Chapter 1

CHOICE, RANKING AND SORTING IN FUZZY MULTIPLE CRITERIA DECISION AID

Patrick Meyer, Marc Roubens
Department of Mathematics
University of Liège
Grande Traverse, 12
4000 Liège
Belgium
patrick.meyer@internet.lu, m.roubens@ulg.ac.be

Abstract In this chapter we survey several approaches to derive a recommendation from some preference models for multiple criteria decision aid. Depending on the specificities of the decision problem, the recommendation can be a selection of the best alternatives, a ranking of these alternatives or a sorting. 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.

Keywords: Aggregation with fuzzy environment, Fuzzy choice, Ordinal ordered sorting, Choquet Integral, TOMASO

1. Introduction

Let $A = \{\ldots, x, y, \ldots\}$ be a finite set of potential alternatives, and $\mathcal{J}$ be a set of n points of view. The Multiple Criteria Decision Problem can often be formulated as comparing and/or discriminating between the alternatives on the basis of several points of view.

As clearly stated by B.Roy in his book on Multicriteria Methodology [34] , Multiple Criteria Decision Aid is an activity that creates models to provide the decision maker (DM) with guidelines with respect to his decision problem. Three basic problems are usually put forward:

- the *choice* problem that aims to select a subset of potential alternatives, as restricted as possible, containing the "satisfactory" actions,
- the *sorting* problem that corresponds to the assignment of each alternative into pre-defined categories. These categories correspond to a set M of classes. If M is just a set of labels we talk about a classification problem. If the labels of M can be ordered, we are dealing with an ordered sorting,
- the *ordering* problem that aims at ranking the alternatives by decreasing order of preference. The prescription may be given in terms of a partial or a complete order.

A first step in the Decision Aiding Process consists in the evaluation of the alternatives on each of the points of view and is possibly followed by the definition of a valued preference relation R_j on A for each dimension $j \in \mathcal{J}$.

A second step consists in either determining a global ranking on the alternatives, a sorting into different classes, or a choice function which results in a subset of alternatives of A. Two different procedures can be used: the pre-ranking methods and the pre-aggregation methods.

The *pre-ranking methods* first determine a score $S(x, R_j) = S_j(x)$ for each alternative $x \in A$ and each point of view $j \in \mathcal{J}$. An aggregation rule M_v then transforms those partial scores into a global score $S(x; R_1, \ldots, R_n)$, where v represents weights linked to the points of view. v is either a vector $(v(1), \ldots, v(n))$ or a monotone set function $v : 2^{\mathcal{J}} \to [0, 1]$ fulfilling $v(\emptyset) = 0$ and $v(\mathcal{J}) = 1$. This procedure will be used in the TOMASO method which deals with ordinal data and interacting points of view. An ordered sorting is obtained and all alternatives are comparable.

The *pre-aggregation methods* first determine a global binary relation R on A using an aggregation rule $M_v : R = M_v(R_1, \ldots, R_n)$. Comparisons of partial evaluations are performed dimension by dimension and their results are then aggregated. Usually this relation is constructed so as to reflect the majoritarian preference among the set of points of view. This approach allows a fine and flexible description of preferences without forcing arbitrarily alternatives to be comparable and allows to take into account not only concordance between pairs of alternatives but also discordance. A global score $S'(x, R)$ transforms the global information on each pair of alternatives into a global rating related to each alternative.

However a global partial order on the alternatives might be obtained if top-down or bottom-up procedures are considered (as the combination of in and out-flows in PROMETHEE [3] and the intersection of direct and inverse complete preorders in ELECTRE II [30]).

This chapter is built around three main subjects. First of all, a general description of the different ways to deal with a multiple criteria decision problem is proposed. In Section 2 we describe the different types of data one may encounter. Section 3 presents the concepts of valued preference relation and outranking relation. Section 4 describes the two possibilities for aggregation: pre-aggregation and pre-scoring. Section 5 deals with the particular multiple criteria decision aiding problematic called the sorting. This is done in view of Section 6. There we focus on a particular sorting procedure called TOMASO . It is a multiple criteria sorting procedure for the assignment of alternatives to ordered classes based on a pre-ranking method. The alternatives are evaluated on different interacting points of view using performance levels (scores). The objective is to aggregate these partial evaluations by the Choquet integral. The basic technique we present is due to Roubens [27]. An evolution to this method is explicited, in case the basic procedure has no solution. The fuzzy measures associated to the Choquet integral can be learnt from a subset of alternatives (called prototypes) which are assigned beforehand to the classes by the DM. This leads in a first stage to solving a linear constraint satisfaction problem whose unknown variables are the coefficients of the fuzzy measure. If a fuzzy measure is found, the boundaries of the classes are calculated, and the alternatives are classified. If no solution is found to this problem, an alternate way is suggested, which can lead to ambiguous assignments of the prototypes.

Both results can be analysed by means of the importance indexes and the interaction indexes of the assessed fuzzy measure. These two parameters give the following indications on the fuzzy measure:

- the importance indexes make it possible to appraise the overall importance of each point of view and each combination of points of view;
- the interaction indexes measure the extent to which the points of view interact (positively or negatively).

Finally, in Section 7, we focus on a choice procedure for the selection of a set of "good" alternatives that includes a fuzzy approach based on a pre-aggregation method. It can be considered as a substitute to the ELECTRE IS ([32], [19]) method or a complement to its prescriptions. The chapter finishes on some conclusions and perspectives.

2. The Data Set

Without any loss of generality, we will suppose hereafter that the higher an evaluation of an alternative on a point of view, the better the alternative is in the eyes of the decision maker.

For each point of view $j \in \mathcal{J}$, the evaluation related to each alternative is possibly given under one of the following forms:

- an *ordinal value* g_j defined on a s_j-point performance scale, that is a totally ordered set $X_j := \{g_1^j \prec_j \ldots \prec_j g_{s_j}^j\}$. It usually corresponds to linguistic ordered data.

- a *fuzzy ordinal value*, i.e. a membership function $\mu_j(u) \in [0,1]$, $\forall u \in X_j$. The degree of membership can be interpreted as the degree of compatibility of the evaluation with u. The fuzzy set is supposed to be normal ($\sup_u \mu_j(u) = 1$) and convex ($\forall u, v, w \in X_j, v \in [u,w], \mu_j(v) \leq \min\{\mu_j(u), \mu_j(w)\}$).

- a *cardinal value* g_j that associates the alternative with a real number indicating its performance. This is the most conventional way of building a preference model and in that case we are talking about a true-criterion.

- a *fuzzy interval*, i.e. a membership function $\mu_j(u) \in [0,1], \forall u \in \mathbb{R}$ that is supposed to be normal and convex. Every λ-cut is a closed interval $I_j^\lambda = \{u : \mu_j(u) \geq \lambda\}$.

  A particular example of a fuzzy interval corresponds to a *trapezoidal fuzzy number* defined by the parameters $(g_j^-, g_j^+, \sigma_j^-, \sigma_j^+)$:

$$\mu_j(u) = \begin{cases} 1 - \frac{g_j^- - u}{\sigma_j^-} \text{ if } g_j^- - \sigma_j^- \leq u \leq g_j^- \\ 1 \text{ if } g_j^- \leq u \leq g_j^+ \\ 1 - \frac{u - g_j^+}{\sigma_j^+} \text{ if } g_j^+ \leq u \leq g_j^+ + \sigma_j^+ \end{cases}$$

  This may correspond to imprecise information on the evaluation of a given alternative: it lies possibly in the support ($g_j^- - \sigma_j^- \leq u \leq g_j^+ + \sigma_j^+$) and belongs certainly to the kernel ($g_j^- \leq u \leq g_j^+$).

  A symmetric trapezoidal fuzzy number is such that $\sigma_j^- = \sigma_j^+$ and may translate the indifferences and preferences that might exist between values that are assessed to an alternative. In that situation we call

$$\begin{cases} g_j^+ - g_j^- = q_j \text{ (indifference threshold)} \\ g_j^+ + \sigma_j^+ - (g_j^- - \sigma_j^-) = p_j \text{ (preference threshold)} \\ \frac{g_j^+ + g_j^-}{2} = g_j \end{cases}$$

These definitions are interesting as a help to understand the concepts of indifference and preference thresholds. All the values between $g_j - \frac{1}{2}q_j$ and $g_j + \frac{1}{2}q_j$ are considered as indifferent. Values greater than $g_j + \frac{1}{2}p_j$ are better than g_j and those lower than $g_j - \frac{1}{2}p_j$ are worse than g_j. Even in the case of complete and precise information, a small positive difference does not always justify the preference.

3. Valued preference relation and outranking relation

Now that we have described the different possible evaluations in Section 2 the goal of this section is to recall the concepts of valued preference relation and outranking relation. We define the degree to which an alternative x is not worse than y for point of view j. Let $R_j(x,y)$ be this degree, for each ordered pair (x,y) of alternatives. We use the same notations as in Section 2 for the different possible evaluations.

Similarly to the different possibilities described in Section 2, the degree $R_j(x,y)$ has different definitions and properties:

- for an *ordinal* or *cardinal value* g_j:

  $$R_j(x,y) = \begin{cases} 1 \text{ if } g_j(x) \geq g_j(y) \\ 0 \text{ otherwise.} \end{cases}$$

  This crisp binary relation is a linear quasiorder.

- for a *fuzzy ordinal value*, $R_j(x,y)$ defines the degree of the preference of x over y and is considered as the possibility that x is not worse than y:

  $$\begin{aligned} R_j(x,y) = \Pi_j(x \geq y) &= \max\nolimits_{u \geq v} \min(\mu_j^x(u), \mu_j^x(v)), u, v \in X_j \\ &= \max\nolimits_u \min(\mu_j^x(u), \mu_j^y(u)), u \in X_j. \end{aligned}$$

  Π_j is a valued binary relation such that $\max(\Pi_j(x,y), \Pi_j(y,x)) = 1, \forall x, y \in A$. Roubens and Vincke [28] have proved that Π_j is a fuzzy interval order and every λ-cut is a crisp interval order.

- for *fuzzy intervals*, $R_j(x, y)$ is also defined as the possibility that x is not worse than y:

$$R_j(x, y) = \Pi_j(x \geq y) = \max_{u \geq v} \min(\mu_j^x(u), \mu_j^y(v)), u, v \in \mathbb{R}.$$

If the kernel of μ_j^x is located to the right of the kernel of μ_j^y, then $\Pi_j(x \geq y) = 1$ and $\Pi_j(y \geq x)$ equals the height of the intersection of μ_j^x and μ_j^y, $h_j(x, y)$ (see figure 1.1). This valued binary relation presents the same properties as the fuzzy ordinal value.

