that in the next step of the progressive method, the DM can focus on these two subsets of alternatives in order to determine which one is the best one. The other subsets of 2 alternatives can already be rejected without any regret at this stage of the progressive decision aiding process.

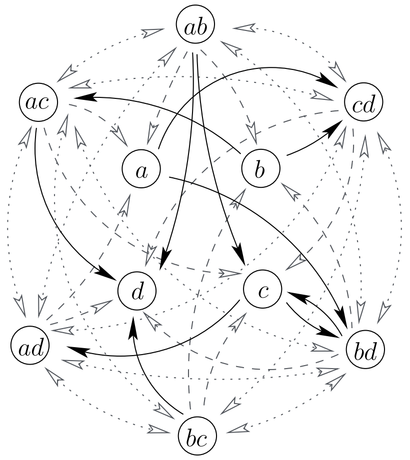

Figure 5.4: Crisp outranking digraph for the best 2-committee problem

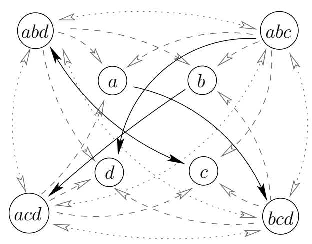

Figure 5.5: Crisp outranking digraph for the best 3-committee problem

The Rubis choice recommendation for the 3-team problem is given by the set $\{\{a, b, c\}, \{a, c, d\}\}$. Again, these two potential candidates for a best 3-team are incomparable and the best candidate might be found in a further step of the decision aiding process.

In case of the search for the best k-committee, again two searches can be performed (namely for $k = 2$ and $k = 3$). Both situations are represented on Figures 5.4 and 5.5. The dotted (resp. dashed) arcs represent the crisp relations of type (1) (resp. (2)) from the definition of $\widetilde{S}^c$. As one can clearly see, they are merely technical arcs to allow the use of the Rubis choice recommendation algorithm.

The Rubis choice recommendation for the 2-committee problem is given by the choice $\{a, b\}$. Indeed, the set $\{a, b\}$ clearly respects the definition of the 2-committee, namely that both alternatives are outranking in their entirety c and d.

For the best 3-committee problem, two potential k-sets could be considered: either the choice $\{a, b, c\}$ or the choice $\{a, b, d\}$. The final selection of either one (or both) of these choices as choice recommendation(s) will depend on their determinateness and / or their strictness. Both concepts directly depend on the precise values of the bipolar-valued characterisations of the two potential choices.

Part III

Preference elicitation in Choquet integral-based Multiattribute Value Theory and exploitation

Part III: Preference elicitation in Choquet integral-based Multiattribute Value Theory and exploitation

Pourquoi faire simple quand on peut faire compliqué ?!

Les Shadoks, Jacques Rouxel[2]

Abstract

The last part of this work focusses on Multiattribute Value Theory based on the Choquet integral as the aggregation operator of the partial evaluations of the alternatives. The Choquet integral can be seen as an extension of the classical weighted arithmetic mean which allows to take into account interactions between the criteria.

This third part is divided into four chapters, each of them containing some personal contributions. The first one is an introductory chapter on MAVT and the Choquet integral. Furthermore, it contains our results on an extension of the Choquet integral to the case of imprecise partial evaluations, inspired from our article [MR05b].

Chapter 7 is a review of different methods for the elicitation of the preferences of a DM in Choquet integral-based MAVT. It is based on our paper [GKM07] and contains the identification method that we first published in a chapter of [MR05a].

In the third chapter of this part we solve different MCDA problems in an MAVT context and present our contributions to both the classical crisp instance, and the case taking into account impreciseness. It is based on our articles [MR05b, MR05a] and [MMR05].

Finally, in Chapter 9 we present the Kappalab [GKM06] R package, to which we contributed, and an application. To do so, we detail the interactive process for the elicitation of a DM's preferences on a fictitious example.

[2] Why should things be easy when they can be tricky?!

Chapter 6

On Multiattribute Value Theory and the Choquet integral

Contents

In this chapter we introduce Multiattribute Value Theory in the first section and focus on a model which allows to take into account interactions among the criteria. Then, in Section 6.2, we introduce the Choquet integral which will be used as an aggregation function in the selected MAVT model. Finally, in the third section we detail our work on an extension of the Choquet integral to the case where the evaluations of the alternatives on the criteria suffer from impreciseness.

Note that Sections 6.1 and 6.2 are inspired from our articles [GKM07] and [MR05a], whereas Section 6.3 presents results from our work published in [MR05b].

6.1 Multiattribute Value Theory

overall value function

The aim of Multiattribute Value Theory (MAVT) [KR76] is to model the preferences of the DM, represented by a binary relation $\succeq$ on X, by means of an *overall value function* $U : X \to \mathbb{R}$ such that,

$$x \succeq y \iff U(x) \geq U(y), \qquad \forall\, x, y \in X.$$

The assessment of the function U can be done via different methods, described for example in [vE86, Chapter 8] in an additive value function model. Roughly speaking, they can be assigned to two classes: on the one hand, methods based on direct numerical estimations (as, e.g., *direct rating* of alternatives on a cardinal scale), and on the other hand, those based on indifference judgements (as, e.g., the *dual standard sequences* which builds a series of equally spaced intervals on the scale of values). To avoid a direct elicitation of the value function, some methods propose to infer a preference model from so-called *holistic information* about the DM's preferences (as, e.g., the UTA method presented in [JS82]).

In this work we will choose this latter option and determine the function U by means of an interactive and incremental process requiring from the DM that he expresses his preferences over a small subset of selected objects. We focus on this *learning* procedure in Chapter 7 of this work. The resulting overall value function can then be seen as a *numerical representation* of the preference relation $\succeq$ on X, and consequently as a synthetic view of the expertise of the DM.

The preference relation $\succeq$ is assumed to be complete and transitive, hence a weak order. As far as the overall value function is considered, the most frequently encountered model is the *additive value model* (see, e.g., [BP05] or Section 1.3.2). In this work, we consider the more general *transitive decomposable model* of Krantz et al. [KLST71, BP04] in which U is defined by

$$U(x) := F(u_1(x_1), \ldots, u_n(x_n)), \qquad \forall\, x = (x_1, \ldots, x_n) \in X, \tag{6.1}$$

where the functions $u_i : X_i \to \mathbb{R}$ are called the *marginal*[1] *value functions* and $F : \mathbb{R}^n \to \mathbb{R}$, non-decreasing in its arguments, is sometimes called the *aggregation function*. For the previous decomposable model to hold, it is necessary that the preference relation is a *weakly separable* weak order (see, e.g., [BP05, KLST71]).

criterion

Recall that in the context of MAVT (see Chapter 1.3), we use the term *criterion* to designate the association of an attribute $i \in N$ with the corre-

[1] In this work we use both adjectives *marginal* and *partial* to describe these value functions.

sponding marginal value function u_i.

The exact form of the overall value function U depends the preferences which are expressed by the DM. When *mutual preferential independence* (see, e.g., [Vin92]) among criteria can be assumed, it is frequent to consider that the function F is additive and takes the form of a weighted sum. The decomposable model given in Equation 6.1 can then be the taken as the classical additive value model, if further conditions are satisfied. In practice however, mutual preferential independence among criteria might sometimes be hardly verified.

In order to be able to take interaction phenomena among criteria into account, it has been proposed to substitute the weight vector involved in the computation of weighted sums by a monotone set function on N, called *capacity* [Cho53] or *fuzzy measure* [Sug74]. Such an approach can be regarded as taking into account not only the importance of each criterion but also the importance of each subset of criteria. A natural extension of the weighted arithmetic mean in such a context is the *Choquet integral* with respect to (w.r.t.) the defined capacity [Gra92, Mar00a, LG03].

capacity

Choquet integral

The use of a Choquet integral as an aggregation function in Equation 6.1 requires the ability to compare the value of an object according to the different criteria. In other words, it is necessary that the marginal value functions are *commensurable*, i.e. $u_i(x_i) = u_j(x_j)$ if and only if, for the DM, the object x is satisfied to the same extent on criteria i and j; see, e.g., [GLV03] for a more complete discussion on commensurability.

commensurable

Consequently, in this Choquet integral framework, as far as the value functions are considered, for any $x \in X$, the quantity $u_i(x_i)$ can then be interpreted as a level of the satisfaction of the value x_i for the DM.

In the considered context, commensurable marginal value functions can be determined by using the extension of the MACBETH methodology [BV99] proposed in [LG03]; see also [GLV03, GL04]. This task is not trivial and can take a large percentage of the time dedicated to an MCDA problem. In this work we do not discuss this problem further. We rather focus on the problem of the elicitation of the capacity (or *capacity identification problem*) and on the exploitation of the overall value function in Chapter 8.

Once the different marginal value functions have been determined, the next step is to determine the parameters of the aggregator. In the case of the weighted sum, this amounts to determining the importance of the n criteria (as trade-offs that a DM would be willing to make). In an additive situation, it might be realistic to ask the DM to provide such a weight vector. Nevertheless, as soon as the Choquet integral is used, the number of parameters of the capacity can become huge (at most $2^n - 2$). It is therefore not reasonable to ask the DM to supply the analyst with such an information.

It is consequently advisable to determine the capacity from some learning data. We call this information the *initial preferences* of the DM and it usually consists of a partial weak order over a (small) subset of the set X of alternatives, a partial weak order over the set of criteria, intuitions about the importance of the criteria, etc. The precise form of these prior preferences are discussed in Chapter 7.

Once the capacity has been determined from the initial preferences of the DM, it is possible to calculate the overall value of each of the alternatives. Through this step, the elements of X become comparable and it is consequently possible to solve the classical MCDA problematiques.

6.2 The Choquet integral as an aggregation operator

In this section we introduce the Choquet integral which will be used as the overall value function in the selected MAVT context.

6.2.1 Capacities and Choquet integral

Note beforehand that in order to avoid a heavy notation, we omit braces for singletons and pairs, e.g., by writing $\mu(i)$, $N \setminus ij$ instead of $\mu(\{i\})$, $N \setminus \{i,j\}$. Furthermore, cardinalities of subsets $S, T, \ldots$, are denoted by the corresponding lower case letters $s, t, \ldots$ Finally, the power set of N will be denoted by $\mathcal{P}(N)$.

As mentioned in the previous section, *capacities* [Cho53], also called *fuzzy measures* [Sug74], can be regarded as generalisations of weighting vectors involved in the calculation of weighted sums.

capacity

Definition 6.2.1. *A capacity on N is a set function $\mu : \mathcal{P}(N) \to [0,1]$ satisfying the following conditions:*

(i) $\mu(\emptyset) = 0$, $\mu(N) = 1$*;*

(ii) for any $S, T \subseteq N$, $S \subseteq T \Rightarrow \mu(S) \leq \mu(T)$.

Furthermore, a capacity μ on N is said to be

- *additive* if $\mu(S \cup T) = \mu(S) + \mu(T)$ for all disjoint subsets $S, T \subseteq N$;

- *cardinality-based* if, for any $T \subseteq N$, $\mu(T)$ depends only on the cardinality of T.
  Formally, there exist $\mu_1, \ldots, \mu_{n-1} \in [0, 1]$ such that $\mu(T) = \mu_t$ for all $T \subsetneq N, T \neq \emptyset$, such that $|T| = t$.

Note that there is only one capacity on N that is both additive and cardinality-based. We call it *the uniform capacity* and denote it by μ^*. It is easy to verify that μ^* is given by

$$\mu^*(T) = t/n, \qquad \forall T \subseteq N.$$

In the framework of aggregation, for each subset of criteria $S \subseteq N$, the number $\mu(S)$ can be interpreted as the *weight* or the *importance* of S. The monotonicity of μ means that the weight of a subset of criteria cannot decrease when new criteria are added to it.

When using a capacity to model the importance of the subsets of criteria, a suitable aggregation operator that generalises the weighted arithmetic mean is the Choquet integral [Gra92, Mar00a, LG03].

Definition 6.2.2. *The Choquet integral of a function $x : N \to \mathbb{R}$ represented by the vector $(x_1, \ldots, x_n)$ w.r.t. a capacity μ on N is defined by* *Choquet integral*

$$C_\mu(x) := \sum_{i=1}^{n} x_{\sigma(i)}[\mu(A_{\sigma(i)}) - \mu(A_{\sigma(i+1)})],$$

where σ is a permutation on N such that $x_{\sigma(1)} \leq \cdots \leq x_{\sigma(n)}$. Also, $A_{\sigma(i)} := \{\sigma(i), \ldots, \sigma(n)\}$, for all $i \in \{1, \ldots, n\}$, and $A_{\sigma(n+1)} := \emptyset$.

Seen as an aggregation operator, the Choquet integral w.r.t. μ can be considered as taking into account interaction phenomena among criteria, that is, *complementarity*[2] or *substitutivity*[3] among elements of N modeled by μ [Mar00a]. *complementarity* *substitutivity*

The Choquet integral generalises the weighted arithmetic mean in the sense that, as soon as the capacity is additive, which intuitively coincides with the independence of the criteria, it collapses into a weighted arithmetic mean.

An intuitive presentation of the Choquet integral is given in [MS00]. An axiomatic characterisation of the Choquet integral as an aggregation operator can be found in [Mar00a]. Note that the first use of the Choquet integral in decision analysis is probably due to Schmeidler in the context of decision under uncertainty; see [Sch89], and see also [Höh82].

[2] The satisfaction of one attribute in a pair is weak compared to the satisfaction of both.
[3] The satisfaction of one attribute in a pair has almost the same effect as the satisfaction of both.

6.2.2 The Möbius transform of a capacity

Any set function $\nu : \mathcal{P}(N) \to \mathbb{R}$ can be uniquely expressed in terms of its *Möbius representation* [Rot64] by

$$\nu(T) = \sum_{S \subseteq T} m_\nu(S), \qquad \forall\, T \subseteq N, \tag{6.2}$$

where the set function $m_\nu : \mathcal{P}(N) \to \mathbb{R}$ is called the *Möbius transform* or *Möbius representation* of ν and is given by

$$m_\nu(S) = \sum_{T \subseteq S} (-1)^{s-t} \nu(T), \qquad \forall\, S \subseteq N. \tag{6.3}$$

Of course, any set of 2^n coefficients $\{m(S)\}_{S \subseteq N}$ does not necessarily correspond to the Möbius transform of a capacity on N. The boundary and monotonicity conditions must be ensured [CJ89], i.e., we must have

$$\begin{cases} m(\emptyset) = \emptyset, \qquad \sum\limits_{T \subseteq N} m(T) = 1, \\ \\ \sum\limits_{\substack{T \subseteq S \\ T \ni i}} m(T) \geq 0, \qquad \forall\, S \subseteq N,\ \forall\, i \in S. \end{cases} \tag{6.4}$$

As shown in [CJ89], in terms of the Möbius representation of a capacity μ on N, for any $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, the Choquet integral of x w.r.t. μ is given by

$$C_{m_\mu}(x) = \sum_{T \subseteq N} m_\mu(T) \bigwedge_{i \in T} x_i, \tag{6.5}$$

where the symbol $\wedge$ denotes the minimum operator. The notation C_{m_μ}, which is equivalent to the notation C_μ, is used to emphasise the fact that the Choquet integral is here computed w.r.t. the Möbius transform of μ.

6.2.3 Analysis of the aggregation

The behaviour of the Choquet integral as an aggregation operator is generally difficult to understand. For a better understanding of the interaction phenomena modeled by the underlying capacity, several numerical indices can be computed [Mar00b, Mar04]. In the sequel, we present two of them in detail.

6.2.3.1 Importance index

Shapley value

The overall importance of a criterion $i \in N$ can be measured by means of its Shapley value [Sha53], which is defined by

$$\phi_\mu(i) := \sum_{T \subseteq N \setminus i} \frac{(n-t-1)!\, t!}{n!} [\mu(T \cup i) - \mu(T)].$$

Having in mind that, for each subset of criteria $S \subseteq N$, $\mu(S)$ can be interpreted as the *importance* of S in the decision problem, the Shapley value of i can be thought of as an average value of the *marginal contribution* $\mu(T \cup i) - \mu(T)$ of criterion i when added to a subset T not containing it. A fundamental property is that the numbers $\phi_\mu(1), \ldots, \phi_\mu(n)$ form a probability distribution over N. In terms of the Möbius representation of μ, the Shapley value of i can be rewritten as

$$\phi_{m_\mu}(i) = \sum_{T \subseteq N \setminus i} \frac{1}{t+1} m_\mu(T \cup i). \tag{6.6}$$

Finally note that $\sum_{i=1}^{n} \phi_\mu(i) = 1$ and $\phi_\mu(i) \in [0,1]$, $\forall i \in N$.

6.2.3.2 Interaction index

In order to intuitively approach the concept of interaction, consider two criteria i and j such that $\mu(ij) > \mu(i) + \mu(j)$. Clearly, the previous inequality seems to indicate a *positive* interaction or *complementary* effect between i and j. Similarly, the inequality $\mu(ij) < \mu(i) + \mu(j)$ suggests that i and j interact in a *negative*, *substitutive* or *redundant* way. Finally, if $\mu(ij) = \mu(i) + \mu(j)$, it seems natural to consider that criteria i and j do not interact, i.e., that they have *independent* roles in the decision problem.

A coefficient measuring the interaction between i and j should therefore depend on the difference $\mu(ij) - [\mu(i) + \mu(j)]$. However, as discussed by Grabisch and Roubens [GR99], the intuitive concept of interaction requires a more elaborate definition. Clearly, one should not only compare $\mu(ij)$ and $\mu(i) + \mu(j)$ but also see what happens when i, j, and ij join other subsets. In other words, an index of interaction between i and j should take into account all the coefficients of the form $\mu(T \cup i)$, $\mu(T \cup j)$, and $\mu(T \cup ij)$, with $T \subseteq N \setminus ij$.

Murofushi and Soneda [MS93] suggested to measure the average interaction between two criteria i and j by means of the following *interaction index*: *interaction index*

$$I_\mu(ij) := \sum_{T \subseteq N \setminus ij} \frac{(n-t-2)!\,t!}{(n-1)!} [\mu(T \cup ij) - \mu(T \cup i) - \mu(T \cup j) + \mu(T)].$$

Note that, given a subset T not containing i and j, the expression

$$\mu(T \cup ij) - \mu(T \cup i) - \mu(T \cup j) + \mu(T)$$

can be regarded as the difference between the marginal contributions $\mu(T \cup ij) - \mu(T \cup i)$ and $\mu(T \cup j) - \mu(T)$. We call this expression the *marginal interaction between i and j in the presence of T*. Indeed, it seems natural to consider that if

$$\mu(T \cup ij) - \mu(T \cup i) > \mu(T \cup j) - \mu(T) \text{ (resp. } <),$$

i and j interact *positively* (resp. *negatively*) in the presence of T.

