University of Catania

Department of Economics and Business

HIERARCHY AND INTERACTION OF CRITERIA IN ROBUST ORDINAL REGRESSION

Salvatore Corrente

PhD thesis

Supervisor

Prof. **Salvatore Greco**

INTERNATIONAL PhD in APPLIED MATHEMATICS

XXV CYCLE

List of original papers

- S. Corrente, S. Greco, and R. Słowiński. Multiple Criteria Hierarchy Process in Robust Ordinal Regression. *Decision Support Systems,* 53(3): 660-674, 2012.

- S. Corrente, S. Greco, and R. Słowiński. Multiple Criteria Hierarchy Process with ELECTRE and PROMETHEE. *OMEGA,* (in press). http://dx.doi.org/10.1016/j.omega.2012.10.009

- S. Angilella, S. Corrente, and S. Greco. SMAA-Choquet: Stochastic Multicriteria Acceptability Analysis for the Choquet Integral. In S. Greco et al., editor, *IPMU 2012, Part IV, CCIS 300,* pages 248-257.

- S. Corrente, J.R. Figueira, and S. Greco. Interaction of Criteria and Robust Ordinal Regression in Bi-polar PROMETHEE Methods. In S. Greco et al., editor, *IPMU 2012, Part IV, CCIS 300,* pages 469-479.

- S. Angilella, S. Corrente, S. Greco, and R.Słowiński. MUSA-INT: Multicriteria customer satisfaction analysis with interacting criteria. Submitted to OMEGA.

- S. Angilella, S. Corrente, S. Greco, and R.Słowiński. Multiple Criteria Hierarchy Process for the Choquet integral. Accepted to the EMO 2013 Conference.

Introduction

All decision making situations we deal with during our lives have multiple criteria structure. That is, several alternatives are evaluated with respect to some point of view, technically evaluation criteria, and then compared in order to make the "best" decision. For example, buying a car can be seen as a decision making problem in which the different cars are the alternatives, while price, comfort, acceleration, maximum speed etc. can be considered as criteria to which the cars are evaluated against. Multiple Criteria Decision Aiding (MCDA) proposes useful methodologies to make decisions considering the preferences of the Decision Maker (DM). Obviously, every one would like to have the ideal alternative, that is the alternative that is the best for all considered criteria, but often a compromise has to be chosen. In fact, generally, when comparing two alternatives one is better compared to some criteria, while the other is better considering other criteria. In this situation, the following question arises: how to compare them?

Two different methodologies aiming to aggregate the evaluations of the alternatives with respect to the considered criteria are known in literature: the Multiple Attribute Utility Theory (MAUT) and the outranking methods. The first aims to assign to each alternative its utility, that is a numerical evaluation being representative of its worth, while the second uses binary relations to compare alternatives pairwise. Both methodologies use several parameters and one can fix these parameters using a direct or a indirect technique. The direct technique consists of asking the DM to directly provide the parameters required by the aggregation model, while the indirect technique consists of asking the DM to provide some preference information regarding some reference alternatives from which one can elicit the preferential parameters.

In general, there could be more than one set of parameters compatible with the preference information provided by the DM. All of them compare in the same way the reference alternatives but they could compare differently the other alternatives not provided as example by the DM. In order to simultaneously take into account all the sets of parameters compatible with the preference information provided by the DM, the Robust Ordinal Regression (ROR) constructs two preference relations, one necessary and one possible. The necessary preference relation holds if the preference of an alternative over another is true for all sets of parameters compatible with the preferences of the DM, while the possible preference relation holds if the preference of an alternative over another is true for at least one set of parameters compatible with the preferences of the DM.

In this thesis we have dealt with two important issues of MCDA, that is the interaction between criteria and the hierarchy of criteria.

The use of MAUT and outranking methods as preference models is based on the mutual independence between criteria. In many real world problems, the criteria are not independent but interacting. This means that it is possible to observe a certain form of synergy or redundancy between the evaluation criteria. For example, considering again the problem of evaluating a car, the criteria maximum speed and acceleration are redundant because often a very fast car also has good acceleration, while price and maximum speed criteria are synergetic because a fast and cheap car is very appreciated. In these cases, one needs to use non-additive integrals, as Choquet and Sugeno integrals, being the most known non-additive integrals in MCDA. Another possibility is to use an "enriched" utility function in which other than a marginal utility for each considered criterion there are some further components representing a bonus or a malus for synergetic or redundant criteria.

On the basis of the concept of interaction between criteria, we have extended two very well known MCDA methods, that are MUSA and PROMETHEE, giving rise to MUSA-int and the bipolar PROMETHEE method.

Because the Choquet integral involves the use of many parameters and the elicitation of these parameters is a very troublesome problem for the DM, we have integrated the Choquet integral and the SMAA methodologies building the SMAA-Choquet method.

In complex decision making problems, the alternatives are evaluated with respect to a family of criteria organized in a hierarchical way. This means that all criteria are not at the same level, but it is possible to fix some root criteria, a set of subcriteria descending from each root criterion, a set of subsubcriteria descending from each subcriterion, and so on. For example, in location problems, one has to take into account economical, environmental and social aspects. Each of these could be considered as a macro-criterion having different subcriteria. For example, the social macro-criterion can have "impact on individuals within site" and "impact on regional demographics" as subcriteria while economic macro-criterion can present "risk of commercial failure" and "employment" as subcriteria and so on.

In the thesis, the problem of aggregating the evaluations of each alternative with respect to all criteria in the hierarchy has been dealt both within MAUT and outranking methods and by also using the Choquet integral in case the hierarchy is composed of interacting criteria.

The thesis is organized as follows. In the first chapter, we shall describe the basic concepts of a MCDA problem. Chapter 2 contains contributions related to the interaction between criteria concept, that are:

- SMAA-Choquet: Stochastic Multicriteria Acceptability Analysis for the Choquet Integral,
- Interaction of Criteria and Robust Ordinal Regression in Bi-polar PROMETHEE Methods,
- MUSA-INT: Multicriteria customer satisfaction analysis with interacting criteria.

In chapter 3 we introduce the Multiple Criteria Hierarchy Process (MCHP) and its applications to MAUT, outranking methods and Choquet integral. Final remarks are contained in the last chapter.

Contents

Chapter 1

Basic concepts of Multiple Criteria Decision Aiding

Making any type of decision, from buying a car to locating a nuclear plant, from choosing the best student deserving a scholarship to ranking the cities of the world according to their liveability, involves the evaluation of several alternatives with respect to different aspects, technically called evaluation criteria. In these cases, we speak of Multiple Criteria Decision Making (MCDM) problems (for a survey on MCDM see [29]) for which Multiple Criteria Decision Aiding (MCDA) methodologies need to be provided to the Decision Maker (DM) in order to make the "best" decision. In a MCDM problem, a set of n alternatives $A = \{a_1, \ldots, a_n\}$ is evaluated with respect to a consistent family of m criteria $G = \{g_1, \ldots, g_m\}$ [103], where consistent means that it is exhaustive (all relevant criteria are taken into account), coherent (if two alternatives a and b have the same evaluations on all but one criteria, and a gets an evaluation better than b on the remaining criterion, than a should be preferred to b) and non-redundant (the removal of one criterion from the family makes the new set of criteria not exhaustive). In general, each criterion $g_j \in G$ can be considered as a function $g_j : A \rightarrow \mathcal{I}_j$, where $\mathcal{I}_j \subseteq \mathbb{R}$ if the criterion is quantitative, that is, it can be described with real numbers, while $\mathcal{I}_j$ is a set of discrete, ordered identifiers if the criterion is qualitative. Each criterion g_j can have an increasing or a decreasing direction of preference. In the first case, the higher the evaluation $g_j(a)$, the better a is with respect to criterion g_j; in the second case, the higher the evaluation $g_j(a)$, the worse a is with respect to criterion g_j (in the following, for the sake of simplicity and without loss of generality, we shall suppose that all evaluation criteria have an increasing direction of preference). Just for example, evaluating a car involves both quantitative and qualitative criteria having increasing or decreasing direction of preference. Price and acceleration are typical quantitative criteria while comfort and

safety are qualitative criteria. Among these, acceleration, comfort and safety have an increasing direction of preference, while price has obviously a decreasing direction of preference.
According to Roy [105], in MCDM the following four different problematics can be distinguished: description, choice, sorting and ranking.

- The *description* problematic consists of elaborating an appropriate set of alternatives A, building a consistent family of criteria G and determining, for all or some $a \in A$, their performances on the considered criteria;

- The *choice* problematic consists of selecting a small number (as small as possible) of "good" alternatives in such a way that a single alternative may finally be chosen;

- The *sorting* problematic consists of assigning each alternative to one of the predefined and ordered categories;

- The *ranking* problematic consists of defining a complete or partial order on A; this preorder is the result of a procedure allowing to put together in classes alternatives which can be judged indifferent, and to rank these classes.

Given two alternatives $a, b \in A$ and considering their evaluations with respect to the m criteria belonging to G, very often a will be better than b for some of the criteria while b will be better than a for the remaining criteria. For this reason, in order to cope with one of the last three problematics mentioned above, the necessity to aggregate the evaluations of the alternatives taking into account the preferences of the DM arises. In the literature, the two most known aggregation approaches are the following:

- assigning to each alternative $a \in A$ a real number synthesizing the evaluations of a with respect to the m criteria and being representative of the desirability of a with respect to the problem at hand,

- building some binary preference relations in order to compare pairs of alternatives in the set A.

In the first approach, the number assigned to $a \in A$ is independent from the evaluations of the other alternatives; it leads us to define a complete preorder on A and it does not allow any incomparability among the alternatives. The MAUT [80], that will be described in the next section, is based on this approach.

In the second approach, the pairwise comparisons can generate some intransitivities and the most appropriate conclusion comparing pairs of alternatives is the incomparability. The outranking methods that will be described in section 1.2 are based on binary preference relations on the set of alternatives.

1.1 Multiple Attribute Utility Theory

According to Dyer [26], preference theory studies the fundamental aspects of individual choice behavior, such as how to identify and quantify an individual's preferences over a set of alternatives and how to construct appropriate preference representation functions for decision making. An important feature of preference theory is that it is based on rigorous axioms which characterize an individual's choice behavior. These preference axioms are essential for establishing preference representation functions, and provide the rationale for the quantitative analysis of preference. In this context, it is possible to distinguish between preferences under conditions of certainty or risk and over alternatives described by a single attribute or by multiple attributes. We shall refer to a preference representation function under certainty as a value function, and to a preference representation function under risk as a utility function [80]. In the following, we shall consider multiple attribute value functions. Given $\mathcal{I} = \prod_{j=1}^{m} \mathcal{I}_j$, and denoted by $\succsim$ the DM's preference relation over $\mathcal{I}$ (where $a \succsim b$ reads a is at least as good as b), we will be interested in conditions allowing to determine the existence of a function $U : \mathcal{I} \to \mathbb{R}$ such that $a \succsim b$ if and only if $U(\overline{g}(a)) \geq U(\overline{g}(b))$, with $\overline{g}(a), \overline{g}(b) \in \mathcal{I}$ and $\overline{g}(a) = (g_1(a), \ldots, g_m(a))$, for all $a \in A$. A necessary condition for the existence of such a function is that $\succsim$ is a weak order (complete and transitive binary relation). A second condition (and then both are necessary and sufficient) is that $A/\sim$ contains a countable order-dense subset (Birkhoff-Milgram theorem) where $\sim$ is the symmetric part of $\succsim$ while $\succ$ is the asymmetric one (see also [52]).

The most common approach for evaluating multiattribute alternatives is to use an additive representation. In an additive representation, a real value is assigned to each alternative a by:

$$U(a) = \sum_{j=1}^{m} u_j(g_j(a))$$

where u_j are single attribute non-decreasing value functions over $\mathcal{I}_j$. Defining $x_j^* = \max_{a \in A} g_j(a)$ and $x_{j,*} = \min_{a \in A} g_j(a)$ the best and the worst evaluations an alternative belonging to A can get on criterion g_j respectively, we normalize the value function U imposing that $\sum_{j=1}^{m} u_j(x_j^*) = 1$ and $u_j(x_{j,*}) = 0$ for all $j = 1, \ldots, m$.

If our interest is in simply rank-ordering the available alternatives, then the key condition for the additive form is the *mutual preference independence* of the set of criteria G. We say that the set of criteria $T \subseteq G$ is *preferentially independent* [129] of $G \setminus T$ if, for all $a_T, b_T \in \prod_{j\in T} \mathcal{I}_j$, and for all $c_{G\setminus T}, d_{G\setminus T} \in \prod_{j\in G\setminus T} \mathcal{I}_j$,

$$(a_T, c_{G\setminus T}) \succsim (b_T, c_{G\setminus T}) \Leftrightarrow (a_T, d_{G\setminus T}) \succsim (b_T, d_{G\setminus T})$$

that is, the preference of $(a_T, c_{G\setminus T})$ over $(b_T, c_{G\setminus T})$ does not depend on $c_{G\setminus T}$. The whole set of criteria G is said to be *mutually preferentially independent* if T is preferentially independent of $G \setminus T$ for every $T \subseteq G$.

Even if the additive value function would seem to be an attractive choice for practical applications of multiattribute decision making, the assessment of the single attribute value functions relies on techniques that are cumbersome in practice, and that force the decision maker to make explicit tradeoffs between two or more criteria. Keeney and Raiffa [80] illustrate two assessment procedures for ordinal additive value functions.

Using a value function U, one gets a complete order among the considered alternatives and therefore it does not generate any incomparability among them.

1.2 Outranking methods

Outranking methods were first developed in France in the late sixties by B. Roy following difficulties experienced with the value function approach in dealing with practical problems. Outranking methods build a preference relation, usually called an outranking relation, among alternatives evaluated on several criteria. An outranking relation is a binary relation S on the set of alternatives A such that aSb means that alternative a is at least as good as alternative b and it holds if a majority of the criteria supports this assertion and the opposition of the other criteria is not too strong. The binary outranking relation is neither complete (it is possible that not(aSb) and not(bSa)) nor transitive (it is possible that aSb and bSc but not(aSc)). Besides, the outranking methods take into account imprecisions, incertitudes, and arbitrarity over the data using the quasi-criteria and pseudo-criteria [110].

The two most known families of outranking methods are ELECTRE [104] and PROMETHEE [15, 16, 17].

1.2.1 ELECTRE methods

ELECTRE methods [104] were proposed by B.Roy for the first time in the late sixties to give a realistic representation of four basic situation of preferences, that is indifference, weak preference, strong preference and incomparability. Based on the binary outranking relation S, four situations may occur:

- a is preferred to b (aPb) iff aSb and not bSa,
- b is preferred to a (bPa) iff bSa and not aSb,
- a is indifferent to b (aIb) iff aSb and bSa,
- a is incomparable to b (aRb) iff not(aSb) and not(bSa).

Remark that using outranking relation to model preferences introduces the incomparability relation (R) being useful in situations in which the DM is not able to compare two actions.
The construction of an outranking relation is based on the *concordance* and the *non-discordance* tests. The concordance test on aSb is verified if a sufficient majority of criteria is in favor of this assertion, while the non-discordance test is verified if none of the criteria in the minority opposes too strongly to the assertion aSb.
Two sets of parameters are meaningful in ELECTRE methods: importance coefficients and thresholds. For each criterion g_j, the importance coefficient w_j represents the weight of the criterion g_j inside the family of criteria G when it is in favor of aSb. For the sake of simplicity, and without loss of generality, it is supposed that $w_j \geq 0$, $\forall j \in G$ and $\sum_{j \in G} w_j = 1$. These weights can be constant or dependent from the evaluations of the alternatives.
For each criterion g_j, three thresholds are taken into account in the ELECTRE methods:

- the *indifference* threshold q_j, being the largest difference $g_j(b) - g_j(a)$ compatible with the indifference on criterion g_j among alternatives a and b,
- the *preference* threshold p_j, being the smallest difference $g_j(b) - g_j(a)$ compatible with the preference on criterion g_j of b over a,
- the *veto* threshold v_j, being the smallest difference $g_j(b) - g_j(a)$ incompatible with the outranking of a over b. This means that, if $g_j(b) - g_j(a) \geq v_j$ then *not*(aSb).

Depending on the type of problem they are able to deal with, several ELECTRE methods can be distinguished: ELECTRE I [100] and IS [109] for choice problems; ELECTRE II [101], III [102] and IV [107] for ranking problems; ELECTRE TRI [133] and Tri-nC [2] for sorting problems. In the following we shall illustrate only the ELECTRE IS and ELECTRE III (for a review on ELECTRE methods see [104]).

The ELECTRE IS method builds for each criterion g_j, $j \in G$, and for each pair of alternatives $(a,b) \in A \times A$, the partial concordance index $\phi_j(a,b)$ and the comprehensive concordance index $C(a,b)$. $\phi_j(a,b)$ represents the degree of outranking of a over b with respect to criterion g_j. It is defined by a non-increasing function of $g_j(b) - g_j(a)$:

$$\phi_j(a,b) = \begin{cases} 1 & \text{if} \quad g_j(b) - g_j(a) \leq q_j, \\ \frac{[g_j(b)-g_j(a)]-q_j}{p_j-q_j} & \text{if} \quad q_j < g_j(b) - g_j(a) < p_j \\ 0 & \text{if} \quad g_j(b) - g_j(a) \geq p_j. \end{cases}$$

$\phi_j(a,b) \in [0,1]$ and the higher the value of $\phi_j(a,b)$, the higher the concordance with the outranking of a over b on criterion g_j is.

The comprehensive concordance index is instead defined by:

$$C(a,b) = \sum_{j=1}^{m} w_j \cdot \phi_j(a,b),$$

and it represents how much alternative a globally outranks alternative b.

In ELECTRE IS,

- the concordance test is verified if $C(a,b) \geq \lambda$, where λ is called concordance cutting level, and $\lambda \in [0.5, 1]$,
- the non-discordance test is verified if there is no criterion putting a veto on the outranking of a over b. Formally, this can be expressed saying that for all criteria $g_j, j \in G$, $g_j(b) - g_j(a) < v_j$.

In ELECTRE III, the construction of an outranking relation S is not based on the concordance and the non-discordance tests. ELECTRE III builds for each criterion $g_j \in G$ and for each pair of alternatives $(a,b) \in A \times A$ the discordance index $d_j(a,b)$ and the credibility index $\rho(a,b)$. $d_j(a,b)$ represents the degree of discordance on the outranking of a over b compared to the criterion g_j. It is defined by a non-decreasing function of $g_j(b) - g_j(a)$:

$$d_j(a,b) = \begin{cases} 1, & \text{if } g_j(b) - g_j(a) \geq v_j, \\ \frac{[g_j(b)-g_j(a)]-p_j}{v_j-p_j}, & \text{if } p_j < g_j(b) - g_j(a) < v_j, \\ 0, & \text{if } g_j(b) - g_j(a) \leq p_j. \end{cases}$$

$d_j(a,b) \in [0,1]$, and the higher the value of $d_j(a,b)$, the higher the discordance with the outranking of a over b on criterion g_j is.

The credibility index expresses the degree of credibility of the outranking of a over b. It is defined as:

$$\rho(a,b) = C(a,b) \prod_{\{j:\, d_j(a,b) > C(a,b)\}} \frac{1 - d_j(a,b)}{1 - C(a,b)}$$

where $C(a,b)$ is the comprehensive concordance index already presented in ELECTRE IS, $\rho(a,b) \in [0, C(a,b)]$, and obviously the higher the value of $\rho(a,b)$, the higher the credibility of the outranking of a over b is.

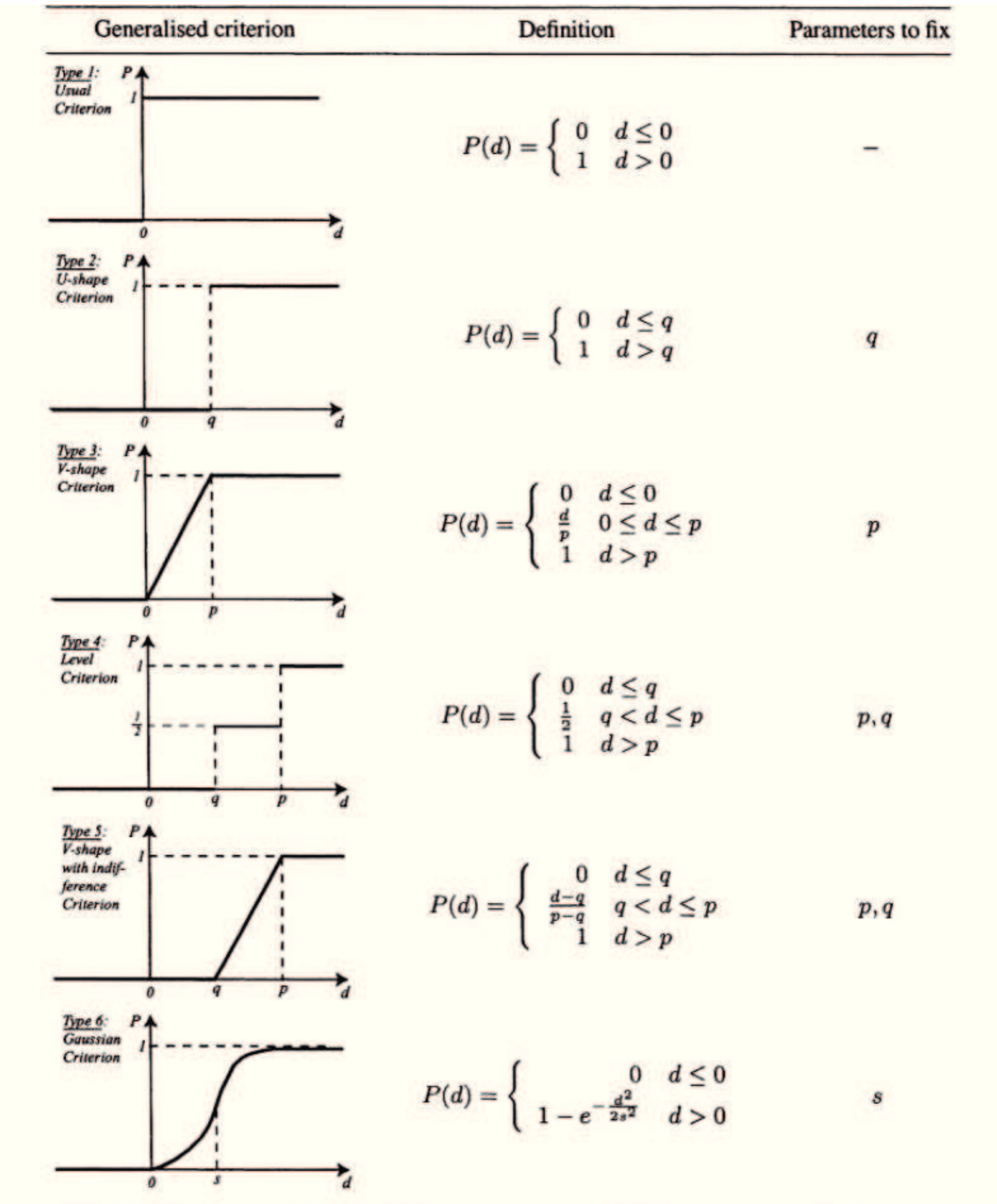
From the definition of $\rho(a,b)$ it follows that if none of the criteria opposes veto to the outranking of a over b (that is $d_j(a,b) = 0$ for all $j \in G$), then $\rho(a,b) = C(a,b)$; if some criterion opposes a veto on the outranking of a over b (that is there exists a criterion g_j, with $j \in G$, such that $d_j(a,b) = 1$), then $\rho(a,b) = 0$ and in all other cases the credibility index $\rho(a,b)$ is lower than the comprehensive concordance index $C(a,b)$.

1.2.2 PROMETHEE methods

The PROMETHEE methods [15, 16, 17] were born in 1982 thanks to J.P. Brans; PROMETHEE methods build an outranking relation in order to reduce the number of incomparabilities among some pairs of alternatives. In PROMETHEE methods, the DM has to provide inter-criteria parameters and intra-criterion parameters. Inter-criteria parameters are the importance of criteria w_j, $j \in G$, representing the importance of criterion g_j inside the family of criteria G as in the ELECTRE methods. Also in this case, we suppose that $w_j \geq 0$, $\forall j \in G$, and $\sum_{j \in G} w_j = 1$. Intra-criterion parameters regard instead the choice, for each criterion $g_j \in G$, of one preference function $P_j(d_j)$ among the six showed in Fig.1.1[1].

[1] This figure has been taken from [15]

Figure 1.1: Types of preference functions



$P_j(d_j(a,b))$ gives the degree of preference of a over b. It is defined as a non-decreasing function of $d_j(a,b) = g_j(a) - g_j(b)$, and it involves 0,1,2 or 3 parameters (the indifference threshold q_j and the preference threshold p_j as in the ELECTRE methods and a parameter s defining the inflection point in the sixth preference function showed in Fig.1.1). In the following, we shall describe PROMETHEE I and II (for a survey on PROMETHEE methods see [15, 16, 11]).

For each pair of alternatives $(a,b) \in A \times A$, PROMETHEE I and II compute $\pi(a,b) = \sum_{j=1}^{m} w_j \cdot P_j(a,b)$ being the equivalent of the comprehensive concordance index $C(a,b)$ in the ELECTRE methods and representing the comprehensive degree of preference of a over b. Obviously the greater the value of

$\pi(a, b)$, the greater the preference of a over b is.

In order to assess the worth of $a \in A$ compared to all other alternatives in A, PROMETHEE I and II compute the positive and the negative outranking flows:

- $\Phi^+(a) = \dfrac{1}{n-1} \displaystyle\sum_{b \in A \setminus \{a\}} \pi(a, b)$,
- $\Phi^-(a) = \dfrac{1}{n-1} \displaystyle\sum_{b \in A \setminus \{a\}} \pi(b, a)$,

Additionally, for each alternative $a \in A$, PROMETHEE II computes also the net outranking flow:

- $\Phi(a) = \Phi^+(a) - \Phi^-(a)$.

The positive outranking flow expresses how much an alternative a is outranking all other alternatives. The higher $\Phi^+(a)$, the better the alternative a is.

The negative outranking flow expresses how much an alternative a is outranked by all the others. The lower $\Phi^-(a)$, the better the alternative a is.

The net outranking flow is a balance between the positive and the negative outranking flows. The higher $\Phi(a)$, the better the alternative a is.

PROMETHEE I provides a partial ranking among alternatives defining preference (P^I), indifference (I^I) and incomparability (R^I) relations based on the positive and negative outranking flows:

- aP^Ib iff $\Phi^+(a) \geq \Phi^+(b)$, $\Phi^-(a) \leq \Phi^-(b)$ and at least one of the two inequalities is strict,
- aI^Ib iff $\Phi^+(a) = \Phi^+(b)$ and $\Phi^-(a) = \Phi^-(b)$,
- aR^Ib otherwise.

Differently from PROMETHEE I, PROMETHEE II defines only a preference (P^{II}) and an indifference (I^{II}) relation based on the net outranking flow:

- $aP^{II}b$ iff $\Phi(a) > \Phi(b)$,
- $aI^{II}b$ iff $\Phi(a) = \Phi(b)$.

In PROMETHEE II, all the alternatives are comparable and the following properties hold:

- $-1 \leq \Phi(a) \leq 1$, $\forall a \in A$,
- $\displaystyle\sum_{a \in A} \Phi(a) = 0$.

1.3 Choquet integral and bi-polar Choquet integral

As stated in subsection 1.1, an additive value function can represent the preferences of the DM only if the set of criteria G is mutually preferentially independent. In the following we provide an example inspired by [43] in which an additive value function can not represent the preferences of the DM. Let us suppose that a Dean has to compare four students whose marks on a common scale [0,20] on the subjects of Literature, Mathematics and Physics are shown in Table 1.1. (S)he thinks that scientific subjects are very important but (s)he does not want to favour students that are good in scientific subjects but who are lacking in literature. Besides (s)he thinks that Mathematics and Physics are redundant because, generally, a student good in Mathematics is also good in Physics. Comparing x and y, the Dean observes that both students have good scores on scientific subjects but

Table 1.1: Student's evaluations

Student	Mathematics (M)	Physics (P)	Literature (L)
x	14	13	6
y	14	11	8
w	5	13	6
z	5	11	8

y has a better mark than x in Literature, so (s)he states that y is preferred to x. Comparing w and z, the Dean observes that their marks on Mathematics are very low, so, because scientific subject are more important than literature, (s)he states that w is preferred to z. We could summarize the preferences of the Dean with the following rules:

R1) For a student good in Mathematics, Literature is more important than Physics,

R2) For a student bad in Mathematics, Physics is more important than Literature.

Using the definition given in subsection 1.1, it is easy to see that the set of criteria {Physics, Literature} is not preferentially independent from criterion {Mathematics} and therefore the whole set of criteria {Mathematics, Physics, Literature} is not mutually preferentially independent. In this case, an additive value function is not able to represent the preferences of the Dean. In fact,

- the preference $y \succ x$, is translated by

$$u_M(14) + u_P(11) + u_L(8) > u_M(14) + u_P(13) + u_L(6), \tag{1.1}$$

- while the preference $w \succ z$ is translated by

$$u_M(5) + u_P(11) + u_L(8) < u_M(5) + u_P(13) + u_L(6). \tag{1.2}$$

The two inequalities lead to the following contradiction:

$$u_P(11) + u_L(8) > u_P(13) + u_L(6) \quad \text{and} \quad u_P(11) + u_L(8) < u_P(13) + u_L(6).$$

In many real world problems, the evaluation criteria are not mutually preferentially independent but it is possible to observe some interaction between them. For example, let us consider the evaluation of a car considering the following criteria: maximum speed, acceleration and price. In this case, there may exist a negative interaction (redundancy) between maximum speed and acceleration because a car with a high maximum speed also has a good acceleration; so, even if each of these two criteria is very important for a DM who likes sport cars, their joint impact on reinforcement of preference of a more speedy and better accelerating car over a less speedy and worse accelerating car will be smaller than a simple addition of the impacts of the two criteria considered separately in validation of this preference relation. In the same decision problem, there may exist a positive interaction (synergy) between maximum speed and price because a car with a high maximum speed and relatively low price is very much appreciated. Thus, the comprehensive impact of these two criteria on the strength of preference of a more speedy and cheaper car over a less speedy and more expensive car is greater than the impact of the two criteria considered separately in validation of this preference relation. To handle the interaction among criteria, one can consider non-additive integrals, such as the Choquet integral [21] and the Sugeno integral [119], or an additive value function augmented by additional components reinforcing the value when there is positive interaction for some pairs of criteria, or penalizing the value when this interaction is negative, like in UTA^{GMS}-INT [58] or in MUSA-INT [5] (for a comprehensive survey on the use of non-additive integrals in MCDA see [39, 44, 43]).

Given the set of criteria G and denoted by 2^G the power set of G, that is the set of all subsets of G, a capacity or fuzzy measure, is a set function $\mu : 2^G \rightarrow [0,1]$ satisfying the following properties:

1a) $\mu(\emptyset) = 0$, $\mu(G) = 1$,

2a) $\mu(A) \leq \mu(B)$, for all $A \subseteq B \subseteq G$.

Roughly speaking, $\mu(A)$ expresses the degree to which the coalition of criteria $A \subseteq G$ is important for making a decision. A capacity is said to be additive if $\mu(A \cup B) = \mu(A) + \mu(B)$ whenever $A \cap B = \emptyset$.

Note that if a capacity is additive, then it suffices to define the m coefficients $\mu(\{g_1\}), \ldots, \mu(\{g_m\})$ to define it.

Let μ be a capacity on G, and $x = (x_1, \ldots, x_m) \in \mathcal{I} : x_j = g_j(x) \geq 0, \forall j = 1 \ldots, m$; then, the Choquet integral of x with respect to the capacity μ is defined by:

$$C_\mu(x) = \int_0^1 \mu(\{j \in G : x_j \geq t\})dt, \tag{1.3}$$

or equivalently, as

$$C_\mu(x) = \sum_{j=1}^{m} x_{(j)} \left[\mu(A_{(j)}) - \mu(A_{(j+1)})\right] = \sum_{j=1}^{m} \left[x_{(j)} - x_{(j-1)}\right] \mu(A_{(j)}), \tag{1.4}$$

where $0 = x_{(0)} \leq x_{(1)} \leq \ldots \leq x_{(m)}$, $A_{(j)} = \left\{i \in G : x_i \geq x_{(j)}\right\}$ and $A_{(m+1)} = \emptyset$.
Considering the example showed above, and using the definition of the Choquet integral as seen in equation (1.4), we get:

$$y \succ x \quad \Leftrightarrow \quad \mu(\{M\}) + 1 > 2\mu(\{M, P\}), \tag{1.5}$$

and

$$w \succ z \quad \Leftrightarrow \quad 2\mu(\{P\}) > \mu(\{P, L\}). \tag{1.6}$$

Because inequalities (1.5) and (1.6) are not in contradiction, then the Choquet integral is able to describe the preferences of the Dean.

The Choquet integral can be redefined in terms of the Möbius representation [37], without reordering the criteria, as:

$$C_\mu(x) = \sum_{T \subseteq G} a(T) \min_{i \in T} g_i\left(x\right).$$

One of the main drawbacks of the Choquet integral is the necessity to elicit and give an adequate interpretation of $2^m - 2$ parameters (because $\mu(\emptyset) = 0$, and $\mu(G) = 1$). In order to reduce the number of parameters to be computed and to eliminate a description of the interactions among criteria that is too strict, which is not realistic in many applications, the concept of fuzzy k-additive measure has been considered [40]; a *fuzzy measure* is called *k-additive* if $a(T) = 0$ with $T \subseteq G$, when $|T| > k$ and there exists at least one $T \subseteq G$, with $|T| = k$, such that $a(T) > 0$. We observe that a 1-additive measure is the common additive fuzzy measure. In many real decision problems, it suffices to consider 2-additive measures. In this case, positive and negative interactions between couples of criteria are modeled without considering the interaction among triples, quadruplets and generally m-tuples, (with $m > 2$) of criteria. From the point of view of MCDA, the use of 2-additive measures is justified by observing that the information on the importance of the single criteria and the interactions between couples of criteria are noteworthy. Moreover, it might not be easy or not straightforward for the DM to provide information on the interactions among three or more criteria during the decision procedure. From a computational point of view, the interest in the 2-additive measures lies in the fact that any decision model needs to evaluate a number $m + \binom{m}{2}$ of parameters (in terms of Möbius representation, a value $a(\{i\})$ for every criterion i and a value $a(\{i,j\})$ for every couple of distinct criteria $\{i,j\}$). With respect to a 2-additive fuzzy measure, the inverse transformation to obtain the fuzzy measure $\mu(R)$ from the Möbius representation is defined as:

$$\mu(R) = \sum_{i\in R} a(\{i\}) + \sum_{\{i,j\}\subseteq R} a(\{i,j\}), \ \forall R \subseteq G. \tag{1.7}$$

With regard to 2-additive measures, properties **1a)** and **2a)** have, respectively, the following formulations:

1b) $a(\emptyset) = 0, \ \sum_{i\in G} a(\{i\}) + \sum_{\{i,j\}\subseteq G} a(\{i,j\}) = 1,$

2b) $\begin{cases} a(\{i\}) \geq 0, \ \forall i \in G, \\ a(\{i\}) + \sum_{j\in T} a(\{i,j\}) \geq 0, \ \forall i \in G \text{ and } \forall\, T \subseteq G \setminus \{i\}, T \neq \emptyset. \end{cases}$

In this case, the representation of the Choquet integral of $x \in A$ is given by:

$$C_\mu(x) \ = \ \sum_{\{i\}\subseteq G} a(\{i\})(g_i(x)) + \sum_{\{i,j\}\subseteq G} a(\{i,j\}) \min\{g_i(x), g_j(x)\}. \tag{1.8}$$

Finally, we recall the definitions of the importance and interaction indices for a couple of criteria. The *importance index* or Shapley value [113] is given by:

$$\varphi(\{i\}) = a(\{i\}) + \sum_{j\in G\setminus\{i\}} \frac{a(\{i,j\})}{2}, \quad \forall i \in G, \tag{1.9}$$

while the *interaction index* [90] for a couple of criteria $\{i,j\} \subseteq G$, in case of two additive capacities, is given by:

$$\varphi(\{i,j\}) = a(\{i,j\}). \tag{1.10}$$

Now, let us suppose that the Dean has to evaluate four students whose marks on (M), (P), and (L) are given in Table 1.2.