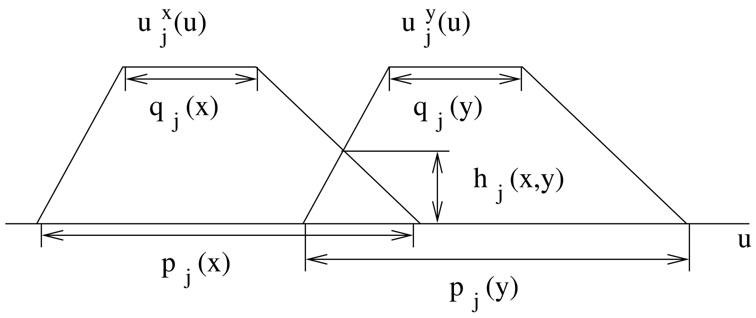

Figure 1.1. Comparing two fuzzy intervals

Starting from the credibility of the preference of x over y it is possible to define [6] [7]:

- the degree of strict preference of x over y as the necessity that x is strictly better than y:

$$P_j(x, y) = 1 - \Pi_j(y \geq x) = 1 - R_j(y, x)$$

- the degree of indifference between x and y as:

$$I_j(x, y) = \min(R_j(x, y), R_j(y, x))$$

- for of a *symmetric trapezoidal number*:

$$R_j(x, y) = \frac{\frac{p_j(x)+p_j(y)}{2} - \min\{g_j(y) - g_j(x), \frac{p_j(x)+p_j(y)}{2}\}}{\frac{p_j(x)+p_j(y)}{2} - \min\{g_j(y) - g_j(x), \frac{q_j(x)+q_j(y)}{2}\}},$$

where $\frac{0}{0}$ should be taken as 0.

If p_j and q_j are linear functions of g_j, then R_j is a fuzzy semiorder and every λ-cut is a crisp semiorder [6] [7].

If $p_j = q_j$, then we obtain a crisp interval order. Let us define $g'_j(x) = g_j(x) - \frac{q_j(x)}{2}$. We then have:

- for the strict preference:

$$\begin{aligned} xP_jy &\iff R_j(x,y)=1, R_j(y,x)=0 \\ &\iff g'_j(x) - g'_j(y) > q_j(y) \end{aligned}$$

- for the indifference:

$$\begin{aligned} xI_jy &\iff R_j(x,y)=R_j(y,x)=1 \\ &\iff |g'_j(x) - g'_j(y)| \leq \min(q_j(x), q_j(y)) \end{aligned}$$

An extra condition of local consistency should be added [34]:

$$g'_j(x) > g'_j(y) \Rightarrow g'_j(x) + q_j(x) \geq g'_j(y) + q_j(y)$$

The criterion function g'_j and the threshold function q_j define a semi-criterion and the structure (P_j, I_j) is a semiorder.

The classical procedures Electre III [31] and Promethee [3] are using this approach based on the intersection of fuzzy sets.

According to Perny [26], the degree of preference of x over y may be considered in very general terms as

$$R_j(x,y) = f_j[g_j(x), \mathcal{N}g_j(y)]$$

where f_j is a non-decreasing function of both arguments, $\mathcal{N}$ is a strong negation and $R_j(x,x) = 1$. Perny proved that such a valued preference relation is a fuzzy semiorder and every λ-cut constitutes a crisp semiorder [26]. As a particular case, we have the concordance index defined by Roy [31]:

$$R_j(x,y) = \min\{1, \max\{0, \frac{g_j(x) - g_j(y) + p_j(g_j(x))}{p_j(g_j(x)) - q_j(g_j(x))}\}\}$$

where p_j and q_j are non-decreasing functions of g_j and correspond respectively to a preference threshold and an indifference threshold. For consistency reasons, $p_j(g_j(x)) \geq q_j(g_j(x))$. The concordance index R_j is meaningful (i.e. is invariant under admissible transformations of g_j) if g_j is defined on an interval scale (admissible transformations are $h_j = r \cdot g_j + s, r > 0$). q_j corresponds to a constant or a proportion of g_j and p_j is expressed as a proportion of g_j [31].

Similarly, according to Perny, we may also define a degree of discredit as

$$D_j(x,y) = h_j[g_j(y), \mathcal{N}g_j(x)]$$

where h_j is a non decreasing function of both arguments, $D_j(x,x) = 0$ and $\min\{R_j(x,y), D_j(x,y)\} = 0$. Under these conditions, D_j is a fuzzy partial order and every λ-cut represents a crisp partial order. As previously, we can consider the particular case of the discordance index defined by Roy [31]:

$$D_j(x,y) = \min\{1, \max\{0, \frac{g_j(y) - g_j(x) - p_j(g_j(x))}{v_j(g_j(x)) - p_j(g_j(x))}\}\},$$

where v_j corresponds to a veto threshold which expresses the existence of a discordant point of view that prohibits to accept the idea that x is globally preferred to y.

4. Aggregation procedures

4.1 Pre-aggregation methods

Let us first consider the methods that propose to merge the marginal information about each pair of alternatives (x,y) in terms of concordance (and possibly discordance) indexes into a global relation that expresses the overall importance of the consensus on the fact that "x is globally not worse than y".

Roy [34] introduces an outranking relation $\mathcal{O}(x,y)$ that corresponds to the "agreement versus discordance" measure linked to the proposition that x is globaly not worse than y. It indicates the importance of the coalition of the points of view that agree with the proposition by taking also into account the discordance.

In general, if v_j represents the relative importance of each point of view j, $(j \in \mathcal{J}, |\mathcal{J}| = n)$, we may consider two aggregation operators M_R and M_D such that

$$R = M_R(R_1, \ldots, R_n; v_1, \ldots, v_n)$$

$$D = M_D(D_1, \ldots, D_n; v_1, \ldots, v_n)$$

M_R is a monotonic function of the first arguments such that $M_R(0, \ldots, 0; v_1, \ldots, v_n) = 0$ and $M_R(1, \ldots, 1; v_1, \ldots, v_n) = 1$. M_D is a monotonic function of the first arguments that should satisfy:

$$(\exists j, j \in \mathcal{J} : D_j(x,y) = 1) \Rightarrow D(x,y) = 1$$

stating that if at least one point of view is totally discordant with the proposition that x is not worse than y, the global discordance should be maximal for that specific pair of alternatives.

We could consider the following approach:

- for R the compensative idempotent operator (weighted sum)

$$R(x,y) = \sum_{j\in\mathcal{J}} v_j R_j(x,y), \sum_{j\in\mathcal{J}} v_j = 1$$

- for $1-D$ the non-discordance index (geometric mean)

$$1 - D(x,y) = \Pi_{j\in\mathcal{J}}(1 - D_j(x,y))^{v_j}.$$

R measures the overall importance of the agreement and D allows to give a bad rating as soon as one important partial evaluation of the discordance is achieved.

Finally the outranking relation is obtained as a combination of concordant and discordant aspects as:

$$\mathcal{O}(x,y) = R(x,y) \cdot [1 - D(x,y)].$$

Roy on his side considered in ELECTRE III:

- for R the compensative weighted sum operator
- for the outranking degree

$$\begin{aligned}\mathcal{O}(x,y) &= R(x,y) \text{ if } D_j(x,y) \leq R(x,y) \text{ for all } j \in \mathcal{J},\\ &= R(x,y)\Pi_{j\in\mathcal{J}(x,y)} \tfrac{1-D_j(x,y)}{1-R(x,y)} \text{ otherwise,}\end{aligned}$$

  where $\mathcal{J}(x,y)$ corresponds to the subset of points of view for which $D_j(x,y) > R_j(x,y)$.

In this case, the outranking degree is thus equal to the concordance index if no point of view is discordant, or if no veto is used and is lowered if the level of discordance $(1 - D_j)$ increases above a threshold value.

Most of the existing proposals linked to pre-aggregation methods simply merge the marginal information related to the agreement on the proposal that x is globally not worse than y. They are thus directly linked to the concordance measures $R_j(x,y)$. The subjectivity of the decision maker with respect to the importance of each of the points of view can be used in different ways to obtain a global compromise. We consider here three of these approaches.

- the weighted sum (good items compensate bad ones with respect to different points of view):

$$R = \sum_{j\in\mathcal{J}} v_j R_j, \sum_{j\in\mathcal{J}} v_j = 1$$

- the weighted minimum (the outranking value is high if the partial evaluations are favorable on each of the points of view)

$$R = \min_{j \in \mathcal{J}} \max(1 - v_j, R_j), \max_{j \in \mathcal{J}} v_j = 1$$

- the weighted maximum (the outranking value is high if at least one of the points of view presents a good evaluation)

$$R = \max_{j \in \mathcal{J}} \min(v_j, R_j), \max_{j \in \mathcal{J}} v_j = 1$$

.

Weighted maximum and minimum can be interpreted as weighted medians (see [5]). The interested reader can refer to [12], [6] and [7] for a more elaborate list of aggregators.

In the case of a *choice* problem the outranking relations $\mathcal{O}$ (initially with crisp outranking relations and later with λ-cuts of the valued outranking relations) were exploited by Roy using the kernel concept (internally stable and dominating subset of A) in Electre I ([29], [19]) and later in Electre IS ([19], [32]). The idea is to track the maximal circuits to transform them into indifference cliques or suppress these circuits by eliminating the less credible outrankings.

Another approach was proposed by Orlovski [22]. He considers the fuzzy set of non dominated elements over A as

$$\mathrm{ND}(x) = \min_{y \in A \setminus \{x\}} \neg P(y, x), \quad \forall x \in A$$

where $P(y, x)$ corresponds to the degree of strict preference associated to $R(y, x)$ (see [6], [7], [24]). The rational choice corresponds to these alternatives giving the maximal value of ND:

$$A^{\mathrm{ND}} := \{x \in A : \max_{x \in A} \mathrm{ND}(x)\}$$

Under certain sufficient conditions (transitivity of R) this subset corresponds to maximal values equal to one; such good alternatives are called *unfuzzy non dominated alternatives* (UND-alternatives) and the corresponding rational choice is

$$A^{\mathrm{UND}} := \{x \in A : P(y, x) = 0 \quad \forall y \in A\}.$$

In Section 7 we consider the case where the valued relations R are ordinal values defined on a discrete finite set L (L-valued binary relations) and we determine the choice set as a kernel with a maximum degree of credibility.

In the case of a *sorting* problem the outranking relations $\mathcal{O}$ are used in procedures where a decision tree is used or by filtering as in ELECTRE TRI ([19], [33]). These procedures use a cutting procedure that transforms the fuzzy outranking relations into a sequence of crisp and nested outranking relations.

In the case of an *ordering* problem the outranking relations $\mathcal{O}$ are used to construct two complete pre-orders, one arising from a ascending distillation procedure and another constructed from a descending distillation procedure. Another prescription consists in the intersection of the two previous pre-orders. These exploitation procedures are described in ELECTRE III ([31], [19]).

4.2 Pre-scoring methods

In this type of approach, the implicit assumption that there exists a complete and transitive comparability of the alternatives is made. The most typical example of such methods corresponds to an ordering or a sorting that is based on the weighted sum of some partial scores $S_i(x)$ $(x \in A)$. The additive representation of the utilities (expressed in terms of the partial scores) however implies preferential independance of the utilities.

One way to avoid this independance condition is to use the Choquet integral [4] as an aggregator.