The quantity $I_\mu(ij)$ can therefore be interpreted as a measure of the *average* marginal interaction between i and j. An important property is that $I_\mu(ij) \in [-1, 1]$ for all $ij \subseteq N$, the value 1 (resp. -1) corresponding to maximum complementarity (resp. substitutivity) between i and j [Gra97b]. In terms of the Möbius representation of μ, $I_\mu(ij)$ can be rewritten as

$$I_{m_\mu}(ij) = \sum_{T \subseteq N \setminus ij} \frac{1}{t+1} m_\mu(T \cup ij). \tag{6.7}$$

Other indices that can help to understand the behaviour of a Choquet integral are *veto* and *favour* indices, *orness* and *andness* degrees, etc; see e.g [Mar00b, Mar04] for a more complete list of behavioural indices.

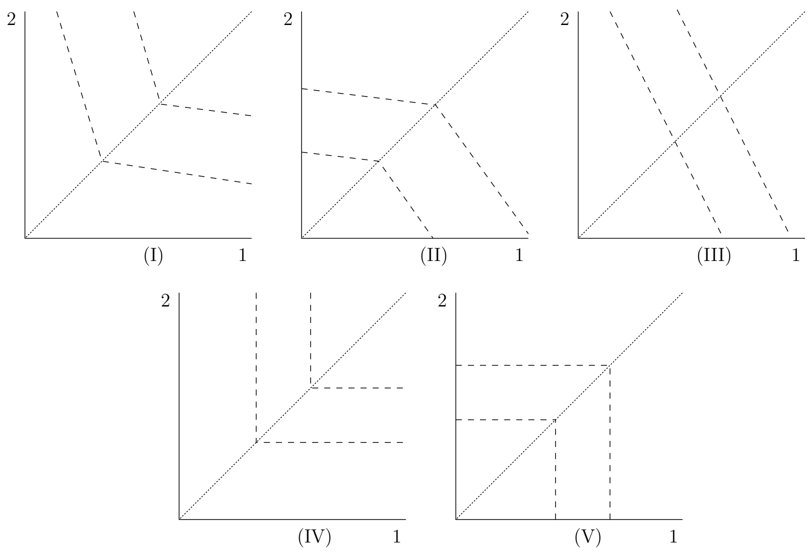

Figure 6.1: Interpretation of the Choquet integral if $|N| = 2$

Before finishing this section, we turn to a representation of the Choquet integral which helps to understand the concepts of positive and negative interactions. Imagine a problem involving 2 criteria and consider the 5 situations of Figure 6.1 (inspired from [GR00] and [MR05a]). The dashed lines represent alternatives having the same overall evaluations through a Choquet integral. The 5 cases can be characterised as follows:

I : $\mu(1) + \mu(2) < \mu(12)$: complementarity (or positive interaction);

II : $\mu(1) + \mu(2) > \mu(12)$: redundancy (or negative interaction);

III : $\mu(1) + \mu(2) = \mu(12)$: additivity;

IV : $\mu(1) = \mu(2) = 0$: maximal complementarity;

V : $\mu(1) = \mu(2) = 1$: maximal redundancy.

Note that for case (IV) (resp. (V)) the Choquet integral corresponds to the min (resp. max) function.

6.2.4 The concept of k-additivity

From the results presented in Sections 6.2.1 and 6.2.2, one can see that a capacity μ on N is completely defined by the knowledge of $2^n - 2$ coefficients, for instance $\{\mu(S)\}_{\emptyset \neq S \subsetneq N}$ or $\{m_\mu(S)\}_{\emptyset \neq S \subsetneq N}$. Such a complexity may be prohibitive in many applications. The fundamental notion of *k-additivity* proposed by Grabisch [Gra97b] enables to find a trade-off between the complexity of the capacity and its expressivity. *k-additivity*

Definition 6.2.3. *Let $k \in \{1, \ldots, n\}$. A capacity μ on N is said to be k-*additive *if its Möbius representation satisfies $m_\mu(T) = 0$ for all $T \subseteq N$ such that $t > k$ and there exists at least one subset T of cardinality k such that $m_\mu(T) \neq 0$.*

As one can easily check, the notion of 1-additivity coincides with that of additivity. Note that, in this case, it follows from Equation (6.7) that the interaction index is zero for any pair of criteria, which is in accordance with the intuition that an additive capacity cannot model interaction. More generally, it can be shown that a k-additive capacity, $k \in \{2, \ldots, n\}$, can model interaction among at most k criteria; see, e.g., [FKM06].

Let $k \in \{1, \ldots, n\}$ and let μ be a k-additive capacity on N. From Equation (6.2), we immediately have that

$$\mu(S) = \sum_{\substack{\emptyset \neq T \subseteq S \\ t \leq k}} m_\mu(T), \qquad \forall\, S \subseteq N,$$

which confirms that a k-additive capacity ($k < n$) is completely defined by the knowledge of $\sum_{l=1}^{k} \binom{n}{l}$ coefficients.

Finally note that for a k-additive capacity m_μ, Equation (6.5) becomes:

$$C_{m_\mu}(x) = \sum_{\substack{T \subseteq N \\ |T| \leq k}} m_\mu(T) \bigwedge_{i \in T} x_i.$$

6.3 Extending the Choquet integral to fuzzy numbers

In this section we present an extension of the classical Choquet integral to the case where the partial evaluations of the alternatives contain impreciseness represented by so-called fuzzy numbers.

6.3.1 Fuzzy numbers and fuzzy sets

fuzzy set

In this section, we first recall general concepts on fuzzy sets, fuzzy numbers and possibility distributions. In 1965, Zadeh [Zad65] introduced the concept of fuzzy set to be able to represent and manipulate data which have nonstatistical uncertainty or in case of impreciseness. Let A be a classical set. A *fuzzy set* $\widetilde{B}$ in A can be defined by its membership function

$$\eta_{\widetilde{B}} : A \to [0,1].$$

For $z \in A$, $\eta_{\widetilde{B}}(z) = 0$ means that z does not belong to $\widetilde{B}$, $\eta_{\widetilde{B}}(z) = 1$ represents the complete membership of z to $\widetilde{B}$, and the values between 0 and 1 stand for intermediate memberships. We write $\widetilde{B}(z) := \eta_{\widetilde{B}}(z)$ for the degree of membership of the element z in the fuzzy set $\widetilde{B}$, for each z in A.

The *support* $\underline{\widetilde{B}}$ of a fuzzy set $\widetilde{B}$ of A is the crisp set of elements of A for which the membership degree to $\widetilde{A}$ is non-zero,

$$\underline{\widetilde{B}} = \{z \in A : \widetilde{B}(z) > 0\}.$$

A fuzzy set $\widetilde{B}$ of A is said to be *normal* if there exists an element z in $\widetilde{B}$ for which $\widetilde{B}(z) = 1$.

Below we will define fuzzy numbers of $\mathbb{R}$. To simplify our discourse, we restrict here to the case where $\widetilde{B}$ is a fuzzy set in $\mathbb{R}$. Let us note $\mathrm{cl}(\underline{\widetilde{B}})$ the closure of the support of $\widetilde{B}$. A $\lambda-$*level set* of $\widetilde{B}$ of A is given by

$$[\widetilde{B}]^\lambda = \begin{cases} \{z \in \mathbb{R} : \widetilde{B}(z) \geq \lambda\} & \text{if } \lambda > 0, \\ \mathrm{cl}(\underline{\widetilde{B}}) & \text{if } \lambda = 0. \end{cases}$$

A fuzzy set $\widetilde{B}$ of $\mathbb{R}$ is said to be *convex* if $[\widetilde{B}]^\lambda$ is a convex subset of $\mathbb{R}$ for all $\lambda \in [0,1]$.

fuzzy number

A *fuzzy number* $\widetilde{z}$ of $\mathbb{R}$ is a fuzzy set in $\mathbb{R}$ that is normal, convex and has a continuous membership function of bounded support. Let $\mathcal{F}$ be the family of all fuzzy numbers. For a fuzzy number $\widetilde{z} \in \mathcal{F}$ we define

$$z_m(\lambda) = \min[\widetilde{z}]^\lambda, \quad z_M(\lambda) = \max[\widetilde{z}]^\lambda.$$

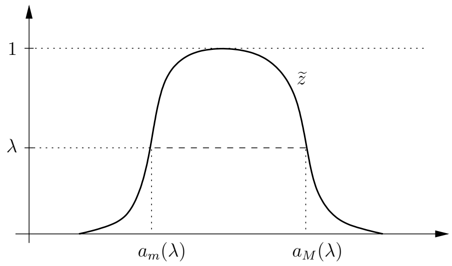

Figure 6.2: A fuzzy number $\widetilde{z}$

We now can introduce the concept of possibility distributions by means of the fuzzy numbers as defined in [Zad65] and [Zad78]. This allows to see fuzzy numbers as *possibility distributions*. Let $a, b \in \mathbb{R} \cup \{-\infty, +\infty\}$ with $a \leq b$. The possibility that $\widetilde{z} \in \mathcal{F}$ takes its value from the interval $[a, b]$ is defined by

possibility distribution

$$\mathrm{Pos}(\widetilde{z} \in [a,b]) = \max_{z \in [a,b]} \widetilde{z}(z).$$

In particular for $\lambda \in [0,1]$,

$$\mathrm{Pos}(\widetilde{z} \leq z_m(\lambda)) = \max_{z \leq z_m(\lambda)} \widetilde{z}(z) = \lambda,$$

$$\mathrm{Pos}(\widetilde{z} \geq z_M(\lambda)) = \max_{z \geq z_M(\lambda)} \widetilde{z}(z) = \lambda.$$

In [DP80] the authors write a fuzzy number in a very general way as

$$\widetilde{z}(z) = \begin{cases} L(\frac{a-z}{\alpha}) & \text{if } a-\alpha \leq z \leq a, \\ 1 & \text{if } a < z \leq b, \\ R(\frac{z-b}{\beta}) & \text{if } b < z \leq b+\beta, \\ 0 & \text{otherwise,} \end{cases}$$

where $[a, b]$ is the *peak* of $\widetilde{z}$, $L, R : [0,1] \to [0,1]$ are upper semi-continuous and non-increasing shape functions with $L(0) = R(0) = 1$ and $L(1) = R(1) = 0$ which are called *side functions*. The support $\underset{\sim}{z}$ is equal to $]a-\alpha, b+\beta[$.

side function

Note that *crisp numbers* are particular cases of fuzzy numbers. Indeed, if $\widetilde{z} \in \mathcal{F}$ with $\underset{\sim}{z} = \{z\}$, then $\widetilde{z}$ is a crisp number.

A particular type of fuzzy numbers are the *trapezoidal* fuzzy numbers. A fuzzy number $\widetilde{z}$ is called a trapezoidal fuzzy number if its membership function can be written as

trapezoidal fuzzy number

$$\widetilde{z}(z) = \begin{cases} 1 - \frac{a-z}{\alpha} & \text{if } a-\alpha \leq z \leq a, \\ 1 & \text{if } a < z \leq b, \\ 1 - \frac{z-b}{\beta} & \text{if } b < z \leq b+\beta, \\ 0 & \text{otherwise.} \end{cases}$$

α and β are called the left and right width. We use the notation $\widetilde{z} = (a, b, \alpha, \beta)$. One can show that $[\widetilde{z}]^\lambda = [a - (1-\lambda)\alpha, b + (1-\lambda)\beta]$, for all $\lambda \in [0,1]$. In the particular case where $a = b$ the fuzzy number is called *triangular* and a is said to be the centre of $\widetilde{z}$.

triangular fuzzy number

In order to use mathematical operations on fuzzy numbers, Zadeh [Zad65] introduced the sup-min extension principle which allows to work consistently with the crisp case. The idea is that each function f on crisp sets induces a corresponding *fuzzy* function on fuzzy sets. If $\widetilde{z_1}, \dots, \widetilde{z_n} \in \mathcal{F}$ and $f : \mathbb{R}^n \to \mathbb{R}$ is a continuous function, then the via the sup-min extension principle we extend f to fuzzy numbers as follows:

extension principle

$$f(\widetilde{z_1}, \dots, \widetilde{z_n})(z) = \sup_{f(z_1,\dots,z_n)=z} \min\{\widetilde{z_1}(z_1), \dots, \widetilde{z_n}(z_n)\}, \forall z \in \mathbb{R}.$$

It is important to note that, as shown in [Ngu78], $f(\widetilde{z_1}, \dots, \widetilde{z_n})$ is a fuzzy number.

An important hypothesis underlies Zadeh's extension principle. It supposes that the fuzzy numbers are *non-interactive*. Interactivity among fuzzy numbers is defined via their joint possibility distribution, defined hereafter.

Let $\widetilde{z_1}, \dots, \widetilde{z_n}$ be fuzzy numbers. An n-dimensional possibility distribution $\widetilde{c}$ is a fuzzy set in $\mathbb{R}^n$ with a normal, continuous membership function of bounded support. $\widetilde{c}$ is called the *joint possibility distribution* of $\widetilde{z_1}, \dots, \widetilde{z_n}$ if

$$\widetilde{z_i}(z_i) = \max_{\substack{z_j \in \mathbb{R} \\ j \neq i}} \widetilde{c}(z_1, \dots, z_n)$$

holds for all $z_i \in \mathbb{R}, i = 1, \dots, n$. $\widetilde{z_i}$ is called the i-th marginal possibility distribution of $\widetilde{c}$. Furthermore, if $\widetilde{z_1}, \dots, \widetilde{z_n} \in \mathcal{F}$ are fuzzy numbers, and $\widetilde{c}$ is their joint possibility distribution then

$$\widetilde{c}(z_1, \dots, z_n) \leq \min\{\widetilde{z_1}(z_1), \dots, \widetilde{z_n}(z_n)\} \quad \forall z_i \in \mathbb{R}, i = 1, \dots, n.$$

Equivalently $[\widetilde{c}]^\lambda \subseteq [\widetilde{z_1}]^\lambda \times \dots \times [\widetilde{z_n}]^\lambda$ for all $\lambda \in [0,1]$. Fuzzy numbers $\widetilde{z_i} \in \mathcal{F}$, $i = 1, \dots, n$ are said to be *non-interactive* if $[\widetilde{c}]^\lambda = [\widetilde{z_1}]^\lambda \times \dots \times [\widetilde{z_n}]^\lambda$ for all $\lambda \in [0,1]$.

In the following section, the task will be to determine the Choquet integral of a function $\widetilde{x} : N \to \mathbb{R}$ represented by a vector of criterial evaluations $(\widetilde{x_1}, \dots, \widetilde{x_n})$ w.r.t. a capacity μ on N. We will suppose that the fuzzy numbers $\widetilde{x_i}$, $i \in N$ are non-interactive in the sense discussed above.

Nevertheless in the context of MCDA, this does not mean that the criteria cannot interact. As presented in Section 6.2.3.2, at the level of the criteria,

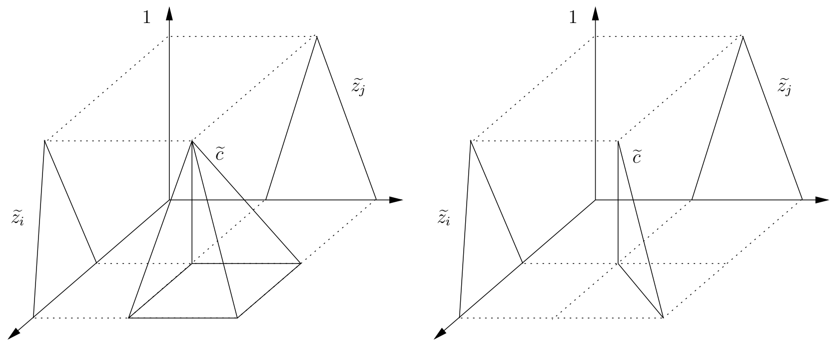

Figure 6.3: Non-interactive (left) and interactive (right) fuzzy numbers $\widetilde{z}_i$ and $\widetilde{z}_j$

different types of interactions can occur, as for example redundancy, complementarity or preferential dependence[4].

6.3.2 A fuzzy extension of the Choquet integral

We can see that Definition (6.5) of the Choquet integral in terms of the Möbius representation of a capacity μ on N is a combination of functions which are continuous on $\mathbb{R} \times \mathbb{R}$, namely the addition $(+)$, the multiplication $(\cdot)$ and the minimum $(\wedge)$ functions. By using the extension principle of Zadeh described in Section 6.3.1 one can extend these three functions to their fuzzy versions as follows.

First of all, the extension $\widetilde{+}$ of the addition of two real numbers to two fuzzy numbers can be defined as

$$\widetilde{z}_1 \widetilde{+} \widetilde{z}_2(z) = \sup_{a+b=z} \min[\widetilde{z}_1(a), \widetilde{z}_2(b)].$$

The result is a fuzzy number according to [Ngu78]. In particular, the sum of two trapezoidal (resp. triangular) fuzzy numbers remains a trapezoidal (resp. triangular) fuzzy number.

Indeed, let us consider two trapezoidal fuzzy numbers $\widetilde{z} = (a_z, b_z, \alpha_z, \beta_z)$ and $\widetilde{y} = (a_y, b_y, \alpha_y, \beta_y)$. According to the definition of the sum of two fuzzy numbers we have $\widetilde{z} \widetilde{+} \widetilde{y} = (a_z + a_y, b_z + b_y, \alpha_z + \alpha_y, \beta_z + \beta_y)$.

[4]Consider $x, y \in A$ for which the evaluations on $S \subseteq N$ are equal. The subset $N \setminus S$ is preferentially independent of S if the preference of x over y is not influenced by their common part on S.

The multiplication, in our case, is a scalar multiplication of a fuzzy number by a crisp number $p \in \mathbb{R}$ (a coefficient of m_μ). $p \,\tilde{\cdot}\, \widetilde{z}$ is such that

$$[p \,\tilde{\cdot}\, \widetilde{z}]^\lambda = p[\widetilde{z}]^\lambda$$

for any $\lambda \in [0,1]$. Again, the result of this scalar multiplication is a fuzzy number. In the particular case of a trapezoidal (resp. triangular) fuzzy number, the output remains a trapezoidal (resp. triangular) fuzzy number. For a trapezoidal number $\widetilde{z} = (a_z, b_z, \alpha_z, \beta_z)$ we have $p\widetilde{z} = (pa_z, pb_z, p\alpha_z, p\beta_z)$.

Finally, the extension $\widetilde{\wedge}$ of the minimum of two crisp numbers becomes

$$\widetilde{\bigwedge}(\widetilde{z}_1, \widetilde{z}_2)(z) = \sup_{\min(a,b)=z} \min[\widetilde{z}_1(a), \widetilde{z}_2(b)].$$

Again the result is a fuzzy number. But in the special case of trapezoidal or triangular fuzzy numbers, the result no longer remains a trapezoidal or triangular fuzzy number. The side functions become piecewise linear functions. Figure 6.4 represents the minimum (bold dashes) of two fuzzy numbers.