Table 1.2: Second student's evaluations

Student	Mathematics (M)	Physics (P)	Literature (L)
x	14	17	5
y	14	15	7
w	8	17	5
z	8	15	7

According to rules **R1)** and **R2)**, the Dean states that y is preferred to x and w is preferred to z. Aggregating using the Choquet integral, from $y \succ x$ we get:

$$5 + 9\mu(\{M,P\}) + 3\mu(\{P\}) < 7 + 7\mu(\{M,P\}) + \mu(\{P\}) \Leftrightarrow \mu(\{M,P\}) + \mu(\{P\}) < 1. \tag{1.11}$$

In the same way, $w \succ z$ is translated by:

$$5 + 3\mu(\{M,P\}) + 9\mu(\{P\}) > 7 + \mu(\{M,P\}) + 7\mu(\{P\}) \Leftrightarrow \mu(\{M,P\}) + \mu(\{P\}) > 1. \tag{1.12}$$

Obviously, inequalities (1.11) and (1.12) are in contradiction. This means that the Choquet integral can not explain the preferences of the Dean. The reason is that the Choquet integral satisfies the comonotonic additive property, that is $C_\mu(a+b) = C_\mu(a) + C_\mu(b)$ for all comonotonic a and b where $a, b \in \mathbb{R}^m$ are comonotonic if $a_{j_1} > a_{j_2} \Rightarrow b_{j_1} \geq b_{j_2}$ for any $j_1, j_2 \in \{1, \ldots, m\}$. In fact, it is a straightforward observation that in our case the evaluation vectors of the four students are comonotonic and therefore

$$y \succ x \Leftrightarrow C_\mu(14,15,7) > C_\mu(14,17,5) \Leftrightarrow C_\mu(14,15,7) - C_\mu(14,17,5) = C_\mu(0,-2,2) > 0$$

$$w \succ z \Leftrightarrow C_\mu(8,17,5) > C_\mu(8,15,7) \Leftrightarrow C_\mu(8,15,7) - C_\mu(8,17,5) = C_\mu(0,-2,2) < 0$$

being obviously in contradiction.

Rules **R1**) and **R2**) make an implicit reference to a neutral level that is neither good nor bad. This suggests that criteria should be considered in a bipolar scale. According to [44], a scale on $\mathcal{I}_j$ is bipolar if there exists in $\mathcal{I}_j$ a particular element or level 0_j, called neutral level, such that the elements of $\mathcal{I}_j$ preferred to 0_j are considered as "good", while the elements of $\mathcal{I}_j$ less preferred than 0_j are considered as "bad" for the DM. Typical examples of bipolar scales are [-1,1] (bounded cardinal), $\mathbb{R}$ (unbounded cardinal) or {very bad, bad, medium, good, excellent} (ordinal). Considering the bipolar scale, the problem to define the importance of coalitions of criteria and to evaluate the overall score of an alternative arises. Let us take for simplicity the $[-1,1]$ scale, with neutral level 0. The simplest way is to say that "positive" and "negative" parts are symmetric, so the overall evaluations of binary alternative $(1_A, 0_{A^c})$ (i.e., the alternative having evaluation equal to one for all criteria in A and evaluation equal to 0 for all other criteria) is the opposite of the one of negative binary alternative $(-1_A, 0_{A^c})$. This leads to the symmetric Choquet integral (also called Šipoš integral [114]):

$$C_\mu(x) = C_\mu(x^+) - C_\mu(x^-)$$

where $x_j^+ = \max\{x_j, 0\}$, $\forall j = 1, \ldots, m$, and $x^- = (-x)^+$.

Considering the evaluations in Table 1.2, the value 10 can be considered as the neutral level for all three criteria. Therefore, we can transform the evaluations in Table 1.2 subtracting 10 to each mark (see Table 1.3).

Table 1.3: Bipolar evaluations

Student	Mathematics (M)	Physics (P)	Literature (L)
x	4	7	-5
y	4	5	-3
w	-2	7	-5
z	-2	5	-3

Trying to explain the preferences of the Dean using the symmetric Choquet integral, we get:

$$y \succ x \quad \Leftrightarrow \quad \mu(\{L\}) > \mu(\{P\}),$$

$$w \succ z \quad \Leftrightarrow \quad \mu(\{L\}) < \mu(\{P\}),$$

being again in contradiction.

A more complex model would consider only the independence between positive and negative parts, that is to say, positive binary alternatives define a capacity μ_1, while negative binary alternatives define a different capacity μ_2. This leads to the well-known Cumulative Prospect Theory (CPT) mode, of Kahnemann and Tversky [127]:

$$CPT\mu_1, \mu_2(x) = C_{\mu_1}(x^+) - C_{\mu_2}(x^-).$$

It is straightforward proving that, also in this case, the translations of the preference information of the Dean lead to a contradiction:

$$y \succ x \quad \Leftrightarrow \quad \mu_1(\{L\}) > \mu_2(\{P\}),$$

$$w \succ z \quad \Leftrightarrow \quad \mu_1(\{L\}) < \mu_2(\{P\}).$$

A more general model considers that independence between positive and negative part does not hold, so that we have to consider ternary alternatives $(1_A, -1_B, 0_{(A\cup B)^c})$, and assign to each of them a number in $[-1, 1]$. We denote this number as $\hat{\mu}(A, B)$, i.e., a two-argument function, whose first argument is the set of totally satisfied criteria, and the second one the set of totally unsatisfied criteria, the remaining criteria being at the neutral level. This function is called *bi-capacity* [41, 42, 54], since it plays the role of a capacity, but with two arguments corresponding to the positive and the negative sides of a bipolar scale.

Formally speaking, considering the set $\mathcal{J} = \{(S, T) : S, T \subseteq G \text{ and } S \cap T = \emptyset\}$, a bicapacity is a function $\hat{\mu} : \mathcal{J} \to [-1, 1]$ such that:

- $\hat{\mu}(\emptyset, \emptyset) = 0$, $\hat{\mu}(G, \emptyset) = 1$, $\hat{\mu}(\emptyset, G) = -1$,
- $\hat{\mu}(C, D) \leq \hat{\mu}(E, F)$, for all $(C, D), (E, F) \in \mathcal{J}$ such that $C \subseteq E$ and $D \supseteq F$.

Given $x = (x_1, \ldots, x_m) \in \mathcal{I}$ such that $x_j = g_j(x)$, and a bicapacity $\hat{\mu}$, the bipolar Choquet integral of x with respect to $\hat{\mu}$ is defined as:

$$Ch^B(x, \hat{\mu}) = \int_0^1 \hat{\mu}(\{i \in G : x_i \geq t\}, \{i \in G : x_i \leq -t\})dt$$

or equivalently, as:

$$Ch^B(x,\hat{\mu}) = \sum_{j\in\mathcal{G}^>} |x_{(j)}| \Big[\hat{\mu}(C_{(j-1)}, D_{(j-1)}) - \hat{\mu}(C_{(j)}, D_{(j)})\Big]$$

where: $|x_{(1)}| \leq \ldots \leq |x_{(m)}|$, $C_{(0)} = \{i \in G : x_i > 0\}$, $D_{(0)} = \{i \in G : x_i < 0\}$,
$\mathcal{G}^> = \{i \in G \ : \ |x_{(i)}| > 0\}$, $C_{(j)} = \{i \in \mathcal{G}^> \ : \ x_i \geq |x_{(j)}|\}$, and $D_{(j)} = \{i \in \mathcal{G}^> \ : \ -x_i \geq |x_{(j)}|\}$.

Trying to represent the preferences of the Dean related to the students' evaluations in Table 1.3 using the bi-polar Choquet integral, we get:

$$y \succ x \quad \Leftrightarrow \quad \hat{\mu}(\{M,P\},\emptyset) > \hat{\mu}(\{P\},\emptyset) + \hat{\mu}(\{M,P\},\{L\}) + \hat{\mu}(\{M\},\{L\}) \tag{1.13}$$

and

$$w \succ z \quad \Leftrightarrow \quad \hat{\mu}(\{P\},\{L\}) > 0. \tag{1.14}$$

Because inequalities (1.13) and (1.14) are not in contradiction, the bipolar Choquet integral is able to explain the preferences of the Dean.
Bicapacities have been introduced to represent complex preferences that cannot be modeled with a capacity.

1.4 Robust Ordinal Regression

Each decision model requires the specification of some parameters. For example, using MAUT, the parameters are related to the formulation of the marginal value functions $u_j(g_j(a)), j = 1,\ldots,m$; using non-additive integrals, the parameters are related to fuzzy measures while using outranking methods the parameters are related to thresholds and importance coefficients. Within MCDA, many methods have been proposed to determine the parameters characterizing the considered decision model in an indirect way, i.e., inducing the values of such parameters from some holistic preference comparisons of alternatives given by the DM. Eliciting direct preference information from the DM can be counterproductive in real-world decision making situations because of a high cognitive effort required. Consequently, asking directly the DM to provide values for the parameters seems to make the DM uncomfortable. Eliciting indirect preference is less demanding of cognitive effort. Indirect preference information is mainly used in the ordinal regression paradigm. According to this paradigm, a holistic preference information on a subset of some reference or training alternatives is known first and then a preference model compatible with the information is built and applied to the whole set of alternatives. The ordinal regression paradigm has been applied within the two main MCDA approaches: those using a value function as preference model [19, 72, 73, 95, 117], and those using an outranking relation as preference model [88, 89].

Usually, from among many sets of parameters of a preference model representing the preference information given by the DM, only one specific set is selected and used to work out a recommendation. Since the selection of one from among many sets of parameters compatible with the preference information given by the DM is rather arbitrary, *Robust Ordinal Regression* (ROR) proposes taking into account all the sets of parameters compatible with the preference information, in order to give a recommendation in terms of necessary and possible consequences of applying all the compatible preference models on the considered set of alternatives. In ROR, using linear programming one obtains two relations in the set A: the *necessary* weak preference relation, which holds for any two alternatives $a, b \in A$ if and only if a is at least as good as b for all compatible preference models, and the *possible* weak preference relation, which holds for this pair if and only if a is at least as good as b for at least one compatible preference model.

The first method applying ROR is a generalization of the UTA method [72], called UTA^{GMS} [55]. Differently from UTA method, UTA^{GMS} requires from the DM a set of pairwise comparisons on a set of reference alternatives $A^R \subseteq A$ as preference information. Besides, it takes into account all monotonic marginal value functions and not only the piecewise linear. Its extension, called GRIP [33], considers also intensities of preference among reference alternatives comprehensively, that is taking into account all criteria simultaneously, or partially, that is considering a single criterion. Other ROR-methodologies using the additive value function as preference model are: Extreme Ranking Analysis [76], UTA^{GMS}-INT [58] and UTADIS^{GMS} [57]. Extreme ranking analysis examines how different can be rankings provided by all compatible value functions determining the highest and the lowest ranks, and the score that an alternative can attain; the UTA^{GMS}-INT deals with interacting criteria considering an additive value function augmented by two components corresponding to "bonus" or "malus" values for positively and negatively interacting criteria, respectively; UTADIS^{GMS} is a sorting method that aims at assigning alternatives to pre-defined and ordered classes. The preference information provided by the DM consists of the assignment of reference alternatives to one or several contiguous classes.

Other methodologies applying the ROR are: PROMETHEE^{GKS} [76], ELECTRE^{GKMS} [47], NAROR [7], UTA^{GMS}-GROUP and UTADIS^{GMS}-GROUP [48] and UTA^{GSS} [59]. PROMETHEE^{GKS} and ELECTRE^{GKMS} extend the PROMETHEE and the ELECTRE methods; NAROR (non-additive ordinal regression) uses the Choquet integral as preference model considering not only intensity of preference on pairs of reference alternatives but also pairwise comparisons on the importance of criteria and the sign and the intensity of interaction among pairs of criteria; UTA^{GMS}-GROUP and UTADIS^{GMS}-GROUP apply ROR to group decision making in ranking and sorting problems respectively while UTA^{GSS} uses ROR in case of interacting criteria on bipolar scales. New extensions of the ROR will be introduced in this thesis.

Even if the recommendations obtained using ROR are "more robust" than a recommendation made using an arbitrarily chosen compatible model, for some decision-making situations a score is needed to assign to different alternatives; for this reason, some users would like to see the "most representative" model among all the compatible ones. This allows assigning a score to each alternative. Based on the ROR concept, the most representative model is the compatible model maximizing the difference of values between alternatives a and b for which a is necessarily preferred to b but b is not necessarily preferred to a, and minimizing the difference of values between alternatives a and b for which neither a is necessarily preferred to b nor b is necessarily preferred to a. The most representative model concept has been introduced for the first time in [32] and then applied to deal with ranking and choice problems [75], outranking methods [77], sorting problems [49] and group decision making [74].

1.5 SMAA methods

Accepted the decision model $M(\xi, w)$ where ξ is the matrix of criteria evaluations and w the vector of preference parameters representing the subjective preferences of the DM, with precise ξ and w, the decision model will produce precise results according to the problem setting. This means that under perfect information about the criteria evaluations and about precise values for the preference parameters, the decision-making problem is trivially solved by applying the decision model and accepting the recommended solution. However, in real-life decision problems, most of the associated information is uncertain or imprecise.

Stochastic multicriteria acceptability analysis (SMAA) is a family of MCDA methodologies for problems where the uncertainty is so significant that it should be considered explicitly. Incomplete criteria and preference information are represented by suitable probability distributions $f_{\chi}(\xi)$ and $f_W(w)$ where $\chi \subseteq \mathbb{R}^{n\times m}$, $\xi \in \chi$, and W is the set of parameters representing the preferences of the DM. Depending on the considered preference model, and consequently on the type of preference information provided, several variants of SMAA methods exist (for a complete survey on SMAA methods see [121]). These methods are based on exploring the weight space in order to describe the preferences that make each alternative the most preferred one, or that would give a certain rank to a specific alternative. In general, they can not determine a unique best solution, but some inferior solutions can be eliminated, because they do not correspond to any possible preferences, and widely accepted solutions favoured by a large variety of different preferences can be identified. In the following, we shall describe the characteristics of the SMAA methods using the weighted sum as preference model:

$$U(a_i, w) = \sum_{j=1}^{m} w_j g_j(a_i).$$

Chosen an alternative $a_i \in A$, a vector of weights $w \in W = \{w \in \mathbb{R}^m : w_j \geq 0, \forall j, \text{ and } \sum_{j=1}^{m} w_j = 1\}$ and a set of alternatives' evaluations $\xi \in \chi$, SMAA computes

$$rank(i, \xi, w) = 1 + \sum_{k \neq i} \rho(u(\xi_k, w) > u(\xi_i, w)),$$

where $\rho(true) = 1$ and $\rho(false) = 0$, and the set of favourable rank weights $W_i^r(\xi)$

$$W_i^r(\xi) = \{w \in W : rank(i, \xi, w) = r\}$$

where $r \in \{1, \ldots, n\}$; $rank(i, \xi, w)$ is the position alternative a_i gets in the final ranking when evaluations ξ and preference information represented by weights vector w are considered, while $W_i^r(\xi)$ measures the variety of weights that give rank r to alternative a_i. Based on $W_i^r(\xi)$, the main results of SMAA analysis are the rank acceptability indices, the central weight vectors and the confidence factors for each alternative.

- The rank acceptability index

  $$b_i^r = \int_{\xi \in \chi} f_X(\xi) \int_{w \in W_i^r(\xi)} f_W(w)\, dw\, d\xi$$

  describes the share of parameters giving to alternative a_i the position r in the final ranking obtained; in particular, b_i^1 measures the variety of weights that make alternative a_i most preferred. Obviously, the best alternatives are those having rank acceptability index greater than zero for the first positions and rank acceptability index close to zero for the lower positions.

- The central weight vector

$$w_i^c = \frac{1}{b_i^1} \int_{\xi \in \chi} f_X(\xi) \int_{w \in W_i^1(\xi)} f_W(w) w \; dw \; d\xi$$

describes the preferences of a typical DM that make alternative a_i the most preferred;

- The confidence factor

$$p_i^c = \int_{\substack{\xi \in \chi:\ u(\xi_i, w_i^c) \geq u(\xi_k, w_i^c) \\ \forall k=1,\dots,m}} f_X(\xi) \; d\xi$$

measures if the criteria measurements are accurate enough to discern the efficient alternatives.

All indices are computed solving multidimensional integrals, while approximations of these integrals are computed via Monte Carlo simulations.
SMAA methods have been applied to deal with many real decision making problems. See [84] for a summary of these applications.

Chapter 2

Interaction between criteria

As stated in section 1.1, the axiom of mutual preference independence between criteria is the basis for the use of the additive preference models. In many real world cases, the criteria are not independent and it is possible observing a certain form of positive (synergy) or negative (redundancy) interaction between them.

We have underlined that the Choquet integral is the most well known non-additive integral used to aggregate the alternatives' evaluations in case the criteria in a unipolar scale are interacting, while its extension, the bipolar Choquet integral, is used to deal with the aggregations of the alternatives' evaluations in case the criteria in a bipolar scale are interacting. In this section we illustrate our contributions aiming to extend the interaction of criteria concept in MCDA. In particular, in section 2.1 we present the SMAA-Choquet method in which the Choquet integral is considered as preference model and the SMAA method is used to investigate the ranking of the alternatives on varying the parameters compatible with some preference information provided by the DM. The bipolar PROMETHEE method extending the outranking method PROMETHEE to the case of interacting criteria in a bipolar scale is presented in section 2.2 while in section 2.3 is discussed MUSA-INT, that is the extension of customer satisfaction analysis method MUSA to the case of interacting criteria.

2.1 SMAA-Choquet: Stochastic Multicriteria Acceptability Analysis for the Choquet Integral

2.1.1 Introduction

In a multiple criteria decision problem (see [29] for a comprehensive state of the art), an alternative a_j, belonging to a finite set of n alternatives $A = \{a_1, a_2, \ldots, a_n\}$, is evaluated on the basis of a consistent family of m criteria $G = \{g_1, g_2, \ldots, g_m\}$. In our approach we make the assumption that each criterion $g_i \colon A \to \mathbb{R}$ is an interval scale of measurement. From here on, we will use the terms criterion g_i or criterion i interchangeably ($i = 1, 2, \ldots, m$). Without loss of generality, we assume that all the criteria have to be maximized.

We define a marginal weak preference relation as follows:

$$a_r \text{ is at least as good as } a_s \text{ with respect to criterion } i \;\Leftrightarrow\; g_i(a_r) \geq g_i(a_s).$$

The purpose of Multi-Attribute Utility Theory (MAUT) [80] is to represent the preferences of a Decision Maker (DM) on a set of alternatives A by an overall value function $U \colon \mathbb{R}^m \to \mathbb{R}$ with $U(g_1(a_r), \ldots, g_m(a_r)) = U(a_r)$:

- a_r is indifferent to $a_s \;\Leftrightarrow\; U(a_r) = U(a_s)$,
- a_r is preferred to $a_s \;\Leftrightarrow\; U(a_r) > U(a_s)$.

The principal aggregation model of value function is the multiple attribute additive utility [80]:

$$U(a_j) = u_1(g_1(a_j)) + u_2(g_2(a_j)) + \ldots + u_m(g_m(a_j)) \quad \text{with} \quad a_j \in A,$$

where u_i are non-decreasing marginal value functions for $i = 1, 2, \ldots, m$.

As it is well-known in literature, the underlying assumption of the preference independence of the multiple attribute additive utility is unrealistic since it does not permit to represent interaction between the criteria under consideration. In a decision problem we, usually, distinguish between positive and negative interaction among criteria, representing synergy and redundancy among criteria respectively. In particular, two criteria are synergic (redundant) when the comprehensive importance of these two criteria is greater (smaller) than the importance of the two criteria considered separately.

Within Multiple Criteria Decision Analysis (MCDA), the interaction of criteria has been considered in a decision model based upon a non-additive integral, *viz.* the Choquet integral [21] (see [39] for a comprehensive survey on the use of non-additive integrals in MCDA).

One of the main drawbacks of the Choquet integral decision model is the elicitation of its parameters representing the importance and interaction between criteria.

In literature, many multicriteria disaggregation procedures have been proposed to infer such parameters from the DM (see for example, [86] and [6]). Recently, an approach based on the determination of necessary and possible preference relations within the so-called *Robust Ordinal Regression* has been extended to the Choquet integral decision model (see [7]).

The principal aim of the work is to include eventual DM's uncertain preference information on the importance and interaction among criteria.

In this direction, we propose an extension of the Stochastic Multicriteria Acceptability Analysis (SMAA) [82, 83] to the Choquet integral decision model.

The section is organized as follows. In Section 2.1.2, we present the basic concepts relative to interaction between criteria and to the Choquet integral. In Section 2.1.3, we briefly describe the SMAA methods. An extension of the SMAA method to the Choquet integral decision model is introduced in Section 2.1.4 and illustrated by a didactic example in Section 2.1.5. Some conclusions and future directions of research are presented in Section 2.1.6.

2.1.2 The Choquet integral decision model

Let 2^G be the power set of G (*i.e.* the set of all subsets of G); a fuzzy measure (capacity) on G is defined as a set function $\mu : 2^G \rightarrow [0,1]$ satisfying the following properties:

1a) $\mu(\emptyset) = 0$ and $\mu(G) = 1$ (boundary conditions),

2a) $\forall\, T \subseteq R \subseteq G,\;\; \mu(T) \leq \mu(R)$ (monotonicity condition).

A fuzzy measure is said to be additive if $\mu(T \cup R) = \mu(T) + \mu(R)$, for any $T, R \subseteq G$ such that $T \cap R = \emptyset$. An additive fuzzy measure is determined uniquely by $\mu(\{1\}), \mu(\{2\}) \ldots, \mu(\{m\})$. In fact, in this case, $\forall\; T \subseteq G$, $\mu(T) = \displaystyle\sum_{i \in T} \mu(\{i\})$. In the other cases, we have to define a value $\mu(T)$ for every subset T of G, obtaining $2^{|G|}$ coefficients values. Therefore, we have to calculate the values of $2^{|G|} - 2$ coefficients, since we know that $\mu(\emptyset) = 0$ and $\mu(G) = 1$.

The Möbius representation of the fuzzy measure μ (see [99]) is defined by the function $a : 2^G \rightarrow \mathbb{R}$ (see [112]) such that:

$$\mu(R) = \sum_{T \subseteq R} a(T).$$

Let us observe that if R is a singleton, *i.e.* $R = \{i\}$ with $i = 1, \cdots, m$ then $\mu(\{i\}) = a(\{i\})$. If R is a couple (non-ordered pair) of criteria, *i.e.* $R = \{i, j\}$, then $\mu(\{i, j\}) = a(\{i\}) + a(\{j\}) + a(\{i, j\})$.

In general, the Möbius representation $a(R)$ is obtained by $\mu(R)$ in the following way:

$$a(R) = \sum_{T \subseteq R} (-1)^{|R-T|} \mu(T).$$

In terms of Möbius representation (see [20]), properties **1a)** and **2a)** are, respectively, formulated as:

1b) $a(\emptyset) = 0, \ \sum_{T \subseteq G} a(T) = 1,$

2b) $\forall\, i \in G$ and $\forall R \subseteq G \setminus \{i\}, \ \sum_{T \subseteq R} a(T \cup \{i\}) \geq 0.$

Let us observe that in MCDA, the importance of any criterion $g_i \in G$ should be evaluated considering all its global effects in the decision problem at hand; these effects can be "decomposed" from both theoretical and operational points of view in effects of g_i as single, and in combination with all other criteria. Therefore, a criterion $i \in G$ is important with respect to a fuzzy measure μ not only when it is considered alone, *i.e.* for the value $\mu(\{i\})$ in itself, but also when it interacts with other criteria from G, *i.e.* for every value $\mu(T \cup \{i\})$, $T \subseteq G \setminus \{i\}$.

Given $x \in A$ and μ being a fuzzy measure on G, then the *Choquet integral* [21] is defined by:

$$C_\mu(x) = \sum_{i=1}^{m} \left[\left(g_{(i)}(x) \right) - \left(g_{(i-1)}(x) \right) \right] \mu(A_i), \tag{2.1}$$

where $_{(\cdot)}$ stands for a permutation of the indices of criteria such that:

$g_{(1)}(x) \leq g_{(2)}(x) \leq ... \leq g_{(m)}(x)$, with $A_i = \{(i), ..., (m)\}$, $i = 1, .., m$, and $g_{(0)}(x) = 0$.

The Choquet integral can be redefined in terms of the Möbius representation [37], without reordering the criteria, as:

$$C_\mu(x) = \sum_{T \subseteq G} a(T) \min_{i \in T} g_i(x). \tag{2.2}$$

One of the main drawbacks of the Choquet integral is the necessity to elicitate and give an adequate interpretation of $2^{|G|} - 2$ parameters. In order to reduce the number of parameters to be computed and to eliminate a too strict description of the interactions among criteria, which is not realistic in many applications, the concept of fuzzy k-additive measure has been considered [40].

A *fuzzy measure* is called *k-additive* if $a(T) = 0$ with $T \subseteq G$, when $|T| > k$ and there exists at least one $T \subseteq G$, with $|T| = k$, such that $a(T) > 0$. We observe that a 1-additive measure is the common additive fuzzy measure. In many real decision problems, it suffices to consider 2-additive measures. In this case, positive and negative interactions between couples of criteria are modeled without considering the interaction among triples, quadruplets and generally m-tuples, (with $m > 2$) of criteria. From the point of view of MCDA, the use of 2-additive measures is justified by observing that the information on the importance of the single criteria and the interactions between couples of criteria are noteworthy. Moreover, it could be not easy or not straightforward for the DM to provide information on the interactions among three or more criteria during the decision procedure. From a computational point of view, the interest in the 2-additive measures lies in the fact that any decision model needs to evaluate a number $m + \binom{m}{2}$ of parameters (in terms of Möbius representation, a value $a(\{i\})$ for every criterion i and a value $a(\{i,j\})$ for every couple of distinct criteria $\{i,j\}$). With respect to a 2-additive fuzzy measure, the inverse transformation to obtain the fuzzy measure $\mu(R)$ from the Möbius representation is defined as:

$$\mu(R) = \sum_{i \in R} a\left(\{i\}\right) + \sum_{\{i,j\} \subseteq R} a\left(\{i,j\}\right), \ \forall R \subseteq G. \tag{2.3}$$

With regard to 2-additive measures, properties **1b)** and **2b)** have, respectively, the following formulations:

1c) $a\left(\emptyset\right) = 0, \ \sum_{i \in G} a\left(\{i\}\right) + \sum_{\{i,j\} \subseteq G} a\left(\{i,j\}\right) = 1,$

2c) $\begin{cases} a\left(\{i\}\right) \geq 0, \ \forall i \in G, \\ a\left(\{i\}\right) + \sum_{j \in T} a\left(\{i,j\}\right) \geq 0, \ \forall i \in G \text{ and } \forall \, T \subseteq G \setminus \{i\}\,, T \neq \emptyset. \end{cases}$

In this case, the representation of the Choquet integral of $x \in A$ is given by:

$$C_\mu(x) \quad = \quad \sum_{\{i\} \subseteq G} a\left(\{i\}\right)\left(g_i\left(x\right)\right) + \sum_{\{i,j\} \subseteq G} a\left(\{i,j\}\right) \min\{g_i\left(x\right), g_j\left(x\right)\}. \tag{2.4}$$

Finally, we recall the definitions of the importance and interaction indices for a couple of criteria. The Shapley value [113] expressing the importance of criterion $i \in G$, is given by:

$$\varphi\left(\{i\}\right) = a\left(\{i\}\right) + \sum_{j \in G \setminus \{i\}} \frac{a\left(\{i,j\}\right)}{2}, \quad i \in G, \tag{2.5}$$

The *interaction index* [90] expressing the sign and the magnitude of the sinergy in a couple of criteria $\{i, j\} \subseteq G$, in case of a 2-additive capacity μ, is given by:

$$\varphi(\{i, j\}) = a(\{i, j\}). \tag{2.6}$$

2.1.3 SMAA

A utility function gives a value to an alternative in order to represent its degree of desirability with respect to the decision problem under consideration; in its easiest form

$$u(a_j, w) = \sum_{i=1}^{m} w_i g_i(a_j)$$

this function depends on two sets of parameters: the set W of weight vectors relative to the set of criteria $G = \{g_1, g_2, \ldots, g_m\}$ and the set of evaluations $g_i(a_j)$ of alternative a_j with respect to the set of the considered criteria $g_i \in G$.
In order to find a vector of weights of the model, in literature two different techniques are used: direct and indirect. The direct technique consists of asking the DM to provide directly all these parameters; the indirect technique consists of asking the DM a set of information from which a set of parameters of the model can be elicited. In this context two situations can occur:

- the DM can not provide or does not want to provide this information,
- different DMs can provide different sets of parameters.

Stochastic Multicriteria Acceptability Analysis (SMAA) [82, 83] is a multicriteria decision support method taking into account this uncertainty or lack of information. These methods are based on exploring the set W of weight vectors and the space of alternatives' evaluations in order to state recommendation regarding a possible ranking obtained by the considered alternatives. For each weight vector $w \in W$, and for each set of alternatives' evaluations $\xi \in \chi$, where χ is the set of all vectors of possible evaluations with respect to considered criteria, SMAA computes the rank of an alternative a_j:

$$rank(j, \xi, w) = 1 + \sum_{k \neq j} \rho\left(u(\xi_k, w) > u(\xi_j, w)\right)$$

where $\rho(false) = 0$ and $\rho(true) = 1$.

Then, for each alternative a_j, for each alternatives' evaluations $\xi \in \chi$ and for each rank r, SMAA computes the set of weights of criteria for which alternative a_j assumes rank r:

$$W_j^r(\xi) = \{w \in W : rank(j, \xi, w) = r\}.$$

Imprecision and uncertainty are represented in SMAA as probability distributions: one f_W regarding weights of the set W and one f_χ regarding evaluations of the set χ.

The SMAA methodology is mainly based on the computation of three indices:

- the rank acceptability index:

$$b_j^r = \int_{\xi \in \chi} f_X(\xi) \int_{w \in W_j^r(\xi)} f_W(w)\, dw\, d\xi$$

  measuring for each alternative a_j and for each rank r the variety of different DM's preferences information giving to a_j the rank r. b_j^r is a real number bounded between 0 and 1; obviously, an alternative presenting large acceptability index for the best ranks will be preferred to an alternative having lower acceptability index for the same best ranks;

- the central weight vector:

$$w_j^c = \frac{1}{b_j^1} \int_{\xi \in \chi} f_X(\xi) \int_{w \in W_j^1(\xi)} f_W(w) w\, dw d\xi$$

  describing the middle preferences of the DMs giving to a_j the best position;

- the confidence factor:

$$p_j^c = \int_{\substack{\xi \in \chi:\ u(\xi_j, w_j^c) \geq u(\xi_k, w_j^c) \\ \forall k = 1, \ldots, m}} f_\chi(\xi)\, d\xi$$

  defined as the probability of an alternative to be the preferred one with the preferences expressed by its central weight vector.

From a computational point of view, the considered indices are evaluated by the multidimensional integrals approximated by using the Monte Carlo method.

All these sets of indices are considered simultaneously in order to help the DM to choose the better solution of the decision problem under consideration.

The first paper on SMAA [82], appeared in 1998, is a generalization to the $n-$dimensional case of two works of Bana E Costa [8, 9] providing an acceptability index only for the first rank; that is it calculates which weights configurations give to each alternative the best rank; its generalization is SMAA 2 [83] in which for each alternative it is computed not only the acceptability index corresponding to the best rank, but also the acceptability indices for all the other possible ranks. For a detailed survey on SMAA methods see [121].

2.1.4 An extension of the SMAA method to the Choquet integral decision model

As explained in Section 2.1.2, adopting the Choquet integral in terms of Möbius representation with a 2-additive measure as utility value of every alternative a_j, we need to estimate $m + \binom{m}{2}$ parameters (the Möbius measures).

Following a multicriteria disaggregation paradigm, preference information about importance and interaction of criteria given by the DM can be represented by the following system of linear constraints on G, denoted by E^{DM}:

$$\begin{cases} \varphi(\{g_i\}) > \varphi(\{g_j\}), & \text{if criterion } i \text{ is more important than criterion } j, \text{ with } i,j \in G, \\ \varphi(\{g_i, g_j\}) > 0, & \text{if criteria } i \text{ and } j \text{ are synergic with } i,j \in G, \\ \varphi(\{g_i, g_j\}) < 0, & \text{if criteria } i \text{ and } j \text{ are redundant with } i,j \in G, \\ \\ \left.\begin{array}{l} a(\{\emptyset\}) = 0, \ \sum\limits_{g_i \in G} a(\{g_i\}) + \sum\limits_{\{g_i,g_j\} \subseteq G} a(\{g_i, g_j\}) = 1, \\ a(\{g_i\}) \geq 0, \ \forall g_i \in G, \\ a(\{g_i\}) + \sum\limits_{g_j \in T} a(\{g_i, g_j\}) \geq 0, \forall g_i \in G \text{ and } \forall\, T \subseteq G \setminus \{g_i\} \end{array}\right\} (*) \end{cases}$$

where $(*)$ denotes the set of boundary and monotonicity constraints on the Möbius measures.

In order to explore the preference given by the Choquet integral within the set of compatible parameters, we integrate Choquet integral with SMAA. The sampling of compatible preference parameters (Möbius measures) is obtained by a Hit-and-Run procedure [116] on the set of constraints E^{DM} (see also [122] for a recent application of the above algorithm in multiple criteria decision analysis).

The rank acceptability index of every alternative is evaluated by considering the variety of different compatible preference parameters (the Möbius measures obtained after each iteration) giving to alternative $a_j \in A$ the rank r on the basis of a utility function expressed in terms of a Choquet integral. At the same time as preference parameters we compute the Möbius measures corresponding to the capacities for which the Choquet integral ranks every alternative a_j as the best.

2.1.5 A didactic example

In this Section, an example inspired from Barba-Romero and Pomerol [10] illustrates the multicriteria model explained in Section 2.1.4. We consider an executive manager (the DM) that has to hire an employee in her company. She evaluates a set $A = \{a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8, a_9\}$ of nine candidates on the basis of the following four criteria: educational degree (criterion g_1), professional experience (criterion g_2), age (criterion g_3), job interview (criterion g_4).
The evaluation performance matrix of the candidates is presented in Table 2.1.
The scores of every criterion are on a $[0, 10]$ scale and are supposed to be maximized.
Let us consider the following DM's partial preference information on the set of criteria G:

- criterion g_1 is more important than criterion g_2;
- criterion g_2 is more important than criterion g_3;
- criterion g_3 is more important than criterion g_4;
- there is a positive interaction between criteria g_1 and g_2;
- there is a positive interaction between criteria g_2 and g_3;
- there is a negative interaction between criteria g_2 and g_4.