Let us consider an alternative x which is described by its partial scores vector $S(x) = (S_1(x), \ldots, S_n(x))$. The Choquet integral of x is then defined by:

$$\mathcal{C}_v(S(x)) := \sum_{i=1}^{n} S_{(i)}(x)[v(A_{(i)}) - v(A_{(i+1)})]$$

where v represents a fuzzy measure on $\mathcal{J}$, that is a monotone set function $v : 2^{\mathcal{J}} \to [0,1]$ fulfilling $v(\emptyset) = 0$ and $v(\mathcal{J}) = 1$. This fuzzy measure merely expresses the importance of each subset of points of view. The parentheses used for indexes represent a permutation on $\mathcal{J}$ such that

$$S_{(1)}(x) \leq \ldots \leq S_{(n)}(x),$$

and $A_{(i)}$ represents the subset $\{(i), \ldots, (n)\}$.

We note that for additive measures ($v(S \cup T) = v(S) + v(T)$, whenever $S \cap T = \emptyset$) the Choquet integral coincides with the usual discrete Lebesgue integral and the set function v is simply determined by the importance of each point of view: $v(1), \ldots, v(n)$. In this particular case

$$\mathcal{C}_v(S(x)) = \sum_{i=1}^{n} v(i) S_i(x) \ \ (x \in A),$$

which is the natural extension of the Borda score as defined in voting theory if alternatives play the role of candidates and points of view represent voters.

If points of view cannot be considered as being independent, the importance of the combinations $S \subseteq \mathcal{J}$, namely $v(S)$ has to be taken into account.

Some combinations of points of view might present a positive interaction or *synergy*. Although the importance of some points of view, members of a combination S, might be low, the importance of a pair, a triple, ..., might be substantially larger and $v(S) > \sum_{i \in S} v(i)$.

In other situations, points of view might exhibit negative interaction or *redundancy*. The union of some points of view do not have much impact on the decision and for such combinations S, $v(S) < \sum_{i \in S} v(i)$.

The Choquet integral presents standard properties for aggregation [15], [39],[18]: it is continuous, non decreasing, located between min and max.

The major advantage linked to the use of the Choquet integral derives from the large number of parameters $(2^n - 2)$ associated with a fuzzy measure. On the other hand, this flexibility can also be considered as a serious drawback when assessing real values to the importance of all possible combinations. We will come back to this important question in Section 6.

Let us present an equivalent definition of the Choquet integral. Let v be a fuzzy measure on $\mathcal{J}$. The Möbius transform of v is a set function $m : 2^{\mathcal{J}} \to \mathbb{R}$ defined by

$$m(S) = \sum_{T \subseteq S} (-1)^{|S|-|T|} v(T) \ \ (S \subseteq \mathcal{J}).$$

This transformation is invertible and thus constitutes an equivalent form of a fuzzy measure and v can be recovered from m by using

$$v(S) = \sum_{T \subseteq S} m(T) \ \ (S \subseteq N).$$

This transformation can be used to redefine the Choquet integral without reordering the partial scores:

$$\mathcal{C}_v(S(x)) = \sum_{T \subseteq \mathcal{J}} m(T) \bigwedge_{i \in T} S_i(x).$$

A fuzzy measure is k-additive ([10]) if its Möbius transform m satisfies $m(S) = 0$ for S such that $|S| > k$ and there exists at least one subest S such that $|S| = k$ and $m(S) \neq \emptyset$.

Thus, k-additive fuzzy measures can be represented by at most $\sum_{j=1}^{k} \binom{n}{j}$ coefficients.

For a k-additive fuzzy measure,

$$\mathcal{C}_v(S(x)) = \sum_{\substack{T\subseteq\mathcal{J}\\|T|\leq k}} m(T) \bigwedge_{j\in T} S_j(x).$$

In order to assure boundary and monotonicity conditions imposed on v, the Möbius transform of a k-additive fuzzy measure must satisfy:

$$\begin{cases} m(\emptyset) = 0, \quad \sum\limits_{\substack{T\subseteq\mathcal{J}\\|T|\leq k}} m(T) = 1 \\ \sum\limits_{\substack{T\,:\,i\in T\subseteq S\\|T|\leq k}} m(T) \geq 0, \quad \forall S \subseteq \mathcal{J},\ \forall j \in S \end{cases}$$

In Section 6 we present a sorting method using the Choquet integral and based on supervised learning. But first let us introduce some general considerations on the problematic of sorting alternatives.

5. The sorting problem

Let A be a set of q potential alternatives which are to be assigned to disjoint ordered classes. Let $F = \{g_1, \ldots, g_n\}$ be a set of points of view. For each index of point of view $j \in \mathcal{J} = \{1, \ldots, n\}$, the alternatives are evaluated according to a s_j-point ordinal performance scale represented by a totally ordered set

$$X_j := \{g_1^j \prec_j \ldots \prec_j g_{s_j}^j\}.$$

Therefore, an alternative $x \in A$ can be identified with its corresponding profile

$$(x_1, \ldots, x_n) \in \prod_{j=1}^{n} X_j =: X,$$

where for any $j \in \mathcal{J}$, x_j is the partial evaluation of x on point of view j.

Let us consider a partition of X into m nonempty increasingly ordered classes $\{Cl_t\}_{t=1}^m$. This means that 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.

The sorting problem we are dealing with consists in partitioning the alternatives of A into the classes $\{Cl_t\}_{t=1}^m$.

In Greco *et al.* [11], a very general theorem states that, under a simple condition of monotonicity, a discriminant function can be found which strictly separates the classes $\{Cl_t\}_{t=1}^m$ by thresholds. In Roubens [27] a restriction to the class of n-place Choquet integrals and normalised scores as criteria functions is made. Hereafter we present the sorting procedure derived from this particular case.

6. The TOMASO method

The TOMASO method (**T**echnique for **O**rdinal **M**ulti**a**ttribute **S**orting and **O**rdering) is mainly based on two techniques (which can lead to the same results under certain conditions). The original method has first been described in [27]. In the following Subsection, we present its basics. In Subsection 6.2 we show how it is possible to deal with a larger set of problems.

6.1 The classical way

The different stages of the original TOMASO are listed below:

1 Modification of the criteria evaluations into scores;

2 Use of a Choquet integral as a discriminant function;

3 Assessment of fuzzy measures by questionning the DM and by solving a linear constraint satisfaction problem;

4 Calculation of the borders of the classes and assignment of the alternatives to the classes;

5 Analysis of the results (interaction, importance, leave one out, visualisation).

In this Section we roughly present these different elements.

One of the most difficult tasks is to modify the original ordinal evaluations of the alternatives on the criteria into some "scores" which can be aggregated by means of a Choquet integral. For example, two ordinal scales X_j and X_k can have a distinct number of evaluation levels and very different intrinsic meanings. The transformations of the scales should take into account these possible characteristics in order to obtain comparable evaluations. Two natural possibilities appear: the scores are built on basis of the data which are to be analysed or the scores are constructed completely out of the context of the problem. In the first case, the scores are solely based on the information which is contained in the set of alternatives which are considered. In the second case, the scales $X_j, j \in \mathcal{J}$ are modified in a general way, without taking into account the particular structure of the analysed set of data. At the present stage of our research, we suggest three possible alternatives for the building of these evaluation scores. In each of the cases, the DM must be aware of the consequences of his choice. Therefore, a deep analysis of the problem is important for its complete understanding.

First of all, in case the problem can be resumed to the set A of potential alternatives and if the DM is a single person, then one possible way to build the scores is to consider pairwise comparisons of the alternatives on each of the points of view. For each point of view $j \in \mathcal{J}$, the order on X_j ($\preceq_j$) can be caracterised by a valuation $R_j : A \times A \to \{0, 1\}$ defined by $R_j(x, y) = 1$ if $x_j \succeq y_j$, 0 otherwise. Starting from this valuation we define a *partial net score* $S_j : A \to \mathbb{R}$ by

$$S_j(x) := \sum_{y \in A} [R_j(x, y) - R_j(y, x)] \quad (x \in A, j \in \mathcal{J}). \tag{1.1}$$

The interpretation of the integer $S_j(x)$ is natural: it represents the number of times that x is preferred to any other alternative of A minus the number of times that any other alternative of A is preferred to x for point of view j. One can show that the partial net scores identify the corresponding partial evaluations. We furthermore normalise these scores so that they range in the unit interval. The highest partial net score which can be obtained corresponds to the following general case:

- one single alternative $x_{\max} \in A$ has $\mathrm{ord}_j(x_{\max}) = i$;
- no alternative $x \in A$ has $\mathrm{ord}_j(x) > i$;
- the remaining alternatives $x \in A \backslash \{x_{\max}\}$ have $\mathrm{ord}_j(x) < i$.

Therefore, the highest possible partial net score is $S_{j,\max}(x) = q - 1$. Similarly, the lowest possible partial net score is $S_{j,\min}(x) = -(q - 1)$. We can therefore write the normalised partial net scores S_j^N as follows:

$$S_j^N(x) := \frac{S_j(x) + (q - 1)}{2(q - 1)} \in [0, 1] \quad \forall j \in \mathcal{J}, \forall x \in A. \tag{1.2}$$

On contrary of the original ordinal partial evaluations, the partial net scores (and the normalised partial net scores) are commensurable. During the whole the chapter we will use the notation $S^N(x) := (S_1^N(x), \ldots, S_n^N(x))$.

Two important questions now arise: how can this choice be motivated, and how can it be interpreted? First of all, the DM must understand that the selection of the set of potential alternatives A will have an influence on the final result. Therefore this choice must be made with much care. Then, his way of thinking must be a comparison of the alternatives on each of the points of view. Let us consider a short example which clearly illustrates this way of obtaining the scores. Suppose that we have to deal with a sorting problem with two qualitative ordinal criteria on a set of cars. The first point of view C_1 expresses the degree of comfort of the alternatives and is evaluated on a 3-point ordinal scale $X_1 = \{\text{Bad} \prec_1 \text{Medium} \prec_1 \text{Good}\}$. The second one

C_2 expresses the fuel consumption of the cars on a 3-point ordinal scale $X_2 = \{\text{High} \prec_2 \text{Normal} \prec_2 \text{Low}\}$. The set of potential alternatives consists in 6 cars. The DM is aware that the results will depend on these 6 alternatives, but he considers that they have been chosen in a right way (for example, they are the only possible cars that he can afford with his tight budget). One can then assume that the absolute value of an alternative on a point of view is not informative, unless considered in relation with the other elements of A. We summarise this short example in table 1.1. It shows the distribution of the alternatives among the different evaluation levels of the two points of view. If one reasons according to the

Table 1.1. Number of alternatives per evaluation level

C_1		C_2	
Good	4	Low	2
Medium	1	Normal	2
Bad	1	High	2

comparison philosophy, it appears clearly that it is less exceptional to be "Good" than to be "Low". In fact, there are many good cars, but fewer cars with a low fuel consumption. Similarly, being "Bad" is worse than being "High". This means that having a high fuel consumption is less exceptional than being an uncomfortable car. The scores, as defined earlier, reflect these properties. They are given in table 1.2. This exam-

Table 1.2. Score of each of the evaluation levels

C_1		C_2	
Good	7/10	Low	9/10
Medium	2/10	Normal	5/10
Bad	0/10	High	1/10

ple shows that it is not senseless to modelise the DM's way of thinking by these scores. Three conditions should be satisfied: the decisions must be taken by a single DM, the set of potential alternatives must be chosen carefully and the DM should evaluate the alternatives by comparisons.