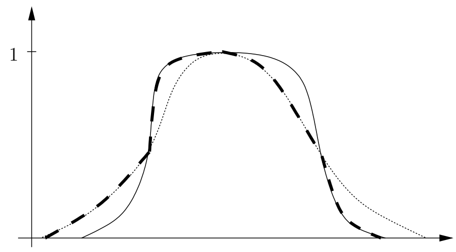

Figure 6.4: Minimum of two fuzzy numbers (bold dashes)

fuzzy Choquet integral

After having presented the extensions of the operators needed to use the Choquet integral as an aggregation operator for fuzzy numbers, we define the *fuzzy extension of the Choquet integral* as

$$\widetilde{C_{m_\mu}}(x) = \widetilde{\sum_{T \subseteq N}} m_\mu(T) \,\tilde{\cdot}\, \widetilde{\bigwedge_{i \in T}} \widetilde{x}_i, \tag{6.8}$$

where m_μ is the Möbius transform of the capacity μ.

The Choquet integral of a vector of fuzzy partial evaluations is a fuzzy number. In case of partial evaluations which are trapezoidal or triangular fuzzy numbers, the resulting fuzzy number has piecewise linear side functions.

It is obvious that this extension remains valid in the case of a k-additive fuzzy measure. The definition can then be written

$$\widetilde{C_{m_\mu}}(x) = \widetilde{\sum_{\substack{T\subseteq N\\|T|\leqslant k}}} m_\mu(T) \,\tilde{\cdot}\, \widetilde{\bigwedge_{i\in T}} \widetilde{x_i}.$$

Formula 6.8 is based on Definition 6.5 of the Choquet integral, obtained via the Möbius transform of Definition 6.2.2. Note that it would be interesting to determine to which extent this fuzzy extension is dependent on the underlying representation and which algebraic transformations lead to equivalent fuzzy definitions of the Choquet integral.

Let us consider a short example to show how the Choquet integral is calculated in case of triangular partial evaluations. This example is inspired from Example 5.1 of [YWHL05].

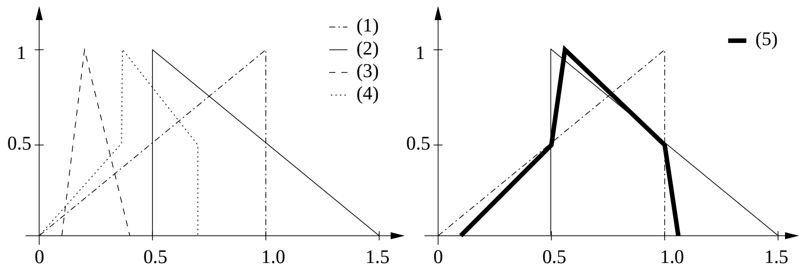

Figure 6.5: Towards the Choquet integral of fuzzy numbers

Example *Let* $N = \{1,2\}$, $\mu(1) = 0.1$, $\mu(2) = 0.2$, $\mu(1,2) = 1$. *The Möbius transform* m_μ *can easily be obtained as* $m_\mu(1) = 0.1$, $m_\mu(2) = 0.2$, $m_\mu(1,2) = 0.7$. *The partial evaluations are represented by the triangular fuzzy numbers* $\widetilde{x_1} = (1,1,0)$ *and* $\widetilde{x_2} = (0.5,0,1)$. *The Choquet integral then becomes:*

$$\begin{aligned}\widetilde{C_{m_\mu}}((\widetilde{x_1},\widetilde{x_2})) &= m_\mu(1) \,\tilde{\cdot}\, \widetilde{x_1} \,\tilde{+}\, m_\mu(2) \,\tilde{\cdot}\, \widetilde{x_2} \\ &\quad \tilde{+}\, m_\mu(1,2) \,\tilde{\cdot}\, \widetilde{\bigwedge}(\widetilde{x_1},\widetilde{x_2}) \\ &= (0.2, 0.1, 0.2) \,\tilde{+}\, 0.7 \,\tilde{\cdot}\, \widetilde{\bigwedge}(\widetilde{x_1},\widetilde{x_2}) \\ &:= \widetilde{s} \,\tilde{+}\, \widetilde{s'}\end{aligned}$$

where $\widetilde{s} = (0.2, 0.1, 0.2)$ *and* $\widetilde{s'} = 0.7 \,\tilde{\cdot}\, \widetilde{\bigwedge}(\widetilde{x_1},\widetilde{x_2})$. *Figure* 6.5 *shows a few steps of these calculations.* (1) *and* (2) *stand respectively for* $\widetilde{x_1}$ *and* $\widetilde{x_2}$. (3) *and* (4) *represent* $\widetilde{s}$ *and* $\widetilde{s'}$. *Finally,* (5) *shows the aggregated value* $\widetilde{C_m}((\widetilde{x_1},\widetilde{x_2}))$.

Note that in [YWHL05] a fuzzy extension of the Choquet integral is presented, based on the classical representation of the Choquet integral by means of a capacity μ. A difficulty in that case is the ordering of the fuzzy numbers which is not needed in the present approach via the Möbius transform m_μ and may be problematic in the case of non-linear fuzzy numbers. A numerical method is also developed to estimate the value of the Choquet integral using trapezoidal fuzzy numbers. The authors of [CM04] present an interval-based Choquet integral to derive preferences on multicriteria alternatives.

Fuzzy set theory was introduced in MCDA in both methodological schools. The authors of [CF96] review a certain number of applications of fuzzy numbers in MCDA, and refer to other surveys.

Chapter 7

The capacity identification problem in Choquet integral-based MAVT

Contents

In this chapter we first introduce the capacity identification problem and show that it can be described as an optimisation problem. In the first section we furthermore detail and formalise the type of information which can be used to represent the DM's preferences. In Section 7.2 we survey the main identification methods presented in the literature and detail our own proposal.

Note that this chapter is based on our work published in [GKM07]. In particular, the method presented in Section 7.2.4 is based on our article [MR05a].

As classically done, and as already mentioned, the asymmetric part of the binary relation $\succeq$, representing the DM's preferences, will be denoted by $\succ$ and its symmetric part by $\sim$.

Finally note that this chapter is situated in a context, where the aggregation performed via the Choquet integral exclusively concerns crisp evaluation.

7.1 The identification problem

As already mentioned in Section 6.1, we assume that the marginal value functions have been determined beforehand. The objective of this section is to present how to identify a capacity, if it exists, such that the Choquet integral w.r.t. this capacity numerically represents the preferences of the DM (see Equation (6.1)).

learning set

Let $O \subset X$ be a reference set of alternatives on which the DM is able to express some preferences. The set O is usually composed either of real options from the DM's expertise domain or of selected, potentially fictitious objects, on which the DM's reasoning may be useful to model his preferences. Using the terminology of artificial intelligence, the set O could be seen as a *learning set*. Its usually small cardinality (rarely more than 20 alternatives) is due to the fact that the expression of the preferences of the DM is generally a very time-consuming and wearying process. See for instance [Vin92, Chapter 3] for a more complete discussion about the subset O.

initial preferences

Once an appropriate subset O has been determined, the DM is asked to express his *initial preferences*. These preferences, from which the capacity is to be determined, can take the form of:

- a partial weak order $\succeq_O$ over O (ranking of the available objects);
- overall evaluations of the alternatives of O;
- a partial weak order $\succeq_N$ over N (ranking of the importance of the criteria);
- quantitative intuitions about the importance of some criteria;
- a partial weak order $\succeq_P$ on the set of pairs of criteria (ranking of interactions);
- intuitions about the type and the magnitude of the interaction between some criteria;
- the behaviour of some criteria as *veto* or *favour* [Mar00b, Mar04, Gra97a];
- the knowledge of an *inter-additive* partition of N, i.e. roughly, a partition composed of pairwise independent subsets of criteria [FM00];
- etc.

In the context of MAVT based on the Choquet integral, it seems natural to translate some of the above prior information as follows:

- $x \succ_O x'$ can be translated as $C_\mu(u(x)) - C_\mu(u(x')) \geq \delta_C$;
- $x \sim_O x'$ can be translated as $-\delta_C \leq C_\mu(u(x)) - C_\mu(u(x')) \leq \delta_C$;

- $i \succ_N j$ can be translated as $\phi_\mu(i) - \phi_\mu(j) \geq \delta_{Sh}$;

- $i \sim_N j$ can be translated as $-\delta_{Sh} \leq \phi_\mu(i) - \phi_\mu(j) \leq \delta_{Sh}$;

- $ij \succ_P kl$ can be translated as $I_\mu(ij) - I_\mu(kl) \geq \delta_I$;

- $ij \sim_P kl$ can be translated as $-\delta_I \leq I_\mu(ij) - I_\mu(kl) \leq \delta_I$;

where $u(x) := (u_1(x_1), \ldots, u_n(x_n))$ for all $x \in X$, and δ_C, δ_{Sh} and δ_I are nonnegative indifference thresholds to be defined by the DM. In other terms, the partial weak orders $\succeq_O$, $\succeq_N$, $\succeq_P$ previously mentioned are translated into partial semiorders with fixed indifference thresholds. Note that in practice a constraint of the form $I_\mu(ij) - I_\mu(kl) \geq \delta_I$ is generally accompanied either by the constraint $I_\mu(ij) \leq 0$ or by the constraint $0 \leq I_\mu(kl)$.

The remaining more quantitative information could be translated as follows (although this is more questionable):

- intuitions about the importance of a criterion i could be translated as $a \leq \phi_\mu(i) \leq b$, where the reals $a, b \in [0,1]$, $a \leq b$, are to be fixed by the DM;

- intuitions about the type and the magnitude of the interaction between two criteria i and j could be translated as $a \leq I_\mu(ij) \leq b$, where $0 \leq a \leq b \leq 1$, in case of complementarity and, where $-1 \leq a \leq b \leq 0$, in case of substitutivity.

Finally, the veto (resp. favour) effect of a criterion i can be directly translated as $\mu(T) = 0$ for all $T \subseteq N$ such that $T \not\ni i$ (resp. $\mu(T) = 1$ for all $T \subseteq N$ such that $T \ni i$) [Mar00b, Prop. 3 and 4].

Most of the identification methods proposed in the literature give rise to an optimisation problem:

$$\text{min or max } f$$

$$\text{subject to} \begin{cases} \mu(S \cup i) - \mu(S) \geq 0, \forall\, i \in N, \forall\, S \subseteq N \setminus i, \\ \mu(N) = 1, \\ C_\mu(u(x)) - C_\mu(u(x')) \geq \delta_C, \\ \vdots \\ \phi_\mu(i) - \phi_\mu(j) \geq \delta_{Sh}, \\ \vdots \\ I_\mu(ij) - I_\mu(kl) \geq \delta_I, \\ I_\mu(kl) \geq 0 \\ \vdots \end{cases}$$

where μ is a capacity on N, and f is an objective function that distinguishes the various identification methods. Its arguments can be of various types, but include in general the parameters of the capacity which have to be determined. The exact form of f will be discussed hereafter in Section 7.2.

The constraints are a selection of the previously listed *initial preferences*, expressed on certain alternatives $(x, x' \in X)$ and criteria $(i, j, k, l \in N)$. They are linear with respect to the parameters of the capacity. Note that the above presentation of the optimisation problem is solely a generic presentation. The constraints depend on the type of preferential information provided by the DM. We consequently adopted this understandable but unformal notation for the sake of simplicity.

A solution to the above problem is a general capacity defined by $2^n - 1$ coefficients. The number of variables involved increases exponentially with n. Consequently the computational time will also increase at least exponentially. For large problems, both for computational and simplicity reasons, it may be preferable to restrict the set of possible solutions to k-additive capacities, $k \in \{1, \ldots, n\}$, typically $k = 2$ or 3. Furthermore, for parsimony reasons, it is possible to start with a small k, and to increase its value if necessary, until a feasible solution is found.

The idea is here simply to rewrite the above optimisation problem in terms of the Möbius transform of a k-additive capacity using Equations (6.3), (6.5), (6.6) and (6.7), which will decrease the number of variables from $2^n - 1$ to $\sum_{l=1}^{k} \binom{n}{l}$ as one can see from Table 7.1. We obtain

$$\text{min or max } f$$

$$\text{subject to} \begin{cases} \sum\limits_{\substack{T \subseteq S \\ t \leq k-1}} m_\mu(T \cup i) \geq 0, \, \forall i \in N, \, \forall S \subseteq N \setminus i, \\ \sum\limits_{\substack{T \subseteq N \\ 0 < t \leq k}} m_\mu(T) = 1, \\ C_{m_\mu}(u(x)) - C_{m_\mu}(u(x')) \geq \delta_C, \\ \vdots \\ \phi_{m_\mu}(i) - \phi_{m_\mu}(j) \geq \delta_{Sh}, \\ \vdots \\ I_{m_\mu}(ij) - I_{m_\mu}(kl) \geq \delta_I, \\ I_{m_\mu}(kl) \geq 0 \\ \vdots \end{cases}$$

where m_μ is the Möbius representation of a k-additive capacity μ on N.

k \ n	2	3	4	5	6	7	8	9	10
1	2	3	4	5	6	7	8	9	10
2	3	6	10	15	21	28	36	45	55
3	-	7	14	25	41	63	92	129	175
n	3	7	15	31	63	127	255	511	1023

Table 7.1: Influence of the order of k-additivity and the number n of criteria on the number of variables.

Of course, the above optimisation problem may be infeasible if the constraints are inconsistent. Such a situation can arise for three main reasons :

- The preferential information provided by the DM is contradictory or violates natural axioms underlying most decision making procedures such as compatibility with dominance, transitivity of strict preferences, etc.
- The number of parameters of the model, i.e. the number of coefficients of the Möbius transform, is too small to have all the constraints satisfied. In this case, in order to increase the number of free parameters, and therefore to improve the expressivity of the model, the approach usually consists in incrementing the order of k-additivity.
- It may happen however that even with a general (n-additive) capacity, the constraints imposed by the DM, still being in accordance with the previously mentioned natural axioms, cannot be satisfied. In such a case, some more specific axioms underlying the Choquet integral model are violated (see e.g. [Wak89], Theorem VI.5.1 concerning comonotonic contradictory tradeoffs, or [MDGP97]) and the Choquet integral cannot be considered as sufficiently flexible for modelling the initial preferences of the DM.

Note that in the latter case, where the Choquet integral-based model is not rich enough to model the DM's preferences, we propose a less constrained approach in Section 7.2.4, where a capacity which violates certain of these preferences is determined.

It is important to note that, in the *interactive* framework as defined in Part I, finding a solution to the above optimisation problem does not necessarily end the identification process. Indeed, the obtained Choquet integral-based representation is then usually analysed by means of the indices presented in Section 6.2.3. If the results are not completely in accordance with the DM's reasoning, his *initial* preferences are enriched by additional constraints and a new identification is performed. This interactive loop continues then until a satisfactory representation of the DM's preference and an adequate recommendation are found. Note that in Chapter 9, we present an application of such an interactive elicitation on a fictitious example by means of the Kappalab [GKM06] package.

7.2 Main methods for capacity identification

As discussed in Section 7.1, most methods for capacity identification proposed in the literature result in optimisation problems. They differ according to their objective function and the preferential information they require as input.

After presenting two methods that could be seen as generalisations of multiple linear regression, we review methods based on *maximum split*, *minimum variance* and *minimum distance* identification principles. We end this section by describing a hybrid method that we developed, providing an approximate solution if there are no capacities compatible with the DM's preferences.

7.2.1 Least-squares based methods

Historically, the first approach that has been proposed can be regarded as a generalisation of classical multiple linear regression [MM89]. It requires the additional knowledge of the desired overall evaluations $y(x)$ of the available objects $x \in O$. The objective function is defined as

$$f_{LS}(m_\mu) := \sum_{x \in O} \left[C_{m_\mu}(u(x)) - y(x)\right]^2,$$

where $u(x) := (u_1(x_1), \ldots, u_n(x_n))$ for all $x \in X$. The aim is to minimise the average quadratic distance between the overall values $\{C_{m_\mu}(u(x))\}_{x \in O}$ computed by means of the Choquet integral and the desired overall scores $\{y(x)\}_{x \in O}$ provided by the DM.

The optimisation problem takes therefore the form of a quadratic program, not necessarily strictly convex [GNW95], which implies that the solution, if it exists, is not necessarily unique (this aspect is investigated in detail in [MG99]). In order to avoid the use of quadratic solvers, heuristic suboptimal versions of this approach have been proposed by Ishii and Sugeno [IS96], Mori and Murofushi [MM89] and Grabisch [Gra95]. Let us detail this latter approach.

Called *Heuristic Least Mean Squares* (HLMS), it is based on a gradient approach starting from a capacity μ defined by the DM that we call the *initial capacity*. This capacity is typically an additive capacity representing the DM's prior idea of what the aggregation function F in Equation (6.1) should be. In the absence of clear requirements on the aggregation function, a very natural choice for μ is the uniform capacity μ^* since the Choquet integral w.r.t. that capacity is nothing else than the arithmetic mean. Once the initial capacity μ has been chosen, for each object $x \in O$, the gradient modifies only the coefficients of μ involved in the computation of $C_\mu(u(x))$ (without violating monotonicity constraints). When all the available objects have been used, unmodified coefficients of μ are modified towards the average value of neighboring coefficients. This forms one iteration, and the process is restarted until a stopping criterion is satisfied. The advantage over the optimal quadratic approach is that only the vector of the coefficients of μ has to be stored, while in the latter, a squared matrix of same dimensions has to be stored. Also, as we will see in Section 9.3.1, this heuristic approach tends to provide less extreme solutions than the optimal approach. However, unlike for the optimal quadratic method, it is not possible to require that the solution be k-additive, $k < n$.

In the context of MAVT based on the Choquet integral, the main inconvenience of these methods is that they require the knowledge of the desired overall values $\{y(x)\}_{x\in O}$, which often can only hardly be obtained from the DM.

7.2.2 A maximum split method

An approach based on linear programming was proposed by Marichal and Roubens [MR00]. The proposed identification method can be stated as follows:

$$\max f_{LP}(\varepsilon) := \varepsilon$$

$$\text{subject to} \begin{cases} \sum_{\substack{T\subseteq S\\ t\leq k-1}} m_\mu(T\cup i) \geq 0,\ \forall i\in N,\ \forall S\subseteq N\setminus i, \\ \sum_{\substack{T\subseteq N\\ 0<t\leq k}} m_\mu(T) = 1, \\ C_{m_\mu}(u(x)) - C_{m_\mu}(u(x')) \geq \delta_C + \varepsilon, \\ \vdots \\ \phi_{m_\mu}(i) - \phi_{m_\mu}(j) \geq \delta_{Sh}, \\ \vdots \\ I_{m_\mu}(ij) - I_{m_\mu}(kl) \geq \delta_I, \\ I_{m_\mu}(kl) \geq 0 \\ \vdots \end{cases}$$

Roughly speaking, the idea of the proposed approach is to maximise the minimal difference between the overall values of objects that have been ranked by the DM through the partial weak order $\succeq_O$ (hence the name *maximum split*). Indeed, if the DM states that $x \succ_O x'$, he may want the overall values to reflect this difference in the most significant way.

The main advantage of this approach is its simplicity. However, as the least squares based approach presented in the previous subsection, this identification method does not necessarily lead to a unique solution, if any. Furthermore, as it will be illustrated in Chapter 9, the provided solution can sometimes be considered as too extreme, since it corresponds to a capacity that maximises the difference between overall values.