The DM's preference constraints jointly with the boundary and monotonicity conditions are summarized in the following system:

$$
\begin{cases}
\varphi(\{g_1\}) > \varphi(\{g_2\}), \\
\varphi(\{g_2\}) > \varphi(\{g_3\}), \\
\varphi(\{g_3\}) > \varphi(\{g_4\}), \\
\varphi(\{g_1, g_2\}) > 0, \\
\varphi(\{g_2, g_3\}) > 0, \\
\varphi(\{g_2, g_4\}) < 0, \\
a(\{\emptyset\}) = 0, \quad \sum\limits_{g_i \in G} a(\{g_i\}) + \sum\limits_{\{g_i, g_j\} \subseteq G} a(\{g_i, g_j\}) = 1, \\
a(\{g_i\}) \geq 0, \ \forall g_i \in G, \\
a(\{g_i\}) + \sum\limits_{g_j \in T} a(\{g_i, g_j\}) \geq 0, \forall g_i \in G \text{ and } \forall\, T \subseteq G \setminus \{g_i\}\,.
\end{cases}
$$

Then, as explained in Section 2.1.4, we apply a Hit-and-Run sampling [116] for a number maximum of iterations denoted by MaxIter.

We evaluate the rank acceptabilities of every alternative, measuring on the basis of a Choquet integral the utility value of each alternative a_j and for each rank r, the variety of preference parameters (the Möbius measures obtained after each iteration) giving to a_j the rank r.

Considering MaxIter equal to 100.000, in the example the rank acceptabilities of the nine alternatives are displayed in Table 2.2.

Table 2.1: Evaluation matrix

	Alternatives								
Criteria	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9
g_1	8	3	10	5	8	5	8	5	0
g_2	6	1	9	9	0	9	10	7	10
g_3	7	10	0	2	8	4	5	9	2
g_4	5	10	5	9	6	7	7	4	8

Then we compute as central preference parameters the Möbius measures which give the best position to every alternative (in our example only a_1, a_3, a_7, see Table 2.3).

According to the obtained first rank acceptability of every alternative and their corresponding central preference parameters summarized in Table 2.3, we can observe that:

- on average, a_1 is the most preferred alternative if *educational degree* (g_1) is the most important criterion and there is a synergy between *professional experience* (g_2) and *age* (g_3),

- on average, a_3 is the most preferred alternative if *educational degree* is the most important criterion and there is synergy between *educational degree* and *professional experience*,

Table 2.2: Rank acceptabilities (b_i) in percentages

Alt	b_j^1	b_j^2	b_j^3	b_j^4	b_j^5	b_j^6	b_j^7	b_j^8	b_j^9
a_1	0.03	51.32	43.64	4.35	0.66	0.00	0.00	0.00	0.00
a_2	0.00	0.29	0.60	1.37	5.00	5.78	34.51	50.04	2.41
a_3	13.28	34.08	20.51	12.25	9.62	5.49	3.94	0.83	0.01
a_4	0.00	0.00	0.16	2.58	19.16	56.33	19.92	1.85	0.00
a_5	0.00	0.00	0.48	4.60	5.29	7.94	36.55	44.50	0.65
a_6	0.00	0.00	5.10	20.90	48.51	21.10	4.37	0.03	0.00
a_7	86.70	13.15	0.15	0.00	0.00	0.00	0.00	0.00	0.00
a_8	0.00	1.15	29.37	53.96	11.76	3.37	0.39	0.01	0.00
a_9	0.00	0.00	0.00	0.00	0.00	0.00	0.32	2.75	96.94

Table 2.3: First rank acceptability (b_1) and central weights

Alt	b_1	$a(\{1\})$	$a(\{2\})$	$a(\{3\})$	$a(\{4\})$	$a(\{1,2\})$	$a(\{1,3\})$	$a(\{1,4\})$	$a(\{2,3\})$	$a(\{2,4\})$	$a(\{3,4\})$
a_1	0.03	0.32799	0.1308	0.055755	0.18084	0.045792	0.1751	-0.10253	0.25613	-0.040672	-0.029203
a_3	13.28	0.4833	0.12818	0.19372	0.16042	0.21504	-0.12405	-0.044048	0.053126	-0.049708	-0.015979
a_7	86.70	0.23032	0.15372	0.14022	0.1788	0.2164	0.053909	0.0067834	0.092624	-0.063965	-0.0088056

- on average, a_7 is the most preferred alternative if all criteria are almost equally important except *educational degree* that is a bit more important than the other criteria, and there is synergy between *professional experience* and *age*, and redundancy between *age* and *job interview*.

2.1.6 Conclusions

In this section, we have considered an extension of the SMAA method to the Choquet integral decision model.

As future work, we plan to extend the SMAA-Choquet by including some DM's preference information on the alternatives such as a_i is preferred to a_j and considering the criteria expressed in possible ranges of evaluations. Moreover, the SMAA-Choquet method could be enriched calculating an index that gives the probability that an alternative a_r is better than an alternative a_s. Finally, the SMAA-Choquet method could be improved by coupling it with the approach of the *Robust Ordinal Regression* [60], and applying it to some more sophisticated fuzzy integrals such as the level dependent Choquet integral [50].

2.2 Interaction of Criteria and Robust Ordinal Regression in Bi-polar PROMETHEE Methods

2.2.1 Introduction

Multiple Criteria Decision Analysis (MCDA) dealing with the comparison of the reasons in favor and against a preference of an alternative a over an alternative b is of the utmost importance in a wide range of decision-making real world situations (for state-of-the-art surveys on MCDA see [29]). This kind of comparison is important, but it is only a part of the question. Indeed, after *recognizing the criteria in favor and the criteria against* of the preference of a over b, there is the very tricky question of comparing them (for a general discussion about bipolar aggregations of pros and cons in MCDA see [45]). In this second step, some important observations must be taken into account.

One element that should be considered is the *synergy* (a strengthening effect) or the *redundancy* (a weakening effect) of criteria in favor of a preference of an action a against an action b. For example, if one likes sport cars, maximum speed and acceleration are very important criteria. However, since in general speedy cars have also a good acceleration, giving a high weight to both criteria can over evaluate some cars. Thus, it seems reasonable to give maximum speed and acceleration considered together a weight smaller than the sum of the two weights given to these criteria when considered separately. In this case we have a redundancy between the criteria of maximum speed and acceleration. On the contrary, we have a synergy effect between maximum speed and price because, in general, speedy cars are also expensive and, therefore, a car which is good on both criteria is very appreciated. In this case, it seems reasonable to give maximum speed and price considered together a weight greater than the sum of the two weights given to these criteria when considered separately. Of course there could be similar effects of synergy and redundancy regarding the criteria against the comprehensive preference of a over b.

We have also to take into account the antagonism effects related to the fact that *the importance of criteria may also depend on the criteria which are opposed* to them. For example, a bad evaluation on aesthetics reduces the importance of maximum speed. Thus, the weight of maximum speed should be reduced when there is a negative evaluation on aesthetics. In this case, we have an antagonism effect between maximum speed and aesthetics.

Those types of interactions between criteria have been already taken into consideration in the ELECTRE methods [30]. In this work, we deal with the same problem using the bipolar Choquet integral [41, 42] applied to the PROMETHEE method [15, 17] (for the original Choquet integral see [21]).

This section is organized as follows. In section 2.2.2 we introduce the application of the bipolar Choquet integral to PROMETHEE method. In section 2.2.3, we discuss elicitation of preference information permitting to fix the value of the preference parameters of the model (essentially the bicapacities of the bipolar Choquet integral). Noting that generally there could be more than one bicapacity compatible with the preference information provided by the Decision Maker (DM), in section 2.2.4 we propose to adopt Robust Ordinal Regression (ROR) [60] that takes into account the whole set of compatible bicapacities. Within ROR we distinguish between necessary preference, in case of an alternative a is at least as good as an alternative b for all the compatible bicapacities, and the possible preference, in case of an alternative a is at least as good as an alternative b for at least one of the compatible bicapacities. In section 2.2.5, we present a didactic example showing how to use the bipolar PROMETHEE method. Besides, in the same example, we apply the concept of the most representative model [32] (see also [49, 75]) and the SMAA methodology [82] to our approach. Section 2.2.6 provides some conclusions and lines for future research.

2.2.2 The Bipolar PROMETHEE method

Let us consider a set of actions or alternatives $A = \{a, b, c, \ldots\}$ evaluated with respect to a set of criteria $G = \{g_1, \ldots, g_m\}$, where $g_j : A \to \mathbb{R}$, $j \in \mathcal{J} = \{1, \ldots, m\}$ and $|A| = n$. PROMETHEE [15, 17] is a well-known MCDA method that aggregates preference information of a DM through an outranking relation. Considering for each criterion g_j a weight w_j (representing the importance of criterion g_j within the family of criteria G), an indifference threshold q_j (being the largest difference $d_j(a,b) = g_j(a) - g_j(b)$ compatible with the indifference between alternatives a and b), and a preference threshold p_j (being the minimum difference $d_j(a,b)$ compatible with the preference of a over b), PROMETHEE builds a function $P_j(a,b)$ non decreasing with respect to $d_j(a,b)$, whose formulation (see [15] for other formulations) can be stated as follows

$$P_j(a,b) = \begin{cases} 0 & \text{if } \ d_j(a,b) \leq q_j \\ \frac{d_j(a,b)-q_j}{p_j-q_j} & \text{if } \ q_j < d_j(a,b) < p_j \\ 1 & \text{if } \ d_j(a,b) \geq p_j \end{cases}$$

It represents a degree of preference of a over b on criterion g_j.

For each ordered pair of alternatives $(a,b) \in A \times A$, PROMETHEE method computes the value

$$\pi(a,b) = \sum_{j \in \mathcal{J}} w_j P_j(a,b)$$

representing how much alternative a is preferred to alternative b taking into account the whole set of criteria. It can assume values between 0 and 1 and obviously the greater the value of $\pi(a,b)$, the greater the preference of a over b.

In order to compare an alternative a with all the other alternatives of the set A, PROMETHEE computes the negative and the positive net flow of a in the following way:

$$\phi^-(a) = \frac{1}{n-1}\sum_{c\in A\setminus\{a\}} \pi(c,a) \qquad \text{and} \qquad \phi^+(a) = \frac{1}{n-1}\sum_{c\in A\setminus\{a\}} \pi(a,c).$$

These net flows represent how much the alternatives of $A \setminus \{a\}$ are preferred to a and how much a is preferred to the alternatives of $A \setminus \{a\}$. Besides, PROMETHEE computes also the net flow $\phi(a) = \phi^+(a) - \phi^-(a)$. Taking into account these net flows, three relations can be built: preference $(\mathcal{P})$, indifference $(\mathcal{I})$, and incomparability $(\mathcal{R})$; PROMETHEE I and PROMETHEE II differ for the way in which these preference relations are defined.

In PROMETHEE I we have:

- $a\mathcal{P}b$ iff $\Phi^+(a) \geq \Phi^+(b)$, $\Phi^-(a) \leq \Phi^-(b)$ and at least one of the two inequalities is strict;
- $a\mathcal{I}b$ iff $\Phi^+(a) = \Phi^+(b)$ and $\Phi^-(a) = \Phi^-(b)$;
- $a\mathcal{R}b$ otherwise.

In PROMETHEE II instead, the comparison between alternatives a and b is done considering their net flows $\phi(a)$ and $\phi(b)$, having:

- $a\mathcal{P}b$ iff $\Phi(a) > \Phi(b)$;
- $a\mathcal{I}b$ iff $\Phi(a) = \Phi(b)$.

Within the bipolar framework, we can consider the bipolar preference functions $P_j^B : A \times A \rightarrow [-1,1], j \in \mathcal{J}$ as follows:

$$P_j^B(a,b) = P_j(a,b) - P_j(b,a) = \begin{cases} P_j(a,b) & \text{if} \quad P_j(a,b) > 0 \\ \\ -P_j(b,a) & \text{if} \quad P_j(a,b) = 0 \end{cases} \tag{2.7}$$

Notice that $P_j^B(a,b) = -P_j^B(b,a)$ for all $j \in \mathcal{J}$ and for all pairs $(a,b) \in A \times A$.

Determining comprehensive preferences

The aggregation of bipolar preference functions P_j^B through the bipolar Choquet integral is based on a bicapacity [41, 42], being a function $\hat{\mu} : P(\mathcal{J}) = \{(C, D) : C, D \subseteq \mathcal{J} \text{ and } C \cap D = \emptyset\} \to [-1, 1]$, such that

- $\hat{\mu}(\emptyset, \mathcal{J}) = -1, \hat{\mu}(\mathcal{J}, \emptyset) = 1, \hat{\mu}(\emptyset, \emptyset) = 0$,
- for all $(C, D), (E, F) \in P(\mathcal{J})$, if $C \subseteq E$ and $D \supseteq F$, then $\hat{\mu}(C, D) \leq \hat{\mu}(E, F)$.

According to [46], we consider the following expression for a bicapacity $\hat{\mu}$:

$$\hat{\mu}(C, D) = \mu^+(C, D) - \mu^-(C, D), \quad \text{for all } (C, D) \in P(\mathcal{J}) \tag{2.8}$$

where $\mu^+, \mu^- : P(\mathcal{J}) \to [0, 1]$ such that:

$$\mu^+(\mathcal{J}, \emptyset) = 1, \qquad \mu^+(\emptyset, B) = 0, \ \forall B \subseteq \mathcal{J}, \tag{2.9}$$

$$\mu^-(\emptyset, \mathcal{J}) = 1, \qquad \mu^-(B, \emptyset) = 0, \ \forall B \subseteq \mathcal{J}, \tag{2.10}$$

$$\mu^+(C, D) \leq \mu^+(E, F), \text{ for all } (C, D), \ (E, F) \in P(\mathcal{J}) : C \subseteq E, \ D \supseteq F, \tag{2.11}$$

$$\mu^-(C, D) \leq \mu^-(E, F), \text{ for all } (C, D), \ (E, F) \in P(\mathcal{J}) : C \supseteq E, \ D \subseteq F. \tag{2.12}$$

The interpretation of the functions μ^+ and μ^- is the following. Given the pair $(a, b) \in A \times A$, let us consider $(C, D) \in P(\mathcal{J})$ where C is the set of criteria expressing a preference of a over b and D the set of criteria expressing a preference of b over a. In this situation, $\mu^+(C, D)$ represents the importance of criteria from C when criteria from D are opposing them, and $\mu^-(C, D)$ represents the importance of criteria from D opposing C. Consequently, $\hat{\mu}(C, D)$ represents the balance of the importance of C supporting a and D supporting b.

Given $(a, b) \in A \times A$, the bipolar Choquet integral of preference functions $P_j^B(a, b)$ with respect to the bicapacity $\hat{\mu}$ can be written as follows

$$Ch^B(P^B(a, b), \hat{\mu}) = \int_0^1 \hat{\mu}(\{j \in \mathcal{J} : P_j^B(a, b) > t\}, \{j \in \mathcal{J} : P_j^B(a, b) < -t\})dt,$$

while the bipolar comprehensive positive preference of a over b and the comprehensive negative preference of a over b with respect to the bicapacity $\hat{\mu}$ are respectively:

$$Ch^{B+}(P^B(a,b),\hat{\mu}) = \int_0^1 \mu^+(\{j \in \mathcal{J} : P_j^B(a,b) > t\}, \{j \in \mathcal{J} : P_j^B(a,b) < -t\})dt,$$

$$Ch^{B-}(P^B(a,b),\hat{\mu}) = \int_0^1 \mu^-(\{j \in \mathcal{J} : P_j^B(a,b) > t\}, \{j \in \mathcal{J} : P_j^B(a,b) < -t\})dt,$$

where μ^+ and μ^- have been defined before.

From an operational point of view, the bipolar aggregation of the preferences $P_j^B(a,b)$ can be computed as follows: for all the criteria $j \in \mathcal{J}$, the absolute values of these preferences should be re-ordered in a non-decreasing way, as follows:

$$|P_{(1)}^B(a,b)| \leq |P_{(2)}^B(a,b)| \leq \ldots \leq |P_{(j)}^B(a,b)| \leq \ldots \leq |P_{(m)}^B(a,b)|$$

The bipolar comprehensive Choquet integral with respect to the bicapacity $\hat{\mu}$ for the pair $(a,b) \in A \times A$ can now be determined:

$$Ch^B(P^B(a,b),\hat{\mu}) = \sum_{j \in \mathcal{J}^>} |P_{(j)}^B(a,b)| \Big[\hat{\mu}(C_{(j)}, D_{(j)}) - \hat{\mu}(C_{(j+1)}, D_{(j+1)})\Big] \tag{2.13}$$

where $P^B(a,b) = \Big[P_j^B(a,b),\ j \in \mathcal{J}\Big]$, $\mathcal{J}^> = \{j \in \mathcal{J} \ : \ |P_{(j)}^B(a,b)| > 0\}$, $C_{(j)} = \{i \in \mathcal{J}^> \ : \ P_i^B(a,b) \geq |P_{(j)}^B(a,b)|\}$, $D_{(j)} = \{i \in \mathcal{J}^> \ : \ -P_i^B(a,b) \geq |P_{(j)}^B(a,b)|\}$ and $C_{(m+1)} = D_{(m+1)} = \emptyset$.
We could give also the following two definitions

$$Ch^{B+}(P^B(a,b),\mu^+) = \sum_{j \in \mathcal{J}^>} |P_{(j)}^B(a,b)| \Big[\mu^+(C_{(j)}, D_{(j)}) - \mu^+(C_{(j+1)}, D_{(j+1)})\Big], \tag{2.14}$$

$$Ch^{B-}(P^B(a,b),\mu^-) = \sum_{j \in \mathcal{J}^>} |P_{(j)}^B(a,b)| \Big[\mu^-(C_{(j)}, D_{(j)}) - \mu^-(C_{(j+1)}, D_{(j+1)})\Big]. \tag{2.15}$$

$Ch^B(P^B(a,b),\hat{\mu})$ gives the comprehensive preference of a over b and it is equivalent to $\pi(a,b) - \pi(b,a) = P^C(a,b)$ in the classical PROMETHEE method while $Ch^{B+}(P^B(a,b),\hat{\mu})$ and $Ch^{B-}(P^B(a,b),\hat{\mu})$ give, respectively, how much a outranks b (considering the reasons in favor of a) and how much a is outranked by b (considering the reasons against a).
From the definitions above, it is straightforward proving that, for all $a, b \in A$,

$$Ch^B(P^B(a,b),\hat{\mu}) = Ch^{B+}(P^B(a,b),\mu^+) - Ch^{B-}(P^B(a,b),\mu^-) \tag{2.16}$$

It is reasonable expecting that, for all $a, b \in A$, $Ch^B(P^B(a,b),\hat{\mu}) = -Ch^B(P^B(b,a),\hat{\mu})$. For this reason, we get the following Proposition:

Proposition 2.2.1. $Ch^B(P^B(a,b),\hat{\mu}) = -Ch^B(P^B(b,a),\hat{\mu})$ *for all possible* a, b, *iff* $\hat{\mu}(C,D) = -\hat{\mu}(D,C)$ *for each* $(C,D) \in P(\mathcal{J})$.

Proof. Let us prove that if $\hat{\mu}(C,D) = -\hat{\mu}(D,C)$, then $Ch^B(P^B(a,b),\hat{\mu}) = -Ch^B(P^B(b,a),\hat{\mu})$. As noticed, $P^B_j(a,b) = -P^B_j(b,a)$ for all $j \in \mathcal{J}$, and consequently $|P^B_{(j)}(a,b)| = |-P^B_{(j)}(b,a)| = |P^B_{(j)}(b,a)|$ for all $j \in \mathcal{J}$.

From this, it follows that:

(α) $C_{(j)}(a,b) = \{i \in \mathcal{J}^> \ : \ P^B_i(a,b) \geq |P^B_{(j)}(a,b)|\} = \{i \in \mathcal{J}^> \ : \ -P^B_i(b,a) \geq |P^B_{(j)}(b,a)|\} =$
$= D_{(j)}(b,a)$;

(β) $D_{(j)}(a,b) = \{i \in \mathcal{J}^> \ : \ -P^B_i(a,b) \geq |P^B_{(j)}(a,b)|\} = \{i \in \mathcal{J}^> \ : \ P^B_i(b,a) \geq |P^B_{(j)}(b,a)|\} =$
$= C_{(j)}(b,a)$.

From (α) and (β) we have that

(γ)
$$\begin{aligned} &Ch^B(P^B(a,b),\hat{\mu}) = \\ &= \sum_{j \in \mathcal{J}^>} |P^B_{(j)}(a,b)| \Big[\hat{\mu}(C_{(j)}(a,b), D_{(j)}(a,b)) - \hat{\mu}(C_{(j+1)}(a,b), D_{(j+1)}(a,b))\Big] = \\ &= \sum_{j \in \mathcal{J}^>} |P^B_{(j)}(b,a)| \Big[\hat{\mu}(D_{(j)}(b,a), C_{(j)}(b,a)) - \hat{\mu}(D_{(j+1)}(b,a), C_{(j+1)}(b,a))\Big]. \end{aligned}$$

Since $\hat{\mu}(C,D) = -\hat{\mu}(D,C)$, $\forall (C,D) \in P(\mathcal{J})$, from ($\gamma$) we have that,

(δ)
$$\begin{aligned} &Ch^B(P^B(b,a),\hat{\mu}) = \\ &= \sum_{j \in \mathcal{J}^>} |P^B_{(j)}(b,a)| \Big[\hat{\mu}(C_{(j)}(b,a), D_{(j)}(b,a)) - \hat{\mu}(C_{(j+1)}(b,a), D_{(j+1)}(b,a))\Big] = \\ &= \sum_{j \in \mathcal{J}^>} |P^B_{(j)}(b,a)| \Big[- \hat{\mu}(D_{(j)}(b,a), C_{(j)}(b,a)) + \hat{\mu}(D_{(j+1)}(b,a), C_{(j+1)}(b,a))\Big] \\ &= -Ch^B(P^B(a,b),\hat{\mu}). \end{aligned}$$

Let us now prove that if $Ch^B(P^B(a,b),\hat{\mu}) = -Ch^B(P^B(b,a),\hat{\mu})$, then $\hat{\mu}(C,D) = -\hat{\mu}(D,C)$. Let us consider the pair (a,b) such that,

$$P^B_j(a,b) = \begin{cases} 1 & \text{if } \ j \in C \\ -1 & \text{if } \ j \in D \\ 0 & \text{otherwise} \end{cases} \tag{2.17}$$

In this case we have that $Ch^B(P^B(a,b),\hat{\mu}) = \hat{\mu}(C,D)$ and $Ch^B(P^B(b,a),\hat{\mu}) = \hat{\mu}(D,C)$. Thus if $Ch^B(P^B(a,b),\hat{\mu}) = -Ch^B(P^B(b,a),\hat{\mu})$, from (iv) we obtain that $\hat{\mu}(C,D) = -\hat{\mu}(D,C)$ and the proof is concluded. $\square$

Analogously, it can be proved the following Proposition;

Proposition 2.2.2. *$Ch^{B+}(P^B(a,b),\mu^+) = Ch^{B-}(P^B(b,a),\mu^-)$ for all possible $a,b,$ iff $\mu^+(C,D) = \mu^-(D,C)$ for each $(C,D) \in P(\mathcal{J})$.*

Corollary 2.2.1. *$Ch^B(P^B(a,b),\hat{\mu}) = -Ch^B(P^B(b,a),\hat{\mu})$ for all possible $a,b,$ if $\mu^+(C,D) = \mu^-(D,C)$ for each $(C,D) \in P(\mathcal{J})$.*

Proof. This can be seen as a Corollary both of Proposition 2.2.1 and Proposition 2.2.2. In fact,

- $\mu^+(C,D) = \mu^-(D,C)$ for each $(C,D) \in P(\mathcal{J})$ implies that $\hat{\mu}(C,D) = -\hat{\mu}(D,C)$ for each $(C,D) \in P(\mathcal{J})$, and by Proposition 2.2.1, it follows the thesis.
- By Proposition 2.2.1, $\mu^+(C,D) = \mu^-(D,C)$ for each $(C,D) \in P(\mathcal{J})$ implies that $Ch^{B+}(P^B(a,b),\mu^+) = Ch^{B-}(P^B(b,a),\mu^-)$ and from this it follows obviously the thesis by equation (2.16).

$\square$

Using equations (2.13), (2.14) and (2.15), we can define:

$$\phi^B(a) = \frac{1}{n-1}\sum_{b \in A\setminus\{a\}} Ch^B(P^B(a,b),\hat{\mu}) \tag{2.18}$$

$$\phi^{B+}(a) = \frac{1}{n-1}\sum_{b \in A\setminus\{a\}} Ch^{B+}(P^B(a,b),\mu^+) \tag{2.19}$$

$$\phi^{B-}(a) = \frac{1}{n-1}\sum_{b \in A\setminus\{a\}} Ch^{B-}(P^B(a,b),\mu^-) \tag{2.20}$$

By equation (2.16), it follows that $\phi^B(a) = \phi^{B+}(a) - \phi^{B-}(a)$ for each $a \in A$.

Determining the importance, the interaction, and the power of the opposing criteria

Several studies dealing with the determination of the relative importance of criteria were proposed in MCDA (see e.g. [108]). The question of the interaction between criteria was also studied in the context of MAUT methods [86]. In this section we present a quite similar methodology for PROMETHEE method, which takes into account also the power of the opposing criteria.

The case of PROMETHEE method

The use of the bipolar Choquet integral is based on a bicapacity which assigns numerical values to each element $P(\mathcal{J})$. Let us remark that the number of elements of $P(\mathcal{J})$ is 3^m. This means that the definition of a bipolar capacity requires a rather huge and unpractical number of parameters. Moreover, the interpretation of these parameters is not always simple for the DM. Therefore, the use of the bipolar Choquet integral in real-world decision-making problems requires some methodology to assist the DM in assessing the preference parameters (bicapacities). In the following we consider only the 2-additive bicapacities [41, 34], being a particular class of bicapacities.

Defining a manageable and meaningful bicapacity measure

According to [46], we give the following decomposition of the functions μ^+ and μ^- previously defined:

- $\mu^+(C,D) = \sum_{j\in C} a^+(\{j\},\emptyset) + \sum_{\{j,k\}\subseteq C} a^+(\{j,k\},\emptyset) + \sum_{j\in C,\, k\in D} a^+(\{j\},\{k\})$
- $\mu^-(C,D) = \sum_{j\in D} a^-(\emptyset,\{j\}) + \sum_{\{j,k\}\subseteq D} a^-(\emptyset,\{j,k\}) + \sum_{j\in C,\, k\in D} a^-(\{j\},\{k\})$

The interpretation of each $a^{\pm}(\cdot)$ is the following:

- $a^+(\{j\},\emptyset)$, represents the power of criterion g_j by itself; this value is always non negative;
- $a^+(\{j,k\},\emptyset)$, represents the interaction between g_j and g_k, when they are in favor of the preference of a over b; when its value is zero there is no interaction; on the contrary, when the value is positive there is a synergy effect when putting together g_j and g_k; a negative value means that the two criteria are redundant;
- $a^+(\{j\},\{k\})$, represents the power of criterion g_k against criterion g_j, when criterion g_j is in favor of a over b and g_k is against to the preference of a over b; this leads always to a reduction or no effect on the value of μ^+ since this value is always non-positive.

An analogous interpretation can be applied to the values $a^-(\emptyset,\{j\})$, $a^-(\emptyset,\{j,k\})$, and $a^-(\{j\},\{k\})$.

In what follows, for the sake of simplicity, we will use a_j^+, a_{jk}^+, $a_{j|k}^+$ instead of $a^+(\{j\},\emptyset)$, $a^+(\{j,k\},\emptyset)$ and $a^+(\{j\},\{k\})$, respectively; and a_j^-, a_{jk}^-, $a_{j|k}^-$ instead of $a^-(\emptyset,\{j\})$, $a^-(\emptyset,\{j,k\})$ and $a^-(\{j\},\{k\})$, respectively, obtaining

$$
\begin{aligned}
\hat{\mu}(C,D) &= \mu^+(C,D) - \mu^-(C,D) = \\
&= \sum_{j\in C} a_j^+ - \sum_{j\in D} a_j^- + \sum_{\{j,k\}\subseteq C} a_{jk}^+ - \sum_{\{j,k\}\subseteq D} a_{jk}^- + \sum_{j\in C,\, k\in D} a_{j|k}^+ - \sum_{j\in C,\, k\in D} a_{j|k}^-
\end{aligned}
$$

We call $\hat{\mu}$, 2-additive decomposable bicapacity. An analogous decomposition has been proposed directly for $\hat{\mu}$ without considering μ^+ and μ^- in [35]. The following conditions should be fulfilled:

Monotonicity conditions

1) $\mu^+(C,D) \leq \mu^+(C \cup \{j\}, D), \;\; \forall\, j \in \mathcal{J}, \; \forall (C \cup \{j\}, D) \in P(\mathcal{J})$

$$\sum_{h \in C} a_h^+ + \sum_{\{h,k\} \subseteq C} a_{hk}^+ + \sum_{h \in C, k \in D} a_{h|k}^+ \leq \sum_{h \in C \cup \{j\}} a_h^+ + \sum_{\{h,k\} \subseteq C \cup \{j\}} a_{hk}^+ + \sum_{h \in C \cup \{j\}, k \in D} a_{h|k}^+ \Leftrightarrow$$

$$\Leftrightarrow \sum_{h \in C} a_h^+ + \sum_{\{h,k\} \subseteq C} a_{hk}^+ + \sum_{h \in C, k \in D} a_{h|k}^+ \leq a_j^+ + \sum_{h \in C} a_h^+ + \sum_{\{h,k\} \subseteq C} a_{hk}^+ + \sum_{k \in C} a_{jk}^+ + \sum_{h \in C, k \in D} a_{h|k}^+ + \sum_{k \in D} a_{j|k}^+ \Leftrightarrow$$

$$\Leftrightarrow a_j^+ + \sum_{k \in C} a_{jk}^+ + \sum_{k \in D} a_{j|k}^+ \geq 0, \;\; \forall\, j \in \mathcal{J}, \; \forall (C \cup \{j\}, D) \in P(\mathcal{J})$$

2) $\mu^+(C,D) \geq \mu^+(C, D \cup \{j\}), \;\; \forall\, j \in \mathcal{J}, \; \forall (C, D \cup \{j\}) \in P(\mathcal{J})$

$$\sum_{h \in C} a_h^+ + \sum_{\{h,k\} \subseteq C} a_{hk}^+ + \sum_{h \in C, k \in D} a_{h|k}^+ \geq \sum_{h \in C} a_h^+ + \sum_{\{h,k\} \subseteq C} a_{hk}^+ + \sum_{h \in C, k \in D \cup \{j\}} a_{h|k}^+ \Leftrightarrow$$

$$\Leftrightarrow \sum_{h \in C, k \in D} a_{h|k}^+ \geq \sum_{h \in C, k \in D} a_{h|k}^+ + \sum_{h \in C} a_{h|j}^+ \Leftrightarrow$$

$$\Leftrightarrow \sum_{h \in C} a_{h|j}^+ \leq 0, \;\; \forall\, j \in \mathcal{J}, \; \forall (C, D \cup \{j\}) \in P(\mathcal{J})$$

being already satisfied because $a_{h|j}^+ \leq 0$, $\forall h, j \in \mathcal{J}, h \neq j$. 1) and 2) are equivalent to the general monotonicity condition for μ^+ (see eq. (2.11)).

The same kind of monotonicity should be satisfied for μ^-.

3) $\mu^-(C,D) \leq \mu^-(C, D \cup \{j\}), \;\; \forall\, j \in \mathcal{J}, \; \forall (C, D \cup \{j\}) \in P(\mathcal{J})$

$$\sum_{h \in D} a_h^- + \sum_{\{h,k\} \subseteq D} a_{hk}^- + \sum_{h \in C, k \in D} a_{h|k}^- \leq \sum_{h \in D \cup \{j\}} a_h^- + \sum_{\{h,k\} \subseteq D \cup \{j\}} a_{hk}^- + \sum_{h \in C, k \in D \cup \{j\}} a_{h|k}^- \Leftrightarrow$$

$$\Leftrightarrow \sum_{h \in D} a_h^- + \sum_{\{h,k\} \subseteq D} a_{hk}^- + \sum_{h \in C, k \in D} a_{h|k}^- \leq \sum_{h \in D} a_h^- + a_j^- + \sum_{\{h,k\} \subseteq D} a_{hk}^- + \sum_{k \in D} a_{jk}^- + \sum_{h \in C, k \in D} a_{h|k}^- + \sum_{h \in C} a_{h|j}^- \Leftrightarrow$$

$$\Leftrightarrow a_j^- + \sum_{k \in D} a_{jk}^- + \sum_{h \in C} a_{h|j}^- \geq 0, \;\; \forall\, j \in \mathcal{J}, \; \forall (C, D \cup \{j\}) \in P(\mathcal{J})$$

4) $\mu^-(C,D) \geq \mu^-(C \cup \{j\}, D), \;\; \forall\, j \in \mathcal{J}, \; \forall (C \cup \{j\}, D) \in P(\mathcal{J})$

$$\sum_{h\in D} a^-_h + \sum_{\{h,k\}\subseteq D} a^-_{hk} + \sum_{h\in C, k\in D} a^-_{h|k} \geq \sum_{h\in D} a^-_h + \sum_{\{h,k\}\subseteq D} a^-_{hk} + \sum_{h\in C\cup\{j\}, k\in D} a^-_{h|k} \Leftrightarrow$$

$$\Leftrightarrow \sum_{h\in C, k\in D} a^-_{h|k} \geq \sum_{h\in C, k\in D} a^-_{h|k} + \sum_{k\in D} a^-_{j|k} \Leftrightarrow$$

$$\Leftrightarrow \sum_{k\in D} a^-_{j|k} \leq 0, \ \ \forall\, j \in \mathcal{J},\ \forall (C\cup\{j\}, D) \in P(\mathcal{J})$$

being already satisfied because $a^-_{h|j} \leq 0$, $\forall h, j \in \mathcal{J}, h \neq j$. 3) and 4) are equivalent to the general monotonicity condition for μ^- (see eq. (2.12)).

Conditions 1), 2), 3) and 4) ensure the monotonicity of the bi-capacity, $\hat{\mu}$, on $\mathcal{J}$, obtained as the difference of μ^+ and μ^-, that is,

$$\forall\ (C,D),\ (E,F)\ \in\ P(\mathcal{J})\ \ \text{such that}\ \ C \supseteq E,\ D \subseteq F,\ \ \hat{\mu}(C,D) \geq \hat{\mu}(E,F).$$

Boundary conditions

1. $\mu^+(\mathcal{J}, \emptyset) = 1$, i.e., $\sum_{j\in\mathcal{J}} a^+_j + \sum_{\{j,k\}\subseteq\mathcal{J}} a^+_{jk} = 1$
2. $\mu^-(\emptyset, \mathcal{J}) = 1$, i.e., $\sum_{j\in\mathcal{J}} a^-_j + \sum_{\{j,k\}\subseteq\mathcal{J}} a^-_{jk} = 1$

The 2-additive bipolar Choquet integral

The corollary of the following theorem expresses the bipolar Choquet integral in terms of the above 2-additive decomposition.

Theorem 2.2.1. *Given a 2-additive decomposable bicapacity $\hat{\mu}$, then for all $x \in \mathbb{R}^m$*

1. $Ch^{B+}(x,\mu^+) = \displaystyle\sum_{j\in\mathcal{J}, x_j>0} a^+_j x_j + \sum_{\substack{j,k\in\mathcal{J}, j\neq k: \\ : x_j, x_k>0}} a^+_{jk} \min\{x_j, x_k\} + \sum_{\substack{j,k\in\mathcal{J}, j\neq k: \\ : x_j>0, x_k<0}} a^+_{j|k} \min\{x_j, -x_k\}$
2. $Ch^{B-}(x,\mu^-) = -\displaystyle\sum_{j\in\mathcal{J}, x_j<0} a^-_j x_j - \sum_{\substack{j,k\in\mathcal{J}, j\neq k: \\ : x_j, x_k<0}} a^-_{jk} \max\{x_j, x_k\} - \sum_{\substack{j,k\in\mathcal{J}, j\neq k: \\ : x_j>0, x_k<0}} a^-_{j|k} \max\{-x_j, x_k\}$

Proof. We shall prove only point 1. Proof of point 2. can be obtained analogously.