Secondly, let us consider the cases where multiple DMs intervene or where the decisions are not taken according to the previously described comparison philosophy. Here, the scoring functions are built "out of the context". This means that the values given to each of the evaluation levels of the ordinal scales don't depend on the set A. If the DM cannot help us with the building of such scores, we can approximate these discrete utility functions by the following formula:

$$S_j^N(x) := \frac{\mathrm{ord}_j(x) - 1}{s_j - 1} \in [0,1] \quad \forall j \in \mathcal{J}, \forall x \in A.$$

where $\mathrm{ord}_j : A \to \{1, \ldots, s_j\}$ is a mapping defined by $\mathrm{ord}_j(x) = r \iff x_j = g_r^j$. $S_j^N(x)$ does not represent a real utility and probably does not correspond to the utility the DM has in mind. We therefore continue to call it a score.

Finally, we would like to point out a particular situation, where the DM considers that any possible alternative which can be built out of the evaluation scales is a potential alternative. In this case, A equals the set of all possible alternatives which can be built from the sets $X_j, j \in \mathcal{J}$, i.e. $A = \prod_{j=1}^n X_i$. The partial net score formula 1.1 then becomes

$$S_j(x) = q(\frac{2\mathrm{ord}_j(x) - 1}{s_j} - 1) \quad (x \in X, j \in \mathcal{J}), \tag{1.3}$$

These partial net scores are normalised according to the formula 1.2.

We now come to the crucial part of the aggregation of the normalised partial net scores of a given alternative x by means of a Choquet integral [4]. The advantage of this aggregator is mainly that it allows to deal with interacting (depending) points of view. According to what has been said in Section 4:

$$\mathcal{C}_v(S^N(x)) := \sum_{j=1}^{n} S_{(j)}^N(x)[v(A_{(j)}) - v(A_{(j+1)})]$$

where v is a fuzzy measure on $\mathcal{J}$; that is a monotone set function $v : 2^{\mathcal{J}} \to [0,1]$ fulfilling $v(\emptyset) = 0$ and $v(\mathcal{J}) = 1$. The parentheses used for indexes stand for a permutation on $\mathcal{J}$ such that

$$S_{(1)}^N(x) \leq \ldots \leq S_{(n)}^N(x),$$

and for any $j \in \mathcal{J}$, $A_{(j)}$ represents the subset $\{(j), \ldots, (n)\}$. The characterisation of the Choquet integral by Marichal ([13], [15]) clearly justifies the way the partial scores are aggregated.

The next step of this method is to assess the fuzzy measures in order to classify the alternatives of A. One can easily understand that it is impossible to ask the DM to give values for the $2^n - 2$ free parameters of the fuzzy measure v. Practically, the assessment of the fuzzy measures is done by asking the DM to provide a set of prototypes $P \subseteq A$ and their assignments to the given classes; that is a partition of P into prototypic classes $\{P_t\}_{t=1}^m$ where $P_t := P \cap Cl_t$ for $t \in \{1, \ldots, m\}$. The values of the fuzzy measure are then derived from this information as described hereafter.

We would like the Choquet integral to strictly separate the classes Cl_t. Therefore, the following necessary condition is imposed

$$\mathcal{C}_v(S^N(x)) - \mathcal{C}_v(S^N(x')) \geq \varepsilon \tag{1.4}$$

for each ordered pair $(x, x') \in P_t \times P_{t-1}$ and each $t \in \{2, \ldots, m\}$, where ε is a given strictly positive thershold.

Due to the increasing monotonicity of the Choquet integral, the number of separation constraints 1.4 can be reduced significantly. Thus, it is enough to consider *border elements* of the classes. To formalise this concept, we first define a dominance relation D (partial order) on X by

$$xDy \iff x_j \succeq_j y_j, \text{ for all } j \in \mathcal{J}.$$

As *upper border* of the prototypic class P_t we use the set of non-dominated alternatives of P_t defined by

$$ND_t := \{x \in P_t \text{ such that } \nexists x' \in P_t \setminus \{x\} : x'Dx\}.$$

Similarly, the *lower border* of the prototypic class is given by the set of non-dominating alternatives of P_t which is defined by

$$Nd_t := \{x \in P_t \text{ such that } \nexists x' \in P_t \setminus \{x\} : xDx'\}.$$

The separation conditions restricted to the prototypes of the subsets $ND_t \cup Nd_t$, $t \in \{1, \ldots, m\}$ put together with the monotonicity constraints on the fuzzy measure, form a linear program [17] whose unknowns are the capacities $v(S)$, $S \subset \mathcal{J}$ and where ε is a non-negative variable to be maximised in order to deliver well separated classes.

We use the principle of parsimony for the resolution of this problem. If there exists a k-additive fuzzy measure v^*, k being kept as low as possible, then we determine the boundaries of the classes as follows:

- lower boundary of Cl_t: $z(t) := \min_{x \in Nd_t} \mathcal{C}_{v^*}(S^N(x))$;
- upper boundary of Cl_t: $Z(t) := \max_{x \in ND_t} \mathcal{C}_{v^*}(S^N(x))$.

At this point, any alternative $x \in A$ can be classified in the following way:

- x is assigned to class Cl_t if $z_t \leq \mathcal{C}_{v^*}(S^N(x)) \leq Z_t$;
- x is assigned to class $Cl_t \cup Cl_{t-1}$ if $Z_{t-1} < \mathcal{C}_{v^*}(S^N(x)) < z_t$.

A final step of the classical TOMASO method concerns the evaluation of the results and the interpretation of the behavior of the Choquet integral. The meaning of the values $v(T)$ is not clear to the DM. They don't immediatly indicate the global importance of the points of view, nor their degree of interaction. It is possible to derive some indexes from the fuzzy measure which are helpful to interpret its behavior. Among them, the TOMASO method proposes to have a closer look at the importance indexes [36] and the interaction indexes [22]. We present the calculation of these indexes in Section 6.3.

6.2 An alternate way

It may happen that the linear program described in Subsection 6.1 has no solution. This occurs when the prototypic elements violate the axioms that are imposed to produce a discriminant function of Choquet type ([15] [39]), in particular the triple cancellation axiom.

In such a case, and in order to present a solution to the DM, we suggest to find a fuzzy measure by solving the following quadratic program

$$\min \sum_{x \in \cup_{t \in \{1,\ldots,m\}} \{ND_t \cup ND_t\}} [\mathcal{C}_v(S^N(x)) - y(x)]^2,$$

where the unkowns are

- the capacities $v(S)$ which determine the fuzzy measure;
- some global evaluations $y(x)$ for each $x \in \cup_{t \in \{1,\ldots,m\}} \{ND_t \cup Nd_t\}$.

The capacities $v(S)$ are constrained by the monotonicity conditions (as previously shown in Section 6.1). The global evaluations $y(x)$ must verify the classification imposed by the DM. In other words, for every ordered pair $(x, x') \in Nd_t \times ND_{t-1}$, $t \in \{2, \ldots, m\}$ the condition $y(x) - y(x') \geq \varepsilon'$, $\varepsilon' > 0$ must be satisfied.

Intuitively, for a given alternative $x \in A$, its Choquet integral $\mathcal{C}_v(S^N(x))$ should be as close as possible to the global evaluation $y(x)$, without being constrained by monotonicity conditions which might violate the triple cancellation axiom for example. On the other hand, the evaluation $y(x)$ is constrained by the these conditions derived from the original classification given by the DM on the prototypes.

Unlike the method described in Section 6, in this case, ε' plays the role of a parameter, which needs to be fixed by the DM. As previously, we use the principle of parsimony when searching for a solution (keep k as low as possible; at worst k equals the number of points of view). A correct choice of ε' remains one of the main challenges of our future research. It is clear that ε' has to be chosen in $]0, 1/n[$.

As in the classical method, the next step is to determine the structure of the classes. We determine an assignment for every alternative of X in terms of intervals of contiguous classes on the basis of the information provided by the Choquet integrals related to the prototypes of $P \subseteq A$.

First of all, let us suppose that $S^N(x^-) := (0, \ldots, 0)$ is classified to the worst class, Cl_1 and that $S^N(x^+) := (1, \ldots, 1)$ is classified to the best class, Cl_m.

To each assignment $I(x)$ correspond a lower class label $\underline{l}(x)$ and an upper class label $\bar{l}(x)$, $\underline{l}, \bar{l} \in \mathcal{J}$. We say that the alternative $x \in X$ is *precisely assigned* to $Cl_{l(x)}$ if for the assignment $I(x)$ we have $\underline{l}(x) = \bar{l}(x) =: l(x)$. Else, the alternative x is said to be *ambiguously assigned* to the interval of labels $I(x) = [\underline{l}(x), \bar{l}(x)]$. The *degree of the assignment* corresponds to the number of contiguous classes contained in $I(x)$, $d(x) = \bar{l}(x) - \underline{l}(x) + 1$.

The assignments are done according to the procedure described hereafter. Starting from the prototypes $x \in P$, their Choquet integrals $\mathcal{C}_v(S^N(x))$ and their original classification label $Cl(x)$ (according to the DM's choice), we define for every $u \in [0, 1]$,

$$m(u) = \max_{x \in P : \mathcal{C}_v(S^N(x)) \le u} Cl(x), \text{and}$$

$$M(u) = \min_{x \in P : \mathcal{C}_v(S^N(x)) \ge u} Cl(x).$$

m (resp. M) is a right (resp. left) continuous stepwise function of argument u with values belonging to the discrete finite set $\mathcal{J}$.

We now define for each $u \in [0, 1]$ an interval of contiguous classes $I(u) = [\underline{l}(u), \bar{l}(u)]$ where

$$\underline{l}(u) = \min\{m(u), M(u)\}$$

$$\bar{l}(u) = \max\{m(u), M(u)\}.$$

Obviously $\underline{l}(u) \le \bar{l}(u)$ and due to monotonicity of m and M we have: $\underline{l}(u) \le \underline{l}(v), \bar{l}(u) \le \bar{l}(v), \forall u, v \in [0, 1]$ with $u \le v$.

The interval $[0, 1]$ is partitioned into (closed, semi-open or open) intervals $I_s, s = 1, \ldots, S$, and each of those intervals of $[0, 1]$ receives an assignment of the type $[\underline{l}(s), \bar{l}(s)]$ (or semi-open or open) in such a way that: if $u, v \in [0, 1], u \le v$ and if u is assigned to $I_r := [\underline{l}(r), \bar{l}(r)]$ and v is assigned to $I_{r'} := [\underline{l}(r'), \bar{l}(r')]$ then $\underline{l}(r) \le \underline{l}(r')$ and $\bar{l}(r) \le \bar{l}(r')$.