Note that we have extended this identification method to handle MCDA ordered sorting problems [MMR05]. Details concerning this topic can be found in Chapter 8.

7.2.3 Minimum variance and minimum distance methods

The idea of the minimum variance method [Koj07] is to favour the *least specific* capacity, if any, compatible with the initial preferences of the DM. The objective function is defined as the *variance* of the capacity, i.e.

$$f_{MV}(m_\mu) := \frac{1}{n}\sum_{i\in N}\sum_{S\subseteq N\setminus i}\gamma_s(n)\left(\sum_{T\subseteq S} m_\mu(T\cup i) - \frac{1}{n}\right)^2,$$

where $\gamma_s(n) = \frac{(n-s-1)!s!}{n!}$.

As shown in [Koj07], minimising this variance is equivalent to maximising the extended Havrda and Charvat entropy of order 2. This method can therefore be equivalently regarded as a maximum entropy approach. The optimisation problem takes the form of the following strictly convex quadratic program:

$$\min f_{MV}(m_\mu)$$

$$\text{subject to}\ \begin{cases} \sum\limits_{\substack{T\subseteq S\\ t\leq k-1}} m_\mu(T\cup i)\geq 0,\ \forall i\in N,\ \forall S\subseteq N\setminus i,\\ \sum\limits_{\substack{T\subseteq N\\ 0<t\leq k}} m_\mu(T)=1,\\ C_{m_\mu}(u(x))-C_{m_\mu}(u(x'))\geq \delta_C,\\ \vdots\\ \phi_{m_\mu}(i)-\phi_{m_\mu}(j)\geq \delta_{Sh},\\ \vdots\\ I_{m_\mu}(ij)-I_{m_\mu}(kl)\geq \delta_I,\\ I_{m_\mu}(kl)\geq 0\\ \vdots \end{cases}$$

As discussed in [Koj07, KMR05], the Choquet integral w.r.t. the minimum variance capacity compatible with the initial preferences of the DM, if it exists, is the one that will exploit the most on average its arguments.

One of the advantages of this approach is that it leads to a unique solution, if any, because of the strict convexity of the objective function. Also, in the case of *poor* initial preferences involving a small number of constraints, this unique solution will not exhibit a too specifical behaviour characterised for instance by very high positive or negative interaction indices or a very uneven Shapley value.

A generalisation of this approach [Koj06] consists in finding, if it exits, the capacity closest to a capacity defined by the DM and compatible with his initial preferences. As already discussed in Section 7.2.1, this *initial capacity* is typically an additive capacity representing the DM's prior idea of what the aggregation function should be. In the absence of clear requirements a very natural choice for μ is the uniform capacity μ^*. In order to practically implement such a minimum distance principle, in [Koj06], three quadratic distances have been studied. In the sequel, we restrict ourselves to the following one defined, for any two capacities μ, μ on N by

$$d^2(m_\mu, m_\mu) := \int_{[0,1]^n} [C_{m_\mu}(x) - C_{m_\mu}(x)]^2 dx. \tag{7.1}$$

This quadratic distance, thoroughly studied in [Mar98, Chap. 7] in the context of the extension of pseudo-Boolean functions, can be interpreted as the expected quadratic difference between overall values computed by C_{m_μ} and C_{m_μ} assuming that the vectors of partial values are uniformly distributed in $[0,1]^n$.

In the absence of clear requirements on the aggregation function, a natural objective function for the above discussed minimum distance principle is thus given by

$$f_{MD}(m_\mu) := \int_{[0,1]^n} \left[C_{m_\mu}(x) - \frac{1}{n} \sum_{i=1}^{n} x_i \right]^2 dx.$$

The resulting optimisation problem is again a strictly convex quadratic program.

7.2.4 A less constrained method

This approach, which we first presented in [MR05a], can be seen as a generalisation of the least squares methods described in Section 7.2.1. The minimal preferential information which has to be provided by the DM is a weak order over the available objects. The objective function, depending on more variables than the least squares methods described earlier, is defined as

$$f_{GLS}(m_\mu, y) := \sum_{x \in O} \left[C_{m_\mu}(u(x)) - y(x) \right]^2,$$

where $y = \{y(x)\}_{x\in O}$ are additional variables of the quadratic program[1] representing overall unknown evaluations of the objects that must verify the weak order imposed by the DM.

The optimisation problem can be written as the following convex quadratic program :

$$\min f_{GLS}(m_\mu, y)$$

$$\text{subject to} \begin{cases} \sum\limits_{\substack{T\subseteq S \\ t\leq k-1}} m_\mu(T\cup i) \geq 0, \ \forall i\in N, \ \forall S\subseteq N\setminus i, \\ \sum\limits_{\substack{T\subseteq N \\ 0<t\leq k}} m_\mu(T) = 1, \\ y(x) - y(x') \geq \delta_y, \\ \vdots \\ \phi_{m_\mu}(i) - \phi_{m_\mu}(j) \geq \delta_{Sh}, \\ \vdots \\ I_{m_\mu}(ij) - I_{m_\mu}(kl) \geq \delta_I, \\ I_{m_\mu}(kl) \geq 0 \\ \vdots \end{cases}$$

where δ_y is an indifference threshold, playing a similar role as δ_C, i.e., it can be interpreted as the desired minimal difference between the overall values of two objects which are considered as significantly different by the DM. A solution of the quadratic program consists of the Möbius representation m_μ of the capacity and the overall evaluations $y = \{y(x)\}_{x\in O}$.

Let us first intuitively explain the main idea of the approach. As discussed at the end of Section 7.1, assuming that the constraints imposed by the DM are not contradictory and do not question natural multiple criteria decision axioms such as compatibility with dominance, it may still happen that the number of parameters (following from the chosen order of k-additivity) is too small so that these constraints can be satisfied. A first possibility consists in increasing k, if possible. A second solution consists in relaxing some of the constraints by translating the desired weak order over the available objects by means of conditions on the unknown overall evaluations $\{y(x)\}_{x\in O}$. In that case, the Möbius transform m_μ is less constrained and there may exist a solution.

The role of the objective function is to minimise the quadratic difference between the numerical representation $\{y(x)\}_{x\in O}$ of the weak order imposed by the DM and the overall values computed by means of the Choquet integral.

[1]The acronym GLS stands for "generalised least squares".

If the objective function is zero, then for each object x, its overall evaluation $y(x)$ equals its aggregated overall value $C_{m_\mu}(u(x))$. In that case, the weak order obtained by ordering the objects according to their aggregated evaluations is consistent with the weak order imposed by the DM and the threshold δ_y is not violated. Two possibilities arise: either there is a unique solution or there exists an infinity of solutions to the problem. In the second case, the solution is chosen by the solver and its characteristics are difficult to predict.

If the objective function is strictly positive, then the aggregated overall evaluations $\{C_{m_\mu}(u(x))\}_{x\in O}$ do not exactly match the overall unknown evaluations $\{y(x)\}_{x\in O}$ numerically representing the weak order imposed by the DM. Two possibilities arise: either the weak order induced by the aggregated overall evaluations $\{C_{m_\mu}(u(x)\}_{x\in O}$ corresponds to $\succeq_O$ but $x \succ_O x'$ does not necessarily imply $C_{m_\mu}(u(x)) \geq C_{m_\mu}(u(x')) + \delta_y$ (see Subsection 9.3.5), or the weak order induced by the $\{C_{m_\mu}(u(x)\}_{x\in O}$ does not correspond to $\succeq_O$. In the former case, the solution does not respect the DM's choice for the minimal threshold δ_y. In the latter case the weak order is violated *on average* which might not be very satisfactory.

The advantage of this approach is that it may provide a solution even if the weak order over the available objects is incompatible with a Choquet integral model because some specific axioms are violated, if the indifference threshold δ_y is too large, or if some of the constraints on the criteria are not compatible with a representation of the weak order by a Choquet integral. It is then up to the DM to decide if the result is satisfactory or not. Nevertheless, as already explained, this approach should be used with care when the objective function is zero, since then, it simply amounts to letting the quadratic solver choose a feasible solution whose characteristics are difficult to predict.

Note that this identification method can also be used to test if a given problem can be represented by a Choquet integral. If the objective function is strictly positive, then there exists no capacity which allows to represent the problem by means of a Choquet integral (for a fixed order of k-additivity).

In this Chapter we have presented several identification methods, among which is situated our proposal, which is the less constrained method. We have also underlined that some of these methods do not necessarily generate unique solutions. Consequently, the representation of a decision problem by means of an overall value function might in such cases not be unique. Consequently, this can produce some uncertainty in the order induced by U. These robustness and stability issues of the identification methods are nevertheless out of the scope of this work.

Chapter 8

Solving common MCDA problems

Contents

In this chapter we show how common MCDA problems can be solved by exploiting the results of Choquet integral-based MAVT.

In the first section, where we focus on the case where the partial evaluations are crisp numbers, we detail how to solve the choice, the ranking and the ordered sorting problems. Then, in Section 8.2 we switch to the case where the partial evaluation of the alternatives are fuzzy numbers, and present how the choice and the ranking problems can be dealt with. Finally, in the third section we return to progressiveness issues in the context of MAVT.

Note that Section 8.1.2 is based on our articles [MMR05] and [MR05a], whereas Section 8.2 presents our research published in [MR05b].

8.1 The crisp case

We suppose here that the partial evaluations of the elements of X are crisp numbers and have been aggregated by means of a Choquet integral into a fuzzy overall evaluation for each alternative. In the following sections we will show how to exploit this information in order to determine a choice among the alternatives, to rank them and to sort them in ordered classes.

8.1.1 The choice and the ranking problem

Recall that the objective of MAVT is to model the preferences of the DM, represented by a binary relation $\succeq$ on X, by means of an overall value function $U : X \to \mathbb{R}$ such that, $x \succeq y \iff U(x) \geq U(y), \ \forall x, y \in X$. Therefore, once the capacity modelling the DM's preferences has been determined, all the alternatives of X become comparable by means of their overall value calculated by a Choquet integral. Note nevertheless that this representation might not be unique (as shown in Chapter 7.

ranking

Consequently the ranking problem, aiming at positioning the alternatives of X from the *best* to the *worst* one, has a natural solution via the order of the overall values of the alternatives.

choice

Similarly, the choice problem, whose objective is to determine a single best alternative, can be solved by considering the alternative with the highest overall evaluation. In case of a tie, either the DM chooses one of the equivalent alternatives, or, in the context of a progressive method, the analysis can be continued further in order refine some evaluations to discriminate between the equivalent alternatives. The k-choice problem is similarly solved by considering the first k positions of the obtained ranking. In case of ties, again, the DM should either select one of the equivalent elements of X, or, continue the evaluation in a progressive framework to increase the discrimination between tied alternatives.

k-choice

8.1.2 The ordered sorting problem

Consider a partition of X into m nonempty classes $\{Cl_t\}_{t=1}^m$, which are ordered in increasing order of preferences; that is, for any $r, s \in \{1, \ldots, m\}$, with $r > s$, the elements of Cl_r are considered as better than the elements of Cl_s.

We also set

$$Cl_r^{\geqslant} := \bigcup_{t \geqslant r} Cl_t \qquad (r = 1, \dots, m).$$

The objective of the ordered sorting problem is to partition the elements of X into the classes $\{Cl_t\}_{t=1}^m$. The following considerations are mainly based on the following result, adapted from [GMS01, Theorem 2.1], which states that, under a simple condition of monotonicity, it is possible to find a discriminant function that strictly separates the classes $Cl_1, \dots, Cl_m$ by ordered numerical thresholds. *ordered sorting*

For any $x_i \in X_i$ and any $y_{-i} \in X_{-i} := \Pi_{j \in N \setminus \{i\}} X_j$, we set

$$x_i y_{-i} := (y_1, \dots, y_{i-1}, x_i, y_{i+1}, \dots, y_n) \in X.$$

Theorem 8.1.1. *The following two assertions are equivalent:*

1. *For all $i \in N$, $t \in \{1, \dots, m\}$, $x_i, x_i' \in X_i$, $y_{-i} \in X_{-i}$, we have*

$$x_i' \succcurlyeq_i x_i \text{ and } x_i y_{-i} \in Cl_t \;\Rightarrow\; x_i' y_{-i} \in Cl_t^{\geqslant}.$$

2. *There exist*

   - *functions $g_i : X_i \to \mathbb{R}$ $(i \in N)$, strictly increasing,*
   - *a function $f : \mathbb{R}^n \to \mathbb{R}$, increasing in each argument, called discriminant function,*
   - *$m-1$ ordered thresholds $\{z_t\}_{t=2}^m$ satisfying*

$$z_2 \leqslant z_3 \leqslant \cdots \leqslant z_m$$

   such that, for any $x \in X$ and any $t \in \{2, \dots, m\}$, we have

$$f[g_1(x_1), g_2(x_2), \dots, g_n(x_n)] \geqslant z_t \;\Leftrightarrow\; x \in Cl_t^{\geqslant}.$$

For a practical use of this result, Roubens [Rou01] restricted the family of possible discriminant functions to the class of n-variable Choquet integrals and the partial value functions to normalised scores.

8.1.2.1 Capacity identification

We again assume that the value functions of the different criteria have been determined beforehand. The goal of this section is to present how to identify a capacity, if it exists, such that the Choquet integral w.r.t. to this capacity allows to represent the preferences of the DM related to an ordered sorting.

$O \subset X$ is once more a reference set of alternatives on which the DM is able to express some preferences (see Section 7.1). For the sorting problem, we will assume that the DM can assign each element of O to one of the classes Cl_t ($t \in \{1,\dots,m\}$). This assignment renders a partition of O into classes $\{O_t\}_{t=1}^m$, where $O_t := O \cap Cl_t$ for all $t \in \{1,\dots,m\}$. Note that we suppose here that the sets O_t are non-empty, for all $t \in \{1,\dots,m\}$.

As the Choquet integral is supposed to strictly separate the classes O_t (and later the classes Cl_t), the following necessary condition is imposed

$$C_\mu(u(x)) - C_\mu(u(y)) \geqslant \varepsilon \,, \tag{8.1}$$

for each ordered pair $(x,y) \in O_t \times O_{t-1}$ and each $t \in \{2,\dots,m\}$, where ε is a given strictly positive threshold and $u(x) := (u_1(x_1),\dots,u_n(x_n))$.

These separation conditions, put together with the boundary and monotonicity constraints on the fuzzy measure, form a linear constraint satisfaction problem whose unknowns are the coefficients of the capacity. Thus at this stage, the sorting problem consists in finding a feasible solution satisfying all these constraints. Note that if ε is chosen too big, the problem might have no solution.

In practice, the identification methods described in Chapter 7 can be applied to identify a capacity which respects the DM's preferences. In [MMR05], we present how the maximum split approach of Section 7.2.2 can be applied to the ordered sorting problem. In such a case, the strictly positive threshold ε, which is meant to strictly separate the classes, is considered as a non-negative variable to be maximised. The minimum variance and minimum distance methods of Section 7.2.3 can also be used here to determine the capacity. In both cases, the separation constraints require that the DM determines a value for the threshold ε (which replaces in that case δ_C of Section 7.2.3).

Let us now define a dominance relation D on X as follows: For each $x, y \in X$,

$$xDy \quad\Leftrightarrow\quad u_i(x_i) \geqslant u_i(y_i) \;\; \forall i \in N.$$

Being an intersection of complete orders, the binary relation D is a partial order, i.e., it is reflexive, antisymmetric, and transitive. Furthermore we clearly have

$$xDy \quad\Rightarrow\quad C_\mu(u(x)) \geqslant C_\mu(u(y)).$$

We can now define, for each $t \in \{1,\dots,m\}$, the set of non-dominating alternatives of O_t,

$$Nd_t := \{x \in O_t \mid \nexists\, x' \in O_t \setminus \{x\} : xDx'\},$$

and the set of non-dominated alternatives of O_t,

$$ND_t := \{x \in O_t \mid \nexists\, x' \in O_t \setminus \{x\} : x'Dx\}.$$

Due to the increasing monotonicity of the Choquet integral, it is sufficient to consider Constraint (8.1) only for each ordered pair $(x, y) \in Nd_t \times ND_{t-1}$ and each $t \in \{2, \ldots, m\}$. Therefore, the total number of separation constraints boils down to

$$\sum_{t=2}^{m} |Nd_t|\,|ND_{t-1}|.$$

8.1.2.2 Assignment

Let μ^* be a solution for the capacity identification problem presented in the previous section (for any of the selected identification methods). Then any alternative $x \in X$ will be assigned to

- the class Cl_t if

$$\min_{y \in Nd_t} C_{\mu^*}(u(y)) \leqslant C_{\mu^*}(u(x)) \leqslant \max_{y \in ND_t} C_{\mu^*}(u(y)),$$

- the union of the classes Cl_t and Cl_{t-1} if

$$\max_{y \in ND_{t-1}} C_{\mu^*}(u(y)) < C_{\mu^*}(u(x)) < \min_{y \in Nd_t} C_{\mu^*}(u(y)).$$

Suppose now that there exists no solution to the capacity identification problem. It is then possible to use the identification method of Section 7.2.4 which allows to solve less constrained problems. If its objective function is strictly positive, this signifies that certain conditions imposed by the DM are violated. We detail hereafter how to determine an assignment for each alternative of O (and of X) in terms of intervals of contiguous classes.

First of all, let us suppose that $u_i : X_i \to [0, 1]$ ($\forall i \in N$) (without being restrictive, this simplifies the notations of the following considerations). Furthermore assume that $u(x^-) := (0, \ldots, 0)$ is assigned to the worst class O_1, and that $u(x^+) := (1, \ldots, 1)$ is assigned to the best class O_m. Let l_x^o denote the index of the class to which the DM has assigned alternative x of O.

In this context we will show that any alternative $x \in X$ can be assigned to an interval of classes of $\{Cl_t\}_{t=1}^m$. As a consequence, to each such assignment corresponds a lower class index l_x^- and an upper class index l_x^+ of $\{1, \ldots, m\}$.

Let us start by defining different types of assignments which we will consider here.

Definition 8.1.1. *An alternative $x \in X$ is said to be* precisely assigned *to Cl_{l_x} if $l_x^- = l_x^+ =: l_x$. Else x is said to be* ambiguously assigned *to the interval of classes $[Cl_{l_x^-}, Cl_{l_x^+}]$.* *precise and ambiguous assignment*

The following definition is useful in the case where the class to which an alternative x of X should belong is known beforehand, and that the objective is to determine the accuracy of an assignement.

correct assignment

Definition 8.1.2. *An alternative $x \in X$ is said to be correctly assigned if $Cl_{l_x^o} \in [Cl_{l_x^-}, Cl_{l_x^+}]$.*

We define the *degree of the assignment* by $d(x) = l_x^+ - l_x^- + 1$, $\forall x \in X$.