If the bicapacity $\hat{\mu}$ is 2−additive decomposable, then

$$
\begin{aligned}
Ch^{B+}(x,\mu^{+}) &= \sum_{j\in\mathcal{J}^{>}} |x_{(j)}|\Big[\mu^{+}(C_{(j)},D_{(j)}) - \mu^{+}(C_{(j+1)},D_{(j+1)})\Big] = \\
&= \sum_{j\in\mathcal{J}^{>}} |x_{(j)}|\Big[\Big(\sum_{k\in\mathcal{J}^{>},x_k\geq|x_{(j)}|} a_k^{+} - \sum_{k\in\mathcal{J}^{>},x_k\geq|x_{(j+1)}|} a_k^{+}\Big) + \\
&+\Big(\sum_{h,k\in\mathcal{J}^{>},h\neq k,x_h,x_k\geq|x_{(j)}|} a_{hk}^{+} - \sum_{h,k\in\mathcal{J}^{>},h\neq k,x_h,x_k\geq|x_{(j+1)}|} a_{hk}^{+}\Big) + \\
&+\Big(\sum_{h,k\in\mathcal{J}^{>},h\neq k,x_h,-x_k\geq|x_{(j)}|} a_{h|k}^{+} - \sum_{h,k\in\mathcal{J}^{>},h\neq k,x_h,-x_k\geq|x_{(j+1)}|} a_{h|k}^{-}\Big)\Big]
\end{aligned}
$$

Let us remark that,

$$
a)\quad \Big(\sum_{k\in\mathcal{J}^{>},x_k\geq|x_{(j)}|} a_k^{+} - \sum_{k\in\mathcal{J}^{>},x_k\geq|x_{(j+1)}|} a_k^{+}\Big) = \begin{cases} \sum_{k\in\mathcal{J}^{>},x_k=|x_{(j)}|} a_k^{+} & \text{if} \quad |x_{(j)}| < |x_{(j+1)}| \\ 0 & \text{otherwise} \end{cases}
$$

$$
b)\quad \Big(\sum_{k\in\mathcal{J}^{>},-x_k\geq|x_{(j)}|} a_k^{-} - \sum_{k\in\mathcal{J}^{>},-x_k\geq|x_{(j+1)}|} a_k^{-}\Big) = \begin{cases} \sum_{k\in\mathcal{J}^{>},-x_k=|x_{(j)}|} a_k^{-} & \text{if} \quad |x_{(j)}| < |x_{(j+1)}| \\ 0 & \text{otherwise} \end{cases}
$$

$$
c)\quad \Big(\sum_{\substack{h,k\in\mathcal{J}^{>},h\neq k,\\ x_h,x_k\geq|x_{(j)}|}} a_{hk}^{+} - \sum_{\substack{h,k\in\mathcal{J}^{>},h\neq k,\\ x_h,x_k\geq|x_{(j+1)}|}} a_{hk}^{+}\Big) = \begin{cases} \sum_{\substack{h,k\in\mathcal{J}^{>},h\neq k,\\ \min\{x_h,x_k\}=|x_{(j)}|}} a_{hk}^{+} & \text{if} \quad |x_{(j)}| < |x_{(j+1)}| \\ 0 & \text{otherwise} \end{cases}
$$

Considering $a) - c)$ we get that:

$$
\chi) \;=\; \sum_{\substack{j\in\mathcal{J}^{>},\\ |x_{(j)}|<|x_{(j+1)}|}} |x_{(j)}|\Big[\sum_{k\in\mathcal{J}^{>},x_k=|x_{(j)}|} a_k^{+} + \sum_{\substack{h,k\in\mathcal{J}^{>},h\neq k,\\ \min\{x_h,x_k\}=|x_{(j)}|}} a_{hk}^{+} + \sum_{\substack{h,k\in\mathcal{J}^{>},h\neq k,\\ \min\{x_h,-x_k\}=|x_{(j)}|}} a_{h|k}^{+}\Big]
$$

and from this it follows the thesis. □

Proposition 2.2.3. *Given a 2-additive decomposable bicapacity $\hat{\mu}$, then $\hat{\mu}(C,D) = -\hat{\mu}(D,C)$ for each $(C,D) \in P(\mathcal{J})$ iff*

1. *for each $j \in \mathcal{J}$, $a_j^{+} = a_j^{-}$,*

2. *for each* $\{j,k\} \subseteq \mathcal{J}$, $a^+_{jk} = a^-_{jk}$,

3. *for each* $j,k \in \mathcal{J}$, $j \neq k$, $a^+_{j|k} - a^-_{j|k} = a^-_{k|j} - a^+_{k|j}$.

Proof. First, let us prove that

(a) $\hat{\mu}(C,D) = -\hat{\mu}(D,C)$

implies 1., 2. and 3. For each $j \in \mathcal{J}$,

(b) $\hat{\mu}(\{j\},\emptyset) = a^+_j$ and $\hat{\mu}(\emptyset,\{j\}) = -a^-_j$

From (a) and (b) we have,

$$a^+_j = \hat{\mu}(\{j\},\emptyset) = -\hat{\mu}(\emptyset,\{j\}) = a^-_j$$

which is 1.

For each $\{j,k\} \subseteq \mathcal{J}$ we have that,

(c) $\hat{\mu}(\{j,k\},\emptyset) = a^+_j + a^+_k + a^+_{jk}$ and $\hat{\mu}(\emptyset,\{j,k\}) = -a^-_j - a^-_k - a^-_{jk}$

Being $\hat{\mu}(\{j,k\},\emptyset) = -\hat{\mu}(\emptyset,\{j,k\})$, and being $a^+_j = a^-_j$ and $a^+_k = a^-_k$ by 1., we have that for each $\{j,k\} \subseteq \mathcal{J}$, $a^+_{jk} = a^-_{jk}$, i.e. 2.

For all $j,k \in \mathcal{J}$ with $j \neq k$, we have:

$$\hat{\mu}(\{j\},\{k\}) = a^+_j - a^-_k + a^+_{j|k} - a^-_{j|k}$$

$$\hat{\mu}(\{k\},\{j\}) = a^+_k - a^-_j + a^+_{k|j} - a^-_{k|j}$$

Being $\hat{\mu}(\{j\},\{k\}) = -\hat{\mu}(\{k\},\{j\})$ and having proved that $a^+_j = a^-_j, \forall j$, we obtain that $a^+_{j|k} - a^-_{j|k} = -a^+_{k|j} + a^-_{k|j}$ i.e. 3.

It is straightforward to prove that 1., 2., and 3. imply $\hat{\mu}(C,D) = -\hat{\mu}(D,C)$.

□

Corollary 2.2.2. *Given a 2-additive decomposable bicapacity* $\hat{\mu}$, $Ch^B(P^B(a,b),\hat{\mu}) = -Ch^B(P^B(b,a),\hat{\mu})$ *for all* $a,b \in A$ *iff*

1. *for each* $j \in \mathcal{J}$, $a^+_j = a^-_j$,

2. *for each* $\{j,k\} \subseteq \mathcal{J}$, $a^+_{jk} = a^-_{jk}$,

3. for each $j,k \in \mathcal{J}$, $j \neq k$, $a^+_{j|k} - a^-_{j|k} = a^-_{k|j} - a^+_{k|j}$.

Proof. It follows by Propositions 2.2.3 and 2.2.1. □

Proposition 2.2.4. *Given a 2-additive decomposable bicapacity* $\hat{\mu}$, *then* $\mu^+(C,D) = \mu^-(D,C)$ *for each* $(C,D) \in P(\mathcal{J})$ *iff*

1. *for each* $j \in \mathcal{J}$, $a^+_j = a^-_j$,
2. *for each* $\{j,k\} \subseteq \mathcal{J}$, $a^+_{jk} = a^-_{jk}$,
3. *for each* $j,k \in \mathcal{J}$, $j \neq k$, $a^+_{j|k} = a^-_{k|j}$.

Proof. Analogous to Proposition 2.2.3. □

Corollary 2.2.3. *Given a 2-additive decomposable bicapacity* $\hat{\mu}$, $Ch^{B+}(P^B(a,b),\mu^+) = Ch^{B-}(P^B(b,a),\mu^-)$ *for all* $a,b \in A$ *iff*

1. *for each* $j \in \mathcal{J}$, $a^+_j = a^-_j$,
2. *for each* $\{j,k\} \subseteq \mathcal{J}$, $a^+_{jk} = a^-_{jk}$,
3. *for each* $j,k \in \mathcal{J}$, $j \neq k$, $a^+_{j|k} = a^-_{k|j}$.

Proof. It follows by Propositions 2.2.4 and 2.2.2. □

2.2.3 Assessing the preference information

On the basis of the considered 2-additive decomposable bicapacity $\hat{\mu}$, and holding the symmetry condition in Corollary 2.2.2, we propose the following methodology which simplifies the assessment of the preference information.

We consider the following information provided by the DM and their representation in terms of linear constraints:

1. *Comparing pairs of actions.* The constraints represent some pairwise comparisons on a set of training actions. Given two actions a and b, the DM may prefer a to b, b to a or be indifferent to both:

   (a) the linear constraint associated with $a\mathcal{P}b$ is $Ch^B(P^B(a,b),\hat{\mu}) > 0$;

   (b) the linear constraint associated with $a\mathcal{I}b$ is $Ch^B(P^B(a,b),\hat{\mu}) = 0$.

2. *Comparison of the intensity of preferences between pairs of actions.* This comparison can be stated as follows:

$$Ch^B(P^B(a,b),\hat{\mu}) > Ch^B(P^B(c,d),\hat{\mu}) \quad \text{if} \quad (a,b)\mathcal{P}(c,d)$$

where, $(a,b)\mathcal{P}(c,d)$ means that the comprehensive preference of a over b is larger than the comprehensive preference of c over d.

3. *Importance of criteria.* A partial ranking over the set of criteria $\mathcal{J}$ may be provided by the DM:

   (a) criterion g_j is more important than criterion g_k, which leads to the constraint $a_j > a_k$;

   (b) criterion g_j is equally important to criterion g_k, which leads to the constraint $a_j = a_k$.

4. *The sign of interactions.* The DM may be able, for certain cases, to provide the sign of some interactions. For example, if there is a synergy effect when criterion g_j interacts with criterion g_k, the following constraint should be added to the model: $a_{jk} > 0$.

5. *Interaction between pairs of criteria.* The DM can provide some information about interaction between criteria:

   a) if the DM feels that interaction between g_j and g_k is greater than the interaction between g_p and g_q, the constraint should be defined as follows: $|a_{jk}| > |a_{pq}|$ where in particular:

      - if both couples of criteria are synergic then: $a_{jk} > a_{pq}$,
      - if both couples of criteria are redundant then: $a_{jk} < a_{pq}$,
      - if (j,k) is a couple of synergic criteria and (p,q) is a couple of redundant criteria, then: $a_{jk} > -a_{pq}$,
      - if (j,k) is a couple of redundant criteria and (p,q) is a couple of synergic criteria, then: $-a_{jk} > a_{pq}$.

   b) if the DM feels that the strength of the interaction between g_j and g_k is the same of the strength of the interaction between g_p and g_q, the constraint will be the following: $|a_{jk}| = |a_{pq}|$ and in particular:

      - if both couples of criteria are synergic or redundant then: $a_{jk} = a_{pq}$,
      - if one couple of criteria is synergic and the other is redundant then: $a_{jk} = -a_{pq}$,

6. *The power of the opposing criteria.* Concerning the power of the opposing criteria several situations may occur. For example:

   a) when the opposing power of g_k is larger than the opposing power of g_h, with respect to g_j, which expresses a positive preference, we can define the following constraint: $a^+_{j|k} < a^+_{j|h}$ (because $a^+_{j|h} \leq 0$ and $a^-_{j|h} \leq 0$ for all j, k with $j \neq k$);

   b) if the opposing power of g_k, expressing negative preferences, is larger with g_j rather than with g_h, the constraint will be $a^+_{j|k} < a^+_{h|k}$.

A linear programming model

All the constraints presented in the previous section along with the symmetry, boundary and monotonicity conditions can now be put together and form a system of linear constraints. Strict inequalities can be converted into weak inequalities by adding a variable ε. It is well-know that such a system has a feasible solution if and only if when maximizing ε, its value is strictly positive [86]. Considering constraints given by Corollary 2.2.3 for the symmetry condition, the linear programming model can be stated as follows (where $j\mathcal{P}k$ means that criterion g_j is more important than criterion g_k; the remaining relations have a similar interpretation):

$$\text{Max } \varepsilon$$

$$\left.\begin{array}{ll}
Ch^B(P^B(a,b),\hat{\mu}) \geq \varepsilon \text{ if } a\mathcal{P}b, & Ch^B(P^B(a,b),\hat{\mu}) = 0 \text{ if } a\mathcal{I}b, \\
Ch^B(P^B(a,b),\hat{\mu}) \geq Ch^B(P^B(c,d),\hat{\mu}) + \varepsilon \text{ if } (a,b)\mathcal{P}(c,d), & Ch^B(P^B(a,b),\hat{\mu}) = Ch^B(P^B(c,d),\hat{\mu}) \text{ if } (a,b)\mathcal{I}(c,d), \\
a_j - a_k \geq \varepsilon \text{ if } j\mathcal{P}k, & a_j = a_k \text{ if } j\mathcal{I}k, \\
|a_{jk}| - |a_{pq}| \geq \varepsilon \text{ if } \{j,k\}\mathcal{P}\{p,q\}, \text{ (see point 5.a) of the previous subsection)} & \\
|a_{jk}| = |a_{pq}| \text{ if } \{j,k\}\mathcal{I}\{p,q\}, \text{ (see point 5.b) of the previous subsection)} & \\
a_{jk} \geq \varepsilon \text{ if there is synergy between criteria } j \text{ and } k, & \\
a_{jk} \leq -\varepsilon \text{ if there is redundancy between criteria } j \text{ and } k, & \\
a_{jk} = 0 \text{ if criteria } j \text{ and } k \text{ are not interacting}, & \\
\text{Power of the opposing criteria of the type 6:} & \\
a^{+}_{j|k} - a^{+}_{j|p} \geq \varepsilon, & a^{-}_{j|k} - a^{-}_{j|p} \geq \varepsilon, \\
a^{+}_{j|k} - a^{+}_{p|k} \geq \varepsilon, & a^{-}_{j|k} - a^{-}_{p|k} \geq \varepsilon, \\
\text{Symmetry condition (point 3. of Proposition 2.2.3):} & \\
a^{+}_{j|k} = a^{-}_{k|j}, \ \forall\, j,k \in \mathcal{J}, j \neq k & \\
\text{Boundary and monotonicity conditions:} & \\
\sum_{j\in\mathcal{J}} a_j + \sum_{\{j,k\}\subseteq\mathcal{J}} a_{jk} = 1, & \\
a_j \geq 0 \ \forall\, j \in \mathcal{J}, & a^{+}_{j|k},\ a^{-}_{j|k} \leq 0 \ \forall\, j,k \in \mathcal{J}, \\
a_j + \sum_{k\in C} a_{jk} + \sum_{k\in D} a^{+}_{j|k} \geq 0, \ \forall\, j \in \mathcal{J}, \ \forall (C\cup\{j\}, D) \in P(\mathcal{J}), & \\
a_j + \sum_{k\in D} a_{jk} + \sum_{h\in C} a^{-}_{h|j} \geq 0, \ \forall\, j \in \mathcal{J}, \ \forall (C, D\cup\{j\}) \in P(\mathcal{J}). &
\end{array}\right\} E^{A^R}$$

Restoring PROMETHEE

The condition which allows to restore PROMETHEE is the following:

1. $\forall j,k \in \mathcal{J}, \ \ a_{jk} = a^{+}_{j|k} = a^{-}_{j|k} = 0.$

If Condition 1. is not satisfied and the following condition holds

2. $\forall j,k \in \mathcal{J}, a^{+}_{j|k} = a^{-}_{j|k} = 0,$

then the comprehensive preference of a over b is calculated as the difference between the Choquet integral of the positive preferences and the Choquet integral of the negative preferences, with a common capacity μ on $\mathcal{J}$ for the positive and the negative preferences, i.e. there exists $\mu : 2^{\mathcal{J}} \rightarrow [0,1]$, with $\mu(\emptyset) = 0$, $\mu(\mathcal{J}) = 1$, and $\mu(A) \leq \mu(B)$ for all $A \subseteq B \subseteq \mathcal{J}$, such that

$$Ch^B(P^B(a,b),\hat{\mu}) = \int_0^1 \mu(\{j \in \mathcal{J} : P^B_j(a,b) > t\})dt - \int_0^1 \mu(\{j \in \mathcal{J} : P^B_j(a,b) < -t\})dt.$$

We shall call this type of aggregation of preferences, the symmetric Choquet integral PROMETHEE method.

If neither 1. nor 2. are satisfied, but the following condition holds

3. $\forall j,k \in \mathcal{J}, a^{+}_{j|k} = a^{-}_{k|j}$,

then we have the Bipolar symmetric PROMETHEE method.

If conditions 1., 2. and 3. are not satisfied, then we have the Bipolar PROMETHEE method.

A constructive learning preference information elicitation process

The previous Conditions 1-2-3 suggest a proper way to deal with the linear programming model in order to assess the interactive bipolar criteria coefficients. Indeed, it is very wise to try before to elicit weights concordant with the classical PROMETHEE method. If this is not possible, one can consider the symmetric Choquet integral PROMETHEE method which aggregates positive and negative preferences using the same capacity. If, by proceeding in this way, we are not able to represent the DM's preferences, then we can take into account a more sophisticated aggregation procedure by using the bipolar PROMETHEE method. This way to progress from the simplest to the most sophisticated model can be outlined in a four steps procedure as follows:

1. Solve the linear programming problem

$$\text{Max } \varepsilon = \varepsilon_1 \\ \left.\begin{array}{l} E^{A^R} \\ a_{jk} = a^{+}_{j|k} = a^{-}_{j|k} = 0, \quad \forall j,k \in \mathcal{J} \end{array}\right\} E_1 \tag{2.21}$$

   adding to E^{A^R} the constraint related to the previous Condition 1. If E_1 is feasible and $\varepsilon_1 > 0$, the obtained preferential parameters are concordant with the classical PROMETHEE method. Otherwise,

2. Solve the linear programming problem

$$\text{Max } \varepsilon = \varepsilon_2 \\ \left.\begin{array}{l} E^{A^R} \\ a^{+}_{j|k} = a^{-}_{j|k} = 0, \quad \forall j,k \in \mathcal{J} \end{array}\right\} E_2 \tag{2.22}$$

adding to E^{A^R} the constraint related to the previous Condition 2. If E_2 is feasible and $\varepsilon_2 > 0$, the information is concordant with the symmetric Choquet integral PROMETHEE method having a unique capacity for the negative and the positive part. Otherwise,

3. Solve the linear programming problem

$$\begin{aligned} &\text{Max } \varepsilon = \varepsilon_3 \\ &E^{A^R} \end{aligned} \tag{2.23}$$

   If E^{A^R} is feasible and $\varepsilon_3 > 0$, then the preference information is concordant with the bipolar PROMETHEE method. Otherwise,

4. We can try to help the DM by providing some information about inconsistent judgments, when it is the case, by using a similar constructive learning procedure proposed in [87]. In fact, in the linear programming model some of the constraints cannot be relaxed, that is, the basic properties of the model (symmetry, boundary and monotonicity conditions). The remaining constraints can lead to an infeasible linear system which means that the DM provided inconsistent information about her/his preferences. The methods proposed in [87] can then be used in this context, providing to the DM some useful information about inconsistent judgments.

2.2.4 ROR applied to Bipolar PROMETHEE method

In above sections we dealt with the problem of finding a bicapacity restoring preference information provided by the DM in case where multiple criteria evaluations are aggregated by bipolar PROMETHEE method. Generally, there could exist more than one model (in our case the model will be a bicapacity, but in other contexts it could be a utility function or an outranking relation) compatible with the preference information provided by the DM on the training set of alternatives. Each compatible model restores the preference information provided by the DM but two different compatible models could compare the other alternatives not provided as examples by the DM in a different way. For this reason, the choice of one of these models among those compatible could be considered arbitrary. In order to take into account not only one but the whole set of models compatible with the preference information provided by the DM, we consider the ROR [60]. This approach considers the whole set of models compatible with preference information provided by the DM building two preference relations: the weak *necessary* preference relation, for which alternative a is necessarily weakly preferred to alternative b (and we write $a \succsim^N b$), if a is at least as good as b for all compatible models, and the weak *possible* preference relation, for which alternative a is possibly weakly preferred to alternative b (and we write $a \succsim^P b$), if a is at least as good as b for at least one compatible model.

In the case of the bipolar PROMETHEE methods and considering a 2-additive decomposable bi-capacity $\hat{\mu}$, we could define one local outranking and two global outrankings (one for the bipola PROMETHEE I and one for the bipolar PROMETHEE II) of an alternative a over an alternative b. We say that, a is at least as good as b,

- locally, if $Ch^B(P^B(a,b),\hat{\mu}) \geq 0$;
- globally and considering the bipolar PROMETHEE I method, if $\Phi^{B+}(a) \geq \Phi^{B+}(b)$, $\Phi^{B-}(a) \leq \Phi^{B-}(b)$ and at least one of the two inequalities is strict;
- globally and considering the bipolar PROMETHEE II method, if $\Phi^{B}(a) \geq \Phi^{B}(b)$.

To check if a is necessarily preferred to b, we look if it is possible that a does not outrank b. Locally, this means that it is possible that there exists a bicapacity $\hat{\mu}$ such that $Ch^B(P^B(b,a),\hat{\mu}) > 0$; globally, considering the bipolar PROMETHEE I method this means that $\Phi^{B+}(a) < \Phi^{B+}(b)$ or $\Phi^{B-}(a) > \Phi^{B-}(b)$, while considering the bipolar PROMETHEE II method this means that $\Phi^{B}(a) < \Phi^{B}(b)$. Given the following set of constraints,

$$\left.\begin{array}{l}
E^{A^R} \\
\text{if one verifies the truth of global outranking:} \\
\quad \text{if exploited in the way of the bipolar PROMETHEE II method, then:} \\
\quad\quad \Phi^B(a) + \varepsilon \leq \Phi^B(b) \\
\quad \text{if exploited in the way of the bipolar PROMETHEE I method, then:} \\
\quad\quad \Phi^{B+}(a) + \varepsilon \leq \Phi^{B+}(b) + 2M_1 \;\text{ and }\; \Phi^{B-}(a) + 2M_2 \geq \Phi^{B-}(b) + \varepsilon \\
\quad\quad\quad \text{where } M_i \in \{0,1\}, i = 1,2, \text{ and } \sum_{i=1}^{2} M_i \leq 1 \\
\text{if one verifies the truth of local outranking:} \\
\quad\quad Ch^B(P^B(b,a),\hat{\mu}) \geq \varepsilon
\end{array}\right\} E^N(a,b)$$

we say that a is weakly necessarily preferred to b if $E^N(a,b)$ is infeasible or $\varepsilon^* \leq 0$ where $\varepsilon^* = \max \varepsilon$ s.t. $E^N(a,b)$.

To check if a is possibly preferred to b, we check if it is possible that a outrank b for at least one bicapacity $\hat{\mu}$. Partially, this means that there exists a bicapacity $\hat{\mu}$ such that $Ch^B(P^B(a,b),\hat{\mu}) \geq 0$; globally, considering the bipolar PROMETHEE I method this means that $\Phi^{B+}(a) \geq \Phi^{B+}(b)$ and $\Phi^{B-}(a) \leq \Phi^{B-}(b)$ and at least one of the two inequalities is strict, while considering the bipolar PROMETHEE II method this means that $\Phi(a) \geq \Phi(b)$. Given the following set of constraints,

$$\left.\begin{array}{l}
E^{A^R} \\
\text{if one verifies the truth of global outranking:} \\
\quad \text{if exploited in the way of the bipolar PROMETHEE II method, then:} \\
\quad\quad \Phi^B(a) \geq \Phi^B(b) \\
\quad \text{if exploited in the way of the bipolar PROMETHEE I method, then:} \\
\quad\quad \Phi^{B+}(a) \geq \Phi^{B+}(b) \;\; \text{and} \;\; \Phi^{B-}(a) \leq \Phi^{B-}(b) \\
\text{if one verifies the truth of local outranking:} \\
\quad\quad Ch^B(P^B(a,b),\hat{\mu}) \geq 0
\end{array}\right\} E^P(a,b)$$

we say that a is weakly possibly preferred to b if $E^P(a,b)$ is feasible and $\varepsilon^* > 0$ where $\varepsilon^* = \max \varepsilon$ s.t. $E^P(a,b)$.

2.2.5 Didactic Example, the Most Representative Model and SMAA method

Inspired by a famous example in literature [39], let us consider the problem of evaluating High School students according to their grades in Mathematics, Physics and Literature. In the following we suppose that the Director is the DM, while we will cover the role of analyst helping and supporting the DM in (her)his evaluations.

The Director thinks that scientific subjects (Mathematics and Physics) are more important than Literature. However, when students a and b are compared, if a is better than b both at Mathematics and Physics but a is much worse than b at Literature, then the Director has some doubts about the comprehensive preference of a over b.

Mathematics and Physics are in some sense *redundant* with respect to the comparison of students, since usually students which are good at Mathematics are also good at Physics. As a consequence,

if a is better than b at Mathematics, the comprehensive preference of the student a over the student b is stronger if a is better than b at Literature rather than if a is better than b at Physics.
Let us consider the students whose grades (belonging to the range $[0, 20]$) are represented in Table 2.4 and the following formulation of the preference of a over b with respect to each criterion g_j, for all $j = (M)$ Mathematics, (Ph) Physics, (L) Literature.

Students	Mathematics	Physics	Literature
s_1	16	16	16
s_2	15	13	18
s_3	19	18	14
s_4	18	16	15
s_5	15	16	17
s_6	13	13	19
s_7	17	19	15
s_8	15	17	16

Table 2.4: Evaluations of the students

$$P_j(a,b) = \begin{cases} 0 & \text{if} \quad g_j(b) \geq g_j(a) \\ (g_j(a) - g_j(b))/4 & \text{if} \quad 0 < g_j(a) - g_j(b) \leq 4 \\ 1 & \text{otherwise} \end{cases}$$

From the values of the partial preferences $P_j(a,b)$, we obtain the positive and the negative partial preferences $P_j^B(a,b)$ with respect to each criterion g_j, for $j = M, Ph, L$ using the definition (2.7). Thus, to each pair of students (s_i, s_j) is associated a vector of three elements:
$P^B(s_i, s_j) = \left[P_M^B(s_i, s_j), P_{Ph}^B(s_i, s_j), P_L^B(s_i, s_j)\right]$; for example, to the pair of students (s_1, s_2) is associated the vector $P^B(s_1, s_2) = [0.25, 0.75, -0.5]$.

In order to apply the Bipolar PROMETHEE method and taking into account the conditions for the symmetry given by the Corollary 2.2.3, we need all the parameters regarding the degree of importance of each criterion (a_1, a_2 and a_3), the intensity of the interaction for each couple of criteria (a_{12}, a_{13} and a_{23}), and the power of the opposing criteria both in the case of positive and negative preferences for each pair of criteria ($a^+_{j|k}$ and $a^-_{j|k}$ with $j, k = (L), (Ph), (M)$ and $j \neq k$).
In a first moment, let us suppose that the Director is able to give a direct preference information, providing the parameters shown in Table 2.5;

For each pair of students (s_i, s_j), using the parameters in Table 2.5 and the formulation of the Bipolar Choquet integral of Theorem 2.2.1 applied to the Bipolar vector $P^B(s_i, s_j)$, we obtain the results shown in Table 2.6 where the element corresponding to the raw s_i and to the column s_j is

a_1	a_2	a_3
0.96	0.83	0.85

a_{12}	a_{13}	a_{23}
-0.61	-0.58	-0.46

$a^+_{1\|2}$	$a^+_{1\|3}$	$a^+_{2\|1}$	$a^+_{2\|3}$	$a^+_{3\|1}$	$a^+_{3\|2}$
-0.23	-0.19	-0.20	-0.17	-0.16	-0.13

$a^-_{1\|2}$	$a^-_{1\|3}$	$a^-_{2\|1}$	$a^-_{2\|3}$	$a^-_{3\|1}$	$a^-_{3\|2}$
-0.23	-0.17	-0.20	-0.10	-0.18	-0.20

Table 2.5: Direct information provided by the DM

$Ch^B(P^B(s_i, s_j), \hat{\mu})$ being the difference $\pi(s_i, s_j) - \pi(s_j, s_i)$ in the classical PROMETHEE method.

	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8
s_1		0.245	-0.36	-0.2625	0.0225	0.18	-0.475	0.0325
s_2	-0.245		-0.24	-0.18	-0.3925	0.2625	-0.305	-0.37
s_3	0.36	0.24		0.2675	0.3825	0.24	0.17	0.5625
s_4	0.2625	0.18	-0.2675		0.285	0.2025	-0.3825	0.41
s_5	-0.0225	0.3925	-0.3825	-0.285		0.3275	-0.3275	0.0225
s_6	-0.18	-0.2625	-0.24	-0.2025	-0.3275		-0.24	-0.305
s_7	0.475	0.305	-0.17	0.3825	0.3275	0.24		0.355
s_8	-0.0325	0.37	-0.5625	-0.41	-0.0225	0.305	-0.355	

Table 2.6: $Ch^B(P^B(s_i, s_j), \hat{\mu})$ obtained applying the bipolar PROMETHEE method with the parameters obtained directly by the Director

In Figure 2.1, an arrow is directed from s_i to s_j if $Ch^B(P^B(s_i, s_j), \hat{\mu}) > 0$ while in Table 2.7 it is shown a complete ranking of the eight students obtained by the bipolar PROMETHEE II method and using the net flow formulation of equation (2.18); from this we get that s_3 is the best student while s_6 is the worst one.

Now, let us suppose that the Director cannot provide the parameters necessary to implement the Bipolar PROMETHEE method and so we ask (her)him to provide an indirect preference information. (S)he gives this information:

- student s_1 is preferred to student s_2 more than student s_3 is preferred to student s_4,
- student s_7 is preferred to student s_8 more than student s_5 is preferred to student s_6.

Using the Bipolar PROMETHEE method, this preference information is translated as:

- $Ch^B(P^B(s_1, s_2), \hat{\mu}) > Ch^B(P^B(s_3, s_4), \hat{\mu})$,

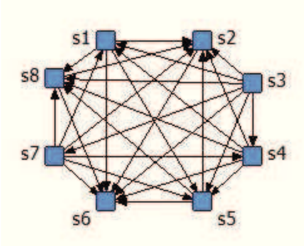


Figure 2.1: Comparison between all pairs of students obtained by the Bipolar PROMETHEE method applied with the parameters provided by the DM

Alternative/Net Flow	$\Phi(s)$
s_3	0.3175
s_7	0.2736
s_4	0.0986
s_5	-0.0393
s_1	-0.0882
s_8	-0.1011
s_2	-0.21
s_6	-0.2511

Table 2.7: Total order of the eight students obtained by the bipolar PROMETHEE II method and the net flows given by equation (2.18)

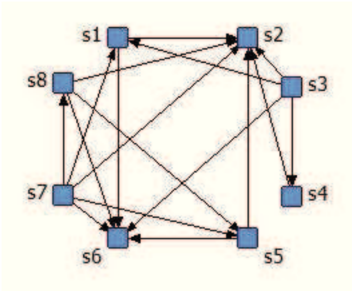


Figure 2.2: Necessary preference relation obtained after the first piece of preference information

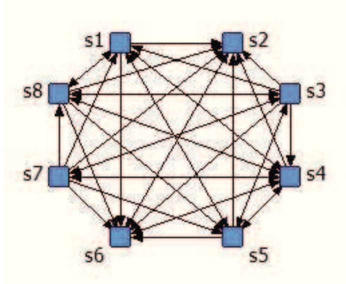


Figure 2.3: Possible preference relation obtained after the first piece of preference information

- $Ch^B(P^B(s_7, s_8), \hat{\mu}) > Ch^B(P^B(s_5, s_6), \hat{\mu})$.

At the beginning, we check if the preferences expressed by the Director can be explained by the classical PROMETHEE method or by the symmetric Choquet integral PROMETHEE method. Solving the linear programming problems (2.21) and (2.22) we obtain $\varepsilon_1 < 0$ and $\varepsilon_2 < 0$ and therefore neither the classical PROMETHEE method nor the symmetric Choquet integral PROMETHEE method are able to explain the preference information provided by the Director. For this reason, we decide to use the Bipolar PROMETHEE method.

Applying the ROR methodology to the Bipolar PROMETHEE II method and considering the preference information provided by the Director, we can show to (her)him the necessary and possible preference relations in Figure 2.2 and 2.3 respectively. In Figure 2.2 an arrow connects the pair of students (s_i, s_j) if student s_i is necessarily preferred to student s_j while in Figure 2.3 an arrow connects the pair of students (s_i, s_j) if student s_i is possibly preferred to student s_j.

Looking at the necessary and possible preference relations shown in Figures 2.2 and 2.3, the Director thinks, without any doubt, that s_3 and s_7 are the best students because s_7 is necessarily preferred to five out of the other seven students while s_3 is necessarily preferred to four out of the

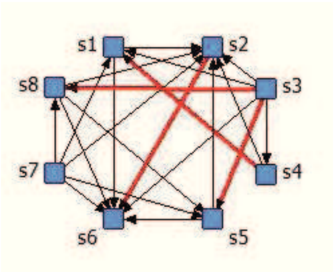


Figure 2.4: Necessary preference relation obtained after the second piece of preference information

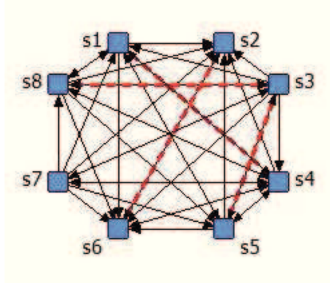


Figure 2.5: Possible preference relation obtained after the second piece of preference information

other seven students. At the same time only two students (s_3 and s_4) are possibly preferred to s_7 while three students (s_5, s_7 and s_8) are possibly preferred to s_3.

In order to enrich (her)his understanding about the problem at hand, the Director decides to provide another preference information regarding students s_4 and s_1 who are not connected by any arrow in the necessary preference relation shown in Figure 2.2, so (s)he states that student s_4 is preferred to student s_1.

Including this new preference information, translated by the constraint $Ch^B(P^B(s_4, s_1), \hat{\mu}) > 0$, we can show to the Director the necessary and possible preference relations in Figures 2.4 and 2.5 respectively.

Because increasing the number of preference information it reduces the dimension of the space of the models compatible with the preference information provided by the Director, the number of pairs in the necessary preference relation will increase (or at least it will be the same of the previous one), while the number of pairs in the possible preference relation will decrease (or at most it will be the same of the previous one). In Figure 2.4, each red marked arrow connects a pair of students that was not present in the previous necessary preference relation while each red dashed marked arrow in Figure 2.5 connects a pair of students that was present in the previous Possible preference relation but it is not present anymore. In this way, we can provide the Director of the information that student s_3 is now necessarily preferred also to students s_5 and s_8 while both of them are not anymore possibly preferred to s_3. At the same time the Director does not get any new information regarding student s_7, so (s)he begins to think that s_3 could be considered the best student.

This procedure could continue until the Director is not satisfied of the obtained result; therefore (s)he could give other preference information obtaining new necessary and possible preference relations.