Moreover if $u = \mathcal{C}_v(S^N(x)), x \in P$ then $\underline{l}(u) \le Cl(x) \le \bar{l}(u)$. This means that each prototype is *correctly classified*, possibly with ambiguity if $d(x) \ge 1$.

The assignment of a prototype a to the intervals of classes leads now to two scenarios:

- a is assigned to a single class (interval of length 0) which corresponds to the original class decided by the DM
- a is assigned to an interval of classes and the original class decided by the DM belongs to this interval.

The quality of a model (classifier) depends on different ratios. A good model has the following *natural* properties:

- a simple model according to parsimony (low k);
- a high number of precise assignments of the elements of P;
- a low number of ambiguous assignments of the elements of P (and the lower the degree of the assignment, the better the model)

For a given ε', the DM has to select a model (k) which seems the best compromise to him in terms of the previously described assignments. The simplest additive model ($k = 1$) can in certain situations be this ideal compromise between simplicity and quality. But in more complex problems, k has to be increased in order to obtain a satisfying number of precisely assigned prototypes.

The next Section briefly presents some indexes (importance, interaction) which give indications on the behaviour of the fuzzy measure.

6.3 Behavioral analysis of aggregation

Now that we have a sorting model for assigning alternatives to classes (based on the linear program or the quadratic program), an important question arises: How can we interpret the behavior of the Choquet integral or that of its associated fuzzy measure? Of course the meaning of the values $v(T)$ is not always clear for the DM. These values do not give immediately the global importance of the points of view, nor the degree of interaction among them.

In fact, from a given fuzzy measure, it is possible to derive some indexes or parameters that will enable us to interpret the behavior of the fuzzy measure. These indexes constitute a kind of *id card* of the fuzzy measure. In this Section, we present two types of indexes: importance and interaction. Other indexes, such as tolerance and dispersion, were proposed and studied by Marichal [13, 16].

6.3.1 Importance indexes. The overall importance of a point of view $j \in \mathcal{J}$ in a decision problem is not solely determined by the value of $v(\{j\})$, but also by all $v(T)$ such that $j \in T$. Indeed, we may have $v(\{j\}) = 0$, suggesting a priori that element j is unimportant, but it may happen that for many subsets $T \subseteq \mathcal{J}$, $v(T \cup \{j\})$ is much greater than $v(T)$, suggesting that j is actually an important element in the decision.

Shapley [36] proposed in 1953 a definition of a coefficient of importance, based on a set of reasonable axioms. The *importance index* or *Shapley value* of point of view j with respect to v is defined by:

$$\phi(v, \{j\}) := \sum_{T \subseteq \mathcal{J} \setminus \{j\}} \frac{(n - |T| - 1)!\,|T|!}{n!}\,[v(T \cup \{j\}) - v(T)]. \qquad (1.5)$$

This index is a fundamental concept in game theory and it expresses a power index. It can be interpreted as a weighted average value of the marginal contribution $v(T \cup \{j\}) - v(T)$ of element j alone in all combinations. To make this clearer, we rewrite the index as follows:

$$\phi(v, \{j\}) = \frac{1}{n} \sum_{t=0}^{n-1} \frac{1}{\binom{n-1}{t}} \sum_{\substack{T \subseteq \mathcal{J} \setminus \{j\} \\ |T| = t}} [v(T \cup \{j\}) - v(T)].$$

Thus, the average value of $v(T \cup \{j\}) - v(T)$ is computed first over the subsets of same size t and then over all the possible sizes. Consequently, the subsets containing about $n/2$ points of view are the less important in the average, since they are numerous and a same point of view j is very often involved into them.

The use of the Shapley value in multicriteria decision making was proposed in 1992 by Murofushi [22]. It is worth noting that a basic property of the Shapley value is

$$\sum_{j=1}^{n} \phi(v, \{j\}) = 1.$$

Note also that, when v is additive, we clearly have $v(T \cup \{j\}) - v(T) = v(\{j\})$ for all $j \in \mathcal{J}$ and all $T \subseteq \mathcal{J} \setminus \{j\}$, and hence

$$\phi(v, \{j\}) = v(\{j\}), \quad j \in \mathcal{J}. \qquad (1.6)$$

If v is non-additive then some points of view are dependent and (1.6) generally does not hold anymore. This shows that it is useful to search for a coefficient of overall importance for each point of view.

6.3.2 Interaction indexes. A further interesting concept is that of *interaction* among points of view. We have seen that when the fuzzy measure is not additive then some points of view interact. Of course, it would be interesting to appraise the degree of interaction among any subset of points of view.

Consider first a pair $\{i,j\} \subseteq \mathcal{J}$ of points of view. It may happen that $v(\{i\})$ and $v(\{j\})$ are small and at the same time $v(\{i,j\})$ is large. The Shapley index $\phi(v,\{j\})$ merely measures the average contribution that point of view j brings to all possible combinations, but it gives no information on the phenomena of interaction existing among points of view.

Clearly, if the marginal contribution of j to every combination of points of view that contains i is greater (resp. less) than the marginal contribution of j to the same combination when i is excluded, the expression

$$[v(T \cup \{i,j\}) - v(T \cup \{i\})] - [v(T \cup \{j\}) - v(T)]$$

is positive (resp. negative) for any $T \subseteq \mathcal{J} \setminus \{i,j\}$. We then say that i and j positively (resp. negatively) interact.

This latter expression is called the *marginal interaction* between i and j, conditioned to the presence of elements of the combination $T \subseteq \mathcal{J} \setminus \{i,j\}$. Now, an interaction index for $\{i,j\}$ is given by an average value of this marginal interaction. Murofushi and Soneda [22] proposed in 1993 to calculate this average value as for the Shapley value. Setting

$$(\Delta_{ij}\, v)(T) := v(T \cup \{i,j\}) - v(T \cup \{i\}) - v(T \cup \{j\}) + v(T),$$

the *interaction index* of points of view i and j related to v is then defined by

$$I(v,\{i,j\}) := \sum_{T \subseteq N \setminus \{i,j\}} \frac{(n-|T|-2)!\,|T|!}{(n-1)!}\,(\Delta_{ij}\, v)(T). \qquad (1.7)$$

It should be mentioned that, historically, the interaction index (1.7) was first introduced in 1972 by Owen (see Eq. (28) in [25]) in game theory to express a degree of complementarity or competitiveness between elements i and j.

6.4 Interpretation of the behaviour of the fuzzy measure

In this Section we briefly show the main advantage to use a Choquet integral rather than the weighted sum as a discriminant function. We therefore take the simple case of two points of view, which can be represented in a plane. Figure 1.2 presents 5 possible ranges of values for the weights v and the corresponding structures of the limits of the classes. One can see that the main difference between the classical weighted sum and the Choquet integral is the greater flexibility of the borders of the classes. The Choquet integral creates piecewise linear borders, which allows to build more precise classes. The different possibilities are summarised by the following list:

- I: $v(1) + v(2) < v(12)$: synergy
- II: $v(1) + v(2) > v(12)$: redundancy
- III: $v(1) + v(2) = v(12) = 1$: additivity
- IV: $v(1) = v(2) = 0$: limit case; maximal synergy
- V: $v(1) = v(2) = 1$: limit case; maximal redundancy

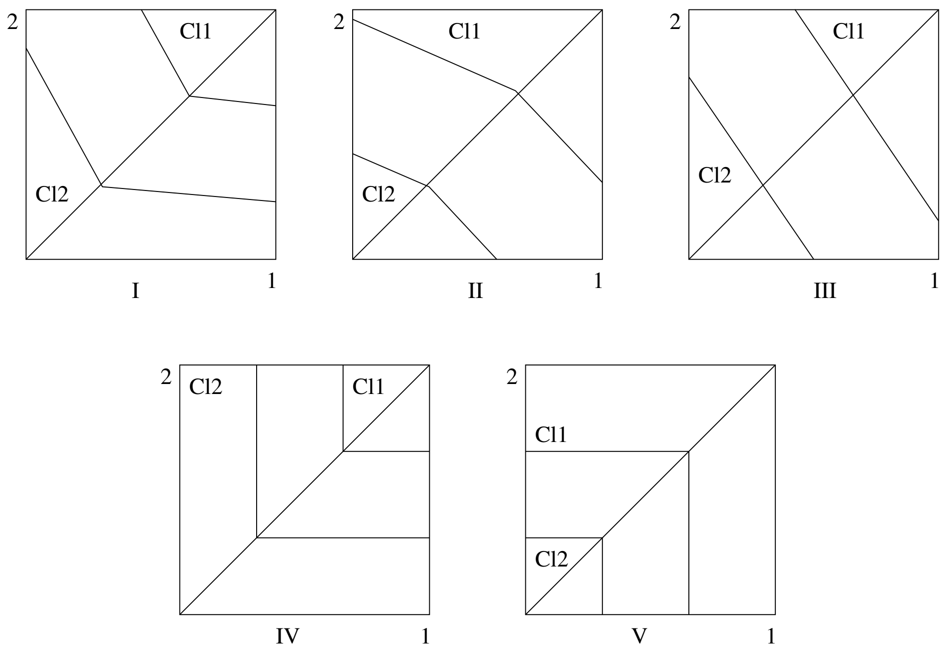

Figure 1.2. Interpretation of the discriminant functions

In [8] the authors give an interpretation to the first two cases. In case of synergy, although the importance of a single criterion for the decision is rather low, the importance of the pair is large. The criteria are said to be *complementary*. In case of redundance, or negative synergy, the union of criteria does not bring much, and the importance of the pair might be roughly the same as the importance of a single criterion.

The limit case (IV) occurs for maximal synergy. In that case, the Choquet integral corresponds to the aggregation by the min function.

Maximal redundancy occurs for case (V), where the Choquet integral is the max function.

In case the number of points of view is larger than two, it becomes quite hard to represent the problem. Nevertheless, the previous short example helps to understand how the borders of the classes are built in such more general examples.

6.5 The software TOMASO

In this short part of the chapter we briefly present the key characteristics of the software TOMASO . It can be downloaded on http://patrickmeyer.tripod.com. It is an implementation of the algorithms which were presented previously. Its name stands for "Tool for Ordinal MultiAttribute Sorting and Ordering". It is written in Visual Basic and uses two external solvers: a free linear program solver (lp_solve 3.0, ftp://ftp.ics.ele.tue.nl/pub/lp_solve/, released under the LGPL license), and a non free quadratic program solver (bpmpd, free trial version at http://www.sztaki.hu/ meszaros/bpmpd/).

It is still under development and many improvements are added on a regular basis. The general steps of the software are outlined hereafter:

- Loading of the ordinal data;
- Choice of a scoring method according to the problem's specificities and calculation of the normalised partial net scores;
- Definition of the prototypes by the DM;
- Search for a fuzzy measure (either by the linear program, or the quadratic program);
- Analysis of the results (classes, Shapley indexes, interaction indexes, accuracies, ...).

A detailed description of the software can be obtained from the authors.

6.6 Testing the method on two problems

In this Section, we apply the previously presented method on two particular problems. The following part describes briefly the two problems. Then they are analysed by the TOMASO method.