For each $s \in [0,1]$, we define:

$$m(s) = \max_{\substack{x \in O: \\ C_\mu(u(x)) \leq s}} l_x^o, \quad \text{and}$$

$$M(s) = \min_{\substack{x \in O: \\ C_\mu(u(x)) \geq s}} l_x^o.$$

m (resp. M) is a right (resp. left) continuous stepwise function of argument s with values belonging to the discrete finite set $\{1, \ldots, m\}$.

To each $s \in [0,1]$ we associate an interval of contiguous classes $[l_s^-, l_s^+]$ such that

$$\begin{aligned} l_s^- &= \min\{m(s), M(s)\} \quad \text{and} \\ l_s^+ &= \max\{m(s), M(s)\}. \end{aligned} \tag{8.2}$$

It can be easily verified that $l_s^- \leq l_s^+$. Furthermore, due to the monotonicity of m and M, for any r and s of $[0,1]$ s.t. $r < s$, we also have:

$$l_r^- \leq l_s^- \text{ and } l_r^+ \leq l_s^+.$$

Formula 8.2 generates a partition of $[0,1]$ into closed, semi-open or open intervals s.t. to each of these intervals can be associated an interval of labels of classes.

Note that for each $x \in O$ we have

$$l_x^- \leq l_x^0 \leq l_x^+.$$

Consequently, either x is correctly and precisely assigned ($d(x) = 1$), or it is correctly and ambiguously assigned ($d(x) > 1$).

The remaining alternatives of $X \setminus O$ are then assigned to intervals of classes, according to their overall evaluations.

student	profile	O_i
A	$(7,5)$	O_2
B	$(6,6)$	O_1
C	$(7,7)$	O_3
D	$(6,8)$	O_4
A’	$(10,7)$	O_6
B’	$(8,8)$	O_5
C’	$(10,5)$	O_3
D’	$(8,6)$	O_4

Table 8.1: Profiles of the students

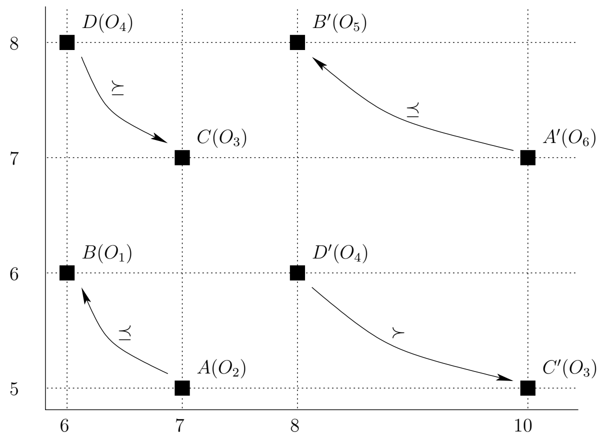

Figure 8.1: Assigning 8 students to 6 classes

8.1.2.3 Illustrative example

Consider a fictitious example where 8 alternatives have to be assigned to 6 ordered classes O_1 to O_6. Each of the alternatives has been evaluated on 2 criteria.

Table 8.1 summarises the value profiles of the 8 alternatives and Figure 8.1, inspired from [MR05a], represents the situation graphically. Note here that the order relation on the classes (arrows) reveals comonotonic contradictory tradeoffs (see [MDGP97]). Hence, the discrimination between the 6 classes is not representable by a Choquet integral. As stated in the previous section, it is nevertheless still possible to find an approximate solution by means of the *less constrained* identification method of Section 7.2.4. We apply it here with $\delta_y = 0.1$.

The capacity determined by the identification method is shown in Table 8.2.

subset	μ
$\emptyset$	0.000
$\{1\}$	0.414
$\{2\}$	0.791
$\{1,2\}$	-0.204

Table 8.2: The capacity for the 8 alternatives problem

The Shapley values are 0.312 for criterion 1 and 0.688 for criterion 2. The interaction index between both criteria equals -0.204.

Figure 8.2 represents a summary of the construction of the classes, as detailed in the previous section. First one can observe that the separation constraint is not respected (between D and C'). Second, one can see that the problem is not suited for a Choquet integral-based MAVT. Nevertheless, with the less constrained identification method, it is still possible to find a satisfactory sorting. 3 alternatives (A', B' and D) are correctly and precisely assigned, whereas the 5 remaining alternatives are correctly and ambiguously assigned.

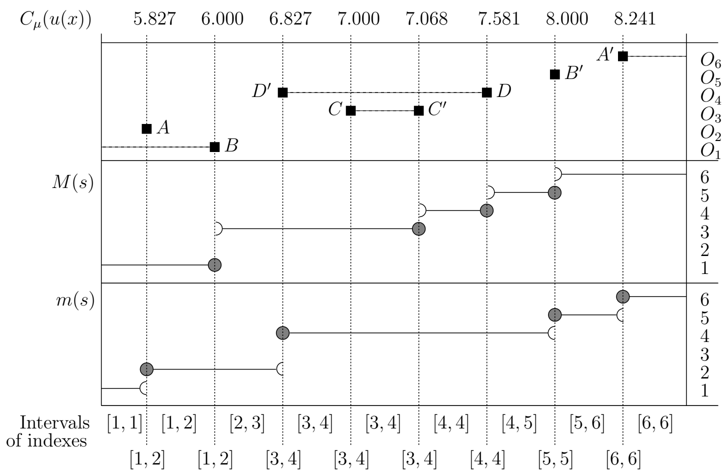

Figure 8.2: Construction of the 6 classes

8.2 The fuzzy case

In this section we suppose that the fuzzy partial evaluations of the elements of X have been aggregated by means of a Choquet integral into a fuzzy overall evaluation for each alternative. In the following sections we will show how to exploit this information in order to determine a choice among the alternatives and to rank them.

8.2.1 The choice problem

Recall that the goal of the choice problematique is to determine an alternative which can be considered as the *best* one. To achieve this, we determine here a subset of alternatives (the *choice recommendation*) among which the best one can be found. Each alternative which does not belong to the choice should be rejected. *choice*

Let us define the degree of plausibility of the preference of one alternative w.r.t. another one. It is given as the possibility Π that an alternative x is not worse than y (let's write $x \succeq y$) in the following way (see e.g. [RV88]): *plausibility of the preference*

$$\Pi(x \succeq y) = \sup_{a \geq b} \big[\min\{[\widetilde{C_{m_\mu}}(\widetilde{u(x)})](a), [\widetilde{C_{m_\mu}}(\widetilde{u(y)})](b)\} \big], \tag{8.3}$$

where $[\widetilde{C_{m_\mu}}(\widetilde{u(x)})](a)$ is a notation for $\eta_{\widetilde{C_{m_\mu}}(\widetilde{u(x)})}(a)$ (see Section 6.3.1), and, where $\widetilde{u(x)} := (\widetilde{u_1(x_1)}, \ldots, \widetilde{u_n(x_n)})$ is the vector of fuzzy partial evaluations of alternative x of X.

Figure 8.3 illustrates the meaning of this degree of plausibility of the preference of x over y ($x, y \in X$). h represents the height at the intersection of the two fuzzy numbers $\widetilde{C_{m_\mu}}(\widetilde{u(x)})$ and $\widetilde{C_{m_\mu}}(\widetilde{u(y)})$ representing alternatives x and y. We have $\Pi(x \succeq y) = h$ and $\Pi(y \succeq x) = 1$.

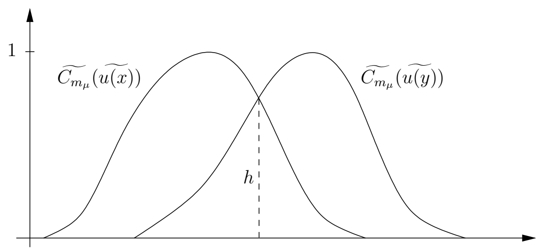

Figure 8.3: Degree of plausibility of the preference of x over y

Roubens and Vincke [RV88] have shown that the plausibility Π as defined in Equation (8.3) is a fuzzy interval order (i.e. a reflexive, complete and Ferrers valued relation) and that it is min − max-transitive.

If these credibilities are computed for each pair of alternatives of X, it is possible to represent the problem as a valued digraph where the nodes represent the alternatives and the arcs the valued relation Π. This graph is called a fuzzy possibility digraph.

Starting from this, one can define a valued strict preference relation by

$$P(x,y) = 1 - \Pi(y \succeq x) \quad \forall (x,y) \in X^2 \text{ [FR94]}.$$

fuzzy preference digraph

The associated digraph is called the *fuzzy preference digraph*.

The objective is to extract information from this digraph, in order to determine a choice among the alternatives of X. Let us first recall the concept of *score of non-domination* of an alternative x of X [Orl78]:

score of non-domination

Definition 8.2.1. *The score of non-domination of an alternative x of X in a fuzzy preference digraph is given by*

$$ND(x) = 1 - \max_{y \neq x} P(y,x).$$

Finally, the choice will be given by the core of X, which is defined as follows:

core

Definition 8.2.2. *The core Y_0 of X is a subset $Y_0 \subseteq X$ such that its elements all have a score of non-domination of 1. Equivalently, $Y_0 = \{x \in X | ND(x) = 1\}$.*

It has been shown in [RV88] that the core of a fuzzy preference digraph is non-empty.

Consequently, Y_0 gives a solution to the choice problem. If Y_0 contains more than one element, in the progressive context, the decision aid may be continued on this restricted set in view of determining a unique best alternative. The following example illustrates a situation where the core is composed of more than one alternative.

> **Example** *Consider the situation presented on Figure 8.4. Alternatives x, y and z of X have respective overall evaluations $\widetilde{C_{m_\mu}}(\widetilde{u(x)})$, $\widetilde{C_{m_\mu}}(\widetilde{u(y)})$ and $\widetilde{C_{m_\mu}}(\widetilde{u(z)})$. Both y and z have a score of non-domination of 1 (one can easily check that $P(y,z) = P(z,y) = 0$). Consequently, both alternatives must be retained for a further analysis step, or the DM has to select one of the equivalently* best *alternatives.*

8.2.2 The ranking problem

In order to determine a ranking on the alternatives of X, we suggest two possibilities in this section: first via a weak order, and second via an interval order.

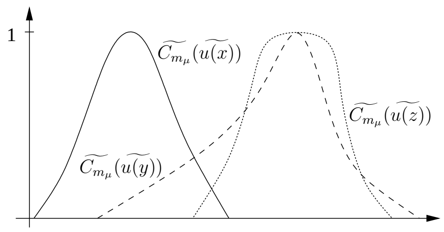

Figure 8.4: Alternatives y and z are held back for a choice recommendation

Let us first introduce the concept of *possibilistic mean* (see [CF01, DP87] for further details). Recall that in Section 6.3.1 we considered the membership function of a fuzzy number $\widetilde{x}$ as a possibility distribution. The *upper possibilistic mean* of $\widetilde{x}$ is then defined by

possibilistic mean

$$\mathcal{M}^+(\widetilde{x}) := \int_0^1 z_M(\lambda)d\lambda = \int_0^1 z_M(\lambda)d\text{Pos}(\widetilde{x} \geq z_M(\lambda))$$

and the *lower possibilistic mean* of $\widetilde{x}$ by

$$\mathcal{M}^-(\widetilde{x}) := \int_0^1 z_m(\lambda)d\lambda = \int_0^1 z_m(\lambda)d\text{Pos}(\widetilde{x} \leq z_m(\lambda)).$$

Intuitively, the upper (resp. lower) possibilistic mean corresponds to the average value of the maxima (resp. minima) of the λ-level sets.

The *possibilistic mean* of the fuzzy number $\widetilde{x}$ is then defined by

$$\mathcal{M}(\widetilde{x}) := \frac{1}{2}[\mathcal{M}^+(\widetilde{x}) + \mathcal{M}^-(\widetilde{x})].$$

Figure 8.5 gives an illustration of the upper possibilistic mean. We represents both the surface and its value on the same figure, which allows to give a better intuition on its semantics. If we consider two fuzzy numbers $\widetilde{x}$ and $\widetilde{y}$ we can define the *upper dominance* $\overset{+}{\geq}$ of $\widetilde{x}$ over $\widetilde{y}$ by

upper dominance

$$\widetilde{x} \overset{+}{\geq} \widetilde{y} \iff \mathcal{M}^+(\widetilde{x}) - \mathcal{M}^+(\widetilde{y}) \geq 0.$$

Similarly one can define the *lower dominance* $\overset{-}{\geq}$.

lower dominance

8.2.2.1 Ranking by a weak order

Let us now apply the concept of possibilistic mean to the overall evaluations of the alternatives of X by means of the fuzzy extension of the Choquet integral.

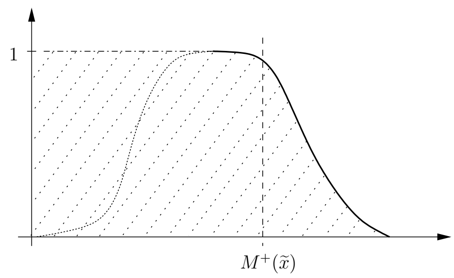

Figure 8.5: Representations of $\mathcal{M}^+(\tilde{x})$: a number and a surface

Let x and y be two alternatives of X. A weak order on X can be defined by the relation $\succeq_{\text{PM}}$ (*is not worse than*):

$$x \succeq_{\text{PM}} y \iff \mathcal{M}(\widetilde{C_m}(\widetilde{u(x)})) \geq \mathcal{M}(\widetilde{C_m}(\widetilde{u(y)})). \tag{8.4}$$

It is possible to give a geometric interpretation to this relation. Consider two alternatives x and y of X with respective overall evaluations $\widetilde{C_{m_\mu}}(\widetilde{u(x)})$ and $\widetilde{C_{m_\mu}}(\widetilde{u(y)})$. These two fuzzy numbers define four areas $A_1, \ldots, A_4$ as shown on Figure 8.6. We can easily see that

$$A_4 - A_3 = \mathcal{M}^+(\widetilde{C_{m_\mu}}(\widetilde{u(x)})) - \mathcal{M}^+(\widetilde{C_{m_\mu}}(\widetilde{u(y)})) \quad \text{(upper dominance) and}$$

$$A_1 - A_2 = \mathcal{M}^-(\widetilde{C_{m_\mu}}(\widetilde{u(x)})) - \mathcal{M}^-(\widetilde{C_{m_\mu}}(\widetilde{u(y)})) \quad \text{(lower dominance).}$$

Definition 8.4 can then be rewritten as

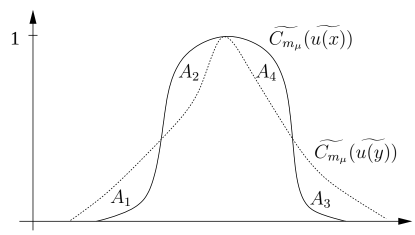

Figure 8.6: Comparing $\widetilde{C_{m_\mu}}(\widetilde{u(x)})$ to $\widetilde{C_{m_\mu}}(\widetilde{u(y)})$

$$x \succeq_{\text{PM}} y \iff A_4 + A_1 \geq A_3 + A_2.$$

This weak order proposal corresponds to the area compensation method of Fortemps and Roubens [FR96].

8.2.2.2 Ranking by an interval order

This second proposal to determine a ranking on X is based on the comparison of the intervals $[\mathcal{M}^-(\widetilde{C_m}(\widetilde{u(x)})), \mathcal{M}^+(\widetilde{C_m}(\widetilde{u(x)}))]$, for each x in X.

Let I be a symmetrical relation, and P be an antisymmetrical relation on $X \times X$ such that

$$\begin{cases} xPy & \iff \mathcal{M}^-(\widetilde{C_{m_\mu}}(\widetilde{u(x)})) > \mathcal{M}^+(\widetilde{C_{m_\mu}}(\widetilde{u(y)})) \\ yPx & \iff \mathcal{M}^-(\widetilde{C_{m_\mu}}(\widetilde{u(y)})) > \mathcal{M}^+(\widetilde{C_{m_\mu}}(\widetilde{u(x)})) \\ xIy & \text{else.} \end{cases}$$

The relation (P, I) is an interval order which can be used to build a ranking on the alternatives of X.

Note that the particularity of such an interval order is that the associated indifference relation I is not transitive. For three alternatives x, y and z of X it is hence possible to have the situation where xIy and yIz, but xPz.

The main difficulty thus arises for indifferent alternatives (or intervals which have a non-empty intersection). In a progressive framework it is possible to focus on these problematic alternatives to make them preferentially more discriminant.

8.2.2.3 On the calculation of the possibilistic means in practise

In real-life MCDA problems, one can suppose that the partial evaluations of the alternatives are given by trapezoidal or triangular fuzzy numbers. Besides the objective is in general to build the simplest possible model, which means in our framework to use a k-additive Choquet integral with k as low as possible.

Let us therefore suppose that the partial valuations of an alternative x of X are trapezoidal fuzzy numbers $\widetilde{u_i(x_i)} = (a_i, b_i, \alpha_i, \beta_i)$ $(i \in N)$ and that we restrict to a 2-additive Choquet integral. The aggregation can then be written as

$$\widetilde{C_{m_\mu}}(\widetilde{u(x)}) = \widetilde{\sum_{i=1}^{n}} m_\mu(i) \tilde{\cdot} \widetilde{u_i(x_i)} \tilde{+} \widetilde{\sum_{\{i,j\}\subseteq N}} m_\mu(i,j) \tilde{\cdot} \widetilde{\min}\{\widetilde{u_i(x_i)}, \widetilde{u_j(x_j)}\}.$$

The minimum of two trapezoidal fuzzy numbers can be summarised by 8 parameters $(a, b, \lambda^-, \alpha', \alpha'', \lambda^+, \beta', \beta'')$ representing two upper piecewise linear shape functions passing through the following points: $(a - \alpha'', 0)$, $(a - \alpha', \lambda^-)$, $(a, 1)$, $(b, 1)$, $(b + \beta', \lambda^+)$, $(b + \beta'', 0)$. Figure 8.7 represents these 8 parameters. Let $\widetilde{m}$ be the minimum of two trapezoidal fuzzy numbers. One can easily obtain that

$$\mathcal{M}^-(\widetilde{m}) = a - \frac{\alpha' + \lambda^-\alpha''}{2} \qquad \text{and}$$

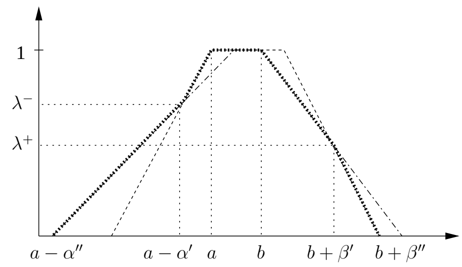

Figure 8.7: The minimum of two trapezoidal fuzzy numbers

$$\mathcal{M}^+(\widetilde{m}) = b + \frac{\beta' + \lambda^+\beta''}{2}.$$

Furthermore, one can see that for any fuzzy numbers $\widetilde{x}$ and $\widetilde{y}$ and any real number p, $\mathcal{M}^\pm(p \tilde{\cdot} \widetilde{x}) = p\mathcal{M}^\pm(\widetilde{x})$ and $\mathcal{M}^\pm(\widetilde{x} \tilde{+} \widetilde{y}) = \mathcal{M}^\pm(\widetilde{x}) + \mathcal{M}^\pm(\widetilde{y})$[1].