The Director is able to give a scholarship to one of the eight students, so (s)he is interested to have a complete order of them. Because we would like to take into account the necessary and possible preference relations that we have obtained using the ROR, we could extend the concept of most representative model to the Bipolar PROMETHEE method. The most representative model is obtained using the results of necessary and possible preference relations of ROR in case the DM is interested to get a unique model representative of all the models compatible with the preference information provided by the DM. The concept of most representative model has been already introduced for ranking and choice problems [32, 75] and also for sorting problems [49]. Based on the necessary and possible preference relations, the most representative model (in our case the model will be a bicapacity) is obtained by maximizing the value of $Ch^B(P^B(a,b),\hat{\mu})$ for each pair of alternatives (a,b) such that a is necessarily preferred to b but b is not necessarily preferred to a, and by minimizing the value of $Ch^B(P^B(a,b),\hat{\mu})$ for each pair of alternatives (a,b) for which neither a is necessarily preferred to b nor b is necessarily preferred to a. The procedure to compute the most representative model is described in the following:

Step 1)

$$\max \varepsilon = \varepsilon^*, \quad \text{s.t.} \left.\begin{array}{l} E^{A^R} \\ Ch(P^B(a,b),\hat{\mu}) \geq \varepsilon \ \text{ if } \ a \succsim^N b \ \text{ and } \ b \not\succsim^N a \end{array}\right\}$$

Step 2)

$$\min \delta = \delta^* \quad \left.\begin{array}{l} \varepsilon = \varepsilon^* \\ E^{A^R} \\ Ch(P^B(a,b),\hat{\mu}) \geq \varepsilon \ \text{ if } \ a \succsim^N b \ \text{ and } \ b \not\succsim^N a \\ \left.\begin{array}{l} Ch(P^B(a,b),\hat{\mu}) \leq \delta \\ Ch(P^B(b,a),\hat{\mu}) \leq \delta \end{array}\right\} \text{if } \ a \not\succsim^N b \ \text{ and } \ b \not\succsim^N a \end{array}\right\}$$

Following the described procedure, we obtain the parameters shown in the Table 2.8. Applying the bipolar PROMETHEE II method with these parameters, we get the values $Ch^B(B^P(a,b),\hat{\mu})$ shown in Table 2.9 while in Figure 2.6 an arrow connects a pair of students (s_i,s_j) if $Ch^B(P^B(s_i,s_j),\hat{\mu}) > 0$. We can apply PROMETHEE II with the net flow formulation of equation (2.18) and getting the total order presented in Table 2.10. From this Table we get that s_7 is the best student, while s_2 is the worst one.

Because the Director would like to be very aware of the final result, we could provide (her)him with other information regarding the eight considered students. In the classical regression method and in the most representative model, we consider only one model compatible with the preference informa-

a_1	a_2	a_3
0.6667	0.5	0.2768

a_{12}	a_{13}	a_{23}
0	-0.1667	-0.2768

$a^+_{1\|2}$	$a^+_{1\|3}$	$a^+_{2\|1}$	$a^+_{2\|3}$	$a^+_{3\|1}$	$a^+_{3\|2}$
0	-0.6667	0	0	0	-0.1101

$a^-_{1\|2}$	$a^-_{1\|3}$	$a^-_{2\|1}$	$a^-_{2\|3}$	$a^-_{3\|1}$	$a^-_{3\|2}$
0	0	0	-0.1101	-0.6667	0

Table 2.8: Parameters obtained using the most representative model

	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8
s_1		0.2917	-0.3333	-0.0975	-0.0692	0.25	-0.3333	0.0417
s_2	-0.2917		-0.3333	-0.25	-0.3333	0.0975	-0.3750	-0.4167
s_3	0.3333	0.3333		0.2083	0.2641	0.3333	0.0417	0.3475
s_4	0.0975	0.25	-0.2083		0.0283	0.1808	-0.2083	0.2083
s_5	0.0692	0.3333	-0.2641	-0.0283		0.2917	-0.2917	-0.0833
s_6	-0.25	-0.0975	-0.3333	-0.1808	-0.2917		-0.3333	-0.3750
s_7	0.3333	0.3750	-0.0417	0.2083	0.2917	0.3333		0.3750
s_8	-0.0417	0.4167	-0.3475	-0.2083	0.0833	0.3750	-0.3750	

Table 2.9: $Ch^B(P^B(s_i, s_j), \hat{\mu})$ obtained applying the bipolar PROMETHEE method with the parameters of the most representative model

tion provided by the DM while in the ROR we take into account simultaneously all the models that are compatible with the preference information provided by the DM.

Another methodology aiming to explore the whole set of models compatible with the preference information provided by the DM is the Stochastic Multiobjective Acceptability Analysis (SMAA). SMAA is a family of MCDA methodologies dealing with imprecision and/or lack of information on the considered data, where with the term data generally we mean the weights of evaluation criteria or the evaluations of the alternatives with respect to the considered criteria. For each choice of the weights and of the alternatives' evaluations, SMAA computes a ranking of the considered alternatives (for the first works on SMAA see [82, 83] while for a survey on the SMAA methodologies see [121]). In our context, in order to explore the set of the models compatible with the preference information provided by the DM, we adapt a methodology presented in [3] inspired by SMAA. We will do a sampling of several models (precisely 100000 models) among those compatible with the DM preferences and for each of these sampled models we obtain a complete order of the eight students using the bipolar PROMETHEE II method with the net flow formulation of equation (2.18). At the end of all

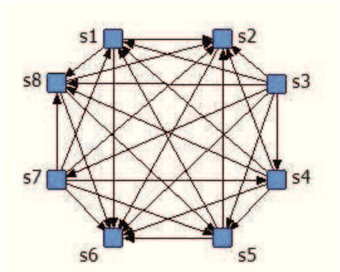


Figure 2.6: Comparison between all pairs of students obtained by the Bipolar PROMETHEE method applied with the parameters of the most representative model

Alternative/Net Flow	$\Phi(s)$
s_7	0.2679
s_3	0.2659
s_4	0.0497
s_5	0.0038
s_8	-0.0139
s_1	-0.0357
s_6	-0.2659
s_2	-0.2718

Table 2.10: Total order of the eight students obtained by the bipolar PROMETHEE II method (applied with the parameters of the most representative model) and the net flows given by equation (2.18)

the iterations, we obtain for each alternative s_j the following values:

- the rank acceptability index b_j^r, representing the proportion of times in which the alternative s_j has obtained the position r, with $r = 1, \ldots, m$,
- the central weight vector w_j^c, representing the mean model giving to alternative s_j the best position in the ranking (obviously this vector will be provided only if alternative s_j reached the first position at least one time in the considered iterations).

Alt/Rank	b^1	b^2	b^3	b^4	b^5	b^6	b^7	b^8
s_1	0	0	0	8.856	19.037	72.099	0.008	0
s_2	0	0	0	0	0	0	56.728	43.272
s_3	44.38	55.59	0.03	0	0	0	0	0
s_4	0	0.005	73.116	9.127	15.83	1.823	0.099	0
s_5	0	0.03	25.185	59.828	13.08	1.877	0	0
s_6	0	0	0	0	0.031	0.077	43.165	56.727
s_7	55.62	44.375	0.005	0	0	0	0	0
s_8	0	0	1.664	22.189	52.022	24.124	0	0.001

Table 2.11: Rank acceptability indices (b^i) in percentages

Applying this methodology to the bipolar PROMETHEE II method, we obtain the results shown in Tables 2.11 and 2.12. Each value b_j^r in Table 2.11 has the following interpretation: picking randomly a model M compatible with the preference information provided by the DM, the probability that student s_j is in the position r in the final ranking obtained applying the bipolar PROMETHEE II method with the net flow formulation of equation 2.18 and the parameters of the model M is the b_j^r%. So, for example, the value 55.62 corresponding to b_7^1 means that picking randolmy a sampled model M, there is a probability of the 55.62% that student s_7 is in the first position in the final ranking obtained applying the bipolar PROMETHEE II method with the model M.

student/parameters	a_1	a_2	a_3	a_{12}	a_{13}	a_{23}	$a^+_{1\|2}$	$a^+_{1\|3}$	$a^+_{2\|1}$
s_3	0.8187	0.6698	0.6056	-0.3852	-0.3619	-0.347	-0.1958	-0.262	-0.1678
s_7	0.6761	0.7801	0.6779	-0.3379	-0.3433	-0.4529	-0.1637	-0.2386	-0.1448

student/parameters	$a^+_{2\|3}$	$a^+_{3\|1}$	$a^+_{3\|2}$	$a^-_{1\|2}$	$a^-_{1\|3}$	$a^-_{2\|1}$	$a^-_{2\|3}$	$a^-_{3\|1}$	$a^-_{3\|2}$
s_3	-0.1222	-0.0974	-0.1304	-0.1677	-0.0982	-0.1959	-0.1277	-0.2612	-0.1249
s_7	-0.2061	-0.0713	-0.1661	-0.1386	-0.0758	-0.1699	-0.1714	-0.234	-0.2009

Table 2.12: Central weight vectors for students s_3 and s_7

From Table 2.11 we observe that only students s_3 and s_7 could reach the first position in the final ranking computed using the bipolar PROMETHEE II method, even if the decision on s_7 as the best student could be considered more robust. From the other side, students s_2, s_6 and s_8 could reach the last position in the final ranking but the probability of s_6 being the worst is the highest. From Table 2.12 we get that student s_3 is the best if Mathematics is the most important criterion, while student s_7 is the best if Physics is the most important criterion.

2.2.6 Conclusions

In this work we have dealt with a generalization of the classical PROMETHEE method. A basic assumption of PROMETHEE method is the independence between criteria which implies that no interaction between criteria is considered. In this work we have developed a methodology permitting to take into account interaction between criteria (synergy, redundancy and antagonism effects) within PROMETHEE method by using the bipolar Choquet integral. In this way we obtained a new method called the Bipolar PROMETHEE method.

The Decision Maker (DM) can give directly the preferential parameters of the method; however, due to their great number, it is advisable using some indirect procedure to elicit the preferential parameters from some preference information provided by the DM.

Since, in general, there is more than one set of parameters compatible with these preference information, we proposed to use the Robust Ordinal Regression (ROR) and the SMAA methodology to consider the whole family of compatible sets of preferential parameters.

We believe that the proposed methodology can be successfully applied in many real world problems where interacting criteria have to be considered.

2.3 MUSA-INT: Multicriteria customer satisfaction analysis with interacting criteria

2.3.1 Introduction

Customer satisfaction evaluation plays a key role in the enterprises' organization, contributing through discovery and representation of customers' preferences to the definition of different salient aspects of companies' strategies.

Among other advantages, customer satisfaction could increase companies' competitiveness [70], identify potential market opportunities, direct new actions to the quality improvement of a product or a service [67], and could also have a positive effect on brand equity [123].

Several approaches have been already developed to evaluate customer satisfaction (see [67] for a detailed list of the existing methods). The most used approaches are the statistical ones: the multiple regression analysis, the discriminant analysis and the conjoint analysis [61],[63] that nowadays is one of the most important marketing research tools (see [68] for an overview and recent developments).

In conjoint analysis, customers are asked to evaluate combinations of different values of the attributes considered for a product or a service. On the basis of customers' answers, conjoint analysis aims at identifying the most desirable attribute values to be implemented in a product or a service.

Customer satisfaction analysis has also been approached using a method based on dominance-based rough set theory [51], which aims at inferring some simple decision rules from the consumers' data [53], differently from the conjoint analysis methods which represent customers' preferences with a comprehensive utility function.

Another interesting approach to customer satisfaction analysis consists in preference learning (see [36] for an updated state-of-the-art) that, given some preferences on a set of objects, searches a function to predict the preferences on a new set of objects. For example, some preference learning applications are provided by a search engine's ranking of web pages according to customers' preferences, or by stores' rankings of particular products according to the preferences expressed on-line by the clients.

Customers' satisfaction evaluation has also been studied from a multicriteria point of view, using the method MUSA (MUlticriteria Satisfaction Analysis [65]). MUSA is a preference disaggregation method that, following the principle of ordinal regression analysis [72], finds an additive utility function representing the satisfaction level of a set of customers based on their expressed preferences collected in a satisfaction survey's data. Using MUSA, the customers are asked to give a comprehensive satisfaction level for a service or a product under consideration, but also a marginal satisfaction level for each one of its features (evaluation criteria). MUSA has many advantages over the traditional customer satisfaction models, since it fully considers the qualitative form of customers' judgments and preferences that are usually expressed in this way in the consumers' questionnaires. The success of MUSA is witnessed by many applications in different fields as, for example, bank sector [64], agricultural marketing [115] and transportation-communication sector [66]. Despite these positive aspects, MUSA is not able to represent positive and negative synergies between specific features of a product or a service, since it considers an additive utility function and, consequently, its underlying hypothesis is preference independence [80],[129].

This is an important issue because it is a common experience that in the evaluation of a product or a service, some features could positively or negatively interact. For example, in the evaluation of a supermarket, prices and special offers have, usually, a negative interaction. In fact, prices and special offers are both important in evaluating a supermarket, however usually a supermarket with low prices has also many special offers and thus, considering together prices and special offers, the total importance is smaller than the sum of their marginal importances. Analogously, one can say that there is a positive synergy between goods' quality and prices, because in general a supermarket with high quality of goods has also high prices and thus a supermarket with a high quality of goods and relatively low prices is well appreciated, such that the total importance of goods' quality and prices considered together is higher than the sum of the importance of their marginal importances.

In Multiple Criteria Decision Aiding (MCDA, see [29] for an updated state-of-the-art) positive and negative interaction among criteria are very often represented using some fuzzy integrals, such as the Choquet integral [21] or some of its generalizations, e.g., the bipolar Choquet integral ([41],[42]; see also [54]) or the level dependent Choquet integral [50], (see [43] for a survey about the use of Choquet integral in MCDA). Fuzzy integrals, and among them the Choquet integral, are aggregation models that, besides other technical assumptions, require a scale of measurement which is cardinal (more precisely, an interval scale [98]) and common to all the criteria (features) taken into consideration. Such a scale permits comparison of evaluations on different criteria, so that, e.g., it becomes possible to say that, for a given supermarket, the level of prices is better than the special offers it proposes and, moreover, these are better than the quality of the goods.

Even if the majority of conjoint analysis methods does not consider interaction among attributes [18], there are several contributions, like [1],[38],[62],[69],[85],[93],[96], that estimate by means of a statistical regression not only a value for each level of each attribute, but also a value for each combination of levels on a set of couples of attributes (possibly all couples of attributes). Another approach proposed to represent interaction among attributes in conjoint analysis is based on the use of the Choquet integral [124],[71],[92],[125],[126],[131].

Since we want to take into account not more than ordinal qualitative aspects of the scales of criteria, we propose MUSA-INT, being a generalization of the multicriteria method MUSA, in which we deal with positive and negative synergies between couples of criteria, using a formulation of the utility function recently proposed in the multicriteria method UTA^{GMS}-INT [58]. Differently from the 2-additive Choquet integral aggregation model, UTA^{GMS}-INT represents positive and negative synergies avoiding any arbitrary transformation of the original ordinal scales into a unique artificial cardinal scale.

This section is organized as follows. In section 2.3.2, we introduce the basic concepts and the relative notation, a brief description of the MUSA method, and the specific utility function adopted in our customer satisfaction model. In section 2.3.3 basic steps of the proposed multicriteria customer satisfaction analysis are described. Section 2.3.4 contains an illustrative example, considering a set of customers' satisfaction questionnaires on which MUSA-INT is applied. Some further extensions of the proposed method are presented in Section 2.3.5. Conclusions and future directions of research are collected in Section 2.3.6.

2.3.2 Basic concepts and the MUSA method

The basic elements of the proposed methodology are the following:

- $C = \{1, \ldots, r\}$ is the set of customers,
- $I = \{1, \ldots, n\}$ is the set of evaluation criteria (features),
- $\mathcal{L}^i = \{\ell^i_1, \ldots, \ell^i_{s_i}\}, i = 1, \ldots, n$, is the set of levels of satisfaction for criterion i: for example, for a given criterion i, the scale could be $\mathcal{L}^i = \{\ell^i_1, \ell^i_2, \ell^i_3\}$, with $\ell^i_1 =$"dissatisfied", $\ell^i_2 =$"satisfied", $\ell^i_3 =$"very satisfied"; the levels $\ell^i_1, \ldots, \ell^i_{s_i}$ are increasingly ordered with respect to the satisfaction level, i.e. the satisfaction represented by ℓ^i_p is greater than the satisfaction represented by $\ell^i_{p-1}, p = 2, \ldots, s_i$,
- $\mathcal{L}^{n+1} = \{\ell^{n+1}_1, \ldots, \ell^{n+1}_{s_{n+1}}\}$ is the set of levels of comprehensive satisfaction: the levels $\ell^{n+1}_1, \ldots, \ell^{n+1}_{s_{n+1}}$ are increasingly ordered with respect to the satisfaction level,
- $sat_{c,i} \in \mathcal{L}^i$ is the satisfaction of customer $c \in C$ with respect to criterion $i \in I$,
- $sat_{c,n+1} \in \mathcal{L}^{n+1}$ is the comprehensive satisfaction of customer $c \in C$,

- $u_i : \mathcal{L}^i \rightarrow [0,1]$ is the marginal utility function of criterion i,
- $U : \mathcal{L}^{n+1} \rightarrow [0,1]$ is the utility of comprehensive levels of satisfaction,
- $Syn^+ \subseteq I^{(2)}$ with $I^{(2)} = \{\{i_1, i_2\} \subseteq I\}$ is the set of all couples of criteria for which there is a positive interaction,
- $Syn^- \subseteq I^{(2)}$ is the set of all couples of criteria for which there is a negative interaction; we have $Syn^+ \cap Syn^- = \emptyset$,
- $syn^+_{ij} : \mathcal{L}^i \times \mathcal{L}^j \rightarrow [0,1]$ is a function non decreasing in both its two arguments representing the strength of the positive interaction between the couples of criteria $\{i,j\} \in Syn^+$,
- $syn^-_{ij} : \mathcal{L}^i \times \mathcal{L}^j \rightarrow [0,1]$ is a function non decreasing in both its two arguments representing the strength of the negative interaction between the couples of criteria $\{i,j\} \in Syn^-$.

Within the MUSA method [65], inspired by the idea of ordinal regression used in the UTA methods [72], one represents customer satisfaction through the following additive utility function,

$$U(sat_{c,n+1}) = \sum_{i=1}^{n} u_i(sat_{c,i}), \ c \in C. \tag{2.24}$$

The utility function (2.24) is obtained by solving the following LP problem [65]:

$$\text{Minimize:} \sum_{c=1}^{r} (\sigma_c^+ + \sigma_c^-), \ \text{s.t.} \tag{2.25}$$

$$\begin{cases} U(sat_{c,n+1}) = \sum_{i=1}^{n} u_i(sat_{c,i}) - \sigma_c^+ + \sigma_c^-, \ \text{for all } c \in C \\ \sigma_c^+ \geq 0, \sigma_c^- \geq 0, \ \text{for all } c \in C \\ \left. \begin{array}{l} u_i(\ell^i_p) \geq u_i(\ell^i_{p-1}), \ p = 2, \ldots, s_i, \ \text{for all } i \in I, \\ U(\ell^{n+1}_p) \geq U(\ell^{n+1}_{p-1}), \ p = 2, \ldots, s_{n+1}, \end{array} \right\} \text{(monotonicity conditions)} \\ \left. \begin{array}{l} u_i(\ell^i_1) = 0, \ \text{for all } i \in I, \\ \sum_{i=1}^{n} u_i(\ell^i_{s_i}) = 1, \\ U(\ell^{n+1}_{s_{n+1}}) = 1, \end{array} \right\} \text{(normalization constraints)} \end{cases}$$

where σ^+ and σ^- are over- and under-estimation errors for every customer's utility function.

Differently from MUSA, the utility function considered in our model is that one proposed in the multicriteria method UTA^{GMS}-INT (see [58]), which takes into account positive and negative synergies between couples of criteria as follows:

$$U(sat_{c,n+1}) = \sum_{i=1}^{n} u_i(sat_{c,i}) + \sum_{\{i,j\}\in Syn^+} \text{syn}^+_{ij}(sat_{c,i}, sat_{c,j}) - \sum_{\{i,j\}\in Syn^-} \text{syn}^-_{ij}(sat_{c,i}, sat_{c,j}),\ c \in C. \tag{2.26}$$

2.3.3 Description of MUSA-INT

In this section, we present a new procedure for finding a utility function representing the overall satisfaction of a set of customers C. The adopted utility function defined by (2.26) considers synergies between satisfaction levels on two criteria, i and j: $sat_{c,i}$ and $sat_{c,j}$.

The multicriteria customer satisfaction analysis, we are proposing, is composed of three main successive phases:

(i) finding a utility function U representing the satisfaction of all customers from set C with a minimal sum of approximation errors;

(ii) identifying a minimal pair (Syn^+, Syn^-) of sets of couples of interacting criteria, where minimality is referred to the inclusion;

(iii) finding a utility function discriminating as much as possible satisfaction levels for both marginal and overall utility functions.

From a computational point of view, each phase consists in solving a specific mixed integer linear program (MILP). Let us examine each phase in detail.

Phase (i): finding a utility function representing the satisfaction of all the customers

Since we want to get a utility function U representing the utility of all customers from set C with a minimal sum of approximation errors, we need to introduce a double error variable ($\sigma_c^+, \sigma_c^- \geq 0$), corresponding to over- and under-estimation, respectively, for every customer's utility as follows:

$$U(sat_{c,n+1}) = \sum_{i=1}^{n} u_i(sat_{c,i}) + \sum_{\{i,j\}\in I^{(2)}} \text{syn}^+_{ij}(sat_{c,i}, sat_{c,j}) - \sum_{\{i,j\}\in I^{(2)}} \text{syn}^-_{ij}(sat_{c,i}, sat_{c,j}) - \sigma_c^+ + \sigma_c^-, \text{ for all } c \in C.$$

The objective function to be minimized is the sum of the error variables for every customer from set C (analogically to the original UTASTAR method [72]):

$$\sum_{c=1}^{r}(\sigma_c^+ + \sigma_c^-) \tag{2.27}$$

Then, we introduce as many binary variables ($\delta_{ij}^+, \delta_{ij}^- \in \{0,1\}$) as twice the couples of criteria, i.e. $2 \times \binom{n}{2}$. The meaning of every binary variable is the following:

$$\delta_{ij}^+(\delta_{ij}^-) = \begin{cases} 1 & \text{if } \{i,j\} \in I^{(2)} \text{ are positively (negatively) interacting,} \\ 0 & \text{if } \{i,j\} \in I^{(2)} \text{ are not positively (negatively) interacting.} \end{cases}$$

For every couple of criteria $\{i,j\} \in I^{(2)}$, three situations can arise:

1) i and j are interacting positively ($\delta_{ij}^+ = 1$),

2) i and j are interacting negatively ($\delta_{ij}^- = 1$),

3) i and j are not interacting ($\delta_{ij}^+ = \delta_{ij}^- = 0$).

In consequence, the following constraints are included in the first MILP problem:

$$\begin{cases} \delta_{ij}^+ + \delta_{ij}^- & \leq & 1, \\ \mathrm{syn}_{ij}^+(\ell_{s_i}^i, \ell_{s_j}^j) & \leq & \delta_{ij}^+, \\ \mathrm{syn}_{ij}^-(\ell_{s_i}^i, \ell_{s_j}^j) & \leq & \delta_{ij}^-. \end{cases} \tag{2.28}$$

Furthermore, a dominance constraint is considered if, for $c,d \in C$, $sat_{c,i} \geq sat_{d,i}$, for all $i \in I$:

$$\begin{aligned} &\sum_{i=1}^{n} u_i(sat_{c,i}) + \sum_{\{i,j\}\in I^{(2)}} syn_{ij}^+(sat_{c,i}, sat_{c,j}) - \sum_{\{i,j\}\in I^{(2)}} syn_{ij}^-(sat_{c,i}, sat_{c,j}) - \sigma_c^+ + \sigma_c^- \geq \\ &\geq \sum_{i=1}^{n} u_i(sat_{d,i}) + \sum_{\{i,j\}\in I^{(2)}} syn_{ij}^+(sat_{d,i}, sat_{d,j}) - \sum_{\{i,j\}\in I^{(2)}} syn_{ij}^-(sat_{d,i}, sat_{d,j}) - \sigma_d^+ + \sigma_d^- . \end{aligned} \tag{2.29}$$

Finally, the MILP formulation includes some technical constraints concerning monotonicity and boundary conditions on the synergies, marginal utilities and overall utility.

Summing up, we get the following MILP problem:

$$\text{Minimize:} \sum_{c=1}^{r} (\sigma_c^+ + \sigma_c^-), \ \text{s.t.} \tag{2.30}$$

$$\left.\begin{array}{l}
U(sat_{c,n+1}) = \displaystyle\sum_{i=1}^{n} u_i(sat_{c,i}) + \sum_{\{i,j\} \in I^{(2)}} \text{syn}_{ij}^+(sat_{c,i}, sat_{c,j}) - \sum_{\{i,j\} \in I^{(2)}} \text{syn}_{ij}^-(sat_{c,i}, sat_{c,j}) - \sigma_c^+ + \sigma_c^-, \\
\text{for all } c \in C, \\
u_i(\ell_1^i) = 0, \forall i \in I, U(\ell_1^{n+1}) = 0, \text{syn}_{ij}^-(\ell_1^i, \ell_1^j) = 0, \text{syn}_{ij}^+(\ell_1^i, \ell_1^j) = 0, \ \text{for all } \ \{i,j\} \in I^{(2)}, \\
\displaystyle\sum_{i=1}^{n} u_i(sat_{c,i}) + \sum_{\{i,j\} \in I^{(2)}} syn_{ij}^+(sat_{c,i}, sat_{c,j}) - \sum_{\{i,j\} \in I^{(2)}} syn_{ij}^-(sat_{c,i}, sat_{c,j}) - \sigma_c^+ + \sigma_c^- \geq \\
\geq \displaystyle\sum_{i=1}^{n} u_i(sat_{d,i}) + \sum_{\{i,j\} \in I^{(2)}} syn_{ij}^+(sat_{d,i}, sat_{d,j}) - \sum_{\{i,j\} \in I^{(2)}} syn_{ij}^-(sat_{d,i}, sat_{d,j}) - \sigma_d^+ + \sigma_d^- \\
\text{if } \ sat_{c,i} \geq sat_{d,i} \text{ for all } i = 1, \ldots, n, \\
\left.\begin{array}{l}
u_i(\ell_p^i) \geq u_i(\ell_{p-1}^i), \ p = 2, \ldots, s_i, \ \text{for all } i \in I, \\
U(\ell_p^{n+1}) \geq U(\ell_{p-1}^{n+1}) + \varepsilon, \ p = 2, \ldots, s_{n+1}, \\
\text{syn}_{ij}^+(\ell_{p_1}^i, \ell_{q_1}^j) \geq \text{syn}_{ij}^+(\ell_{p_2}^i, \ell_{q_2}^j), \\
\text{syn}_{ij}^-(\ell_{p_1}^i, \ell_{q_1}^j) \geq \text{syn}_{ij}^-(\ell_{p_2}^i, \ell_{q_2}^j), \\
u_i(\ell_{p_1}^i) + u_j(\ell_{q_1}^j) - \text{syn}_{ij}^-(\ell_{p_1}^i, \ell_{q_1}^j) \geq u_i(\ell_{p_2}^i) + u_j(\ell_{q_2}^j) - \text{syn}_{ij}^-(\ell_{p_2}^i, \ell_{q_2}^j), \\
\text{with } p_1 \geq p_2 \ \text{ and } q_1 \geq q_2, \\
p_1, p_2 = 1, \ldots, s_i, \ \ q_1, q_2 = 1, \ldots, s_j, \ \text{for all } \{i,j\} \in I^{(2)},
\end{array}\right\} \text{(monotonicity conditions)} \\
\left.\begin{array}{l}
\displaystyle\sum_{i=1}^{n} u_i(\ell_{s_i}^i) + \sum_{\{i,j\} \in I^{(2)}} \text{syn}_{ij}^+(\ell_{s_i}^i, \ell_{s_j}^j) - \sum_{\{i,j\} \in I^{(2)}} \text{syn}_{ij}^-(\ell_{s_i}^i, \ell_{s_j}^j) = 1, \\
\text{syn}_{ij}^+(\ell_{s_i}^i, \ell_{s_j}^j) \leq \delta_{ij}^+, \text{ for all } \{i,j\} \in I^{(2)}, \\
\text{syn}_{ij}^-(\ell_{s_i}^i, \ell_{s_j}^j) \leq \delta_{ij}^-, \text{ for all } \{i,j\} \in I^{(2)},
\end{array}\right\} \text{(boundary conditions)}, \\
\delta_{ij}^+ + \delta_{ij}^- \leq 1, \ \text{for all } \{i,j\} \in I^{(2)}, \\
\delta_{ij}^+, \delta_{ij}^- \in \{0,1\}, \ \text{for all } \{i,j\} \in I^{(2)}, \sigma_c^+ \geq 0, \sigma_c^- \geq 0, \ \text{for all } c \in C,
\end{array}\right\} (E_1)$$

where ε is an arbitrary small positive quantity.

If the objective function of program (2.30) can be minimized to zero, then there exists at least one utility function U, having the form of (2.26), representing the satisfaction of customers expressed by set of constraints (E_1); otherwise, if the minimum value of objective function of (2.30) is positive, then there is no utility function U, having the form of (2.26), compatible with satisfaction of customers expressed by set of constraints (E_1).

The above mixed-integer linear program gives as solution the utility function U and two sets, Syn^+ and Syn^-, of couples of positively and negatively interacting criteria, defined, respectively, as follows:

$$Syn^+ = \{\{i,j\} \in I^{(2)} : \delta_{ij}^+ = 1\}, \;\; Syn^- = \{\{i,j\} \in I^{(2)} : \delta_{ij}^- = 1\}.$$

Let us remark that if program (2.30) gives $\delta_{ij}^+ = \delta_{ij}^- = 0$, for all $\{i,j\} \in I^{(2)}$, i.e. when there are no interactions, the obtained utility function is the same as the one supplied by the MUSA method [65]. For this reason, MUSA-INT can be considered as a generalization of MUSA.

Phase (ii): identifying of a minimal pair (Syn^+, Syn^-)

The pair (Syn^+, Syn^-) of sets of couples of interacting criteria obtained from the mixed-integer linear program (2.30) is not necessarily minimal, in the sense that there could be other sets Syn'^+ and Syn'^- of couples of positively or negatively interacting criteria that could represent the utility of all customers with the same or similar approximation error $\sum_{c=1}^{r}(\sigma_c^+ + \sigma_c^-)$, and such that $Syn'^+ \subseteq Syn^+$ and $Syn'^- \subseteq Syn^-$, with at least one of the two inclusions being strict.

In order to identify a minimal pair (Syn^+, Syn^-) of sets of couples of interacting criteria while possibly accepting a small deterioration of the approximation error resulting from the previously considered mixed-integer linear program (2.30), the following mixed-integer linear programming problem has to be solved:

$$\text{Minimize: } \sum_{\{i,j\} \in I^{(2)}} (\delta_{ij}^+ + \delta_{ij}^-), \text{ s.t.} \tag{2.31}$$

$$\left. \begin{array}{c} (E_1), \\ \sum_{c=1}^{r}(\sigma_c^+ + \sigma_c^-) \leq opterr \times (1+\alpha), \end{array} \right\} (E_2)$$

where *opterr* is the optimal value of the total approximation error $\sum_{c=1}^{r}(\sigma_c^+ + \sigma_c^-)$ resulting from the solution of (2.30), and $0 \leq \alpha < 1$ is a tolerance parameter controlling the possible deterioration of the total optimal approximation error (in fact, we accept a deterioration of $\sum_{c=1}^{r}(\sigma_c^+ + \sigma_c^-)$ by no more than $\alpha \times opterr$).

In result of solving MILP problem (2.31), one gets a utility function U (possibly different from the utility function resulting from (2.30)) and a minimal pair (Syn^+, Syn^-) of sets of couples of positively and negatively interacting criteria, in the sense of inclusion.

Phase (iii): finding the most discriminating utility function

In order to find a utility function U (possibly different from the one obtained in the previous phase) discriminating as much as possible all levels of satisfaction by the marginal utility functions $u_i(\cdot)$, or by the overall utility function $U(\cdot)$, while keeping the same number of interacting couples of criteria, as obtained from (2.31), one has to solve two mixed-integer linear programming problems. The first one tends to discriminate as much as possible the satisfaction levels of the comprehensive utility function:

$$\text{Maximize: } \varepsilon = \varepsilon_{comprehensive}, \quad \text{s.t.} \tag{2.32}$$

$$\left.\begin{array}{c} (E_2), \\ \sum\limits_{\{i,j\} \in I^{(2)}} (\delta_{ij}^+ + \delta_{ij}^-) \leq optsyn, \end{array}\right\} (E_3)$$

where ε is a variable present in the constraint $U(\ell_{p+1}^{n+1}) \geq U(\ell_p^{n+1}) + \varepsilon$, $opterr$ and α have the same meaning as in program (2.31), while $optsyn$ is the optimal value of the objective function of program (2.31).

The solution of MILP problem (2.32) gives a utility function maximizing the minimal difference $U(\ell_p^{n+1}) - U(\ell_{p-1}^{n+1}), p = 2, \ldots, s_{n+1}$. In fact, the minimum of those differences is equal to $\varepsilon_{comprehensive}$, i.e the optimal value of ε given by program (2.32).

The analyst could be interested in finding the most discriminating function not only with respect to the comprehensive utility, but also with respect to the marginal utilities. In order to find such discriminating marginal utility functions, one has to solve the following MILP problem:

$$\text{Maximize: } \varepsilon = \varepsilon_{marginal}, \quad \text{s.t.} \tag{2.33}$$

$$\left.\begin{array}{c}(E_2'),\\ \sum_{\{i,j\}\in I^{(2)}}\left(\delta_{ij}^{+}+\delta_{ij}^{-}\right)\leq optsyn,\end{array}\right\}(E_3')$$

where (E_2') is composed of the same constraints as (E_2), apart from constraints

- $u_i(\ell_p^i) \geq u_i(\ell_{p-1}^i),\ p = 2,\ldots,s_i,$ for all $i \in I$,
- $U(\ell_p^{n+1}) \geq U(\ell_{p-1}^{n+1}) + \varepsilon,\ p = 2,\ldots,s_{n+1},$

that are replaced by

- $u_i(\ell_p^i) \geq u_i(\ell_{p-1}^i) + \varepsilon,\ p = 2,\ldots,s_i,$ for all $i \in I$,
- $U(\ell_p^{n+1}) \geq U(\ell_{p-1}^{n+1}) + \varepsilon_{comprehensive} \times (1-\beta),\ p = 2,\ldots,s_{n+1},$

with $\beta \in [0,1]$ representing the percentage of comprehensive discrimination threshold that the analyst is ready to lose in order to gain on discrimination with respect to the marginal utilities.

As it will be shown in the next section, in some cases, in order to find a discriminating utility function, both with respect to the comprehensive and the marginal satisfaction levels, it may be necessary to increase the admissible total approximation error or increase the number of interacting couples of criteria. In this case, one needs to solve again the two optimization problems (2.32) and (2.33), increasing the chosen value of α in (E_2) or substituting the constraint $\sum_{\{i,j\}\in I^{(2)}}\left(\delta_{ij}^{+}+\delta_{ij}^{-}\right) \leq optsyn$ with $\sum_{\{i,j\}\in I^{(2)}}\left(\delta_{ij}^{+}+\delta_{ij}^{-}\right) \leq (optsyn + \gamma)$, where γ represents the number of additional interactions accepted by the analyst.

In Section 2.3.5, we shall describe how to identify alternative minimal pairs (Syn^{+}, Syn^{-}) of sets of couples of interacting criteria being compatible with a fixed tolerance parameter α. In case of a plurality of minimal pairs (Syn^{+}, Syn^{-}), it is interesting to compute the intersection of all the sets Syn^{+} on one hand, and of all the sets Syn^{-} on the other hand, without deteriorating the approximation error. Let us observe that this interpretation of alternative minimal pairs (Syn^{+}, Syn^{-}) is analogous to the concept of reducts in rough set theory [94]. Moreover, the intersection of all the sets Syn^{+} on one hand, and of all the sets Syn^{-} on the other hand, is analogous to the concept of core in rough set theory [94].