6.6.1 Description of the problems.

The students problem This small example clearly illustrates the procedure when no solution can be found to the linear program. We consider a set of 8 students evaluated on 2 courses (C1, C2). For each matter, the evaluation scale has 10 ordered qualitative levels (1-10). In total, this makes 100 possible different ratings. Besides, for each student, the DM has given a global evaluation on a 6-levelled qualitative ordinal scale (the classes): (very good (6)> good (5)> above average (4)> below average (3)> bad (2)> very bad(1)). A summary of the problem is given in table 1.3.

Table 1.3. Profiles of the students

student	profile	class
A	$(7,5)$	2
B	$(6,6)$	1
C	$(7,7)$	3
D	$(6,8)$	4
A'	$(10,7)$	6
B'	$(8,8)$	5
C'	$(10,5)$	3
D'	$(8,6)$	4

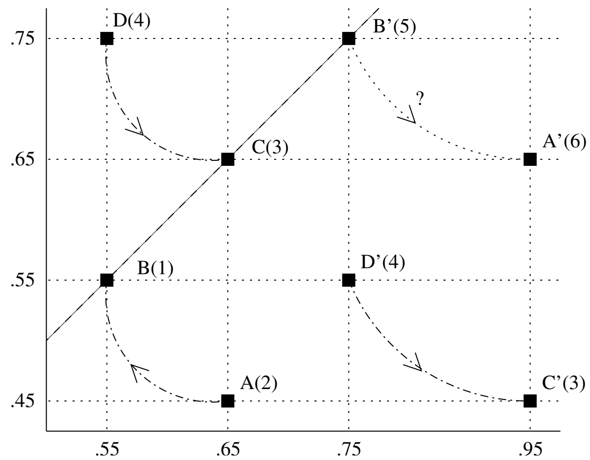

Figure 1.3. Representation of the students problem

The noise annoyance problem This real-life example concerns noise annoyance caused by different sources. Details on these data can be found in [2] and [37]. It was obtained by a survey performed on 2661 persons (alternatives). They were asked to give an estimation of their annoyance level (not at all annoyed (n) $\preceq$ slightly annoyed (s) $\preceq$ moderately annoyed (m) $\preceq$ very annoyed (v) $\preceq$ extremely annoyed(e)) on 21 different potential noise sources (points of view) (noise annoyance caused by road traffic, by rail traffic, ...). This ordering and the exact wording of the questions is in accordance with international standards. Besides, the questionned persons had to give an overall noise annoyance level on the same scale.

The original dataset contains 2661 alternatives and 21 points of view. But unfortunately, its structure is not proper for the TOMASO method as it contains a lot of inconsistencies (aDb but a is in a worse class than b). For the purpose of this chapter, we restrict ourselves to a consistent subset of 155 alternatives and 6 points of view (road traffic (cars, busses,...), air traffic, truck loading and unloading, factories, dance halls, agricultural equipment).

The goal is therefore to find a Choquet integral as a discriminant function which can reproduce the overall noise annoyance level by using the separate noise annoyances as an input.

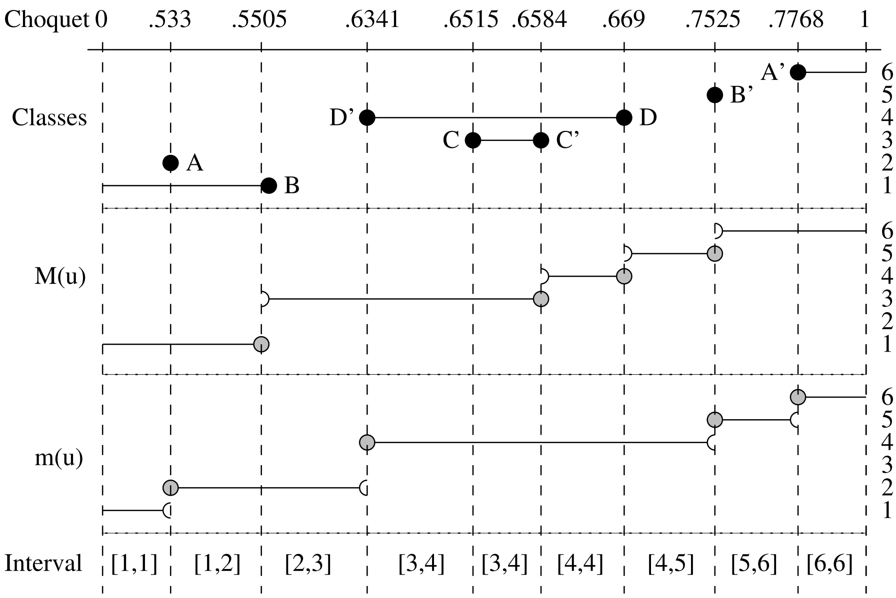

Figure 1.4. Classes for $\varepsilon = .1, k = 1$

6.6.2 Solving the students problem. For this problem, the scales on the points of view are quite rich, but we only have a few alternatives. We therefore suppose that information can be extracted from X. This means that any possible student which can be built out of the evaluation scales on C_1 and C_2 is a potential alternative. A representation of the 2-dimensional problem is given on figure 1.3. It helps to understand why the linear program has no solution. If the triple cancellation property [39] is violated, there exists no Choquet integral which satisfies the constraints imposed by the classification of the prototypes. If triple cancellation was verified in this example, we would have:

$$\left\{ \begin{array}{c} \textsc{Class}(B) \leq \textsc{Class}(A) \\ \& \\ \textsc{Class}(D) \geq \textsc{Class}(C) \\ \& \\ \textsc{Class}(D') \geq \textsc{Class}(C') \end{array} \right\} \Rightarrow \textsc{Class}(B') \geq \textsc{Class}(A'),$$

where $\textsc{Class}(X)$ stands for the index of the class to which X belongs (the higher the better). But in this particular example, we clearly have $\textsc{Class}(D') \leq \textsc{Class}(C')$. Therefore, no solution can be found to the linear program. In other words, this problem cannot be described by the classical TOMASO method by means of a Choquet integral. We therefore go over to the method based on the quadratic probgram.

In this case, a solution can be found for different values of ε and k. For $k = 1$, the best solution is found with $\varepsilon' = 0.01$. 3 out of 8 alternatives are precisely assigned to their classes. The other 5 elements are ambiguously assigned with a degree 2.

Figure 1.4 explains how the assignment described in 6.2 to the intervals of classes works for this particular simplest model ($k = 1$). The original classes, which are shown by means of the 8 prototypes look somewhat chaotic. Both functions $m(u)$ and $M(u)$ help to build the final ordered intervals of classes. It shows that the alternatives A', B' and D are assigned precisely.

A better, but more complex model can be found for $k = 2$. A good solution can be found for $\varepsilon' = .10$. 6 out of 8 alternatives are assigned precisely, whereas 2 out of 8 students (C' and D') are assigned ambiguously with degree 2. This shows the better performance of a Choquet integral over the weighted sum aggregator. A representation of the solution in this case ($k = 2, \varepsilon = .10$) is given on figure 1.5. We observe that the borders of the classes are piecewise linear, and that this allows to cope with a larger set of problems. We can also observe the overlapping zone between classes 3 and 4, which induce the ambiguous assignments of C' and D'.

To conclude this example, in table 1.4 we present the importance indexes for this example, in both models. We can see that the values

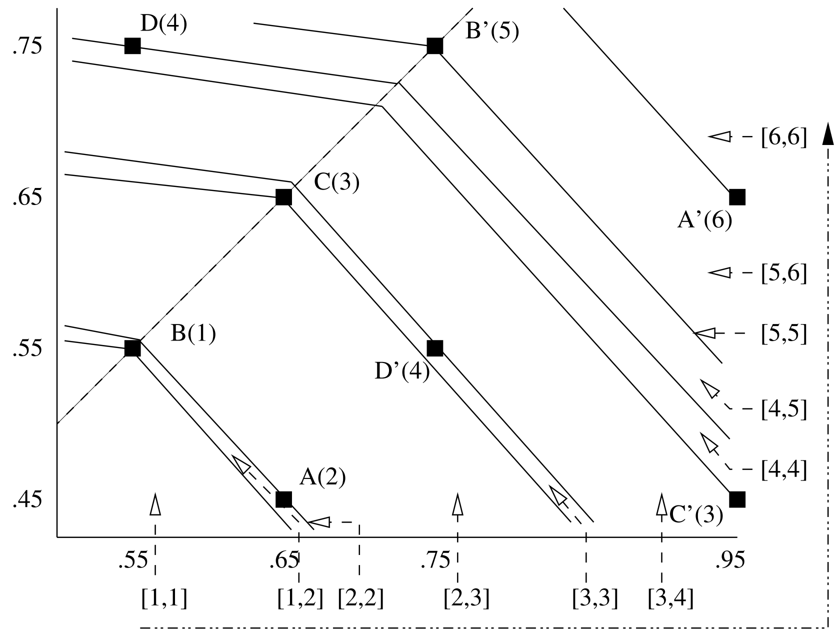

Figure 1.5. Classes for $\varepsilon = .1, k = 1$

are quite identical; the order on the importance of the criteria is clearly respected.

Table 1.4. Importance indexes for the students problem.

Model	Course 1	Course 2
$k = 1$	.414	.586
$k = 2$	.325	.675

6.6.3 Solving the noise annoyance problem. Let us get back to the second problem described in 6.6.1. The data set of 155 prototypes violates the triple cancellation axiom. Therefore, as no solution can be found to the linear program, we switch to the resolution of the quadratic program.

This problem is not adapted for the comparison philosophy for the scores. This is clearly not a decision problem with a single DM. In fact, each of the 155 decisions has been taken by a separate person. One of these persons could not compare his profile to the other ones, before giving a global noise annoyance level. Furthermore, as we cannot ask these 155 people to give us hints on the shape of the discrete utility functions linked to the evaluation scales, we have no other option than considering formula 1.2.

Let us first start with the best possible solution that we can find, for $k = 6$, which means a non-additive fuzzy measure. Quite similar solutions exist for values of ε' between 0.05 and 0.18. We chose an average value of $\varepsilon' = 0.1$.

Table 1.5. Global accuracy

$\varepsilon' = 0.1, k = 6$	
precise	ambiguous $d = 2$
46.45%	53.55%

We can also analyse the accuracy of this discriminant function for each class separately. Table 1.6 shows its performance for each of the 5 ordered classes.