In the present context we therefore have for any alternative $x \in X$:

$$\begin{aligned}\mathcal{M}^\pm[\widetilde{C_{m_\mu}}(\widetilde{u(x)})] &= \sum_{i=1}^{n} m_\mu(\{i\})\mathcal{M}^\pm(\widetilde{u_i(x_i)}) \\ &\quad + \sum_{\substack{i,j=1\\ i\neq j}}^{n} m_\mu(\{i,j\})\mathcal{M}^\pm(\widetilde{\min}\{\widetilde{u_i(x_i)}, \widetilde{u_j(x_j)}\}).\end{aligned}$$

This shows that in this particular case where the partial evaluations are trapezoidal fuzzy numbers and where we restrict to a 2-additive fuzzy number, the possibilistic mean of the overall evaluations of the alternatives can be very conveniently calculated.

8.3 Further considerations on progressiveness in Multiattribute Value Theory

In this section, let us briefly return to the concept of progressiveness, and to how it can be used in the framework of MAVT models. As already mentionned in Chapter 2, two main reasons motivate the use of progressive methods, namely economical constraints and prudence.

Progressiveness can conveniently be used if the partial evaluations of the alternatives suffer from impreciseness, which can be represented by fuzzy numbers. Due to economical limitations, or even prudence, in such a framework, the DM will give imprecise values to the alternatives in a first step. Via the

[1] $\mathcal{M}^\pm$ is an abbreviation to avoid the writing of two similar formulae for $\mathcal{M}^+$ and for $\mathcal{M}^-$

methods described in Section 8.2, it is possible to present a first partial conclusion to the DM for the choice and the ranking problems.

The DM might not be satisfied with this partial conclusion, as it may contain too many alternatives which are considered as indifferent. In such a situation, he can focus on these problematic alternatives in order to remove some impreciseness from their evaluations.

In practice this is done by *narrowing* the fuzzy numbers which are used as partial evaluations of the alternatives. Due either to financial or to time constraints, in a previous step of the progressive analysis, the DM may have evaluated certain alternatives quite roughly on some criteria. The intermediate conclusion indicates him on which alternatives he has to focus in order to obtain the desired recommendation. In the choice problematique, he can restrict his analyses to the elements of the core, whereas in the ranking problematique he may focus on alternatives considered as indifferent in the weak order or the interval order.

Such a *filtering* allows the DM to focus on a few alternatives, for which he can try to obtain less imprecise information and narrow their evaluations. If necessary, this step is repeated a few times, until the DM is satisfied with the final recommendation.

Chapter 9

Application of the Kappalab R package

Contents

In this chapter we present the Kappalab package that we helped to develop with I. Kojadinovic and M. Grabisch. Kappalab, which stands for “laboratory for capacities”, is a package for the GNU R statistical system [R D05] for capacity and integral manipulation on a finite setting and which can be used in the context of MAVT.

In the first section we introduce the framework around Kappalab and its limitations. Then, in Section 9.2 we switch to the description of a fictitious example which we analyse and solve by means of Kappalab in Section 9.3.

Note that this Chapter is based on our article [GKM07] and the manual of Kappalab [GKM06].

9.1 On Kappalab

The identification methods discussed in Chapter 7 have been implemented within the Kappalab package [GKM06] for the GNU R statistical system. The package is distributed as free software and can be downloaded from the *Comprehensive R Archive Network* (`http://cran.r-project.org`) or from `http://www.polytech.univ-nantes.fr/kappalab`. To solve linear programs, the LpSolve R package [B+05] is used; strictly convex quadratic programs are solved using the Quadprog R package [TW04]; finally, not necessarily strictly convex quadratic programs are solved using the `ipop` routine of the Kernlab R package [KSHZ04].

As far as the maximum number of criteria is considered, Kappalab allows to work comfortably with up to $n = 10$ criteria if n-additive capacities are considered and with up to $n = 32$ criteria if 2 or 3-additive capacities are considered.

In the following section, we present a fictitious problem which allows to easily understand the use of the Kappalab package and the output of the different identification methods. We intentionally restrict to the crisp case, as the fuzzy case is not yet implemented in Kappalab. Note also that the sorting problematique is not treated here.

9.2 Description of the problem

We consider an extended version of the fictitious problem presented in [Koj07] concerning the evaluation of students in an institute training econometricians. The students are evaluated w.r.t. five subjects: statistics (S), probability (P), economics (E), management (M) and English (En). The marginal values of seven students a, b, c, d, e, f, g on a $[0, 20]$ scale are given in Table 9.1.

Student	S	P	E	M	En	Mean
a	18	11	11	11	18	13.80
b	18	11	18	11	11	13.80
c	11	11	18	11	18	13.80
d	18	18	11	11	11	13.80
e	11	11	18	18	11	13.80
f	11	11	18	11	11	12.40
g	11	11	11	11	18	12.40

Table 9.1: Partial evaluations of the seven students.

Assume that the institute is slightly more oriented towards statistics and probability and suppose that the DM considers that there are three groups of subjects: statistics and probability, economics and management, and English.

In each of the first two groups, subjects are considered to have the same importance. Furthermore, the DM considers that within those groups, subjects are somewhat substitutive, i.e. they overlap to a certain extent. Finally, if a student is good in statistics or probability (resp. bad in statistics and probability), it is better that he is good in English (resp. economics or management) rather than in economics or management (resp. English). This reasoning, applied to the profiles of Table 9.1, leads to the following ranking:

$$a \succ_{\mathcal{O}} b \succ_{\mathcal{O}} c \succ_{\mathcal{O}} d \succ_{\mathcal{O}} e \succ_{\mathcal{O}} f \succ_{\mathcal{O}} g. \quad (9.1)$$

Furthermore we will assume that the DM considers that two students are significantly different if their overall values differ by at least half a unit.

By considering students a and b, and f and g, it is easy to see that the criteria do not satisfy mutual preferential independence, which implies that there is no additive model that can numerically represent the above weak order.

In order to use the identification methods reviewed in Chapter 7 and implemented in Kappalab, we first create 7 R vectors representing the students:

```
> a <- c(18,11,11,11,18)
> b <- c(18,11,18,11,11)
> c <- c(11,11,18,11,18)
> d <- c(18,18,11,11,11)
> e <- c(11,11,18,18,11)
> f <- c(11,11,18,11,11)
> g <- c(11,11,11,11,18)
```

The symbol `>` represents the prompt in the R shell, the symbol `<-` the assignment operator, and `c` is the R function for the creation of vectors. Let us now analyse the outputs of the different identification methods presented earlier in this work.

9.3 Solving the problem

In this section, the objective is to elicit the preferences of the DM in the context of the fictitious problem presented in Section 9.2. This resolution is done interactively, in accordance with the developments of Chapter 2.

9.3.1 The least squares methods

In order to apply the least squares methods presented in Section 7.2.1, the 7 vectors previously defined and representing the students need first to be concatenated into a 7 row matrix, called `C` here, using the `rbind` (*row bind*) matrix creation function:

```
> C <- rbind(a,b,c,d,e,f,g)
```

Then, the DM needs to provide overall values for the seven students. Although it is unrealistic to consider that this information can always be given, we assume in this subsection that the DM is able to provide it. He respectively assigns 15, 14.5, 14, 13.5, 13, 12.5 and 12 to a, b, c, d, e, f and g. These desired overall values are encoded into a 7 element R vector:

```
> overall <- c(15,14.5,14,13.5,13,12.5,12)
```

The least squares identification routine based on quadratic programming (providing an optimal but not necessarily unique solution) can then be called by typing:

```
> ls <- least.squares.capa.ident(5,2,C,overall)
```

in the R terminal. The first argument sets the number of criteria, the second fixes the desired order of k-additivity, and the last two represent the matrix containing the partial values and the vector containing the desired overall values respectively. The result is stored in an R list object, called here `ls`, containing all the relevant information for analysing the results.

The solution, a 2-additive capacity given under the form of its Möbius representation, can be obtained by typing:

```
> m <- ls$solution
```

and visualised by entering `m` after the prompt:

```
> m
                Mobius.capacity
{}              0.000000
{1}             0.311650
{2}             0.176033
...             ...
{4,5}           0.001752
```

As discussed in Section 7.2.1, for the considered example, the obtained solution is probably not unique [MG99].

The Choquet integral for instance of a w.r.t. the solution can be obtained by typing:

```
> Choquet.integral(m,a)
[1] 15
```

To use the least squares identification routine implementing the heuristic approach proposed in [Gra95], we first need to create the *initial capacity* as discussed in Section 7.2.1. Here, in the absence of clear requirements on the form of the Choquet integral, we take the uniform capacity on the set of criteria:

```
> mu.unif <- as.capacity(uniform.capacity(5))
```

The heuristic least squares identification routine can then be called by typing:

```
> hls <- heuristic.ls.capa.ident(5,mu.unif,C,overall,alpha=0.05)
```

The first argument sets the number of criteria, the second contains the initial capacity, the third represents the matrix of partial evaluations, the fourth the vector containing the desired overall values, and the last the parameter controlling the gradient descent.

The overall values computed using the Choquet integral w.r.t. the two obtained solutions are given in the last two columns of the table below, the sixth column containing the desired overall evaluations, and the seventh the mean value of the evaluations of each alternative:

```
   S  P  E  M En Given Mean   LS  HLS
a 18 11 11 11 18  15.0 13.8 15.0 15.0
b 18 11 18 11 11  14.5 13.8 14.5 14.5
c 11 11 18 11 18  14.0 13.8 14.0 14.0
d 18 18 11 11 11  13.5 13.8 13.5 13.5
e 11 11 18 18 11  13.0 13.8 13.0 13.0
f 11 11 18 11 11  12.5 12.4 12.5 12.5
g 11 11 11 11 18  12.0 12.4 12.0 12.0
```

As one can see, both the optimal and the heuristic methods enable to recover the overall values provided by the DM. Recall however that the solution returned by the optimal quadratic method is 2-additive whereas that returned by the heuristic method is 5-additive.

The Shapley values of the solutions can be computed by means of the `Shapley.value` function taking as argument a capacity and are given in the following table :

```
       S    P    E    M   En
LS  0.29 0.14 0.21 0.13 0.24
HLS 0.24 0.18 0.20 0.16 0.21
```

As one could have expected, the Shapley value of the solution obtained by the heuristic approach is less contrasted than that returned by the optimal quadratic method.

9.3.2 The LP, minimum variance and minimum distance methods

As discussed earlier, the least squares methods applied in the previous section are not well adapted to MAVT since they rely on information that a DM cannot always provide. The LP, the minimum variance and the minimum distance methods require only a partial weak order over the available objects, such as the one provided by the DM in Equation (9.1). This weak order is naturally translated as

$$C_{m_\mu}(a) > C_{m_\mu}(b) > C_{m_\mu}(c) > C_{m_\mu}(d) > C_{m_\mu}(e) > C_{m_\mu}(f) > C_{m_\mu}(g).$$

with indifference threshold $\delta_C = 0.5$.

Practically, this threshold is stored in an R variable:

```
> delta.C <- 0.5
```

and the weak order over the students is encoded into a 6 row R matrix:

```
> Acp <- rbind(c(a,b,delta.C), c(b,c,delta.C), c(c,d,delta.C),
               c(d,e,delta.C), c(e,f,delta.C), c(f,g,delta.C))
```

each row containing a constraint of the form $C_{m_\mu}(u(x)) \geq C_{m_\mu}(u(y)) + \delta_C$.

The LP approach is then invoked by typing:

```
> lp <- lin.prog.capa.ident(5,2,A.Choquet.preorder = Acp)
```

The first argument fixes the number of criteria, the second sets the desired order of k-additivity for the solution, and the last contains the partial weak order provided by the DM. All the relevant information to analyse the solution is stored in the R object `lp`.

The minimum variance approach is called similarly:

```
> mv <- mini.var.capa.ident(5,2,A.Choquet.preorder = Acp)
```

To use the minimum distance approach, we first need to create the *initial capacity*. In the absence of clear requirements from the DM, we choose the uniform capacity on the set of criteria, which can be created by entering:

```
> m.mu <- additive.capacity(c(0.2,0.2,0.2,0.2,0.2))
```

The capacity closest to the uniform capacity compatible with the initial preferences of the DM is then obtained by typing:

```
> md <- mini.dist.capa.ident(m.mu,2,"global.scores",
                             A.Choquet.preorder = Acp)
```

The second argument sets the desired order of k-additivity for the solution, while the third one indicates which of the 3 available quadratic distances between capacities should be used [Koj06]. The character string `"global.scores"` refers to the distance given in Equation (7.1).

The overall values computed using the Choquet integral w.r.t. the 2-additive solutions are given in the following table:

	S	P	E	M	En	Mean	LP	MV	MD
a	18	11	11	11	18	13.8	18.00	15.25	14.95
b	18	11	18	11	11	13.8	17.36	14.75	14.45
c	11	11	18	11	18	13.8	16.73	14.25	13.95
d	18	18	11	11	11	13.8	16.09	13.75	13.45
e	11	11	18	18	11	13.8	15.45	13.25	12.95
f	11	11	18	11	11	12.4	14.82	12.75	12.45
g	11	11	11	11	18	12.4	14.18	12.25	11.95

Note that, as expected, the LP approach leads to more dispersed values, reaching the maximum value (18) that a Choquet integral can take for the seven students.

Note also that, for the minimum variance and the minimum distance methods, the differences between the overall values of two consecutive students in the weak order provided by the DM equal exactly δ_C. The latter observation follows from the fact that, in this example, the aim of both methods is roughly to find the Choquet integral that is the closest to the simple arithmetic mean while being in accordance with the preferential information provided by the DM.

The Shapley values of the 2-additive solutions are:

	S	P	E	M	En
LP	0.45	0.00	0.27	0.05	0.23
MV	0.27	0.16	0.21	0.14	0.22
MD	0.24	0.18	0.20	0.16	0.22

As one can see, all three solutions designate statistics (S) as the most important criterion. Note that the LP solution is very extreme, since the overall importance of probability (P) and management (M) is very small and that of S is close to one half.

However, the overall importances of the criteria are not in accordance with the orientation of the institution. Indeed, one would have expected to obtain that statistics (S) and probability (P), and economics (E) and management (M), have the same importances. This is due to the fact that until now, the preferential information which we used was limited to a small number of students which were ranked by the DM.

In order to build a more accurate model, we can impose additional constraints as we shall see in the next section. This clearly justifies the use of an interactive approach to model the DM's preferences in MAVT.

This last table gives the Möbius representations of the three 2-additive solutions:

```
         LP    MV    MD
{}      0.00  0.00  0.00
{1}     0.73  0.34  0.33
{2}     0.00  0.18  0.19
{3}     0.55  0.25  0.21
{4}     0.09  0.15  0.16
{5}     0.45  0.18  0.14
{1,2}  0.00 -0.13 -0.17
{1,3} -0.36 -0.06 -0.05
{1,4}  0.00 -0.04 -0.06
{1,5} -0.18  0.09  0.10
{2,3}  0.00  0.02  0.03
{2,4}  0.00  0.11  0.15
{2,5}  0.00 -0.03 -0.02
{3,4}  0.00 -0.08 -0.08
{3,5} -0.18  0.04  0.08
{4,5} -0.09 -0.01  0.00
```

As expected, the Möbius representation of the LP solution appears to be the least similar to the Möbius representation of the uniform capacity.

9.3.3 Additional constraints on the Shapley value

As discussed in the previous section, assume now that by considering the Shapley values of the 2-additive solutions obtained above, the DM explicitly requires that statistics (S) and probability (P), and economics (E) and management (M), have the same overall importances, i.e. $S \sim_N P$ and $E \sim_N M$.

These additional constraints are translated as

$$-\delta_\phi \leq \phi_{m_\mu}(S) - \phi_{m_\mu}(P) \leq \delta_\phi \quad \text{and}$$
$$-\delta_\phi \leq \phi_{m_\mu}(E) - \phi_{m_\mu}(M) \leq \delta_\phi,$$

where the indifference threshold δ_ϕ is supposed to have been set to 0.01 by the DM. To encode them, an R variable representing the indifference threshold is first created:

```
> delta.phi <- 0.01
```

The inequalities discussed above are then encoded into a 4 row R matrix:

```
> Asp <- rbind(c(1,2,-delta.phi), c(2,1,-delta.phi),
               c(3,4,-delta.phi), c(4,3,-delta.phi))
```

each row corresponding to a constraint of the form $\phi_{m_\mu}(i) - \phi_{m_\mu}(j) \geq c$, $c \in [0, 1]$.

The LP approach is then invoked by typing

```
> lp2 <- lin.prog.capa.ident(5,2,A.Choquet.preorder = Acp,
                             A.Shapley.preorder = Asp)
```

into the R terminal. The minimum variance and minimum distance routines are called similarly.

The Shapley values of the 2-additive solutions are:

```
      S    P    E    M   En
LP 0.23 0.23 0.18 0.18 0.18
MV 0.22 0.21 0.18 0.17 0.22
MD 0.22 0.21 0.18 0.17 0.22
```

As expected, the solutions satisfy the constraints additionally imposed by the DM.

The overall values computed using the Choquet integral w.r.t. the 2-additive solutions are given in the following table:

```
   S  P  E  M En Mean    LP    MV    MD
a 18 11 11 11 18 13.8 16.03 15.12 14.84
b 18 11 18 11 11 13.8 15.52 14.62 14.34
c 11 11 18 11 18 13.8 15.01 14.12 13.84
d 18 18 11 11 11 13.8 14.50 13.62 13.34
e 11 11 18 18 11 13.8 13.99 13.12 12.84
f 18 18 11 11 11 12.4 13.48 12.62 12.34
g 11 11 18 11 11 12.4 12.97 12.12 11.84
```

This time, the three methods give more similar overall values. This was to be expected as the problem is more constrained.

The interaction indices of the 2-additive capacities obtained by means of the LP, minimum variance and minimum distance methods are respectively given in the three tables below:

```
[LP]    S     P     E     M    En
S      NA -0.27 -0.17  0.00 -0.03
P   -0.27    NA  0.00  0.16 -0.04
E   -0.17  0.00    NA -0.12 -0.06
M    0.00  0.16 -0.12    NA -0.07
En  -0.03 -0.04 -0.06 -0.07    NA

[MV]    S     P     E     M    En
S      NA -0.21 -0.05 -0.06  0.10
P   -0.21    NA  0.01  0.15 -0.03
E   -0.05  0.01    NA -0.12  0.05
M   -0.06  0.15 -0.12    NA -0.01
En   0.10 -0.03  0.05 -0.01    NA
```

```
[MD]    S     P     E     M    En
S      NA -0.21 -0.04 -0.07  0.10
P   -0.21    NA  0.03  0.18 -0.01
E   -0.04  0.03    NA -0.10  0.09
M   -0.07  0.18 -0.10    NA  0.00
En   0.10 -0.01  0.09  0.00    NA
```

As one can see, statistics (S) negatively interacts with almost all the subjects, which again is not in accordance with the orientation of the institution. Indeed, one would expect statistics (S) to be complementary with all subjects except probability (P). Once more, in the perspective of an interactive approach, this can be corrected by imposing additional constraints on the interaction indices as we will see in the next section.