2.3.4 Illustrative example

We will illustrate MUSA-INT using an example originally considered by Grigoroudis and Siskos [65], concerning 20 customers evaluating a service provided by an enterprise. In order to show the full potential of our method, we have augmented the customer dataset presented in [65] by 4 customers, denoted by x, y, w, and z. The main features of our illustrative example are listed hereafter:

1) evaluation of the service involves three criteria concerning: product (1), purchase process (2) and additional service (3);

2) three levels of satisfaction (Very Satisfied (V), Satisfied (S), Dissatisfied (D)) are considered with respect to both, every criterion and comprehensive satisfaction of the service;

3) the customer's satisfaction survey is composed of 24 customers displayed in Table 2.13.

In the following, we suppose us to be the analyst supporting the customer satisfaction expert.

Table 2.13: Consumers' satisfaction survey

Customer	Comprehensive satisfaction	Product (1)	Purchase process (2)	Additional service (3)
1	Satisfied	Very Satisfied	Satisfied	Dissatisfied
2	Dissatisfied	Dissatisfied	Dissatisfied	Dissatisfied
3	Very Satisfied	Very Satisfied	Very Satisfied	Very Satisfied
4	Satisfied	Very Satisfied	Dissatisfied	Satisfied
5	Dissatisfied	Dissatisfied	Dissatisfied	Dissatisfied
6	Very Satisfied	Very Satisfied	Very Satisfied	Very Satisfied
7	Satisfied	Very Satisfied	Dissatisfied	Very Satisfied
8	Satisfied	Very Satisfied	Dissatisfied	Very Satisfied
9	Satisfied	Satisfied	Satisfied	Satisfied
10	Dissatisfied	Dissatisfied	Dissatisfied	Dissatisfied
11	Satisfied	Satisfied	Very Satisfied	Dissatisfied
12	Dissatisfied	Dissatisfied	Dissatisfied	Dissatisfied
13	Very Satisfied	Very Satisfied	Very Satisfied	Very Satisfied
14	Satisfied	Satisfied	Very Satisfied	Dissatisfied
15	Dissatisfied	Dissatisfied	Dissatisfied	Dissatisfied
16	Very Satisfied	Very Satisfied	Very Satisfied	Satisfied
17	Very Satisfied	Very Satisfied	Very Satisfied	Very Satisfied
18	Very Satisfied	Very Satisfied	Very Satisfied	Satisfied
19	Satisfied	Satisfied	Satisfied	Satisfied
20	Dissatisfied	Satisfied	Dissatisfied	Dissatisfied
x	Very Satisfied	Satisfied	Very Satisfied	Satisfied
y	Satisfied	Satisfied	Satisfied	Very Satisfied
w	Dissatisfied	Dissatisfied	Very Satisfied	Satisfied
z	Satisfied	Dissatisfied	Satisfied	Very Satisfied

For customers x, y, w, and z, it is easy to show that the axiom of the preferential independence is violated [80].

Supposing that the utility function of all the customers has an additive form and does not handle synergies between criteria, we can observe the following:

1) since customers x and y have the same levels of satisfaction with respect to criterion Product, and the comprehensive satisfaction level of x is greater than the comprehensive satisfaction level of y, we get:

$$u_1(x) + u_2(x) + u_3(x) > u_1(y) + u_2(y) + u_3(y) \Rightarrow u_2(x) + u_3(x) > u_2(y) + u_3(y);$$

2) since customers w and z have the same levels of satisfaction with respect to criterion Product, and the comprehensive satisfaction level of w is lower than the comprehensive satisfaction level of z, we get:

$$u_1(w) + u_2(w) + u_3(w) < u_1(z) + u_2(z) + u_3(z) \Rightarrow u_2(w) + u_3(w) < u_2(z) + u_3(z);$$

3) since customers x and w have the same levels of satisfaction with respect to criteria Purchase process and Additional service, and customers y and z have the same levels of satisfaction with respect to criteria Purchase process and Additional service, we obtain:

$$u_2(x) + u_3(x) = u_2(w) + u_3(w) \quad \text{and} \quad u_2(y) + u_3(y) = u_2(z) + u_3(z).$$

From 1), 2) and 3) we get a contradiction since at the same time it should be true that $u_2(x)+u_3(x) > u_2(y) + u_3(y)$ and $u_2(x) + u_3(x) < u_2(y) + u_3(y)$.

As a result, we conclude that the customers' overall satisfaction cannot be represented by an additive utility function, and thus, the MUSA method using this type of utility function is not able to fully represent the comprehensive satisfaction of the customers shown in Table 2.13.

For this reason, to represent the customers' comprehensive satisfaction shown in Table 2.13, we apply MUSA-INT, adopting a utility function with positive and negative synergy components (2.26).

According to phase (i), we need to solve MILP problem (2.30). Fixing $\varepsilon = 0.1$, we get $\sigma_c^+ = \sigma_c^- = 0$ for all $c \in C$, while the binary variables, the marginal and comprehensive utilities, as well as the synergies, are displayed in Table 2.14.

In phase (ii), we solve MILP problem (2.31) to determine a minimal pair (Syn^+, Syn^-) of sets of couples of interacting criteria. Fixing $\alpha = 0$ in order to maintain the same approximation error as obtained in the previous phase, we get the same results as shown in Table 2.14. This means that the pair (Syn^+, Syn^-) of sets of couples of interacting criteria found in phase (i) is the minimal one.

Table 2.14: Parameters of the utility function resulting from optimal solution of MILP problem (2.30)

(a) Marginal utilities and interactions

	u_1	u_2	u_3	U
D	0	0	0	0
S	0	0	0	0.5
V	0	0	0	1

δ_{12}^+	δ_{12}^-	δ_{13}^+	δ_{13}^-	δ_{23}^+	δ_{23}^-
1	0	1	0	0	0

(b) Synergies

	syn_{12}^+	syn_{12}^-	syn_{13}^+	syn_{13}^-	syn_{23}^+	syn_{23}^-
VV	0.5	0	0.5	0	0	0
VS	0	0	0.5	0	0	0
VD	0	0	0.5	0	0	0
SV	0.5	0	0.5	0	0	0
SS	0	0	0.5	0	0	0
SD	0	0	0	0	0	0
DV	0	0	0.5	0	0	0
DS	0	0	0	0	0	0
DD	0	0	0	0	0	0

In phase (iii), we proceed in two steps to find the most discriminating utility function. In the first step, when maximizing the discrimination of satisfaction levels of the comprehensive utility, we find the same utility function as in phase (ii), shown in Table 2.14. At this point, in the second step, we accept to lose on the comprehensive discrimination in order to gain on the discrimination of satisfaction levels of marginal utilities; for this reason in problem (2.33) we fix $\beta = 0.6$, and after maximizing the discrimination of satisfaction levels of the marginal utilities, we find the utility function shown in Table 2.15. Remark that satisfaction levels of all but one marginal utilities are well discriminated.

Table 2.15: Parameters of the most discriminating utility function resulting from optimal solution of MILP problem (2.33)

(a) Marginal utilities and interactions

	u_1	u_2	u_3	U
D	0	0	0	0
S	0.5	0.25	0	0.5
V	0.75	0.75	0	1

δ_{12}^+	δ_{12}^-	δ_{13}^+	δ_{13}^-	δ_{23}^+	δ_{23}^-
0	1	1	0	0	0

(b) Synergies

	syn_{12}^+	syn_{12}^-	syn_{13}^+	syn_{13}^-	syn_{23}^+	syn_{23}^-
VV	0	1	0.5	0	0	0
VS	0	0.75	0.5	0	0	0
VD	0	0.75	0.25	0	0	0
SV	0	0.75	0.5	0	0	0
SS	0	0.75	0.5	0	0	0
SD	0	0.5	0	0	0	0
DV	0	0.75	0.5	0	0	0
DS	0	0.25	0	0	0	0
DD	0	0	0	0	0	0

This result is still not satisfactory since we would like to see some discrimination of satisfaction levels for all marginal utilities. In this situation, we would like to know what is the value of the total approximation error that should be accepted in order to obtain a discrimination of at least 0.2 both on the marginal and comprehensive utilities in two separate cases: allowing any number of interactions and allowing a limited number of interactions. Thus, fixing such a minimum discrimination on all the marginal utilities and on the comprehensive utility, we obtain the utility functions shown, respectively, in Tables 2.16 and 2.17.

Table 2.16: Parameters of a utility function ensuring the discrimination of at least 0.2 on both marginal and comprehensive utilities, and accepting an increased number of interactions

(a) Utilities and Interactions

	u_1	u_2	u_3	U
D	0	0	0	0
S	0.5	0.2	0.2	0.5
V	0.7	0.7	0.4	1

δ^+_{12}	δ^-_{12}	δ^+_{13}	δ^-_{13}	δ^+_{23}	δ^-_{23}
0	1	1	0	0	1

(b) Synergies

	syn^+_{12}	syn^-_{12}	syn^+_{13}	syn^-_{13}	syn^+_{23}	syn^-_{23}
VV	0	0.9	0.5	0	0	0.4
VS	0	0.9	0.5	0	0	0.2
VD	0	0.7	0.5	0	0	0
SV	0	0.7	0.5	0	0	0.4
SS	0	0.7	0.5	0	0	0.2
SD	0	0.5	0	0	0	0
DV	0	0.7	0.5	0	0	0.4
DS	0	0.2	0	0	0	0.2
DD	0	0	0	0	0	0

Table 2.17: Parameters of the utility function discriminating all marginal and comprehensive levels of satisfaction by at least 0.2, obtained for total approximation error equal to 0.2, and the number of interactions limited to 2

(a) Marginal utilities and interactions

	u_1	u_2	u_3	U
D	0	0	0	0
S	0.8	0.2	0.6	0.4
V	1	1	0.8	1

δ^+_{12}	δ^-_{12}	δ^+_{13}	δ^-_{13}	δ^+_{23}	δ^-_{23}
0	1	0	0	0	1

(b) Synergies

	syn^+_{12}	syn^-_{12}	syn^+_{13}	syn^-_{13}	syn^+_{23}	syn^-_{23}
VV	0	1	0	0	0	0.8
VS	0	0.8	0	0	0	0.6
VD	0	0.8	0	0	0	0.6
SV	0	0.8	0	0	0	0.6
SS	0	0.8	0	0	0	0.4
SD	0	0.8	0	0	0	0
DV	0	0.8	0	0	0	0.6
DS	0	0	0	0	0	0.4
DD	0	0	0	0	0	0

In the first case, the results of which are shown in Table 2.16, we obtain the desired discrimination for the total approximation error equal to zero (i.e. $\sigma_c^+ = \sigma_c^- = 0$, for all $c \in C$) and for the following sets of couples of interacting criteria: $Syn^+ = \{1,3\}$ and $Syn^- = \{\{1,2\},\{2,3\}\}$. In the second case, the results of which are shown in Table 2.17, the desired discrimination is achieved for the total approximation error $\sum_{c=1}^{r}(\sigma_c^+ + \sigma_c^-) = 0.2$ and the number of couples of interacting criteria limited to two; in this case: $Syn^+ = \emptyset$ and $Syn^- = \{\{1,2\},\{2,3\}\}$.

2.3.5 Further extensions

The three-phase method described until now can be considered a standard procedure; first, we check the existence of a utility function of type (2.26) compatible with the customers' answers (phase (i)), then we look for a minimal pair of sets of couples of interacting criteria (phase (ii)), and finally we look for a utility function having the maximum discrimination power (phase (iii)). In this section, some interesting extensions of this procedure are presented.

Finding other minimal couples of sets of interacting criteria

In phase (ii), in result of solving problem (2.31), we find a minimal pair (Syn_1^+, Syn_1^-), where Syn_1^+ and Syn_1^- are sets of couples of positively and negatively interacting criteria.

In general, there may exist more than one minimal pair (Syn^+, Syn^-) and for this reason, it could be interesting to find them all. In order to find another minimal pair (Syn_2^+, Syn_2^-), one has to solve the following optimization problem:

$$\text{Minimize:} \sum_{\{i,j\}\in I^{(2)}} \left(\delta_{ij}^+ + \delta_{ij}^-\right), \text{ s.t.} \tag{2.34}$$

$$\left.\begin{array}{c} (E_2), \\ \displaystyle\sum_{\{i,j\}\in\left\{Syn_1^+ \cup Syn_1^-\right\}} (\delta_{ij}^+ + \delta_{ij}^-) \leq |Syn_1^+ \cup Syn_1^-| - 1. \end{array}\right\} (E_{M_2})$$

The last constraint in (E_{M_2}) ensures that a newly found pair of sets of indices of couples of interacting criteria is different from the previous one.

Let us suppose that at the $(k-1)^{th}$ iteration we found the minimal pair $(Syn_{k-1}^+, Syn_{k-1}^-)$. In order to check if there exists another minimal pair (Syn_k^+, Syn_k^-), it will be sufficient to solve the following optimization problem:

$$\text{Minimize:} \sum_{\{i,j\}\in I^{(2)}} \left(\delta_{ij}^+ + \delta_{ij}^-\right), \text{ s.t.} \tag{2.35}$$

$$\left.\begin{array}{r}(E_2),\\ \displaystyle\sum_{\{i,j\}\in\left\{Syn_1^+\cup Syn_1^-\right\}}(\delta_{ij}^+ + \delta_{ij}^-) \leq |Syn_1^+ \cup Syn_1^-| - 1,\\ \displaystyle\sum_{\{i,j\}\in\left\{Syn_2^+\cup Syn_2^-\right\}}(\delta_{ij}^+ + \delta_{ij}^-) \leq |Syn_2^+ \cup Syn_2^-| - 1,\\ \cdots\\ \displaystyle\sum_{\{i,j\}\in\left\{Syn_{k-1}^+\cup Syn_{k-1}^-\right\}}(\delta_{ij}^+ + \delta_{ij}^-) \leq |Syn_{k-1}^+ \cup Syn_{k-1}^-| - 1.\end{array}\right\}(E_{M_k})$$

If problem (2.35) is infeasible, then there is no minimal pair (Syn_k^+, Syn_k^-), so that the set M of all minimal pairs (Syn^+, Syn^-) is given by

$$M = \left\{(Syn_1^+, Syn_1^-), \cdots, (Syn_{k-1}^+, Syn_{k-1}^-)\right\}.$$

If, instead, problem (2.35) is feasible, then (Syn_k^+, Syn_k^-) is a new minimal pair with

$$Syn_k^+ = \left\{\{i,j\} \in I^{(2)} \colon \delta_{ij}^+ = 1 \text{ in the solution of problem (2.35)}\right\}$$

and

$$Syn_k^- = \left\{\{i,j\} \in I^{(2)} \colon \delta_{ij}^- = 1 \text{ in the solution of problem (2.35)}\right\}.$$

In phase (ii) of the illustrative example presented in Section 2.3.4, we found the first minimal pair (Syn_1^+, Syn_1^-), in which $Syn_1^+ = \{1,3\}$ and $Syn_1^- = \{1,2\}$. Searching for other minimal pairs, we get:

- the solution of optimization problem (2.35) with constraints (E_{M_2}) is shown in Table 2.18, where $Syn_2^+ = \{1,2\}$ and $Syn_2^- = \{1,3\}$;

δ_{12}^+	δ_{12}^-	δ_{13}^+	δ_{13}^-	δ_{23}^+	δ_{23}^-
1	0	0	1	0	0

Table 2.18: Binary variables defining the second minimal pair

- the solution of optimization problem (2.35) with constraints (E_{M_3}), is shown in Table 2.19, where $Syn_3^+ = \emptyset$ and $Syn_3^- = \{\{1,2\},\{1,3\},\{2,3\}\}$;

δ_{12}^{+}	δ_{12}^{-}	δ_{13}^{+}	δ_{13}^{-}	δ_{23}^{+}	δ_{23}^{-}
0	1	0	1	0	1

Table 2.19: Binary variables defining the third minimal pair

- finally, when solving optimization problem (2.35) with constraints (E_{M_4}) we don't find any other minimal pair, and thus the searching procedure stops. Therefore, the set M of minimal pairs (Syn^{+}, Syn^{-}) is

$$M = \left\{(Syn_1^{+}, Syn_1^{-}), (Syn_2^{+}, Syn_2^{-}), (Syn_3^{+}, Syn_3^{-})\right\}.$$

Customer satisfaction evaluation using a set of compatible preference models

When analyzing the customers' survey presented in Table 2.13, the experts of the company could be interested to know what action should be made in order to improve the customer satisfaction of the service provided. For this reason, the experts could be interested in answering the following question: "is customer satisfaction of profile P_1 = (S, V, S) better than customer satisfaction of profile P_2 = (V, S, V)?" To answer this question, the experts have to consider set $\mathcal{U}$ of utility functions of type (2.26) satisfying set of constraints (E_1) related to the considered customers' survey. Such utility functions are called *compatible* with customers' preferences. Then, a natural question arises whether customer satisfaction of profile P_1 is at least as good as customer satisfaction of P_2 for at least one or for all compatible utility functions from $\mathcal{U}$. By answering this type of questions, the experts can get an additional insight into more meaningful decision investments concerning the service provided by the company. For example, if the experts find that customer satisfaction of profile P_1 is better than customer satisfaction of profile P_2 for all compatible utility functions from $\mathcal{U}$, they can conclude that an action directed to increase satisfaction level from S to V on Purchase process is more appreciated by the customers than another action directed to an analogous improvement on both Product and Additional service. To perform this type of analysis, one can use the Robust Ordinal Regression (ROR) methodology being a family of MCDA methods introduced in [55] (for a recent survey on the topic see [60]). ROR has been applied to ranking problems (see UTA^{GMS}[55], GRIP [33]) and sorting problems (see UTADIS^{GMS} [57]), and also in methods using outranking relations (ELECTRE^{GKMS} [47]) or Choquet integral (NAROR [7]) as preference models.

Given an initial set of preference information provided by a Decision Maker (DM), the ROR aims at obtaining a final recommendation for the decision problem at hand, taking into account not only one preference model compatible with this preference information, but the whole set of compatible preference models simultaneously. In fact, as it is often the case in the inference procedures, several decision models could be compatible with the information provided by the DM, but each one of them could lead to different preferences on the remaining alternatives, not considered by the DM at the stage of expressing the preference information. The choice of one particular preference model among all compatible ones could be considered arbitrary, and so it is more meaningful to take into account the whole set of compatible preference models simultaneously. Supposing the preference model in the form of a set of compatible utility functions, the conclusions drawn by ROR are based on two preference relations:

- the *necessary* preference relation for which alternative a is necessarily preferred to alternative b, if a is at least as good as b for all utility functions compatible with the preference information provided by the DM,
- the *possible* preference relation for which alternative a is possibly preferred to alternative b if a is at least as good as b for at least one utility function compatible with the preference information provided by the DM.

In the following, we adapt the concept of ROR to MUSA-INT considering the following binary relations on the set of profiles $\mathcal{L} = \prod_{i=1}^{n} \mathcal{L}_i$. Given any two profiles, $P_1, P_2 \in \mathcal{L}$:

- profile P_1 is possibly preferred to profile P_2, ($P_1 \succsim^P P_2$), if customer satisfaction of P_1 is at least as good as customer satisfaction of P_2 for at least one compatible utility function of $\mathcal{U}$,
- profile P_1 is necessarily preferred to profile P_2, ($P_1 \succsim^N P_2$), if customer satisfaction of P_1 is at least as good as customer satisfaction of P_2 for all compatible utility functions of $\mathcal{U}$.

Let us stress that in our context, there is no DM but only an analyst supporting the experts of the company in the analysis of customer satisfaction; for this reason the preference information provided by the DM in the ROR context is replaced by the answers to the customers' survey (expressed set of constraints (E_1)) in our method.

In order to compute the necessary and possible preference relations between two profiles of satisfaction $(sat_{a,1}, sat_{a,2}, \dots, sat_{a,n})$ and $(sat_{b,1}, sat_{b,2}, \dots, sat_{b,n})$, one should proceed in the following way.

Considering the sets of constraints,

$$\left.\begin{array}{l}\displaystyle\sum_{i=1}^{n} u_i(sat_{b,i}) + \sum_{(i,j)\in I^{(2)}} syn^{+}_{ij}(sat_{b,i}, sat_{b,j}) - \sum_{(i,j)\in I^{(2)}} syn^{-}_{ij}(sat_{b,i}, sat_{b,j}) \geq \\ \displaystyle\geq \sum_{i=1}^{n} u_i(sat_{a,i}) + \sum_{(i,j)\in I^{(2)}} syn^{+}_{ij}(sat_{a,i}, sat_{a,j}) - \sum_{(i,j)\in I^{(2)}} syn^{-}_{ij}(sat_{a,i}, sat_{a,j}) + \varepsilon, \\ (E_1), \\ \displaystyle\sum_{c\in C}\left(\sigma^{+}_{c} + \sigma^{-}_{c}\right) \leq \sigma^{*}, \end{array}\right\} E^{N}(a,b)$$

and

$$\left.\begin{array}{l}\displaystyle\sum_{i=1}^{n} u_i(sat_{a,i}) + \sum_{(i,j)\in I^{(2)}} syn^{+}_{ij}(sat_{a,i}, sat_{a,j}) - \sum_{(i,j)\in I^{(2)}} syn^{-}_{ij}(sat_{a,i}, sat_{a,j}) \geq \\ \displaystyle\geq \sum_{i=1}^{n} u_i(sat_{b,i}) + \sum_{(i,j)\in I^{(2)}} syn^{+}_{ij}(sat_{b,i}, sat_{b,j}) - \sum_{(i,j)\in I^{(2)}} syn^{-}_{ij}(sat_{b,i}, sat_{b,j}), \\ (E_1), \\ \displaystyle\sum_{c\in C}\left(\sigma^{+}_{c} + \sigma^{-}_{c}\right) \leq \sigma^{*}, \end{array}\right\} E^{P}(a,b)$$

where σ^* is the maximum accepted total approximation error, and ε in the constraint $U(\ell_p^{n+1}) \geq U(\ell_{p-1}^{n+1}) + \varepsilon$ is a variable, we can conclude the following:

- profile $(sat_{a,1}, sat_{a,2}, \ldots, sat_{a,n})$ is necessarily preferred to profile $(sat_{b,1}, sat_{b,2}, \ldots, sat_{b,n})$ if $E^N(a,b)$ is infeasible or $\varepsilon^N \leq 0$, where $\varepsilon^N = \max \varepsilon$, subject to $E^N(a,b)$,

- profile $(sat_{a,1}, sat_{a,2}, \ldots, sat_{a,n})$ is possibly preferred to profile $(sat_{b,1}, sat_{b,2}, \ldots, sat_{b,n})$ if $E^P(a,b)$ is feasible and $\varepsilon^P > 0$, where $\varepsilon^P = \max \varepsilon$, subject to $E^P(a,b)$.

Hereafter, we report some results obtained for our example:

- profile (S,S,V) is *necessarily* preferred to profile (D,V,D), that is profile (S,S,V) is at least as good as profile (D,V,D) for all utility functions compatible with the customers' preferences,

- profile (V,V,D) is *possibly*, but not *necessarily* preferred to profile (D,D,V), that is profile (V,V,D) is at least as good as profile (D,D,V) for at least one utility function compatible with the customers' preferences, however, there exists at least one utility function for which profile (D,D,V) is better than profile (V,V,D).

Analogical conclusions can be drawn using a set of approximately compatible utility functions. Then, one has to consider the following two sets of constraints:

$$\left.\begin{array}{l}
\sum_{i=1}^{n} u_i(sat_{b,i}) + \sum_{\{i,j\}\in I^{(2)}} syn_{ij}^{+}(sat_{b,i}, sat_{b,j}) - \sum_{\{i,j\}\in I^{(2)}} syn_{ij}^{-}(sat_{b,i}, sat_{b,j}) + \sigma_b^{+} - \sigma_b^{-} \geq \\
\geq \sum_{i=1}^{n} u_i(sat_{a,i}) + \sum_{\{i,j\}\in I^{(2)}} syn_{ij}^{+}(sat_{a,i}, sat_{a,j}) - \sum_{\{i,j\}\in I^{(2)}} syn_{ij}^{-}(sat_{a,i}, sat_{a,j}) + \sigma_a^{+} - \sigma_a^{-} + \varepsilon, \\
(E_1), \\
\sum_{c\in C} \left(\sigma_c^{+} + \sigma_c^{-}\right) \leq opterr, \\
\sigma_a^{+} + \sigma_a^{-} \leq opt_1, \\
\sigma_b^{+} + \sigma_b^{-} \leq opt_2
\end{array}\right\} E_1^N(a,b)$$

and

$$\left.\begin{array}{l}
\sum_{i=1}^{n} u_i(sat_{a,i}) + \sum_{\{i,j\}\in I^{(2)}} syn_{ij}^{+}(sat_{a,i}, sat_{a,j}) - \sum_{\{i,j\}\in I^{(2)}} syn_{ij}^{-}(sat_{a,i}, sat_{a,j}) + \sigma_a^{+} - \sigma_a^{-} \geq \\
\geq \sum_{i=1}^{n} u_i(sat_{b,i}) + \sum_{\{i,j\}\in I^{(2)}} syn_{ij}^{+}(sat_{b,i}, sat_{b,j}) - \sum_{\{i,j\}\in I^{(2)}} syn_{ij}^{-}(sat_{b,i}, sat_{b,j}) + \sigma_b^{+} - \sigma_b^{+}, \\
(E_1), \\
\sum_{c\in C} \left(\sigma_c^{+} + \sigma_c^{-}\right) \leq opterr, \\
\sigma_a^{+} + \sigma_a^{-} \leq opt_1, \\
\sigma_b^{+} + \sigma_b^{-} \leq opt_2,
\end{array}\right\} E_1^P(a,b)$$

where σ_a^{+}, σ_a^{-}, σ_b^{+}, and σ_b^{-} are, error variables of the utility values relative to profiles a and b, while opt_1 and opt_2 represent the maximum accepted errors in each one of the considered profile's utility.

Since the new set of constraints, $E_1^N(a,b)$ and $E_1^P(a,b)$, enlarge the decision space of the utility functions compatible with the customers' preferences, the following two preference relations, analogical to the ones introduced above, are defined:

- a *strong necessary* preference relation, for which a is strongly necessarily preferred to b if a is at least as good as b for all utility functions approximately compatible with the customers' preferences,

- a *weak possible* preference relation, for which a is weakly possibly preferred to b if a is at least as good as b for at least one utility function approximately compatible with the customers' preferences.

Let us remark that we have considered two different qualifications (strong and weak) for the necessary and possible preference representation. In fact, taking into account the double-error variables referring to the compared profiles of satisfaction enlarges the decision space of the compatible utility functions. Consequently, if a is at least as good as b with respect to all approximately compatible utility functions, which are many more than the ones in the basic decision space, then a is *strongly necessarily* preferred to b.

On the contrary, if a is not at least as good as b for any compatible model in the enlarged decision space, then this means that even considering utility functions admitting some error, we can't find one utility function for which a is at least as good as b. Among the preference relations obtained for our example, we report the following:

- profile (S,S,V) is *strongly necessarily* preferred to profile (D,V,D),
- profile (S,V,S) is only *necessarily* preferred, but not *strongly necessarily* preferred to profile (V,S,S).

2.3.6 Conclusions

In this work, we proposed MUSA-INT, a new multicriteria customer satisfaction analysis method able to take into account positive and negative interactions among criteria, even if the customers' judgments are qualitative and not quantitative. To explain the customer's preferences, the method employs an additive utility function augmented with components representing positive and negative synergies between two satisfaction levels of two criteria.

Some strong points of our method are listed hereafter:

- the criteria are expressed on ordinal scales, without the necessity of expressing all the criteria on a common scale, as this is the case of the Choquet integral or some other fuzzy integrals;
- the model reveals the synergies among criteria in the customer satisfaction evaluation of a product or a service;
- the synergies among criteria have a meaningful interpretation for the DM as a bonus (for positive interaction) or penalty (for negative interaction), added or subtracted from the sum of marginal utility values;

- there is a parsimonious representation of the interactions by considering minimal pairs of sets of couples of interacting criteria;
- one can identify all minimal pairs of sets of interacting criteria;
- as the preference model (utility function) representing the customers' satisfaction is, in general, not unique, it is possible to take into account the whole set of compatible preference models adopting the Robust Ordinal Regression methodology.

We envisage some possible directions of future research:

(1) Consideration of positive or negative interaction not only between couples of criteria, but also triples, quadruples and, generally, sets of criteria of cardinality greater than 2. Using the example of a supermarket, it may be reasonable to admit that there is a specific surplus in the appreciation due to the presence at the same time of low prices, special offers and good quality. In this case, the considered utility function will become

$$U(sat_{c,n+1}) = \sum_{i=1}^{n} u_i(sat_{c,i}) + \sum_{A \in Syn_G^+} \mathrm{syn}_A^+(sat_{c,i}, i \in A) - \sum_{A \in Syn_G^-} \mathrm{syn}_A^-(sat_{c,i}, i \in A),\ c \in C \tag{2.36}$$

where $Syn_G^+, Syn_G^- \subseteq 2^I$ are the families of all the subsets of criteria for which there is a positive synergy and a negative synergy, respectively. Considerations of synergy among criteria in subsets with cardinality greater than 2 requires to pay a specific attention to the trade-off between the better knowledge one gets about customer satisfaction and the additional computational effort required to get this knowledge.

(2) Consideration of a hierarchal structure of criteria in the customer survey. Indeed, very often the customer is required to evaluate features of a product or a service organized in a given hierarchy. For instance, taking into account our illustrative example, product satisfaction could be split in satisfaction with respect to aspects A_1, A_2 and A_3, so that we have an evaluation on the three aspects and a comprehensive evaluation with respect to "Product". A similar level of detail can be considered for "Purchase product" and "Additional service". In this case, we could consider an ordinal regression approach concordant with the principle of Multiple Criteria Hierarchy Process introduced in [23].

(3) The representation of customers' preferences using an outranking model instead of a utility function; in this case, the interaction can be represented taking into account the concordance index of ELECTRE method presented in [30], or the bipolar PROMETHEE proposed in [22].

(4) Application of all ROR extensions, such as extreme ranking analysis [76] and SMAA applied to Robust Ordinal Regression [78].

Chapter 3

Multiple Criteria Hierarchy Process

Complex real-world decision problems, such as choosing a new product pricing strategy, deciding where to locate manufacturing plants, or forecasting the future of a country, involve factors of different nature. These factors may be political, economic, cultural, environmental, technological, or managerial. Obviously, it is difficult for the DMs to consider so many different points of view simultaneously when assessing the quality of the alternatives. In fact, practical applications are often explicitly imposing a hierarchical structure of criteria. In this case, the preference model may refer to all levels of the hierarchy, representing values of particular scores of the alternatives on indicators, sub-indicators, sub-sub-indicators, etc.

Several contributions in which the evaluation of alternatives is done with respect to a set of criteria structured in a hierarchical way are known in literature. In the following we shall cite only some of them. Belton and Stewart [12] present many real world examples in which evaluation criteria have a hierarchical structure. They assign to each criterion a cumulative and a relative weight. The cumulative weights of criteria at the bottom are computed by pairwise comparisons between them, while the cumulative weight of a criterion in a particular node of the hierarchy is obtained as the sum of the cumulative weights of criteria descending from it and in the subsequent level. Relative weights are assessed within families of criteria, i.e. criteria sharing the same parent so that the weights within each family are normalised to sum up one; Stillwell et al. [118] use a value tree in order to compare three energy options; Keeney et al. [81] construct a list of criteria structured in a hierarchical way in order to evaluate energy systems; Weber et al. [130] proved that the detail of attribute specification enhances attribute weights while Poyhonen et al. [97] observe that the division of criteria in a hierarchical structure can either increase or decrease the weight of an attribute; Mustajoki [91] studies the difference between hierarchical and non hierarchical models in which the preferences are constituted by imprecise weight ratios; Sugeno et al. [120] provide necessary and sufficient conditions so that the Choquet integral can be computed as the sum of several Choquet integrals at different levels.

In this chapter we introduce the Multiple Criteria Hierarchy Process (MCHP) and we apply this concept to MAUT, outranking methods ELECTRE and PROMETHEE and Choquet integral considering also the ROR in the first two cases.

The basic idea of MCHP relies on consideration of preference relations at each node of the hierarchy tree of criteria. These preference relations concern both the phase of eliciting preference information, and the phase of analyzing a final recommendation by the DM. For example, in the phase of eliciting preference information, in a decision problem related to the evaluation of students, one can say not only that student a is comprehensively preferred to student b, but also that a is comprehensively preferred to b because a is preferred to b on subsets of subjects related to Mathematics and Physics, even if b is preferred to a on subjects related to Humanities. Moreover, one can also say that, a is preferred to b on the subset of subjects related to Mathematics because, considering Analysis and Algebra as subjects related to Mathematics, a is preferred to b on Analysis, and this is enough to compensate the fact that b is preferred to a on Algebra.

Putting together MCHP and ROR, permits us to define necessary and possible preference relations at each node of the hierarchy tree. This gives insight into the evolution of the necessary and possible preference relations along the hierarchy tree. In fact, if we know that an alternative a is not necessarily comprehensively preferred to alternative b, with MCHP we can find at which level a particular criterion opposes to the conclusion that a is necessarily preferred to b. All the properties that hold for the "flat" version of ROR methods are also valid in the hierarchical context, and other properties that are characteristic to the context are given in this chapter. In section 3.1, we applied the MCHP to MAUT; in section 3.2, MCHP is extended to the outranking methods and in particular to ELECTRE and PROMETHEE methods while in section 3.3 we presented the application of MCHP to the Choquet integral.

3.1 Multiple Criteria Hierarchy Process in Robust Ordinal Regression

3.1.1 Introduction

It is well known that the dominance relation established in the set of alternatives evaluated on multiple criteria is the only objective information that comes out from a formulation of a multiple criteria decision problem (including sorting, ranking and choice). While dominance relation permits to eliminate many irrelevant (i.e. dominated) alternatives, it does not compare completely all of them, resulting in a situation where many alternatives remain incomparable. This situation may be addressed by taking into account preferences of a Decision Maker (DM). Therefore, all Multiple Criteria Decision Aiding (MCDA) methods (for state-of-the-art surveys on MCDA see [29]) require some preference information elicited by a DM. Information provided by a DM is used within a MCDA process to build a preference model which is then applied on a non-dominated (Pareto-optimal) set of alternatives to arrive at a recommendation.