Table 1.6. Per class accuracy: precise assignments

$\varepsilon' = 0.1, k = 6$				
extremely	very	moderately	slightly	not at all
45.83%	2.33%	66.67%	69.05%	71.43%

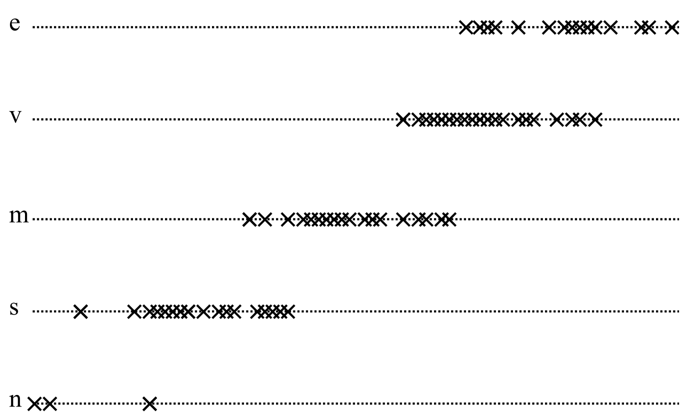

Figure 1.6. Visual representation of the classes, $k = 6$, $\varepsilon' = 0.1$

We can see that the class *very annoyed* is nearly unpredictable in a precise way with this discriminant function. This is due to the fact that it overlaps strongly with the classes *extremely annoyed* and *moderately annoyed*. This phenomenon can be observed on figure 1.6. It represents the assignments to the 5 ordered classes. A cross can represent more than one alternative (if they have equal Choquet integrals).

The assignments of the alternatives to the classes according to their Choquet integral is shown in table 1.7.

Table 1.7. Assignments of the elements to intervals of classes

$\varepsilon' = 0.1, k = 6$	
Interval of Choquets	Interval of classes
$[0, 0.0375[$	$[1, 1]$
$[0.0375, 0.1011]$	$[1, 2]$
$]0.1011, 0.1903[$	$[2, 2]$
$[0.1903, 0.2182]$	$[2, 3]$
$]0.2182, 0.3139[$	$[3, 3]$
$[0.3139, 0.3642[$	$[3, 4]$
$[0.3642, 0.3754]$	$[4, 4]$
$]0.3754, 0.4565]$	$[4, 5]$
$]0.4565, 1]$	$[5, 5]$

The importance indexes of the 6 criteria are given in table 1.8. Taking

Table 1.8. Importance indexes

$\varepsilon' = 0.1, k = 6$					
street	air	truck	factory	dancing	agriculture
.166	.158	.164	.0938	.272	.146

into account our scoring method, we can state that the noise annoyance caused by dancing halls is the most important, followed by street noises, truck loading, air traffic disturbances, agricultural annoyances and finally factory noises.

One should note that the global results of table 1.5 are not too bad. They should not be misinterpreted: there are no erroneous classifications, but only ambiguous ones. In order to obtain a classification of the prototypes into single classes, we suggest to use a l-nearest neighbourhood algorithm to force an assignment. Each prototype $a \in P$ and its Choquet integral is presented to the remaining set $P \backslash \{a\}$ of elements and their Choquet integral. The l closest neighbours of a in terms of the Choquet integral are then selected. Among these l elements, we search for the original class (as decides by the DM) which appears most often. In case of identity, the class is chosen randomly among the equally present classes. a is then assigned to this majority class. A global accuracy and a weighted accuracy are then computed. The global accuracy is simply the ratio of correctly assigned alternatives over the total number of alternatives. The weighted accuracy is the average of the separate accuracies of each class.

Figure 1.7 shows these accuracies for different values of l (let us notice here that the axis for the separate class accuracies is on the left side of the figure, and the axis for both the global and the weighted accuracies is on the right side of the figure). Let us make a few observations. Classes 1 (not at all annoyed) and 5 (extremely annoyed) only contain a few alternatives. A consequence is that if we select l too high, no alternative will be assigned to one of these two classes. As a consequence, the weighted accuracy strongly depends on the right choice of l. On figure 1.7 one can see this influence for both of these classes. As an example, above a value of $l = 11$, the accuracy of class 1 is equal to 0. The choice of l remains a critical one on which the resulting accuracies strongly depend. Table 1.9

Table 1.9. Global and weighted accuracy in %

$\varepsilon' = 0.1, k = 6$		
l	Global	Weighted
1	79.35	81.12
2	72.25	71.00
3	72.90	72.09
4	72.90	71.36
5	74.19	71.91
6	74.84	72.69
7	77.42	74.98
8	76.13	71.23
9	73.55	60.21
10	78.71	69.10
11	78.71	75.54
12	78.71	75.90
average	75.81	72.26

These results can be compared to those obtained by the methods described in [37], [38] and [2]. On this same data set of 155 alternatives and 6 points of view, with a genetic optimization of a Choquet integral with a possibility measure (1-maxitive [20]), their performance is 76.77% for the global accuracy and 80.57% for the weighted accuracy. Globally, the results are comparable. But unfortunately, we have worse results on the weighted accuracy. This is due to the nonuniform distribution of the alternatives among the 5 ordered classes, which is a big disadvantage for the use of the l-nearest neighbour method for the forced classification.

We would like to point out that very similar results (accuracies, shapley indexes) can be obtained with a 3-additive fuzzy measure ($k = 3$). For $k < 3$, the discriminating power of the Choquet integral is quite low for this particular example. In particular, the number of precisely assigned alternatives (before the l-NN procedure) becomes very low. This means that the classes overlap a lot.

7. The Choice Problem

In this Section we consider a way to select a subset of alternatives to consider as a good choice.

Consider a binary relation R whose credibility is evaluated as follows:

$R(x, y) = C_v[R_1(x, y), \ldots, R_i(x, y), \ldots, R_n(x, y)] \in [0, 1]$, for all $x, y \in A$. In the sequel we will only use the ordering of $R(x, y)$ and not their cardinality and we will obtain a L-valued binary relation R (see [1]).

For all $x, y \in A$, $R(x, y)$ belongs to a finite set $L = \{c_0 = 0, c_1, \ldots, c_m = .5, \ldots, c_{2m} = 1$ that constitutes a $(2m+1)$-element chain $c_0 \prec \ldots \prec c_{2m}$. $R(x, y)$ may be understood as the credibility that "x is at least as good as y". The set L is built using the values of R taking into consideration an antitone unary contradiction operator $\neg$ such that $\neg c_l = c_{2m-l}$ for $l = 0, \ldots, 2m$.

If $R(x, y)$ is one of the elements of L, then automatically, $\neg R(x, y)$ belongs to L. We call such a relation an L-valued binary relation.

We denote $L^{\succ m} := \{c_{m+1}, \ldots, c_{2m}\}$ and $L^{\prec m} := \{c_0, \ldots, c_{m-1}\}$.

If $R(x, y) \in L^{\succ m}$ we say that the proposition "$(x, y) \in R$" is L-true. If however $R(x, y) \in L^{\prec m}$, we say that the proposition is L-false. If $R(x, y) = c_m$, the median level (a fix point of the negation operator), then the proposition "$(x, y) \in R$" is L-undetermined.

In the classical case, where R is a crisp binary relation we define a digraph $G(A, R)$ with vertex set A and arc family R. A choice in $G(A, R)$ is a non-empty set Y of A.

A (dominant) kernel is a choice that is stable in G, i.e. $\forall x \neq y \in Y, (x, y) \notin R$ and dominant, i.e. $\forall x \notin Y, \exists y \in Y$ such that $(x, y) \in R$.

We now denote $G^L = G^L(A, R)$ a digraph with vertices set A and a valued arc family that corresponds to the L-valued binary relation R.

We define the level of stability qualification of subset Y of X as

$$\Delta^{\text{sta}}(Y) = \begin{cases} c_{2m} \text{ if } Y \text{ is a singleton,} \\ \\ \min_{y \neq x} \max_{x \neq y}\{\neg R(x, y)\} \text{ otherwise;} \end{cases}$$

and the level of dominance qualification of Y as

$$\Delta^{\text{dom}}(Y) = \begin{cases} c_{2m} \text{ if } Y = A, \\ \\ \min_{x \notin Y} \max_{y \in Y} R(x, y) \text{ otherwise.} \end{cases}$$

Y is considered to be an L-good choice, i.e. L-stable and L-dominant, if $\Delta^{\text{sta}}(Y) \in L^{\succ m}$ and $\Delta^{\text{dom}}(Y) \in L^{\succ m}$. Its qualification corresponds to $Q^{\text{good}}(L) = \min\{\Delta^{\text{sta}}(Y), \Delta^{\text{dom}}(Y)\}$.

We denote $C^{\text{good}}(G^L)$ the possibly empty set of L-good choices in G^L.

The determination of this set in an NP-complete problem even if, following a result of Kitainik [12], we do not have to enumerate the elements of the power set of A, but only have to consider the kernels of the corresponding crisp strict median-level cut relation $R^{\succ m}$ associated to R, i.e. $(x, y) \in R^{\succ m}$ if $R(x, y) \in L^{\succ m}$.

As the kernel in $G(X, R^{\succ m})$ is by definition a stable and dominant crisp subset of A, we consider the possibly empty set of kernels of $G^{\succ m} = G^{\succ m}(A, R^{\succ m})$ which we denote $C^{\text{good}}(G^{\succ m})$.

Kitiainik proved that

$$C^{\text{good}}(G^L) \subseteq C^{\text{good}}(G^{\succ m}).$$

The determination of crisp kernels has been extensively described in the literature (see, for example [35]) and the definition of $C^{\text{good}}(G^L)$ is reduced to the enumeration of the elements of $C^{\text{good}}(G^{\succ m})$ and the calculation of their qualification.

The decision maker might also be interested in bad choices. These choices correspond to absorbent kernels with a qualification greater than c_m. In the classical Boolean framework (see [35]) an (absorbent) kernel is a choice that is stable and absorbent, i.e. $\forall x \notin Y, \exists y \in Y$ such that $(x, y) \in R$. As $(x, y) \in R$ is equivalent to $(y, x) \in R^t$, where matrix R^t represents the transpose of matrix R, all the results obtained for dominant kernels can be immediately transposed for absorbent kernels and definitions like Δ^{bad} and Q^{bad} are obviously and straightforwardly obtained from Δ^{good} and Q^{good}.

Indeed the level of absorbance qualification of Y is defined as

$$\Delta^{\text{abs}}(Y) = \begin{cases} c_{2m} \text{ if } Y = A, \\ \\ \min_{x \notin Y} \max_{y \in Y} R(y, x) \text{ otherwise.} \end{cases}$$

In order to determine a unique rational choice (if any), we first compute dominant kernels in G^L (see [35], [1]) and determine their qualification as not being bad choices, i.e. $\neg Q^{\text{bad}}(Y)$ where $Q^{\text{bad}}(Y) = \min(\Delta^{\text{sta}}(Y), \Delta^{\text{abs}}(Y))$. The selection is based on

$$\max Q^{\text{good}}(Y)$$

If more than one candidate remain, other discriminant functions may be added as minimal absorbancy, lowest cardinality, ...

8. Conclusion

In this chapter we have presented a few approaches to multiple criteria decision aiding. In particular, we have focussed on fuzzy methods for choice, sorting and ordering. We have also described in details the sorting procedure TOMASO which can deal with interacting criteria. Some tests on examples (theoretical and real-life) have shown the interestingness of this method. Further investigations have to be done on the validation of the models. We intend to implement a cross-validation procedure to make stability tests on the data and the method.