9.3.4 Additional constraints on the interaction indices

Assume finally that, in order to be in accordance with the orientation of the institution, the DM imposes that subjects within the same group[1] have to interact in a substitutive way, whereas two subjects from different groups have to interact in a complementary way. This additional preferential information is translated by means of the following constraints:

P	E	M	En	
$-1 \leq I_{m_\mu}(SP) \leq -\delta_I$	$\delta_I \leq I_{m_\mu}(SE) \leq 1$	$\delta_I \leq I_{m_\mu}(SM) \leq 1$	$\delta_I \leq I_{m_\mu}(SEn) \leq 1$	S
	$\delta_I \leq I_{m_\mu}(PE) \leq 1$	$\delta_I \leq I_{m_\mu}(PM) \leq 1$	$\delta_I \leq I_{m_\mu}(PEn) \leq 1$	P
		$-1 \leq I_{m_\mu}(EM) \leq -\delta_I$	$\delta_I \leq I_{m_\mu}(EEn) \leq 1$	E
			$\delta_I \leq I_{m_\mu}(MEn) \leq 1$	M

where δ_I, supposed set to 0.05, is a threshold defined by the DM to be interpreted as the minimal absolute value of an interaction index to be considered as significantly different from zero.

To encode this additional preferential information, an R variable representing the threshold is first created:

```
> delta.I <- 0.05
```

The constraints discussed above are then encoded into a 10 row R matrix:

```
>  Aii <- rbind(c(1,2,-1,-delta.I), c(1,3,delta.I,1),
                c(1,4,delta.I,1), c(1,5,delta.I,1),
                c(2,3,delta.I,1), c(2,4,delta.I,1),
                c(2,5,delta.I,1), c(3,4,-1,-delta.I),
                c(3,5,delta.I,1), c(4,5,delta.I,1))
```

each row corresponding to a constraint of the form $a \leq I_{m_\mu}(ij) \leq b$, $a, b \in [-1, 1]$.

[1]Recall that the three groups of subjects are $\{S, P\}$, $\{E, M\}$, and $\{En\}$.

There are no 2-additive capacities compatible with these additional constraints. The order of k-additivity is then incremented and the LP method is invoked by typing:

```
> lp3 <- lin.prog.capa.ident(5,3, A.Choquet.preorder = Acp,
                              A.Shapley.preorder = Asp,
                              A.interaction.interval = Aii)
```

The minimum variance and minimum distance routines are called similarly.

The Shapley values and the interaction indices of the three 3-additive solutions are given in the four following tables:

	S	P	E	M	En
LP	0.23	0.23	0.16	0.16	0.22
MV	0.23	0.22	0.18	0.18	0.20
MD	0.22	0.21	0.18	0.19	0.21

[LP]	S	P	E	M	En
S	NA	-0.30	0.05	0.05	0.12
P	-0.30	NA	0.07	0.14	0.05
E	0.05	0.07	NA	-0.24	0.05
M	0.05	0.14	-0.24	NA	0.05
En	0.12	0.05	0.05	0.05	NA

[MV]	S	P	E	M	En
S	NA	-0.13	0.05	0.05	0.05
P	-0.13	NA	0.05	0.05	0.05
E	0.05	0.05	NA	-0.05	0.05
M	0.05	0.05	-0.05	NA	0.05
En	0.05	0.05	0.05	0.05	NA

[MD]	S	P	E	M	En
S	NA	-0.21	0.05	0.05	0.05
P	-0.21	NA	0.05	0.05	0.05
E	0.05	0.05	NA	-0.12	0.05
M	0.05	0.05	-0.12	NA	0.05
En	0.05	0.05	0.05	0.05	NA

As expected, the constraints additionally imposed by the DM are satisfied. The overall values computed using the Choquet integral w.r.t. the 3-additive solutions are given in the following table:

```
   S  P  E  M En Mean    LP    MV    MD
a 18 11 11 11 18 13.8 14.06 14.26 14.45
b 18 11 18 11 11 13.8 13.55 13.76 13.95
c 11 11 18 11 18 13.8 13.04 13.26 13.45
d 18 18 11 11 11 13.8 12.53 12.76 12.95
e 11 11 18 18 11 13.8 12.02 12.26 12.45
f 18 18 11 11 11 12.4 11.51 11.76 11.95
g 11 11 18 11 11 12.4 11.00 11.26 11.45
```

We hereby conclude the process of the modelling of the DM's preferences. In the following section we imagine a scenario where the DM considers that the 3-additive solutions above are too complex and where he prefers to have a simpler description of his preferences. In such a case, the DM has to take into account that some of his preferences will be violated.

9.3.5 A simpler solution

Assume that for the sake of simplicity the DM absolutely wants a 2-additive solution for the problem described in the previous subsections. In that case, it is possible to use the generalised least squares method described in Section 7.2.4 to obtain an approximate solution.

First of all, the weak order over the students has to be encoded into a 6 row R matrix:

```
> rk.proto <- rbind(c(1,2), c(2,3), c(3,4), c(4,5), c(5,6), c(6,7))
```

The integers correspond to the line indices of the alternatives a, b, c, d, e, f and g in the matrix `C` defined in Section 9.3.1.

The generalised least squares method can then be called by typing:

```
> gls <- ls.ranking.capa.ident(5, 2, C, rk.proto, 0.5,
                               A.Shapley.preorder = Asp,
                               A.interaction.interval = Aii)
```

The first argument sets the number of criteria, the second the desired order of k-additivity, the third the matrix containing the partial evaluations, the fourth the matrix containing the weak order and the fifth argument is the value of the threshold δ_y. The last two arguments contain the matrices encoding the additional constraints on the Shapley value and on the interaction indices respectively.

Although we know from the previous subsections that there are no 2-additive capacities compatible with the imposed constraints, this method provides a solution with a non zero objective function as we could have expected. The following table gives the aggregated overall values $\{C_{m_\mu}(u(x))\}_{x\in O}$ in the last column and the overall values $\{y(x)\}_{x\in O}$ in the last but one column:

	S	P	E	M	En	Mean	y	GLS
a	18	11	11	11	18	13.8	13.94	13.67
b	18	11	18	11	11	13.8	13.44	13.44
c	11	11	18	11	18	13.8	12.94	12.81
d	18	18	11	11	11	13.8	12.44	12.57
e	11	11	18	18	11	13.8	11.94	11.81
f	18	18	11	11	11	12.4	11.44	11.57
g	11	11	18	11	11	12.4	10.94	11.21

As one can see, the ranking provided by the DM is not violated but the minimal threshold δ_y is not always respected (for example $C_{m_\mu}(f) - C_{m_\mu}(g) < \delta_y$). The Shapley value and the interaction indices of this 2-additive solution are:

S	P	E	M	En
0.23	0.22	0.17	0.16	0.22

	S	P	E	M	En
S	NA	-0.22	0.05	0.05	0.14
P	-0.22	NA	0.06	0.09	0.06
E	0.05	0.06	NA	-0.08	0.15
M	0.05	0.09	-0.08	NA	0.05
En	0.14	0.06	0.15	0.05	NA

which as expected satisfy the constraints imposed by the DM. It is now up to the DM to evaluate if the violation of δ_y does not deteriorate significantly the overall quality of the numerical representation of his preferences.

Last but not least ...

Tout avantage a ses inconvénients et réciproquement.

Les Shadoks, Jacques Rouxel[2]

Contents

In the final chapter of this work, our objective is to give a synthesised view of the results presented in this work, to underline again the different connections between the three parts and to show the new questions which have arisen during our research activities.

Consequently, this chapter is built as follows. In the first section we summarise the structure of the work and present our achievements. Then, in the second section, we show perspectives for our future research activities, and we draw some conclusions in the last section.

Summary of the main achievements

In the first part of the work, we introduced the reader to a particular plan of attack to MCDA problems, namely the *constructive approach*. Via interactions between the main actors of the decision process, it helps to construct and uncover the DM's preferences simultaneously with the determination of the recommendation. Then, we delimited our field of research in a constructive framework, to what we call *progressive methods*. Such methods allow to build the final recommendation via *intermediate partial conclusions*, requiring further investigations of the problem. In particular we showed the consequences

[2] Any advantage has its inconvenients, and reciprocally.

of such methods in both methodological schools of MCDA.

The second part deals with *outranking methods* for the *choice* and the *k-choice* problematique. We presented a new and innovative method, called RUBIS , based on the pioneering research around the classical ELECTRE techniques. Via five extensively discussed *pragmatic principles*, we determined the mathematical construct, namely the outranking *hyperkernel*, which gives a choice recommendation in a general outranking digraph. RUBIS fits well in the framework of progressiveness, as it determines the choice recommendation very prudently and requires further interactions from the DM to determine the final best alternative. In particular, it can be used in problems which contain *missing information* or constrained by time or limited economical resources. We also have discussed the different formulations of the *k*-choice problematique and have shown how they can be solved by means of the RUBIS method.

In the third part, we focussed on Multiattribute Value Theory based on the *Choquet integral*. First we presented a *fuzzy extension* of the Choquet integral, which allows to take into account *impreciseness* in MCDA problems. Then, we presented the *capacity identification problem* along with one of our contributions, which allows to find a capacity which is "as close as possible" to the DM's preferences. We then detailed how different MCDA problematiques can be solved by means of the previously presented techniques. In particular, we put our considerations in a progressive context, which permits to refine the evaluations at later stages of the decision process, if necessary. We finally introduced the Kappalab package which we contributed to develop.

It is worth underlining a few main threads which have guided us through our work. First of all, the concepts of *constructive* and *progressive* MCDA are inherent to our whole discourse. They have been defined in the first part, but are quite regularly highlighted in our methodological discussions of Parts II and III. Then, we have often focussed on the central role of the DM, to show that he is an incontrovertible element of the MCDA. Finally, we have regularly included some practical considerations in our work, which we regard as necessary, in order to give a pragmatic justification to our research activities.

Perspectives

Very regularly during our work, new problems and questions arose, some of which remain completely untouched. Let us mention the most important ones here, as they can be seen as future perspectives for our research activities:

- In a progressive method, an important condition is the availability of the DM. It may happen that, for some reasons, at a given moment of the process, he decides that the final recommendation *must* be obtained in the next stage. This requires that the next step of the process is less *prudent* in order to reach the desired objective with absolute certainty. In particular, it will be interesting to adapt the RUBIS method for the special *one-step* choice problem, where the final recommendation needs to be obtained in a single step.

- After the quite theoretical presentation made in this work, giving a *recipe* on how to use the RUBIS method in a real-world situation will be one of our future tasks.

- As already mentionned, our work on the k-choice problematique has produced only early results, which require further theoretical and algorithmic investigations.

- The ranking and the sorting problematiques can probably also be handled by means of a bipolar-valued outranking relation. A stimulating work would be to see how these typologies of problems can be dealt within a progressive framework with such a relation.

- The properties and structure of an outranking relation, built in a bipolar-valued context, need to be studied in details. It would allow to determine what types of outranking digraphs can really be generated from performance tables.

- We have not discussed *robustness* issues in this work. Investigating the *quality* and the *reliability* of the obtained solution might be a stimulating challenge, in both the outranking and the value function framework.

- In particular, the capacity identification methods do not always produce a unique solution. Investigating the degree of indeterminacy of the results and the robustness of those methods would be a challenging task, which we would like to undertake in the future.

- It would be interesting to extend our considerations on impreciseness modelled by fuzzy numbers to the sorting problematique. In such a situation, progressiveness could allow to give a more and more precise structure of the ordered classes.

- Concerning the capacity identification problem, one of our main interests is linked to the case where the Choquet integral cannot represent the ranking imposed by the DM. We are currently working on a way to determine a capacity, respecting his initial preferences on the alternatives, as accurately as possible.

Concluding remarks

In the preface, we have mentionned the difficulty of solving MCDA problems suffering from impreciseness, missing information and limited economical resources.

We hopefully have convinced the reader, that *progressiveness* could be an appropriate answer to these issues. We have shown that it can be considered as a general framework, guiding a decision analysis, and which can be put around different MCDA methods. Besides, we have presented the generic properties of a progressive context, which clearly depend on the underlying mathematical algorithms, the methodological school and the problematique which has to be solved.

Nevertheless, there exists a certain amount of problems which cannot be solved by a progressive decision analysis. For example, as soon as the recommendation needs to be determined in a *single* step, progressive decision analysis methods are not appropriate. Consequently, situations involving automatic decisions generated by a software, might not be considered in a progressive context.

Any MCDA scholar may already have wondered which method he would use if he had to solve a progressive decision problem and to which methodological school he should stick. In fact, the reader of this work should now agree that the practice of MCDA is not about adhering to either the American or the European stream of thought. With this work we have been able to show that the way of solving a decision problem depends on the available information, on the objectives, on the possible and potential interactions with the DM and on the type of process which is eligible.

We think that the selection of a resolution methodology should therefore be guided by a thorough study of the underlying decision situation, taking into account all the stakeholders of the process. Consequently, sticking at any cost to a given MCDA method is certainly a bad option. Either the decision problem might not be suitable for the selected method or the DM might not be prepared to answer some preferential questions required by the method. Therefore, this work's objective was also an attempt to rub off the dichotomy between the two methodological MCDA schools by presenting our research in both fields from a common point of view.

Bibliography

[Arr51] K.J. Arrow. *Social choice and individual values.* J. Wiley, New York, 1951. 2nd edition, 1963. 7

[B+05] Michel Berkelaar et al. `lpSolve`*: Interface to LPSolve v.5 to solve linear/integer programs*, 2005. R package version 1.1.9. 130

[Bel99] V. Belton. Multicriteria problem structuring and analysis in a value theory framework. In T. Stewart, T. Gal, and T. Hanne, editors, *Multicriteria Decision Making: Advances in MCDM Models, Algorithms, Theory and Applications*, chapter 12. Springer, 1999. 15

[Ber70] C. Berge. *Graphes et hypergraphes.* Dunod, Paris, 1970. 48

[Bis97] R. Bisdorff. On computing kernels from l-valued simple graphs. In *Proceedings of the 5th European Congress on Intelligent Techniques and Soft Computing*, volume 1, pages 97–103. EUFIT'97, Aachen, 1997. 42, 59

[Bis99] R. Bisdorff. Bipolar ranking from pairwise fuzzy outrankings. *Belgian Journal of Operations Research, Statistics and Computer Science*, 37 (4) 97:379–387, 1999. 68

[Bis00] R. Bisdorff. Logical foundation of fuzzy preferential systems with application to the electre decision aid methods. *Computers & Operations Research*, 27:673–687, 2000. 34, 46

[Bis02] R. Bisdorff. Logical foundation of multicriteria preference aggregation. In D. Bouyssou et al., editor, *Essay in Aiding Decisions with Multiple Criteria*, pages 379–403. Kluwer Academic Publishers, 2002. 12, 34, 35

[Bis04] R. Bisdorff. Concordant outranking with multiple criteria of ordinal significance. *4OR, Quarterly Journal of the Belgian, French and Italian Operations Research Societies, Springer-Verlag*, 2:4:293–308, 2004. 41

[Bis06a] R. Bisdorff. On enumerating the kernels in a bipolar-valued outranking digraph. In D. Bouyssou, F. Roberts, and A. Tsoukiás, editors, *Proceedings of the DIMACS-LAMSADE workshop on Voting Theory and Preference Modelling, Paris, 25–28 October 2006*, volume 6 of *Annales du LAMSADE*, pages 1–38. Université Paris-Dauphine, CNRS, Paris, 2006. 59

[Bis06b] R. Bisdorff. The Python digraph implementation for RuBy: User Manual. University of Luxembourg, http://sma.uni.lu/bisdorff/Digraph, 2006. 64

[BM94] J.P. Brans and B. Mareschal. The PROMETHEE-GAIA decision support system for multicriteria investigations. *Investigation Operativa*, 4(2):102–117, 1994. 68

[BM02] J.P. Brans and B. Mareschal. *PROMETHEE-GAIA. Une Méthodologie d'Aide à la Décision en Présence de Critères Multiples*. Ellipses, Paris, France, 2002. 13

[BM05] J.-P. Brans and B. Mareschal. *Multiple criteria decision analysis: state of the art surveys*, chapter PROMETHEE Methods, pages 163–195. Springer – International series in operations research and management science. Figueira, J. and Greco, S. and Ehrgott, M. eds., 2005. 68

[BMP+00] D. Bouyssou, T. Marchant, M. Pirlot, P. Perny, A. Tsoukias, and P. Vincke. *Evaluation and decision models, A critical Perspective*. Kluwer's International Series. Kluwer, Massachusetts, 2000.