A great majority of methods designed for MCDA, assume that all evaluation criteria are considered at the same level, however, it is often the case that a practical application is imposing a hierarchical structure of criteria. For example, in economic ranking, alternatives may be evaluated on indicators which aggregate evaluations on several sub-indicators, and these sub-indicators may aggregate another set of sub-indicators, etc. In this case, the marginal value functions may refer to all levels of the hierarchy, representing values of particular scores of the alternatives on indicators, sub-indicators, sub-sub-indicators, etc. Considering hierarchical, instead of flat, structure of criteria, permits decomposition of a complex decision problem into smaller problems involving less criteria. To handle the hierarchy of criteria, we introduce in this section a Multiple Criteria Hierarchy Process (MCHP). The basic idea of MCHP relies on consideration of preference relations at each node of the hierarchy tree of criteria. These preference relations concern both the phase of eliciting preference information, and the phase of analyzing a final recommendation by the DM. Let us consider a very simple and well known preference model, the linear value function, which assigns to each alternative $a \in A$ the value $U(a) = w_1 g_1(a) + \ldots + w_n g_n(a)$, $w_i \geq 0, i = 1, \ldots n$, where $g_i(a)$ is an evaluation of alternative a on criterion $g_i, i = 1, \ldots, n$. If in the phase of eliciting preference information, the DM declares that alternative a is preferred to alternative b with respect to a criterion which, in a node of the hierarchy tree, groups a set of sub-criteria $\mathcal{G}_{\mathbf{r}}$, this can be modeled as

$$\sum_{i \in \mathcal{G}_{\mathbf{r}}} w_i g_i(a) > \sum_{i \in \mathcal{G}_{\mathbf{r}}} w_i g_i(b),$$

which puts some constraints on the values of admissible weights w_i. In the phase of analyzing a final recommendation, even more important, MCHP shows preference relations $\succsim_{\mathbf{r}}$ on A with respect to the set of subcriteria $\mathcal{G}_{\mathbf{r}}$, such that, for all $a, b \in A$,

$$a \succsim_{\mathbf{r}} b \Leftrightarrow \sum_{i \in \mathcal{G}_{\mathbf{r}}} w_i g_i(a) > \sum_{i \in \mathcal{G}_{\mathbf{r}}} w_i g_i(b),$$

where $a \succsim_{\mathbf{r}} b$ reads alternative a is at least as good as alternative b on the set of subcriteria $\mathcal{G}_{\mathbf{r}}$. Analyzing the preference relation $\succsim_{\mathbf{r}}$ is very useful in any decision aiding process because it permits to look into structural elements of the overall preference relation $\succsim$ taking into account the whole set of criteria, and justify better the final recommendation. For example, in a decision problem related to evaluation of students, one can say not only that student a is comprehensively preferred to student b, i.e. $a \succ b$ (where $\succ$ is the asymmetric part of $\succsim$; analogously, in the following, $\succ_{\mathbf{r}}$ is the asymmetric part of $\succsim_{\mathbf{r}}$), but also that a is comprehensively preferred to b because a is preferred to b on subsets of subjects (subcriteria) related to Mathematics and Physics, i.e. $a \succ_{Mathematics} b$ and $a \succ_{Physics} b$, even if b is preferred to a on subjects related to Humanities, i.e. $b \succ_{Humanities} a$. Moreover, one can also say that, for example, a is preferred to b on the subset of subjects related to Mathematics because, considering Analysis and Algebra as subjects (sub-criteria) related to Mathematics, a is preferred to b on Analysis, i.e. $a \succ_{Analysis} b$, and this is enough to compensate the fact that b is preferred to a on Algebra, i.e. $b \succ_{Algebra} a$. Since partial preference relations $\succsim_{Mathematics}$, $\succsim_{Physics}$, $\succsim_{Humanities}$, $\succsim_{Analysis}$, $\succsim_{Algebra}$, and so on, can be constructed using any MCDA methodology, this shows the universal character of MCHP.

In this section, in order to show the useful features of MCHP, we apply this methodology to a recently proposed family of MCDA methods, called Robust Ordinal Regression (ROR) ([55],[33],[57],[60]). Basic ideas of ROR can be summarized as follows. To deal with a multiple criteria decision problem, Multiple Attribute Utility Theory (MAUT) ([80]) constructs a value function which assigns to each alternative a real number representing its degree of preferability. The first MCDA methods using the ordinal regression approach ([19],[117],[132]), aimed at finding one value function compatible with preference information provided by the DM (see, e.g., [72],[117],[95],[25]). Most frequently additive value functions have been considered, i.e. functions obtained by summing up marginal value functions corresponding to particular criteria. For example, in [72], each marginal value function is a piecewise-linear one. Remark that in case of ordinal regression the preference information is always indirect.

In ordinal regression, and also in ROR, the preference information elicited by the DM is indirect, i.e. the DM provides decision examples, like preferential pairwise comparisons of some selected alternatives. This type of preference information is opposed to the direct one, which is composed of values of parameters of the assumed preference model, like weights or trade-off rates of the weighted sum model. Research indicates that indirect preference elicitation requires less cognitive effort from the DM than the direct one, and thus, it becomes more and more popular.

When building, via ordinal regression, a value function compatible with indirect preference information given as pairwise comparisons of some selected alternatives, one encounters a problem of plurality of compatible value functions. Until recently, the usual practice was to select only one of the compatible value functions, either by the DM or using some mathematical tools for finding a "central" value function. In general, however, each compatible value function gives a different ranking of the considered set of alternatives, and thus, it is reasonable to investigate what is the consequence of applying all compatible value functions on the whole set of considered alternatives. For this reason, ROR takes into account all compatible value functions simultaneously. In this context, two preference relations are considered:

- possible preference relation, for which alternative a is possibly preferred to alternative b if a is at least as good as b for at least one compatible value function, and
- necessary preference relation, for which alternative a is necessarily preferred to alternative b if a is at least as good as b for all compatible value functions.

The first method that applied the concept of ROR was UTA^{GMS} [55]: it takes into account pairwise comparisons of alternatives provided by a DM; GRIP [33] was its generalization taking into account not only pairwise comparisons, but also intensities of preference; ROR has been also applied to sorting problems [57], and it has been adapted to other preference models, like outranking relation [47],[76] and non additive integrals [7].

Applying MCHP to ROR, permits to consider preference information at each level of the hierarchy in the phase of eliciting preference information. Moreover, putting together MCHP and ROR, permits to define necessary and possible preference relations at each node of the hierarchy tree. This gives an insight into evolution of the necessary and possible preference relations along the hierarchy tree. In fact, if we know that a is not necessarily comprehensively preferred to b, with MCHP we can find at which level a particular subcriterion opposes to the conclusion that a is necessarily preferred to b. All the properties that hold for the "flat" version of ROR methods are also valid in the hierarchical context, and other properties that are characteristic to the hierarchical context are given in this section.

Thie section is structured in this way: section 3.1.2 describes some basic concepts of the MCHP; section 3.1.3 describes the GRIP method adapted to the hierarchical context; in section 3.1.4 we present the properties of necessary and possible preference relations; sections 3.1.5 and 3.1.6 describe the concepts of intensity of preference and most representative value function; in section 3.1.7 we present a didactic example; in section 3.1.8 we present some extensions of the hierarchical ROR; conclusions are collected in section 3.1.9.

3.1.2 Multiple Criteria Hierarchy Process (MCHP)

In MCHP, we consider a set $\mathcal{G}$ of hierarchically ordered criteria, i.e. all criteria are not considered at the same level, but they are distributed over l different levels (see Figure 3.1). At level 1, there are first level criteria called root criteria. Each root criterion has its own hierarchy tree. The leafs of each hierarchy tree are at the last level l and they are called elementary subcriteria. Thus, in graph theory terms, the whole hierarchy is a forest. We will use the following notation:

- $A = \{a, b, c, \ldots\}$ is the finite set of alternatives,
- l is the number of levels in the hierarchy of criteria,
- $\mathcal{G}$ is the set of all criteria at all considered levels,
- $\mathcal{I}_{\mathcal{G}}$ is the set of indices of particular criteria representing position of criteria in the hierarchy,
- m is the number of the first level criteria, $G_1, \ldots, G_m$,
- $G_{\mathbf{r}} \in \mathcal{G}$, with $\mathbf{r} = (i_1, \ldots, i_h) \in \mathcal{I}_{\mathcal{G}}$, denotes a subcriterion of the first level criterion G_{i_1} at level h; the first level criteria are denoted by G_{i_1}, $i_1 = 1, \ldots, m$,
- $n(\mathbf{r})$ is the number of subcriteria of $G_{\mathbf{r}}$ in the subsequent level, i.e. the direct subcriteria of $G_{\mathbf{r}}$ are $G_{(\mathbf{r},1)}, \ldots, G_{(\mathbf{r},n(\mathbf{r}))}$,
- $g_{\mathbf{t}} : A \rightarrow \mathbb{R}$, with $\mathbf{t} = (i_1, \ldots, i_l) \in \mathcal{I}_{\mathcal{G}}$, denotes an elementary subcriterion of the first level criterion G_{i_1}, i.e a criterion at level l of the hierarchy tree of G_{i_1},

- EL is the set of indices of all elementary subcriteria:

$$EL = \{\mathbf{t} = (i_1, \dots, i_l) \in \mathcal{I}_{\mathcal{G}}\} \quad \text{where} \quad \begin{cases} i_1 = 1, \dots, m \\ i_2 = 1, \dots, n(i_1) \\ \dots\dots \\ i_l = 1, \dots, n(i_1, \dots, i_{l-1}) \end{cases}$$

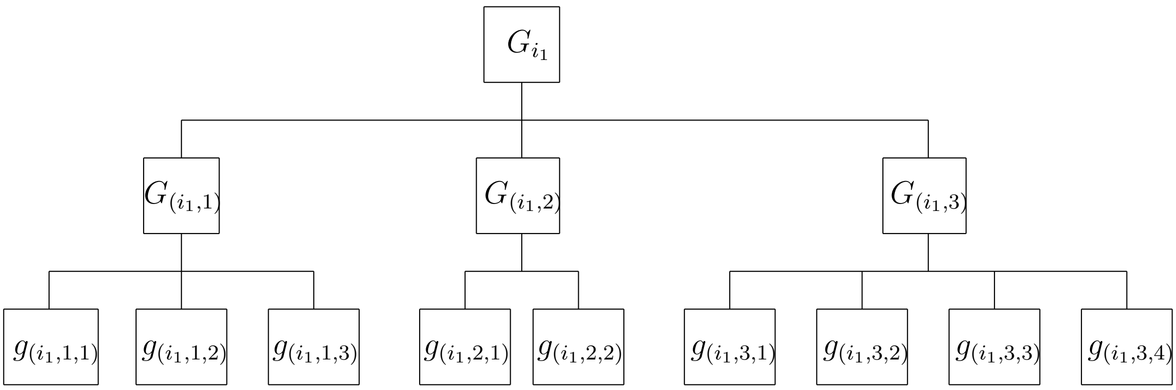
- $E(G_{\mathbf{r}})$ is the set of indices of elementary subcriteria descending from $G_{\mathbf{r}}$, i.e.

$$E(G_{\mathbf{r}}) = \{(\mathbf{r}, i_{h+1}, \dots, i_l) \in \mathcal{I}_{\mathcal{G}}\} \quad \text{where} \quad \begin{cases} i_{h+1} = 1, \dots, n(\mathbf{r}) \\ \dots\dots \\ i_l = 1, \dots, n(\mathbf{r}, i_{h+1}, \dots, i_{l-1}) \end{cases}$$

  thus, $E(G_{\mathbf{r}}) \subseteq EL$,

- when $\mathbf{r} = 0$, then by $G_{\mathbf{r}} = G_0$, we mean the entire set of criteria and not a particular criterion or subcriterion; in this particular case, we have $E(G_0) = EL$.

Figure 3.1: Hierarchy of criteria for the first level (root) criterion G_{i_1}



Without loss of generality we suppose that each elementary subcriterion $g_{\mathbf{t}}, \mathbf{t} \in EL$, maps alternatives to real numbers $g_{\mathbf{t}} : A \to \mathbb{R}$, such that for all $a, b \in A$, $g_{\mathbf{t}}(a) \geq g_{\mathbf{t}}(b)$ means that a is at least as good as b with respect to elementary criterion $g_{\mathbf{t}}$. If criterion $g_{\mathbf{t}}$ has, originally, an ordered qualitative scale, e.g., very bad, bad, medium, good, very good, one can number code such linguistic labels in a way maintaining the preference order. Each alternative $a \in A$ is evaluated directly on the elementary subcriteria only, such that to each alternative $a \in A$ there corresponds a vector of evaluations:

$$(g_{\mathbf{t_1}}(a), \dots, g_{\mathbf{t}_n}(a)), \qquad n = |EL|.$$

Within MCHP, in each node $G_{\mathbf{r}} \in \mathcal{G}$ of the hierarchy tree there exists a preference relation $\succsim_{\mathbf{r}}$ on A, such that for all $a, b \in A$, $a \succsim_{\mathbf{r}} b$ means "a is at least as good as b on subcriterion $G_{\mathbf{r}}$". In the particular case where $G_{\mathbf{r}} = g_{\mathbf{t}}, \mathbf{t} \in EL$, $a \succsim_{\mathbf{t}} b$ holds if $g_{\mathbf{t}}(a) \geq g_{\mathbf{t}}(b)$.

A minimal requirement that preference relations $\succsim_{\mathbf{r}}$ have to satisfy is a dominance principle for hierarchy of criteria, stating that if alternative a is at least as good as alternative b for all subcriteria $G_{(\mathbf{r},j)}$ of $G_{\mathbf{r}}$ of the level immediately below, then a is at least as good as b on $G_{\mathbf{r}}$. For example, if student a is at least as good as student b on Algebra and Analysis, being subcriteria of Mathematics, then a is at least as good as b on Mathematics. Formally, this dominance principle can be stated as follows: given $G_{\mathbf{r}}$, $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, if $a \succsim_{(\mathbf{r},j)} b$ for all $j = 1, \ldots, n(\mathbf{r})$ then $a \succsim_{\mathbf{r}} b$.

In this article, we will aggregate the evaluations of alternative $a \in A$ with respect to the elementary subcriteria $g_{\mathbf{t}}, \mathbf{t} \in EL$, using an additive value function:

$$U(g_{\mathbf{t_1}}(a), \ldots, g_{\mathbf{t}_n}(a)) = \sum_{\mathbf{t} \in EL} u_{\mathbf{t}}(g_{\mathbf{t}}(a)), \tag{3.1}$$

where $u_{\mathbf{t}}$ are marginal value functions, non-decreasing with respect to the evaluation expressed by its argument. Analogously, the marginal value function of alternative $a \in A$ on criterion $G_{\mathbf{r}} \in \mathcal{G}$, is given by:

$$U_{\mathbf{r}}(g_{\mathbf{t}}(a), \mathbf{t} \in E(G_{\mathbf{r}})) = \sum_{\mathbf{t} \in E(G_{\mathbf{r}})} u_{\mathbf{t}}(g_{\mathbf{t}}(a)), \tag{3.2}$$

such that for all $a, b \in A$, $a \succsim_{\mathbf{r}} b$ iff $U_{\mathbf{r}}(a) \geq U_{\mathbf{r}}(b)$.
In the following, to simplify the notation, we shall write $U(a)$ instead of $U(g_{\mathbf{t_1}}(a), \ldots, g_{\mathbf{t}_n}(a))$, $U_{\mathbf{r}}(a)$ instead of $U_{\mathbf{r}}(g_{\mathbf{t}}(a), \mathbf{t} \in E(G_{\mathbf{r}}))$, and $u_{\mathbf{t}}(a)$ instead of $u_{\mathbf{t}}(g_{\mathbf{t}}(a))$.

3.1.3 Multiple Criteria Hierarchy Process applied to a Robust Ordinal Regression method

When aggregating evaluations of alternatives on multiple elementary subcriteria, we will take into account some preference information provided by the DM. This preference information concerns a subset of alternatives $A^R \subseteq A$, called reference alternatives, on which the DM is relatively more confident than on the others. The DM is expected to provide the following preference information:

- a partial preorder $\succsim$ on A^R, whose meaning is: for $a^*, b^* \in A^R$

$$a^* \succsim b^* \;\Leftrightarrow\; \text{“ } a^* \text{ is at least as good as } b^* \text{ ”}.$$

Denoting by $\succsim^{-1}$ the inverse of $\succsim$, i.e. if $a^* \succsim b^*$ then $b^* \succsim^{-1} a^*$, $\sim$ (indifference) is the symmetric part of $\succsim$ given by $\succsim \cap \succsim^{-1}$, i.e. if $a^* \sim b^*$ then $a^* \succsim b^*$ and $a^* \succsim^{-1} b^*$, and $\succ$ (preference) is the asymmetric part given by $(\succsim \setminus \sim)$, i.e. if $a^* \succ b^*$ then $a^* \succsim b^*$ and not $a^* \sim b^*$;

- a partial preorder $\succsim^*$ on $A^R \times A^R$, whose meaning is: for $a^*, b^*, c^*, d^* \in A^R$,

$$(a^*, b^*) \succsim^* (c^*, d^*) \Leftrightarrow \text{“}a^* \text{ is preferred to } b^* \text{ at least as much as } c^* \text{ is preferred to } d^* \text{”}.$$

Analogously to $\succsim$, $\succ^*$ and $\sim^*$ are the asymmetric and the symmetric part of $\succsim^*$;

- given $\mathbf{r} \in \mathcal{I}_{\mathcal{G}}$, a partial preorder $\succsim_{\mathbf{r}}$ on A^R, whose meaning is: for $a^*, b^* \in A^R$,

$$a^* \succsim_{\mathbf{r}} b^* \Leftrightarrow \text{“}a^* \text{ is at least as good as } b^* \text{ with respect to subcriterion } G_{\mathbf{r}}.\text{”}$$

Analogously to $\succsim$, $\succ_{\mathbf{r}}$ and $\sim_{\mathbf{r}}$ are the asymmetric and the symmetric part of $\succsim_{\mathbf{r}}$;

- given $\mathbf{r} \in \mathcal{I}_{\mathcal{G}}$, a partial preorder $\succsim^*_{\mathbf{r}}$ on $A^R \times A^R$, whose meaning is: for $a^*, b^*, c^*, d^* \in A^R$,

$$(a^*, b^*) \succsim^*_{\mathbf{r}} (c^*, d^*) \Leftrightarrow \text{“}a^* \text{ is preferred to } b^* \text{ at least as much as } c^* \text{ is preferred to } d^* \text{ with respect to subcriterion } G_{\mathbf{r}}\text{”}.$$

Analogously to $\succsim$, $\succ^*_{\mathbf{r}}$ and $\sim^*_{\mathbf{r}}$ are the asymmetric and the symmetric part of $\succsim^*_{\mathbf{r}}$.

An additive value function is called *compatible* if it is able to restore the preference information supplied by the DM. Therefore, an additive value function (3.1) is compatible if it satisfies the following set of linear constraints:

$$
\left.\begin{array}{l}
\left.\begin{array}{l}
U(a^*) > U(b^*) \quad \text{if} \;\; a^* \succ b^* \\
U(a^*) = U(b^*) \quad \text{if} \;\; a^* \sim b^* \\
U(a^*) - U(b^*) > U(c^*) - U(d^*) \quad \text{if} \;\; (a^*, b^*) \succ^* (c^*, d^*) \\
U(a^*) - U(b^*) = U(c^*) - U(d^*) \quad \text{if} \;\; (a^*, b^*) \sim^* (c^*, d^*) \\
U_{\mathbf{r}}(a^*) > U_{\mathbf{r}}(b^*) \quad \text{if} \;\; a^* \succ_{\mathbf{r}} b^* \\
U_{\mathbf{r}}(a^*) = U_{\mathbf{r}}(b^*) \quad \text{if} \;\; a^* \sim_{\mathbf{r}} b^* \\
U_{\mathbf{r}}(a^*) - U_{\mathbf{r}}(b^*) > U_{\mathbf{r}}(c^*) - U_{\mathbf{r}}(d^*) \quad \text{if} \;\; (a^*, b^*) \succ^*_{\mathbf{r}} (c^*, d^*) \\
U_{\mathbf{r}}(a^*) - U_{\mathbf{r}}(b^*) = U_{\mathbf{r}}(c^*) - U_{\mathbf{r}}(d^*) \quad \text{if} \;\; (a^*, b^*) \sim^*_{\mathbf{r}} (c^*, d^*)
\end{array}\right\} \begin{array}{l} a^*, b^*, c^*, d^* \in A^R, \\ \mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL \end{array} \\
u_{\mathbf{t}}(x_{\mathbf{t}}^k) - u_{\mathbf{t}}(x_{\mathbf{t}}^{k-1}) \geq 0, \;\; \forall \mathbf{t} \in EL, \; k = 2, ..., m_{\mathbf{t}}(A^R) \\
u_{\mathbf{t}}(x_{\mathbf{t}}^1) \geq 0, \; u_{\mathbf{t}}(x_{\mathbf{t}}^{m_{\mathbf{t}}(A^R)}) \leq u_{\mathbf{t}}(x_{\mathbf{t}}^{m_{\mathbf{t}}}), \;\; \forall \mathbf{t} \in EL \\
u_{\mathbf{t}}(x_{\mathbf{t}}^0) = 0, \;\; \forall \mathbf{t} \in EL \\
\sum_{\mathbf{t} \in EL} u_{\mathbf{t}}(x_{\mathbf{t}}^{m_{\mathbf{t}}}) = 1.
\end{array}\right\} \left(E^{A^R}\right)
$$

where, $x_{\mathbf{t}}^0 = min_{a \in A} \; g_{\mathbf{t}}(a)$, and $x_{\mathbf{t}}^{m_{\mathbf{t}}} = max_{a \in A} \; g_{\mathbf{t}}(a)$; $x_{\mathbf{t}}^k \in X_{\mathbf{t}}(A^R), k = 1, ..., m_{\mathbf{t}}(A^R)$, with $X_{\mathbf{t}}(A^R) \subseteq X_{\mathbf{t}}$, is the set of all different evaluations of reference alternatives from A^R on elementary subcriteria $g_{\mathbf{t}}, \mathbf{t} \in EL$, and $m_{\mathbf{t}}(A^R) = \left|X_{\mathbf{t}}(A^R)\right|$. The values $x_{\mathbf{t}}^k, k = 1, ..., m_{\mathbf{t}}(A^R)$, are increasingly ordered, i.e.,

$$
x_{\mathbf{t}}^1 < x_{\mathbf{t}}^2 < ... < x_{\mathbf{t}}^{m_{\mathbf{t}}(A^R)-1} < x_{\mathbf{t}}^{m_{\mathbf{t}}(A^R)}.
$$

In order to check the existence of a compatible value function, one has to transform first the strict inequalities of E^{A^R} by adding an auxiliary variable ε. Then, we have to solve the following linear programming problem where the variables are the marginal value functions $u_{\mathbf{t}}(x_{\mathbf{t}}^k)$, $k = 1, \ldots, m_{\mathbf{t}}(A^R)$, and $u_{\mathbf{t}}(x_{\mathbf{t}}^{m_{\mathbf{t}}}), \mathbf{t} \in EL$, as well as ε :

maximize ε, subject to the constraints:

$$\left.\begin{array}{l}\left.\begin{array}{l}U(a^*) \geq U(b^*) + \varepsilon \quad \text{if} \;\; a^* \succ b^* \\ U(a^*) = U(b^*) \quad \text{if} \;\; a^* \sim b^* \\ U(a^*) - U(b^*) \geq U(c^*) - U(d^*) + \varepsilon \quad \text{if} \;\; (a^*, b^*) \succ^* (c^*, d^*) \\ U(a^*) - U(b^*) = U(c^*) - U(d^*) \quad \text{if} \;\; (a^*, b^*) \sim^* (c^*, d^*) \\ U_{\mathbf{r}}(a^*) \geq U_{\mathbf{r}}(b^*) + \varepsilon \quad \text{if} \;\; a^* \succ_{\mathbf{r}} b^* \\ U_{\mathbf{r}}(a^*) = U_{\mathbf{r}}(b^*) \quad \text{if} \;\; a^* \sim_{\mathbf{r}} b^* \\ U_{\mathbf{r}}(a^*) - U_{\mathbf{r}}(b^*) \geq U_{\mathbf{r}}(c^*) - U_{\mathbf{r}}(d^*) + \varepsilon \quad \text{if} \;\; (a^*, b^*) \succ^*_{\mathbf{r}} (c^*, d^*) \\ U_{\mathbf{r}}(a^*) - U_{\mathbf{r}}(b^*) = U_{\mathbf{r}}(c^*) - U_{\mathbf{r}}(d^*) \quad \text{if} \;\; (a^*, b^*) \sim^*_{\mathbf{r}} (c^*, d^*)\end{array}\right\} \begin{array}{l} a^*, b^*, c^*, d^* \in A^R, \\ \mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL \end{array} \\ u_{\mathbf{t}}(x^k_{\mathbf{t}}) - u_{\mathbf{t}}(x^{k-1}_{\mathbf{t}}) \geq 0, \;\; \forall \mathbf{t} \in EL, \;\; k = 2, ..., m_{\mathbf{t}}(A^R) \\ u_{\mathbf{t}}(x^1_{\mathbf{t}}) \geq 0, \;\; u_{\mathbf{t}}(x^{m_{\mathbf{t}}(A^R)}_{\mathbf{t}}) \leq u_{\mathbf{t}}(x^{m_{\mathbf{t}}}_{\mathbf{t}}), \;\; \forall \mathbf{t} \in EL \\ u_{\mathbf{t}}(x^0_{\mathbf{t}}) = 0, \;\; \forall \mathbf{t} \in EL \\ \sum_{\mathbf{t} \in EL} u_{\mathbf{t}}(x^{m_{\mathbf{t}}}_{\mathbf{t}}) = 1.\end{array}\right\} \left(E^{A^{R'}}\right)$$

If $\varepsilon(E^{A^{R'}}) > 0$, where $\varepsilon(E^{A^{R'}}) = \max \varepsilon$, s.t. constraints $E^{A^{R'}}$, then there exists at least one compatible value function $U(\cdot)$; if instead $\varepsilon(E^{A^{R'}}) \leq 0$, then there does not exist any compatible value function $U(\cdot)$.

Supposing that there exists at least one compatible value function, each of these functions may induce a different ranking on the whole set A; for this reason the ROR methods, (see [55],[33],[57],[47],[76],[7]), do not take into consideration only one compatible value function but all the compatible value functions simultaneously (we shall denote the set of all compatible value functions by $\mathcal{U}$). Application on the ROR involves the following definitions:

Definition 3.1.1. *Given two alternatives $a, b \in A$, we say that a is weakly necessarily preferred to b, and we write $a \succsim^N b$, if a is at least as good as b for all compatible value functions:*

$$a \succsim^N b \Leftrightarrow U(a) \geq U(b) \;\; \forall U \in \mathcal{U}.$$

Definition 3.1.2. *Given two alternatives $a, b \in A$, we say that a is weakly possibly preferred to b, and we write $a \succsim^P b$, if a is at least as good as b for at least one compatible value function:*

$$a \succsim^P b \Leftrightarrow \exists U \in \mathcal{U} : \;\; U(a) \geq U(b).$$

Definition 3.1.3. *Given two alternatives $a, b \in A$, we say that a is weakly necessarily preferred to b with respect to subcriterion $G_{\mathbf{r}}$, $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, and we write $a \succsim_{\mathbf{r}}^{N} b$, if a is at least as good as b with respect to subcriterion $G_{\mathbf{r}}$ for all compatible value functions:*

$$a \succsim_{\mathbf{r}}^{N} b \Leftrightarrow U_{\mathbf{r}}(a) \geq U_{\mathbf{r}}(b) \;\; \forall U \in \mathcal{U}.$$

Definition 3.1.4. *Given two alternatives $a, b \in A$, we say that a is weakly possibly preferred to b with respect to criterion $G_{\mathbf{r}}$, $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, and we write $a \succsim_{\mathbf{r}}^{P} b$, if a is at least as good as b with respect to criterion $G_{\mathbf{r}}$ for at least one compatible value function:*

$$a \succsim_{\mathbf{r}}^{P} b \Leftrightarrow \exists U \in \mathcal{U} : \;\; U_{\mathbf{r}}(a) \geq U_{\mathbf{r}}(b).$$

Note that for $\mathbf{r} \in EL$, we have:

$$\succsim_{\mathbf{r}}^{N} = \succsim_{\mathbf{r}}^{P} = \{(x_{\mathbf{r}}, y_{\mathbf{r}}) \in X_{\mathbf{r}} \times X_{\mathbf{r}} : x_{\mathbf{r}} \leq y_{\mathbf{r}}\}.$$

Let us remark that we need both the possible and the necessary preference relation $\succsim^{P}$ and $\succsim^{N}$. In fact, considering the necessary preference relation $\succsim^{N}$ only, we loose some important information given by the ROR methodology. For example, for $a, b \in A$, let us consider the two following cases:

case 1) $a \succsim^{N} b$ and $b \succsim^{P} a$,

case 2) $a \succsim^{N} b$ and $b \not\succsim^{P} a$.

In both, case 1) and case 2), $\forall U \in \mathcal{U}$, $U(a) \geq U(b)$. However, in case 1) there is at least one compatible value function $U \in \mathcal{U}$ such that $U(b) \geq U(a)$, while this does not happen in case 2). If we consider only the necessary preference relation $\succsim^{N}$ we are not able to distinguish case 1) from case 2), while, this is not the case if we use also the possible preference relation $\succsim^{P}$.

Necessary weak preference relations ($\succsim^{N}$ and $\succsim_{\mathbf{r}}^{N}$), and possible weak preference relations ($\succsim^{P}$ and $\succsim_{\mathbf{r}}^{P}$) can be calculated as follows. For all alternatives $a, b \in A$, let $X_{\mathbf{t}}(A^R \cup \{a, b\}) \subseteq X_{\mathbf{t}}$ be the set of all different evaluations of alternatives from $A^R \cup \{a, b\}$ on criterion $g_{\mathbf{t}}, \mathbf{t} \in EL$, and $m_{\mathbf{t}}(A^R \cup \{a, b\}) = |X_{\mathbf{t}}(A^R \cup \{a, b\})|$. The values $x_{\mathbf{t}}^{k} \in X_{\mathbf{t}}(A^R \cup \{a, b\}),\; k = 1, \ldots, m_{\mathbf{t}}(A^R \cup \{a, b\})$, are increasingly ordered, i.e.,

$$x_{\mathbf{t}}^{1} < x_{\mathbf{t}}^{2} < \ldots < x_{\mathbf{t}}^{m_{\mathbf{t}}(A^R \cup \{a,b\})-1} < x_{\mathbf{t}}^{m_{\mathbf{t}}(A^R \cup \{a,b\})}.$$

Then, the characteristic points of $u_{\mathbf{t}}(\cdot)$, $\mathbf{t} \in EL$, are in $x_{\mathbf{t}}^0,\ x_{\mathbf{t}}^k,\ k = 1, \ldots, m_{\mathbf{t}}(A^R \cup \{a, b\}),\ x_{\mathbf{t}}^{m_{\mathbf{t}}}$.

Let us consider the following ordinal regression constraints $E(a, b)$, with $u_{\mathbf{t}}(x_{\mathbf{t}}^k)$, $\mathbf{t} \in EL$, $k = 1, \ldots, m_{\mathbf{t}}(A^R \cup \{a, b\})$, $u_{\mathbf{t}}(x_{\mathbf{t}}^{m_{\mathbf{t}}}), \mathbf{t} \in EL$, and ε as variables:

$$\left.\begin{array}{l}
\left.\begin{array}{l}
U(a^*) \geq U(b^*) + \varepsilon \quad \text{if} \quad a^* \succ b^* \\
U(a^*) = U(b^*) \quad \text{if} \quad a^* \sim b^* \\
U(a^*) - U(b^*) \geq U(c^*) - U(d^*) + \varepsilon \quad \text{if} \quad (a^*, b^*) \succ^* (c^*, d^*) \\
U(a^*) - U(b^*) = U(c^*) - U(d^*) \quad \text{if} \quad (a^*, b^*) \sim^* (c^*, d^*) \\
U_{\mathbf{r}}(a^*) \geq U_{\mathbf{r}}(b^*) + \varepsilon \quad \text{if} \quad a^* \succ_{\mathbf{r}} b^* \\
U_{\mathbf{r}}(a^*) = U_{\mathbf{r}}(b^*) \quad \text{if} \quad a^* \sim_{\mathbf{r}} b^* \\
U_{\mathbf{r}}(a^*) - U_{\mathbf{r}}(b^*) \geq U_{\mathbf{r}}(c^*) - U_{\mathbf{r}}(d^*) + \varepsilon \quad \text{if} \quad (a^*, b^*) \succ_{\mathbf{r}}^* (c^*, d^*) \\
U_{\mathbf{r}}(a^*) - U_{\mathbf{r}}(b^*) = U_{\mathbf{r}}(c^*) - U_{\mathbf{r}}(d^*) \quad \text{if} \quad (a^*, b^*) \sim_{\mathbf{r}}^* (c^*, d^*)
\end{array}\right\} \begin{array}{l} a^*, b^*, c^*, d^* \in A^R; \\ \mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL \end{array} \\
u_{\mathbf{t}}(x_{\mathbf{t}}^k) - u_{\mathbf{t}}(x_{\mathbf{t}}^{k-1}) \geq 0,\ \mathbf{t} \in EL,\ k = 2, ..., m_{\mathbf{t}}(A^R \cup \{a, b\}) \\
u_{\mathbf{t}}(x_{\mathbf{t}}^1) \geq 0,\ u_{\mathbf{t}}(x_{\mathbf{t}}^{m_{\mathbf{t}}(A^R \cup \{a,b\})}) \leq u_{\mathbf{t}}(x_{\mathbf{t}}^{m_{\mathbf{t}}}),\ \mathbf{t} \in EL \\
u_{\mathbf{t}}(x_{\mathbf{t}}^0) = 0,\ \mathbf{t} \in EL \\
\sum_{\mathbf{t} \in EL} u_{\mathbf{t}}(x_{\mathbf{t}}^{m_{\mathbf{t}}}) = 1.
\end{array}\right\} (E(a, b))$$

The above constraints depend also on the pair of alternatives $a, b \in A$ because their evaluations $g_{\mathbf{t}}(a)$ and $g_{\mathbf{t}}(b)$ give coordinates for two of $m_{\mathbf{t}}(A^R \cup \{a, b\})$ characteristic points of marginal value function $u_{\mathbf{t}}(\cdot)$, for each $\mathbf{t} \in EL$.

For all $a, b \in A$, and $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, let us consider the following sets of constraints:

$$\left.\begin{array}{l} U(b) \geq U(a) + \varepsilon \\ E(a, b) \end{array}\right\} (E^N(a, b)), \qquad \left.\begin{array}{l} U(a) \geq U(b) \\ E(a, b) \end{array}\right\} (E^P(a, b)),$$

$$\left.\begin{array}{l} U_{\mathbf{r}}(b) \geq U_{\mathbf{r}}(a) + \varepsilon \\ E(a, b) \end{array}\right\} (E_{\mathbf{r}}^N(a, b)), \qquad \left.\begin{array}{l} U_{\mathbf{r}}(a) \geq U_{\mathbf{r}}(b) \\ E(a, b) \end{array}\right\} (E_{\mathbf{r}}^P(a, b)).$$

Thus we get:

- $a \succsim^N b$ iff $E^N(a, b)$ is infeasible or $\varepsilon^N(a, b) \leq 0$, where $\varepsilon^N(a, b) = \max \varepsilon$, s.t. constraints $E^N(a, b)$;
- $a \succsim^P b$ iff $E^P(a, b)$ is feasible and $\varepsilon^P(a, b) > 0$, where $\varepsilon^P(a, b) = \max \varepsilon$, s.t. constraints $E^P(a, b)$;

- $a \succsim^N_{\mathbf{r}} b$ iff $E^N_{\mathbf{r}}(a,b)$ is infeasible or $\varepsilon^N_{\mathbf{r}}(a,b) \leq 0$, where $\varepsilon^N_{\mathbf{r}}(a,b) = \max \varepsilon$, s.t. constraints $E^N_{\mathbf{r}}(a,b)$;
- $a \succsim^P_{\mathbf{r}} b$ iff $E^P_{\mathbf{r}}(a,b)$ is feasible and $\varepsilon^P_{\mathbf{r}}(a,b) > 0$, where $\varepsilon^P_{\mathbf{r}}(a,b) = \max \varepsilon$, s.t. constraints $E^P_{\mathbf{r}}(a,b)$.

3.1.4 Properties of necessary and possible preference relations

The necessary and possible preference relations satisfy some interesting properties presented in the following propositions:

Proposition 3.1.1.

- $\succsim^N \subseteq \succsim^P$; *[55]*
- $\succsim^N$ *is a partial preorder (i.e. reflexive and transitive); [55]*
- $\succsim^P$ *is strongly complete (i.e. for all* $a, b \in A$, $a \succsim^P b$ *or* $b \succsim^P a$*) and negatively transitive; [55]*
- $a \succsim^N b$ *or* $b \succsim^P a$, $\forall a, b \in A$; *[55]*
- $a \succsim^N b$ *and* $b \succsim^P c$*, then* $a \succsim^P c$, $\forall a, b, c \in A$; *[33]*
- $a \succsim^P b$ *and* $b \succsim^N c$*, then* $a \succsim^P c$, $\forall a, b, c \in A$. *[33]*

In case of the hierarchy of criteria, some further properties hold, as showed by the following proposition.