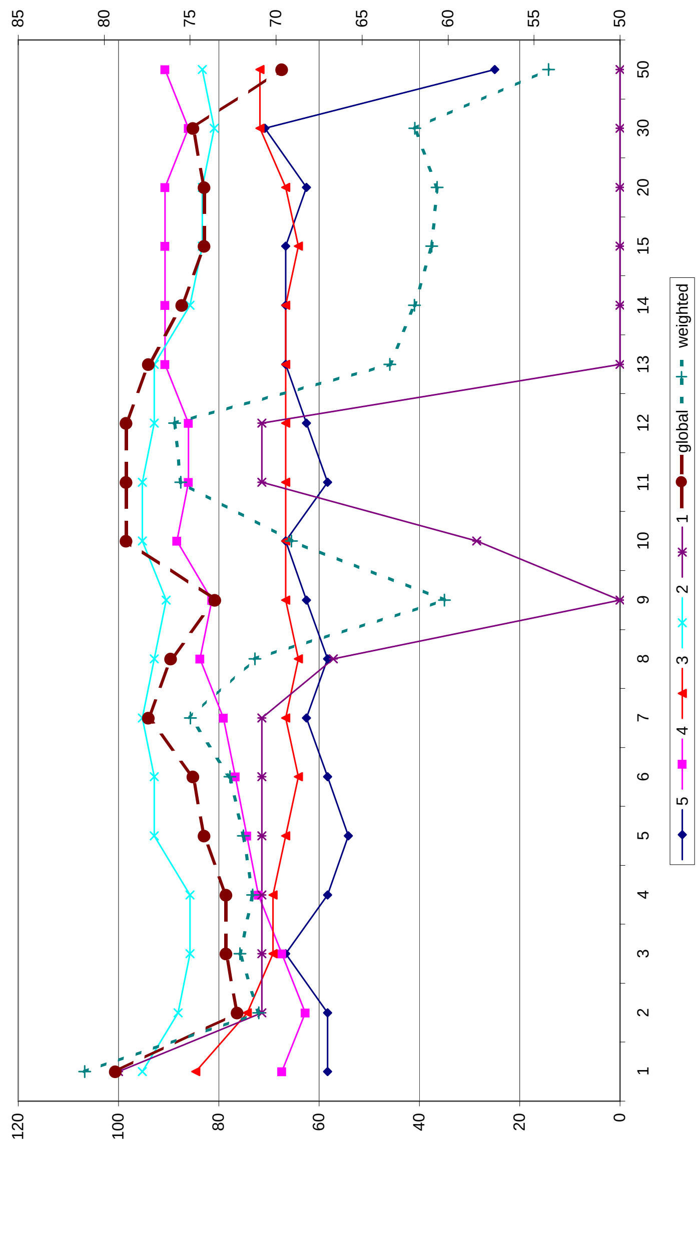

Figure 1.7. Results for the l-nearest neighbour algorithm, $\varepsilon' = 0.1$, $k = 6$

References

[1] R. Bisdorff, M. Roubens, On defining and computing fuzzy kernels and L-valued simple graphs, Intelligent Systems and Soft Computing, In: D. Ruan, P. D'hondt, P; Govaerts and E. Kerre (eds.): *Intelligent Systems and Soft Computing for Nuclear Science and Industry*, World Scientific, Singapore, (1996) 113-122.

[2] D. Botteldooren, A. Verkeyn, Fuzzy models for accumulation of reported community noise annoyance from combined sources, *Journal of the Acoustical Society of America*, 112, (2002) 1496 - 1508.

[3] J.P. Brans, P. Vincke, *A preference ranking organization method, Management Sciences*, 31, (1984) 647-656.

[4] G. Choquet, Theory of capacities, *Annales de l'Institut Fourier*, 5, (1953) 131-295.

[5] D. Dubois, H. Prade, Weighted minimum and maximum in fuzzy set theory, *Information Science*, 39, (1986) 205-210.

[6] J. Fodor, S.A. Orlovski, P. Perny, M. Roubens, The use of fuzzy preference models in multiple criteria: choice, ranking and sorting. In: D. Dubois, H. Prade (eds.): chapter 3 of volume 5 (Operations Research and Statistics), *Handbooks of Fuzzy Sets*, Kluwer Academic Publishers, Dordrecht, Boston, London, (1998) 69-101.

[7] J. Fodor, P. Perny, M. Roubens, Decision Making and Optimization, In: E. Ruspini, P. Bonissone and W. Pedrycz (eds.): Chapter 5.1 of *Handbook of Fuzzy Computation*, Institute of Physics Publ. and Oxford University Press, Bristol, (1998) F.5.1:1-14.

[8] M. Grabisch, M. Roubens, Application of the Choquet Integral in Multicriteria Decision Making, In: M. Grabisch, T. Murofushi, M. Sugeno (eds.): *Fuzzy Measures and Integrals*, Physica Verlag, Heidelberg, (2000) 348-374.

[9] M. Grabisch, J.-L. Marichal, and M. Roubens, Equivalent representations of set functions, *Mathematics of Operations Research*, 25, (2000) 157–178.

[10] M. Grabisch, k-order additive discrete fuzzy measure and their representation, *Fuzzy Sets and Systems*, 92, (1997) 167-189.

[11] 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 (eds.), *AMCDA: Aide Multicritère à la Décision (Multiple Criteria Decision Aiding)*, European Commission Report, EUR 19808 EN, Joint Research Centre, Ispra, (2001) 117-144.

[12] L. Kitainik, *Fuzzy Decision Procedures with Binary Relations. Towards a Unified Theory*, Kluwer Academic Publishers, Dordrecht, (1993).

[13] J.-L. Marichal, *Aggregation operators for multicriteria decision aid*, Ph.D. thesis, Institute of Mathematics, University of Liège, Liège, Belgium, (1998).

[14] J.-L. Marichal, The influence of variables on pseudo-Boolean functions with applications to game theory and multicriteria decision making, *Discrete Applied Mathematics*, 107, (2000) 139–164.

[15] J-L. Marichal, An axiomatic approach of the discrete Choquet integral as a tool to aggregate interacting criteria, *IEEE Transactions on Fuzzy Systems*, 8, (2000) 800-807.

[16] J.-L. Marichal, Behavioral analysis of aggregation in multicriteria decision aid. In: J. Fodor, B. De Baets, and P. Perny (eds.): *Preferences and Decisions under Incomplete Knowledge.* Series: Studies in Fuzziness and Soft Computing Vol. 51. Physica Verlag, Heidelberg, (2000) 153–178.

[17] J-L. Marichal, M. Roubens, On a sorting procedure in the presence of qualitative interacting points of view. In: J. Chojean, J. Leski (eds.):*Fuzzy Sets and their Applications.* Silesian University Press, Gliwice, (2001) 217-230.

[18] J-L. Marichal, Aggregation of interacting criteria by means of the discrete Choquet integral. In: T. Calvo, G. Mayor, R. Mesiar, (eds.), *Aggregation operators: new trends and applications.* Series: Studies in Fuzziness and Soft Computing, Vol 97. Physica-Verlag, Heidelberg, (2002) 224-244.

[19] L.Y. Maystre, J. Pictet, J. Simos, *Méthodes multicritères* ELECTRE, Presses polytechniques et universitaires romandes, Lausanne, (1994).

[20] R. Mesiar, Generalizations of k-order additive discrete fuzzy measures, *Fuzzy Sets and Systems*, 102, (1999) 423 – 428.

[21] T. Murofushi, A technique for reading fuzzy measures (I): the Shapley value with respect to a fuzzy measure, *2nd Fuzzy Workshop*, Nagaoka, Japan, (1992) 39–48. In Japanese.

[22] T. Murofushi, S. Soneda, Techniques for reading fuzzy measures (III): interaction index, *9th Fuzzy Sytem Symposium*, Sapporo, Japan, (1993) 693-696. In Japanese.

[23] S. Orlovski, Decision Making with a fuzzy preference relation, *Fuzzy Sets and Systems*, 1, (1978) 155-167.

[24] S.V. Ovchinnikov, M. Roubens, On strict preference relations, *Fuzzy Sets and Systems* 43, (1991) 319-326.

[25] G. Owen, Multilinear extensions of games, *Management Sciences*, 18, (1972) 64–79.

[26] P. Perny, Modélisation, Agrégation et exploitation de préférences floues dans une problématique de rangement: bases axiomatiques, procédures et logiciel, PhD thesis, Paris-Dauphine University and LAFORIA document 9 (1992).

[27] 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 (eds.): *Aiding Decisions with Multiple Criteria: Essays in Honour of Bernard Roy*. Kluwer Academic Publishers, Dordrecht, (2001) 229-246.

[28] M. Roubens, P. Vincke, Fuzzy preferences in an optimisation perspective. In: J. Kacprzyk and S.A. Orlovski (eds.): *Optimization Models Using Fuzzy Sets and Possibility Theory*, D. Reidel Publ. Co, Dordrecht, Boston, Lancaster, Tokyo, (1987) 77-90.

[29] B. Roy, Classement et Choix en présence de points de vue multiples (la méthode ELECTRE), *Revue d'Information et de Recherche Opérationnelle*, 8, (1968) 57–75.

[30] B. Roy, P. Bertier, La méthode Electre II. Une application au media planning. In: M. Ross (eds.): *Operational Research*, 72, North Holland Publishing Co., (1973) 291-302.

[31] B. Roy, Electre III: Un algorithme de classement fondé sur une représentation floue des préférences en présence de critères multiples, *Cahiers du CERO*, 20, 1 (1978) 3-24.

[32] B. Roy, J.-M. Skalka, Electre IS - Aspects méthodologiques et guide d'utilisation, Université de Paris-Dauphine, *Document du LAMSADE*, no 30, (1985).

[33] B. Roy, D. Bouyssou, Aide à la décision fondée sur une PAMC de type Electre, Université Paris-Dauphine, *Document du LAMSADE*, no 69, (1991).

[34] B. Roy, *Multicriteria Methodology for Decision Aiding*, Kluwer Academic Publishers, Dordrecht (1996).

[35] G. Schmidt, T. Ströhlein, Relations and Graphs, *Discrete mathematics for Computer Scientists*, Springer Verlag, Berlin, Heidelberg, New York (1999).

[36] L.S. Shapley, A value for n-person games. In: H.W. Kuhn, A.W. Tucker (eds.): *Contributions to the Theory of Games*, Vol. II, Annals of Mathematics Studies, 28, Princeton University Press, Princeton, NJ, (1953) 307-317.

[37] A. Verkeyn, D. Botteldooren, B. De Baets, G. De Tre, Modeling annoyance aggregation with Choquet integrals, *Proceedings of Eurofuse Workshop on Information Systems, Villa Monastero*, Varenna, Italy, (2002) 259 - 264.

[38] A. Verkeyn, D. Botteldooren, B. De Baets, G. de Tre, Sugeno Integrals for the Modelling of Noise Annoyance Aggregation, In: T. Bilgic et al. (eds.): *IFSA 2003*, LNAI 2715, (2003) 277–284.

[39] P. Wakker, *Additive Representations of Preferences: A new Foundation of Decision Analysis*, Kluwer Academic Publishers, Dordrecht, Boston, London, (1989).