[BMP+06] D. Bouyssou, T. Marchant, M. Pirlot, A. Tsoukias, and P. Vincke. *Evaluation and decision models with multiple criteria, Stepping stones for the analyst*. Springer's International Series. Springer, New York, 2006. 8, 22

[BMR07] R. Bisdorff, P. Meyer, and M. Roubens. Rubis: a bipolar-valued outranking method for the best choice decision problem. *4OR, Quaterly Journal of the Belgian, French and Italian Operations Research Societies*, 2007. in press. xxi, xxiii, 31, 33, 46

[Bou90] D. Bouyssou. Building criteria: A prerequisite for MCDA. In C.A. Bana e Costa, editor, *Readings in Multiple Criteria Decision Aid*, pages 58–80. Springer-Verlag, Berlin, 1990. 11, 12

[Bou96] D. Bouyssou. Outranking relations: do they have special properties? *Journal of Multi-Criteria Decision Analysis*, 5:99–111, 1996. 40

[Bou06] D. Bouyssou. Notes on bipolar outranking. Personal communication, August 2006. 40

[BP04] D. Bouyssou and M. Pirlot. Preferences for multi-attributed alternatives: Traces, dominance, and numerical representations. *J. of Mathematical Psychology*, 48:167–185, 2004. 86

[BP05] D. Bouyssou and M. Pirlot. Conjoint measurement tools for MCDM: A brief introduction. In J. Figuera, S. Greco, and M. Ehrgott, editors, *Multiple Criteria Decision Analysis: State of the Art Surveys*, pages 73–132. Springer, 2005. 86

[BPR06] R. Bisdorff, M. Pirlot, and M. Roubens. Choices and kernels in bipolar valued digraphs. *European Journal of Operational Research*, 175:155–170, 2006. 33, 42

[BR96] R. Bisdorff and M. Roubens. On defining and computing fuzzy kernels from l-valued simple graphs. In Da Ruan et al., editor, *Intelligent Systems and Soft Computing for Nuclear Science and Industry, FLINS'96 workshop*, pages 113–123. World Scientific Publishers, Singapoure, 1996. 46

[BR03] R. Bisdorff and M. Roubens. Choice procedures in pairwise comparison multiple-attribute decision making methods. In R. Berghammer, B. Möller, and G. Struth, editors, *Relational and Kleene-Algebraic Methods in Computer Science: 7th International Seminar on Relational Methods in Computer Science and 2nd International Workshop on Applications of Kleene Algebra*, volume 3051 of *Lecture Notes in Computer Science*, pages 1–7, Heidelberg, 2003. Springer-Verlag. 43

[BRT88] D. Bell, H. Raiffa, and A. Tversky, editors. *Decision making: descriptive, normative, and prescriptive interactions*. Cambridge university press, Cambridge, 1988. 8

[BS01] V. Belton and T. Stewart. *Muliple Criteria Decision Analysis: An Integrated Approach*. Kluwer Academic Publishers, 2001. 9

[BS03] V. Belton and Th. Stewart. *Multiple Criteria Decision Analysis*. Springer, 2003. 9

[BV99] C.A. Bana e Costa and J.C. Vansnick. Preference relations in MCDM. In T. Gal and T. Hanne T. Steward, editors, *MultiCriteria Decision Making: Advances in MCDM models, algorithms, theory and applications*. Kluwer, 1999. 87

[CF96] C. Carlsson and R. Fullér. Fuzzy multiple criteria decision making: Recent developments. *Fuzzy Sets and Systems*, 78:139–153, 1996. 100

[CF01] C. Carlsson and R. Fullér. On possibilistic mean value and variance of fuzzy numbers. *Fuzzy Sets and Systems*, 122:315–326, 2001. 123

[Cho53] G. Choquet. Theory of capacities. *Annales de l'Institut Fourier*, 5:131–295, 1953. 87, 88

[Chv73] V. Chvaátal. On the computational complexity of finding a kernel. Technical report, Report No. CRM-300, Centre de recherches mathématiques, Université de Montréal, 1973. 61

[CJ89] A. Chateauneuf and J-Y. Jaffray. Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion. *Mathematical Social Sciences*, 17(3):263–283, 1989. 90

[CM04] M. Ceberio and F. Modave. An interval-valued 2-additive choquet integral for multicriteria decision making. In *Proceedings of the 10th Conference IPMU, July 2004, Perugia, Italy*, pages 1567–1574, 2004. 100

[DP80] D. Dubois and H. Prade. *Fuzzy Sets and Systems: Theory and Application*. Academic Press, New York, 1980. 95

[DP87] D. Dubois and H. Prade. The mean value of a fuzzy number. *Fuzzy Sets and Systems*, 24:279–300, 1987. 123

[DT04] L. Dias and A. Tsoukiàs. On the constructive and other approaches in decision aiding. In C.H. Antunes, J. Figueira, and J. Clímaco, editors, *Aide multicritère à la décision : Multiple criteria decision aiding*, pages 13–28. CCDRC/INESCC/FEUC, 2004. 8, 9

[Fis70] P.C. Fishburn. *Utility Theory for Decision Making*. Wiley, New York, 1970. 7

[FKM06] K. Fujimoto, I. Kojadinovic, and J-L. Marichal. Axiomatic characterizations of probabilistic and cardinal-probabilistic interaction indices. *Games and Economic Behavior*, 55:72–99, 2006. 93

[FM00] K. Fujimoto and T. Murofushi. Hierarchical decomposition of the Choquet integral. In M. Grabisch, T. Murofushi, and M. Sugeno, editors, *Fuzzy Measures and Integrals: Theory and Applications*, pages 95–103. Physica-Verlag, 2000. 102

[FR94] J. Fodor and M. Roubens. *Fuzzy preference modelling and multicriteria decision support*. Theory and decision library, Series D: System theory, knowledge engineering and problem solving. Kluwer Academic Publishers, Dordrecht, Boston, London, 1994. 122

[FR96] P. Fortemps and M. Roubens. Ranking and defuzzification methods based on area compensation. *Fuzzy Sets and Systems*, 82:319–330, 1996. 124

[GKM06] M. Grabisch, I. Kojadinovic, and P. Meyer. `kappalab`*: Non additive measure and integral manipulation functions*, 2006. R package version 0.3. 83, 106, 129, 130

[GKM07] M. Grabisch, I. Kojadinovic, and P. Meyer. A review of capacity identification methods for Choquet integral based multi-attribute utility theory: Applications of the Kappalab R package. *European Journal of Operational Research*, 2007. in press. xxi, xxv, 83, 85, 101, 129

[GL04] M. Grabisch and C. Labreuche. Fuzzy measures and integrals in MCDA. In J. Figueira, S. Greco, and M. Ehrgott, editors, *Multiple Criteria Decision Analysis*, pages 563–608. Springer, 2004. 87

[GLV03] M. Grabisch, Ch. Labreuche, and J.C. Vansnick. On the extension of pseudo-Boolean functions for the aggregation of interacting criteria. *European Journal of Operational Research*, 148:28–47, 2003. 87

[GM98] A. Guitouni and J.M. Martel. Tentative guidelines to help choosing an appropriate MCDA method. *European Journal of Operational Research*, 109(2):501–521, September 1998. 18

[GMS01] S. Greco, B. Matarazzo, and R. Slowinski. Conjoint measurement and rough set approach for multicriteria sorting problems in presence of ordinal criteria. In A. Colorni, M. Paruccini, , and B. Roy, editors, *AMCDA: Aide Multicritère à la Décision (Multiple Criteria Decision Aiding)*, number EUR 19808 EN, pages 117–144. European Commission Report, Joint Research Centre, Ispra, 2001. 115

[GNW95] M. Grabisch, H.T. Nguyen, and E.A. Walker. *Fundamentals of uncertainty calculi with applications to fuzzy inference*. Kluwer Academic, Dordrecht, 1995. 106

[GR99] M. Grabisch and Marc Roubens. An axiomatic approach to the concept of interaction among players in cooperative games. *Internat. J. Game Theory*, 28(4):547–565, 1999. 91

[GR00] M. Grabisch and M. Roubens. Application of the choquet integral in multicriteria decision making. In M. Grabisch, T. Murofushi, and M. Sugeno, editors, *Fuzzy Measures and Integrals*, pages 348–374. Physica Verlag, Heidelberg, 2000. 92

[Gra92] M. Grabisch. The application of fuzzy integrals in multicriteria decision making. *European Journal of Operational Research*, 89:445–456, 1992. 14, 87, 89

[Gra95] M. Grabisch. A new algorithm for identyfing fuzzy measures and its application to pattern recognition. In *Int. 4th IEEE Conf. on Fuzzy Systems*, pages 145–150, Yokohama, Japan, March 1995. 106, 132

[Gra97a] M. Grabisch. Alternative representations of discrete fuzzy measures for decision making. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 5(5):587–607, 1997. 102

[Gra97b] M. Grabisch. k-order additive discrete fuzzy measure and their representation. *Fuzzy Sets and Systems*, 92:131–295, 1997. 92, 93

[Höh82] U. Höhle. Integration with respect to fuzzy measures. In *Proceedings IFAC Symposium on Theory and Applications of Digital Control*, New Delhi, 1982. 89

[IS96] K. Ishii and M. Sugeno. A model of human evaluation process using fuzzy measure. *Int. J. Man-Machine Studies*, 67:242–257, 1996. 106

[JS82] E. Jacquet-Lagrèze and Y. Siskos. Assessing a set of additive utility functions for multicriteria decision making: the UTA method. *European Journal of Operational Research*, 10:151–164, 1982. 86

[Kit93] L. Kitainik. *Fuzzy decision procedures with binary relations: towards a unified theory.* Kluwer Academic Publisher, Boston, 1993. 46

[KLST71] D.H. Krantz, R.D. Luce, P. Suppes, and A. Tversky. *Foundations of measurement, volume 1: Additive and polynomial representations.* Academic Press, 1971. 86

[KMR05] I. Kojadinovic, J-L. Marichal, and M. Roubens. An axiomatic approach to the definition of the entropy of a discrete Choquet capacity. *Information Sciences*, 172:131–153, 2005. 108

[Koj06] I. Kojadinovic. Quadratic distances for capacity and bi-capacity approximation and identification. *4OR: A Quarterly Journal of Operations Research*, page in press, 2006. 109, 134

[Koj07] I. Kojadinovic. Minimum variance capacity identification. *European Journal of Operational Research*, 177(1):498–514, 2007. 108, 130

[KR76] R. L. Keeney and H. Raiffa. *Decision with multiple objectives.* Wiley, New-York, 1976. 8, 13, 86

[KSHZ04] Alexandros Karatzoglou, Alex Smola, Kurt Hornik, and Achim Zeileis. `kernlab`: An S4 package for kernel methods in R. *Journal of Statistical Software*, 11(9):1–20, 2004. 130

[Lam07] C. Lamboray. Supporting the search for a group ranking with robust conclusions on prudent orders. *Annales du Lamsade*, (7):145 – 171, 2007. 27

[LG03] C. Labreuche and M. Grabisch. The Choquet integral for the aggregation of interval scales in multicriteria decision making. *Fuzzy Sets and Systems*, 137:11–16, 2003. 14, 87, 89

[LMO83] M. Landry, J.-L. Malouin, and M. Oral. Model validation in operations research. *European Journal of Operational Research*, 92:443 – 457, 1983. 15

[LPB83] M. Landry, D. Pascot, and D. Briolat. Can DSS evolve without changing our view of the concept of problem? *Decision Support Systems*, 1:25–36, 1983. 9

[LR57] R.D. Luce and H. Raiffa. *Games and Decisions*. J. Wiley, New York, 1957. 7, 8

[Mar78] J.G. March. Bounded rationality, ambiguity and the engineering of choice. *Bell Journal of Economics*, 9:587–608, 1978. 15

[Mar98] J-L. Marichal. *Aggregation operators for multicriteria decison aid*. PhD thesis, University of Liège, Liège, Belgium, 1998. 109

[Mar00a] J.-L. Marichal. An axiomatic approach of the discrete Choquet integral as a tool to aggregate interacting criteria. *IEEE Transactions on Fuzzy Systems*, 8(6):800–807, 2000. 14, 87, 89

[Mar00b] J.-L. Marichal. Behavioral analysis of aggregation in multicriteria decision aid. In J. Fodor, B. De Baets, and P. Perny, editors, *Preferences and Decisions under Incomplete Knowledge*, pages 153–178. Physica-Verlag, 2000. 90, 92, 102, 103

[Mar04] J-L. Marichal. Tolerant or intolerant character of interacting criteria in aggregation by the Choquet integral. *European Journal of Operational Research*, 155(3):771–791, 2004. 90, 92, 102

[MB07] P. Meyer and R. Bisdorff. Exploitation of a bipolar-valued outranking relation for the choice of k best alternatives. In *FRANCORO V / ROADEF 2007: Conférence scientifique conjointe en Recherche Opérationnelle et Aide à la Décision*, pages 193–206, Grenoble, France, 20-23 February 2007. xxi, xxiii, 31, 67

[MDGP97] F. Modave, D. Dubois, M. Grabisch, and H. Prade. A Choquet integral representation in multicriteria decision making. In *Working Notes of the AAAI Workshop Frontiers in Soft Computing and Decision Systems*, pages 30–39, Boston, 8-10 November 1997. 105, 119

[MG99] P. Miranda and M. Grabisch. Optimization issues for fuzzy measures. *Int. J. of Uncertainty, Fuzziness, and Knowledge Based Systems*, 7:545–560, 1999. 106, 132

[MM89] T. Mori and T. Murofushi. An analysis of evaluation model using fuzzy measure and the Choquet integral. In *5th Fuzzy System Symposium*, pages 207–212, Kobe, Japan, 1989. In Japanese. 106

[MMR05] J-L. Marichal, P. Meyer, and M. Roubens. Sorting multiattribute alternatives: The tomaso method. *International Journal of Computers & Operations Research*, 32:861–877, 2005. xxi, xxiv, 83, 108, 114, 116

[MR00] J-L. Marichal and M. Roubens. Determination of weights of interacting criteria from a reference set. *European Journal of Operational Research*, 124:641–650, 2000. 107

[MR05a] P. Meyer and M. Roubens. Choice, ranking and sorting in fuzzy multiple criteria decision aid. In J. Figuera, S. Greco, and M. Ehrgott, editors, *Multiple Criteria Decision Analysis: State of the Art Surveys*, pages 471–506. Springer, 2005. xxi, xxiv, 83, 85, 92, 101, 109, 114, 119

[MR05b] P. Meyer and M. Roubens. On the use of the choquet integral with fuzzy numbers in multiple criteria decision support. *Fuzzy Sets and Systems*, pages 927–938, 2005. xxi, xxiv, 83, 86, 114

[MS93] T. Murofushi and S. Soneda. Techniques for reading fuzzy measures (iii): Interaction index. In *9th Fuzzy System Symposium*, pages 693–696, Saporo, Japan, 1993. 91

[MS00] T. Murofushi and M. Sugeno. Fuzzy measures and fuzzy integrals. In M. Grabisch, T. Murofushi, and M. Sugeno, editors, *Fuzzy Measures and Integrals: Theory and Applications*, pages 3–41. Physica-Verlag, 2000. 89

[Ngu78] H. T. Nguyen. A note on the extension principle for fuzzy sets. *Journal of Mathematical Analysis an Applications*, 64:369–380, 1978. 96, 97

[Orl78] S.A. Orlovski. Decision-making with a fuzzy preference relation. *Fuzzy Sets and Systems*, 1:155–167, 1978. 122

[R D05] R Development Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2005. ISBN 3-900051-00-3. 129

[RB86] B. Roy and D. Bouyssou. Comparison of two decision-aid models applied to a nuclear plant siting example. *European Journal of Operational Research*, 25:200–215, 1986. 15

[RB93] B. Roy and D. Bouyssou. *Aide multicritère à la décision: Méthodes et cas*. Economica, Paris, 1993. 12, 13, 15, 36, 46, 47, 48, 56, 68

[Ric53] M. Richardson. Solution of irreflective relations. *Ann. Math.*, 58:573–580, 1953. 54

[Ros89] J. Rosenhead. *Rational analysis of a problematic world*. J. Wiley, New York, 1989. 2nd revised edition in 2001. 9

[Rot64] G-C. Rota. On the foundations of combinatorial theory. I. Theory of Möbius functions. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 2:340–368 (1964), 1964. 90

[Rou01] M. Roubens. Ordinal multiattribute sorting and ordering in the presence of interacting points of view. In D. Bouyssou, E. Jacquet-Lagrèze, P. Perny, R. Slowinsky, D. Vanderpooten, and P. Vincke, editors, *Aiding Decisions with Multiple Criteria: Essays in Honour of Bernard Roy*, pages 229–246. 2001. 115

[Roy68] B. Roy. Classement et choix en présence de points de vue multiples (la méthode ELECTRE). *Revue française d'informatique et de recherche operationelle (RIRO)*, 2:57–75, 1968. 7, 8, 41, 46, 48

[Roy85] B. Roy. *Méthodologie multicritère d'aide à la décision*. Ed. Economica, collection Gestion, 1985. 9, 10, 12, 36, 41, 46, 47, 48

[Roy87a] B. Roy. Des critères multiples en recherche opérationnelle: pourquoi? *Document du LAMSADE*, (80), 1987. 7

[Roy87b] B. Roy. Meaning and validity of interactive procedures as tools for decision making. *European Journal of Operational Research*, 31(3):297–303, 1987. 19

[Roy90] B. Roy. The outranking approach and the foundations of ELECTRE methods. In C.A Bana e Costa, editor, *Readings in Multiple Criteria Decision Aid*, pages 155–183. Springer-Verlag, Berlin, 1990. 11, 12

[Roy93] B. Roy. Decision science or decision-aid science? *European Journal of Operational Research*, 66:184–203, 1993. 8

[RV88] M. Roubens and P. Vincke. Fuzzy possibility graphs and their application to ranking fuzzy numbers. In M. Roubens and J. Kacprzyk, editors, *Non-Conventional Preference Relations in Decision Making*, pages 119–128. Springer-Verlag, Berlin, 1988. 121, 122

[RV96] B. Roy and D. Vanderpooten. The European school of MCDA: Emergence, basic features and current works. *Journal of Multi-Criteria Decision Analysis*, 5:22–38, 1996. 46

[Sam38] P.A. Samuelson. A note on the pure theory of consumer's behavior. *Economica*, 5(17):61 – 71, 1938. 7

[Sch88] G. Schaffer. Savage revisited. In D. Bell, H. Raiffa, and A. Tversky, editors, *Decision Making: descriptive, normative and prescriptive interactions*, pages 193–235. Cambridge University Press, Cambridge, 1988. 8

[Sch89] D. Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, 57(3):571–587, 1989. 89

[Sha53] L. S. Shapley. A value for n-person games. In *Contributions to the theory of games, vol. 2*, Annals of Mathematics Studies, no. 28, pages 307–317. Princeton University Press, Princeton, N. J., 1953. 90

[Sim57] H.A. Simon. A behavioural model of rational choice. In H.A. Simon, editor, *Models of man: social and rational; mathematical essays on rational human behavior in a social setting*, pages 241–260. J. Wiley, New York, 1957. 7

[Sim79] H.A. Simon. Rational decision making in business organisations. *American Economic Review*, 69:493–513, 1979. 15

[SL03] T. Stewart and F.B. Losa. Towards reconciling outranking and value measurement practice. *European Journal of Operational Research*, 145:645–659, 2003. 8, 16, 21

[Sug74] M. Sugeno. *Theory of fuzzy integrals and its applications*. PhD thesis, Tokyo Institute of Technology, Tokyo, Japan, 1974. 87, 88

[TW04] B.A. Turlach and A. Weingessel. `quadprog`*: Functions to solve quadratic programming problems.*, 2004. R package version 1.4-7. 130

[vE86] D. von Winterfeldt and W. Edwards. *Decision Analysis and Behavorial Research*. Cambridge University Press, Cambridge, 1986. 14, 15, 19, 20, 86

[Vin92] P. Vincke. *Multicriteria Decision-aid*. Wiley, 1992. 14, 15, 87, 102

[vM44] J. von Neumann and O. Morgenstern. *Theory of games and economic behavior*. Princeton University Press, Princeton, 1944. Second edition in 1947, third in 1954. 7, 8

[VV89] D. Vanderpooten and Ph. Vincke. Description and analysis of some representative interactive multicriteria procedures. *Mathematical and Computer Modelling*, 12:1221–1238, 1989. 18

[Wak89] P.P. Wakker. *Additive Representations of Preferences: A new Foundation of Decision Analysis.* Kluwer Academic Publishers, Dordrecht, Boston, London, 1989. 8, 105

[Win84] W. Windelband. Beiträge zur Lehre vom negativen Urteil. In *Straßburger Abhandlungen zur Philosophie: Eduard Zeller zu seinem 70. Geburtstage*, pages 165–195. Strasbourg, 1884. 34

[YWHL05] R. Yang, Z. Wang, P-A. Heng, and K-S. Leung. Fuzzy numbers and fuzzification of the Choquet integral. *Fuzzy Sets and Systems*, 153:95–113, 2005. 99, 100

[Zad65] L.A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965. 94, 95, 96

[Zad78] L.A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1:3–28, 1978. 95

Index of margin notes