Proposition 3.1.2. *For every* $\mathbf{r} \in \mathcal{I}_{\mathcal{G}}$,

1. $\succsim^N_{\mathbf{r}} \subseteq \succsim^P_{\mathbf{r}}$;
2. $\succsim^N_{\mathbf{r}}$ *is a partial preorder (i.e. reflexive and transitive);*
3. $\succsim^P_{\mathbf{r}}$ *is strongly complete (i.e. for all* $a, b \in A$, $a \succsim^P_{\mathbf{r}} b$ *or* $b \succsim^P_{\mathbf{r}} a$*) and negatively transitive;*
4. $a \succsim^N_{\mathbf{r}} b$ *or* $b \succsim^P_{\mathbf{r}} a$, $\forall a, b \in A$;
5. $a \succsim^N_{\mathbf{r}} b$ *and* $b \succsim^P_{\mathbf{r}} c$*, then* $a \succsim^P_{\mathbf{r}} c$, $\forall a, b, c \in A$;
6. $a \succsim^P_{\mathbf{r}} b$ *and* $b \succsim^N_{\mathbf{r}} c$*, then* $a \succsim^P_{\mathbf{r}} c$, $\forall a, b, c \in A$.

Proof. See Appendix. □

Let us observe that if we consider the comprehensive preference represented by the value function U at a "zero" level of the hierarchy, where $\mathbf{r} = 0$, we can consider Proposition (3.1.1) as a specific case of Proposition (3.1.2), e.g., we can write $\succsim_0^N \subseteq \succsim_0^P$ instead of $\succsim^N \subseteq \succsim^P$. The next proposition presents some results which are specific for the ROR in case of the hierarchy of criteria.

Proposition 3.1.3. *For every* $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$,

1. *given two alternatives* $a, b \in A$,

$$a \succsim_{(\mathbf{r},j)}^N b \ \ \forall j = 1, \dots, n(\mathbf{r}) \ \Rightarrow \ a \succsim_{\mathbf{r}}^N b;$$

2. *given two alternatives* $a, b \in A$ *such that:*

   $\alpha)$ $a \succsim_{(\mathbf{r},j)}^N b, \quad \forall j \in \{1, \dots, n(\mathbf{r})\} \setminus \{w\}$

   $\beta)$ $a \succsim_{(\mathbf{r},w)}^P b,$

   then $a \succsim_{\mathbf{r}}^P b$;

3. *given two alternatives* $a, b \in A$,

$$a \not\succsim_{(\mathbf{r},j)}^P b \ \ \forall j \in \{1, \dots, n(\mathbf{r})\} \ \Rightarrow \ a \not\succsim_{\mathbf{r}}^P b.$$

Proof. See Appendix. □

3.1.5 Intensity of preference

As in the GRIP method [33], also in case of the hierarchy of criteria it is possible to define quaternary relations $\succsim^{*^N}$, $\succsim^{*^P}$, $\succsim_{\mathbf{t}}^{*^N}$ and $\succsim_{\mathbf{t}}^{*^P}$, $\mathbf{t} \in EL$, related to intensity of preference, as follows:

- for each $a, b, c, d \in A$, we say that a is necessarily preferred to b at least as strongly as c is preferred to d, and we write $(a, b) \succsim^{*^N} (c, d)$, if a is preferred to b at least as strongly as c is preferred to d for all compatible value functions:

$$(a, b) \succsim^{*^N} (c, d) \Leftrightarrow U(a) - U(b) \geq U(c) - U(d), \forall U \in \mathcal{U};$$

- for each $a,b,c,d \in A$, we say that a is possibly preferred to b at least as strongly as c is preferred to d, and we write $(a,b) \succsim^{*^P} (c,d)$, if a is preferred to b at least as strongly as c is preferred to d for at least one compatible value function:

$$(a,b) \succsim^{*^P} (c,d) \Leftrightarrow \exists U \in \mathcal{U} : U(a) - U(b) \geq U(c) - U(d);$$

- for each $a,b,c,d \in A$, we say that a is necessarily preferred to b at least as strongly as c is preferred to d with respect to elementary subcriterion $g_{\mathbf{t}}$, and we write $(a,b) \succsim^{*^N}_{\mathbf{t}} (c,d)$, if a is preferred to b at least as strongly as c is preferred to d with respect to $g_{\mathbf{t}}$ for all compatible value functions:

$$(a,b) \succsim^{*^N}_{\mathbf{t}} (c,d) \Leftrightarrow u_{\mathbf{t}}(a) - u_{\mathbf{t}}(b) \geq u_{\mathbf{t}}(c) - u_{\mathbf{t}}(d), \forall U \in \mathcal{U};$$

- for each $a,b,c,d \in A$, we say that a is possibly preferred to b at least as strongly as c is preferred to d with respect to elementary subcriterion $g_{\mathbf{t}}$, and we write $(a,b) \succsim^{*^P}_{\mathbf{t}} (c,d)$, if a is preferred to b at least as strongly as c is preferred to d with respect to $g_{\mathbf{t}}$ for at least one compatible value function:

$$(a,b) \succsim^{*^P}_{\mathbf{t}} (c,d) \Leftrightarrow \exists U \in \mathcal{U} : u_{\mathbf{t}}(a) - u_{\mathbf{t}}(b) \geq u_{\mathbf{t}}(c) - u_{\mathbf{t}}(d);$$

In case of the hierarchy of criteria, we can further consider quaternary relations $\succsim^{*^N}_{\mathbf{r}}$ and $\succsim^{*^P}_{\mathbf{r}}$, related to intensity of preference with respect to subcriterion $G_{\mathbf{r}} \in \mathcal{G}$ at an intermediate level of the hierarchy, as follows:

- for each $a,b,c,d \in A$, and for each $\mathbf{r} \in \mathcal{I}_{\mathcal{G}}$, we say that a is necessarily preferred to b at least as strongly as c is preferred to d with respect to subcriterion $G_{\mathbf{r}}$, and we write $(a,b) \succsim^{*^N}_{\mathbf{r}} (c,d)$, if a is preferred to b at least as strongly as c is preferred to d with respect to subcriterion $G_{\mathbf{r}}$ for all compatible value functions:

$$(a,b) \succsim^{*^N}_{\mathbf{r}} (c,d) \Leftrightarrow U_{\mathbf{r}}(a) - U_{\mathbf{r}}(b) \geq U_{\mathbf{r}}(c) - U_{\mathbf{r}}(d), \forall U \in \mathcal{U};$$

- for each $a,b,c,d \in A$, and for each $\mathbf{r} \in \mathcal{I}_{\mathcal{G}}$, we say that a is possibly preferred to b at least as strongly as c is preferred to d with respect to subcriterion $G_{\mathbf{r}}$, and we write $(a,b) \succsim^{*^N}_{\mathbf{r}} (c,d)$, if a is preferred to b at least as strongly as c is preferred to d with respect to subcriterion $G_{\mathbf{r}}$ for at least one compatible value function:

$$(a,b) \succsim_{\mathbf{r}}^{*^P} (c,d) \Leftrightarrow \exists U \in \mathcal{U} : U_{\mathbf{r}}(a) - U_{\mathbf{r}}(b) \geq U_{\mathbf{r}}(c) - U_{\mathbf{r}}(d).$$

Observe that quaternary relations $\succsim_{\mathbf{t}}^{*^N}$ and $\succsim_{\mathbf{t}}^{*^P}$, $\mathbf{t} \in EL$, are a particular case of quaternary relations $\succsim_{\mathbf{r}}^{*^N}$ and $\succsim_{\mathbf{r}}^{*^P}$, $\mathbf{r} \in \mathcal{I}_{\mathcal{G}}$, in case $\mathbf{r} \in EL$.

Quaternary relations $\succsim^{*^N}$ and $\succsim^{*^P}$, $\succsim_{\mathbf{r}}^{*^N}$ and $\succsim_{\mathbf{r}}^{*^P}$, and $\succsim_{\mathbf{t}}^{*^N}$ and $\succsim_{\mathbf{t}}^{*^N}$ can be computed as follows. For all alternatives $a, b, c, d \in A$, let $X_{\mathbf{t}}(A^R \cup \{a,b,c,d\}) \subseteq X_{\mathbf{t}}$ be the set of all different evaluations of alternatives from $A^R \cup \{a,b,c,d\}$ on elementary subcriterion $g_{\mathbf{t}}, \mathbf{t} \in EL$, and $m_{\mathbf{t}}(A^R \cup \{a,b,c,d\}) = |X_{\mathbf{t}}(A^R \cup \{a,b,c,d\})|$. The values $x_{\mathbf{t}}^k \in X_{\mathbf{t}}(A^R \cup \{a,b,c,d\})$, $k = 1, \ldots, m_{\mathbf{t}}(A^R \cup \{a,b,c,d\})$, are increasingly ordered, i.e.,

$$x_{\mathbf{t}}^1 < x_{\mathbf{t}}^2 < \ldots < x_{\mathbf{t}}^{m_{\mathbf{t}}(A^R \cup \{a,b,c,d\})-1} < x_{\mathbf{t}}^{m_{\mathbf{t}}(A^R \cup \{a,b,c,d\})}.$$

Then, the characteristic points of $u_{\mathbf{t}}(\cdot)$, $\mathbf{t} \in EL$, are in $x_{\mathbf{t}}^0,\ x_{\mathbf{t}}^k,\ k = 1, \ldots, m_{\mathbf{t}}(A^R \cup \{a,b,c,d\}),\ x_{\mathbf{t}}^{m_{\mathbf{t}}}$.

Let us consider the following ordinal regression constraints $E(a,b,c,d)$, with $u_{\mathbf{t}}(x_{\mathbf{t}}^k)$, $\mathbf{t} \in EL,\ k = 1, \ldots, m_{\mathbf{t}}(A^R \cup \{a,b,c,d\})$, $u_{\mathbf{t}}(x_{\mathbf{t}}^{m_{\mathbf{t}}}), \mathbf{t} \in EL$, and ε as variables:

$$\left.\begin{array}{l}
\left.\begin{array}{l}
U(a^*) \geq U(b^*) + \varepsilon \quad \text{if} \quad a^* \succ b^* \\
U(a^*) = U(b^*) \quad \text{if} \quad a^* \sim b^* \\
U(a^*) - U(b^*) \geq U(c^*) - U(d^*) + \varepsilon \quad \text{if} \quad (a^*, b^*) \succ^* (c^*, d^*) \\
U(a^*) - U(b^*) = U(c^*) - U(d^*) \quad \text{if} \quad (a^*, b^*) \sim^* (c^*, d^*) \\
U_{\mathbf{r}}(a^*) \geq U_{\mathbf{r}}(b^*) + \varepsilon \quad \text{if} \quad a^* \succ_{\mathbf{r}} b^* \\
U_{\mathbf{r}}(a^*) = U_{\mathbf{r}}(b^*) \quad \text{if} \quad a^* \sim_{\mathbf{r}} b^* \\
U_{\mathbf{r}}(a^*) - U_{\mathbf{r}}(b^*) \geq U_{\mathbf{r}}(c^*) - U_{\mathbf{r}}(d^*) + \varepsilon \quad \text{if} \quad (a^*, b^*) \succ_{\mathbf{r}}^* (c^*, d^*) \\
U_{\mathbf{r}}(a^*) - U_{\mathbf{r}}(b^*) = U_{\mathbf{r}}(c^*) - U_{\mathbf{r}}(d^*) \quad \text{if} \quad (a^*, b^*) \sim_{\mathbf{r}}^* (c^*, d^*)
\end{array}\right\} \begin{array}{l} a^*, b^*, c^*, d^* \in A^R; \\ \mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL \end{array} \\
u_{\mathbf{t}}(x_{\mathbf{t}}^k) - u_{\mathbf{t}}(x_{\mathbf{t}}^{k-1}) \geq 0,\ \mathbf{t} \in EL,\ k = 2, ..., m_{\mathbf{t}}(A^R \cup \{a,b,c,d\}) \\
u_{\mathbf{t}}(x_{\mathbf{t}}^1) \geq 0,\ u_{\mathbf{t}}(x_{\mathbf{t}}^{m_{\mathbf{t}}(A^R \cup \{a,b,c,d\})}) \leq u_{\mathbf{t}}(x_{\mathbf{t}}^{m_{\mathbf{t}}}),\ \mathbf{t} \in EL \\
u_{\mathbf{t}}(x_{\mathbf{t}}^0) = 0,\ \mathbf{t} \in EL \\
\sum_{\mathbf{t} \in EL} u_{\mathbf{t}}(x_{\mathbf{t}}^{m_{\mathbf{t}}}) = 1.
\end{array}\right\} (E(a,b,c,d))$$

The above constraints depend also on the alternatives $a, b, c, d \in A$ because their evaluations $g_{\mathbf{t}}(a)$, $g_{\mathbf{t}}(b)$, $g_{\mathbf{t}}(c)$ and $g_{\mathbf{t}}(d)$ give coordinates to four of $m_{\mathbf{t}}(A^R \cup \{a,b,c,d\})$ characteristic points of marginal value function $u_{\mathbf{t}}(\cdot)$, for each $\mathbf{t} \in EL$.

For all $a, b, c, d \in A$, and $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, let us consider the following sets of constraints:

$$\left.\begin{array}{l} U(c) - U(d) \geq U(a) - U(b) + \varepsilon \\ E(a,b,c,d) \end{array}\right\} (E^N(a,b,c,d)), \qquad \left.\begin{array}{l} U(a) - U(b) \geq U(c) - U(d) \\ E(a,b,c,d) \end{array}\right\} (E^P(a,b,c,d)),$$

$$\left.\begin{array}{l} U_{\mathbf{r}}(c) - U_{\mathbf{r}}(d) \geq U_{\mathbf{r}}(a) - U_{\mathbf{r}}(b) + \varepsilon \\ E(a,b,c,d) \end{array}\right\} (E^N_{\mathbf{r}}(a,b,c,d)), \qquad \left.\begin{array}{l} U_{\mathbf{r}}(a) - U_{\mathbf{r}}(b) \geq U_{\mathbf{r}}(c) - U_{\mathbf{r}}(d) \\ E(a,b,c,d) \end{array}\right\} (E^P_{\mathbf{r}}(a,b,c,d)),$$

$$\left.\begin{array}{l} U_{\mathbf{t}}(c) - U_{\mathbf{t}}(d) \geq U_{\mathbf{t}}(a) - U_{\mathbf{t}}(b) + \varepsilon \\ E(a,b,c,d) \end{array}\right\} (E^N_{\mathbf{t}}(a,b,c,d)), \qquad \left.\begin{array}{l} U_{\mathbf{t}}(a) - U_{\mathbf{t}}(b) \geq U_{\mathbf{t}}(c) - U_{\mathbf{t}}(d) \\ E(a,b,c,d) \end{array}\right\} (E^P_{\mathbf{t}}(a,b,c,d)).$$

Thus we get:

- $(a,b) \succsim^{*^N} (c,d)$ iff $E^N(a,b,c,d)$ is infeasible or $\varepsilon^N(a,b,c,d) \leq 0$, where $\varepsilon^N(a,b,c,d) = \max \varepsilon$, s.t. constraints $E^N(a,b,c,d)$;
- $(a,b) \succsim^{*^P} (c,d)$ iff $E^P(a,b,c,d)$ is feasible and $\varepsilon^P(a,b,c,d) > 0$, where $\varepsilon^P(a,b,c,d) = \max \varepsilon$, s.t. constraints $E^P(a,b,c,d)$;
- $(a,b) \succsim^{*^N}_{\mathbf{r}} (c,d)$ iff $E^N_{\mathbf{r}}(a,b,c,d)$ is infeasible or $\varepsilon^N_{\mathbf{r}}(a,b,c,d) \leq 0$, where $\varepsilon^N_{\mathbf{r}}(a,b,c,d) = \max \varepsilon$, s.t. constraints $E^N_{\mathbf{r}}(a,b,c,d)$;
- $(a,b) \succsim^{*^P}_{\mathbf{r}} (c,d)$ iff $E^P_{\mathbf{r}}(a,b,c,d)$ is feasible and $\varepsilon^P_{\mathbf{r}}(a,b,c,d) > 0$, where $\varepsilon^P_{\mathbf{r}}(a,b,c,d) = \max \varepsilon$, s.t. constraints $E^P_{\mathbf{r}}(a,b,c,d)$;
- $(a,b) \succsim^{*^N}_{\mathbf{t}} (c,d)$ iff $E^N_{\mathbf{t}}(a,b,c,d)$ is infeasible or $\varepsilon^N_{\mathbf{t}}(a,b,c,d) \leq 0$, where $\varepsilon^N_{\mathbf{t}}(a,b,c,d) = \max \varepsilon$ s.t. constraints $E^N_{\mathbf{t}}(a,b,c,d)$;
- $(a,b) \succsim^{*^P}_{\mathbf{t}} (c,d)$ iff $E^P_{\mathbf{t}}(a,b,c,d)$ is feasible and $\varepsilon^P_{\mathbf{t}}(a,b,c,d) > 0$, where $\varepsilon^P_{\mathbf{t}}(a,b,c,d) = \max \varepsilon$, s.t. constraints $E^P_{\mathbf{t}}(a,b,c,d)$.

Most of the properties of quaternary relations $\succsim^{*N}$ and $\succsim^{*P}$, $\succsim_{\mathbf{r}}^{*N}$ and $\succsim_{\mathbf{r}}^{*P}$, and $\succsim_{\mathbf{t}}^{*N}$ and $\succsim_{\mathbf{t}}^{*N}$ are the same of those of the GRIP method presented in [33]. However, there are some properties specific to the case of the hierarchy of criteria, which are presented in the following proposition.

Proposition 3.1.4. *For all $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$,*

1. *given four alternatives $a, b, c, d \in A$,*

$$(a,b) \succsim_{(\mathbf{r},j)}^{*N} (c,d), \forall j = 1, \ldots, n(\mathbf{r}) \Rightarrow (a,b) \succsim_{\mathbf{r}}^{*N} (c,d);$$

2. *given four alternatives $a, b, c, d \in A$ such that:*

   (a) $(a,b) \succsim_{(\mathbf{r},j)}^{*N} (c,d) \quad \forall j \in \{1, \ldots, n(\mathbf{r})\} \setminus \{w\}$,

   (b) $(a,b) \succsim_{(\mathbf{r},w)}^{*P} (c,d)$,

   then $(a,b) \succsim_{\mathbf{r}}^{*P} (c,d)$;

3. *given four alternatives $a, b, c, d \in A$,*

$$(a,b) \not\succsim_{(\mathbf{r},j)}^{*P} (c,d) \;\; \forall j \in \{1, \ldots, n(\mathbf{r})\} \;\Rightarrow\; (a,b) \not\succsim_{\mathbf{r}}^{*P} (c,d).$$

Proof. See Appendix. □

3.1.6 The representative value function

The ROR in case of the hierarchy of criteria builds a set of additive value functions compatible with preference information provided by the DM and leads to two preference relations, $\succsim_{\mathbf{r}}^{N}$ and $\succsim_{\mathbf{r}}^{P}$, for each subcriterion $G_{\mathbf{r}}$, $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, from the hierarchy. Such preference relations answer to robustness concerns, since they are in general "more robust" than a preference relation determined by an arbitrarily chosen compatible value function. However, in practice, in some decision-making situations it is required to assign a score to considered alternatives. Moreover, possible and necessary preference relations may be not easy to interpret, even by a DM with some experience in MCDA. Thus, it is useful to determine a value function which represents well all the information contained in necessary and possible preference relations in an easily understandable way. For these reasons, a method for finding among all compatible value functions resulting from ROR a "representative" value function has been proposed in [49],[75]. It is based on the principle of "one for all, all for one", i.e. we look for one value function representing the set of all compatible value functions, and all compatible value functions contribute to define this representative value function.

In case of the hierarchy of criteria, the DM can be interested in a value function representing not only comprehensive necessary and possible preference relations, $\succsim^N$ and $\succsim^P$, but also necessary and possible preference relations $\succsim^N_{\mathbf{r}}$ and $\succsim^P_{\mathbf{r}}$, $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, at intermediate levels. In general, the idea of the "representative value function" is to select from among compatible value functions that one which better highlights the necessary preference by maximizing the difference of values between alternatives $a, b \in A$ for which $a \succ^N b$, i.e. $a \succsim^N b$ and $b \not\succsim^N a$. As secondary objective, one can consider minimizing the difference of values between alternatives $a, b \in A$ for which $a \not\succsim^N b$ and $b \not\succsim^N a$. In case of the hierarchy of criteria one can imagine that the DM gives a sequence of criteria $G_{\mathbf{r}_1}, \ldots, G_{\mathbf{r}_f} \in \mathcal{G}$, ordered with respect to his/her interest. In this case, the representative value function is the one maximizing the difference of values between alternatives $a, b \in A$ for which $a \succ^N_{\mathbf{r}_i} b$, and minimizing the difference of values between alternatives $a, b \in A$ for which $a \not\succsim^N_{\mathbf{r}_i} b$ and $b \not\succsim^N_{\mathbf{r}_i} a$, starting from the most interesting subcriterion $G_{\mathbf{r}_1}$ and proceeding in the above sequence until subcriterion $G_{\mathbf{r}_f}$. In this way, the discrimination power of the "representative value function" is maximal for the most interesting subcriterion $G_{\mathbf{r}_1}$, and it is decreasing, step by step, until subcriterion $G_{\mathbf{r}_f}$. Summing up, the "representative" value function can be found via the following procedure:

1. Consider the set of constraints E^A including constraints representing preference information provided by the DM, and monotonicity constraints on marginal value functions $u_{\mathbf{t}}(\cdot)$, $\mathbf{t} \in EL$, whose characteristic points correspond to all different evaluations of alternatives from set A

(and not only from the reference subset $A^R \subseteq A$) on particular elementary criteria:

$$\left.\begin{array}{l}
\left.\begin{array}{l}
U(a^*) \geq U(b^*) + \varepsilon \quad \text{if} \;\; a^* \succ b^* \\
U(a^*) = U(b^*) \quad \text{if} \;\; a^* \sim b^* \\
U(a^*) - U(b^*) \geq U(c^*) - U(d^*) + \varepsilon \quad \text{if} \;\; (a^*, b^*) \succ^* (c^*, d^*) \\
U(a^*) - U(b^*) = U(c^*) - U(d^*) \quad \text{if} \;\; (a^*, b^*) \sim^* (c^*, d^*) \\
U_{\mathbf{r}}(a^*) \geq U_{\mathbf{r}}(b^*) + \varepsilon \quad \text{if} \;\; a^* \succ_{\mathbf{r}} b^* \\
U_{\mathbf{r}}(a^*) = U_{\mathbf{r}}(b^*) \quad \text{if} \;\; a^* \sim_{\mathbf{r}} b^* \\
U_{\mathbf{r}}(a^*) - U_{\mathbf{r}}(b^*) \geq U_{\mathbf{r}}(c^*) - U_{\mathbf{r}}(d^*) + \varepsilon \quad \text{if} \;\; (a^*, b^*) \succ^*_{\mathbf{r}} (c^*, d^*) \\
U_{\mathbf{r}}(a^*) - U_{\mathbf{r}}(b^*) = U_{\mathbf{r}}(c^*) - U_{\mathbf{r}}(d^*) \quad \text{if} \;\; (a^*, b^*) \sim^*_{\mathbf{r}} (c^*, d^*)
\end{array}\right\} \begin{array}{l} a^*, b^*, c^*, d^* \in A^R, \\ \mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL, \end{array} \\
u_{\mathbf{t}}(x^k_{\mathbf{t}}) - u_{\mathbf{t}}(x^{k-1}_{\mathbf{t}}) \geq 0, \;\; \forall \mathbf{t} \in EL, \;\; k = 1, ..., m_{\mathbf{t}} \\
u_{\mathbf{t}}(x^1_{\mathbf{t}}) = 0, \;\; \forall \mathbf{t} \in EL \\
\sum_{\mathbf{t} \in EL} u_{\mathbf{t}}(x^{m_{\mathbf{t}}}_{\mathbf{t}}) = 1,
\end{array}\right\} (E^A)$$

where, $x^1_{\mathbf{t}} = min_{a \in A} \; g_{\mathbf{t}}(a)$, and $x^{m_{\mathbf{t}}}_{\mathbf{t}} = max_{a \in A} \; g_{\mathbf{t}}(a)$; $x^k_{\mathbf{t}} \in X_{\mathbf{t}}, k = 1, ..., m_{\mathbf{t}}$, with $X_{\mathbf{t}}$ the set of all different evaluations of alternatives from A on elementary subcriteria $g_{\mathbf{t}}, \mathbf{t} \in EL$, and $m_{\mathbf{t}} = |X_{\mathbf{t}}|$. The values $x^k_{\mathbf{t}}, k = 1, ..., m_{\mathbf{t}}$, are increasingly ordered, i.e.,

$$x^1_{\mathbf{t}} < x^2_{\mathbf{t}} < ... < x^{m_{\mathbf{t}}-1}_{\mathbf{t}} < x^{m_{\mathbf{t}}}_{\mathbf{t}}.$$

2. Calculate $\varepsilon^* = \max \; \varepsilon, \;$ s.t. E^A. If $\varepsilon^* > 0$, then there exists at least one value function satisfying constraints of E^A, so go to step 3. If $\varepsilon^* \leq 0$, then there is no value function satisfying E^A, which means that the information provided by the DM cannot be faithfully represented by any additive value function. If the DM accepts to work with not fully compatible value functions, then go to step 3; if the DM decides to remove a part of preference information causing the incompatibility, then after this removal (see section 3.1.8), go to step 3,

3. $i = 1$; $E = E^A$,

4. Determine the necessary preference relation $\succsim^N_{\mathbf{r}_i}$ and the possible preference relation $\succsim^P_{\mathbf{r}_i}$ with respect to subcriterion $G_{\mathbf{r}_i} \in \mathcal{G}$, considering the sets of constraints:

$$\left.\begin{array}{l} U_{\mathbf{r}_i}(b) \geq U_{\mathbf{r}_i}(a) + \varepsilon \\ E^A \end{array}\right\} (E^N_{\mathbf{r}_i}(a,b)), \qquad \left.\begin{array}{l} U_{\mathbf{r}_i}(a) \geq U_{\mathbf{r}_i}(b) \\ E^A \end{array}\right\} (E^P_{\mathbf{r}_i}(a,b)).$$

- $a \succsim^N_{\mathbf{r}_i} b \Leftrightarrow \varepsilon^{*,N}_{\mathbf{r}_i} \leq 0$, where $\varepsilon^{*,N}_{\mathbf{r}_i} = \max \varepsilon$, s.t. constraints $E^N_{\mathbf{r}_i}(a,b)$,
- $a \succsim^P_{\mathbf{r}_i} b \Leftrightarrow \varepsilon^{*,P}_{\mathbf{r}_i} > 0$, where $\varepsilon^{*,P}_{\mathbf{r}_i} = \max \varepsilon$, s.t. constraints $E^P_{\mathbf{r}_i}(a,b)$.

5. For all pairs of alternatives (a,b), such that $a \succ^N_{\mathbf{r}_i} b$, add the following constraint to E: $U_{\mathbf{r}_i}(a) \geq U_{\mathbf{r}_i}(b) + \varepsilon_{\mathbf{r}_i}$; if $i = 1$, then go to step 6, otherwise go to step 7,

$$\left.\begin{array}{l} E \\ U_{\mathbf{r}_i}(a) \geq U_{\mathbf{r}_i}(b) + \varepsilon_{\mathbf{r}_i} \quad \text{if } \ a \succ_{\mathbf{r}_i} b \end{array}\right\} \rightarrow (E)$$

6. Add constraint $\varepsilon_{\mathbf{r}_i} = \varepsilon$ to E,

$$\left.\begin{array}{l} E \\ \varepsilon_{\mathbf{r}_i} = \varepsilon \end{array}\right\} \rightarrow (E)$$

7. Maximize $\varepsilon_{\mathbf{r}_i}$, subject to constraints E.

8. Add the constraint $\varepsilon_{\mathbf{r}_i} = \varepsilon^*_{\mathbf{r}_i}$ to E, with $\varepsilon^*_{\mathbf{r}_i} = \max \varepsilon_{\mathbf{r}_i}$ computed in step 7,

$$\left.\begin{array}{l} E \\ \varepsilon_{\mathbf{r}_i} = \varepsilon^*_{\mathbf{r}_i} \end{array}\right\} \rightarrow (E)$$

9. For all pairs of alternatives (a,b), such that $a \not\succsim^N_{\mathbf{r}_i} b$ and $b \not\succsim^N_{\mathbf{r}_i} a$ (already computed in step 4), add the following constraints to E: $U_{\mathbf{r}_i}(a) - U_{\mathbf{r}_i}(b) \leq \delta_{\mathbf{r}_i}$ and $U_{\mathbf{r}_i}(b) - U_{\mathbf{r}_i}(a) \leq \delta_{\mathbf{r}_i}$,

$$\left.\begin{array}{l} E \\ \left.\begin{array}{l} U_{\mathbf{r}_i}(a) - U_{\mathbf{r}_i}(b) \leq \delta_{\mathbf{r}_i} \\ U_{\mathbf{r}_i}(b) - U_{\mathbf{r}_i}(a) \leq \delta_{\mathbf{r}_i} \end{array}\right\} \text{ if } \ a \not\succsim^N_{\mathbf{r}_i} b \ \text{ and } \ b \not\succsim^N_{\mathbf{r}_i} a \end{array}\right\} \rightarrow (E)$$

10. Minimize $\delta_{\mathbf{r}_i}$, subject to constraints E.

11. Add the constraint $\delta_{\mathbf{r}_i} = \delta^*_{\mathbf{r}_i}$ to E, with $\delta^*_{\mathbf{r}_i} = \min \delta_{\mathbf{r}_i}$ computed in step 10,

$$\left.\begin{array}{l} E \\ \delta_{\mathbf{r}_i} = \delta^*_{\mathbf{r}_i} \end{array}\right\} \rightarrow (E)$$

12. If $i < f$ then go to step 4 with $i := i + 1$, otherwise stop.

Observe that the above procedure takes into account the preference information given by the DM by maximizing the value of auxiliary variable ε in the first iteration. This ensures that the DM's preferences are represented with a maximal discrimination possible. If the DM does not want to express a sequence of subcriteria $G_{\mathbf{r}_1}, \ldots, G_{\mathbf{r}_f} \in \mathcal{G}$, but (s)he wants to compute the representative value function considering only the comprehensively necessary preference relation, it will be enough to perform a single iteration of the procedure described until step 10, considering $i = 1$ and $\mathbf{r}_1 = 0$.

Let us mention that other methods proposed for finding a representative value function in ordinal regression [13, 14], not referring to necessary and possible preference relations, can also be adapted to the case of hierarchy of criteria.

3.1.7 A didactic example

In this section, we apply the procedure described in the previous sections to cope with a hierarchical multiple criteria decision problem which is very frequent in the scholar system, and in the academic sector in particular. Let us suppose that each year a faculty of natural sciences has the economic possibility to give a scholarship to one of its best students; to make the choice, the Dean is considering fifteen students who attended the courses and passed the test of two macro subjects: Mathematics and Chemistry. Mathematics has two sub-subjects: Algebra and Analysis, while Chemistry has two sub-subjects: Analytical Chemistry and Organic Chemistry; each of these sub-subjects has other two sub-subjects for a total of eight elementary sub-subjects shown in Figure 3.2. Using the terminology introduced in Section 3.1.2, the set of alternatives $A = \{\mathbf{A}, \mathbf{B}, \ldots, \mathbf{R}\}$ is composed of 15 alternatives; the number of levels $l = 3$; the set of all criteria $\mathcal{G} = \{G_1, G_2,\ G_{(1,1)}, G_{(1,2)},\ G_{(2,1)}, G_{(2,2)},\ g_{(1,1,1)}, g_{(1,1,2)},\ g_{(1,2,1)}, g_{(1,2,2)},\ g_{(2,1,1)}, g_{(2,1,2)}, g_{(2,2,1)}, g_{(2,2,2)}\}$ is composed of criteria and subcriteria whose names are given in Figure 3.2; the set of indices of all criteria is $\mathcal{I}_{\mathcal{G}} = \{1, 2,\ (1,1), (1,2),\ (2,1), (2,2),\ (1,1,1), (1,1,2),\ (1,2,1), (1,2,2), (2,1,1), (2,1,2),\ (2,2,1), (2,2,2)\}$; the number of first level criteria $m = 2$; if we consider $G_{\mathbf{r}} = G_1$ then $n(\mathbf{r}) = 2$, while if we consider $G_{\mathbf{r}} = G_{(1,1)}$ then $n((1,1)) = 2$; $g_{(1,1,1)}, g_{(1,1,2)},\ g_{(1,2,1)}, g_{(1,2,2)}, g_{(2,1,1)}, g_{(2,1,2)},\ g_{(2,2,1)}, g_{(2,2,2)}$ are the elementary subcriteria; the set of indices of elementary subcriteria is $EL = \{(1,1,1), (1,1,2),\ (1,2,1), (1,2,2),\ (2,1,1), (2,1,2),\ (2,2,1), (2,2,2)\}$; if we consider $G_{\mathbf{r}} = G_1$ then $E(G_{(1)}) = \{(1,1,1), (1,1,2),\ (1,2,1), (1,2,2)\}$ while if we consider $G_{\mathbf{r}} = G_{(2,1)}$ then $E(G_{(2,1)}) = \{(2,1,1), (2,1,2)\}$.

As it was declared in Section 3.1.2, the students are evaluated directly on the elementary subcriteria only, and thus, they are evaluated with respect to the eight elementary sub-subjects; these evaluations are shown in Table 3.1. Each elementary subcriterion has five qualitative levels of evaluation that go from very bad to very good, increasingly ordered.

Figure 3.2: Hierarchical structure of criteria

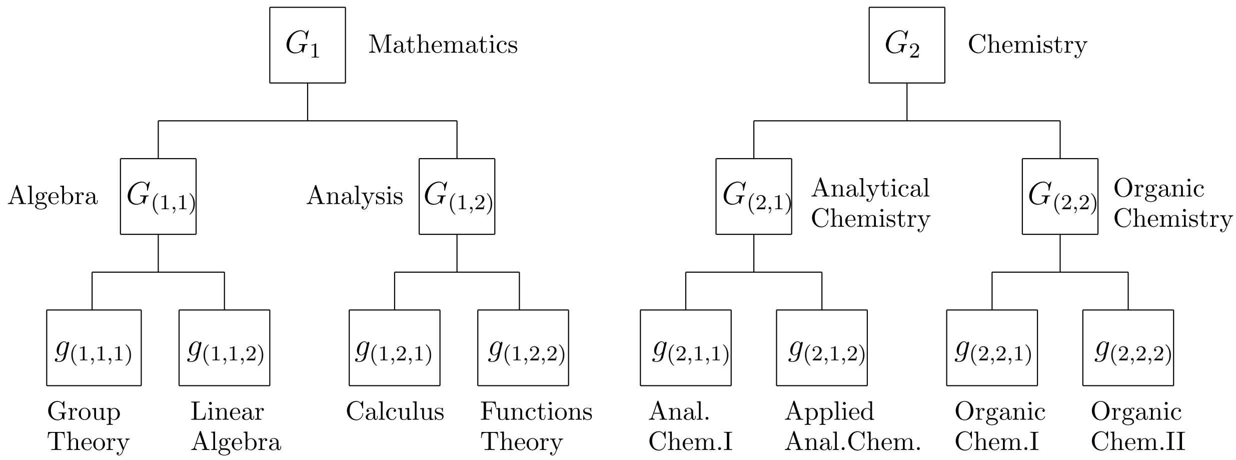
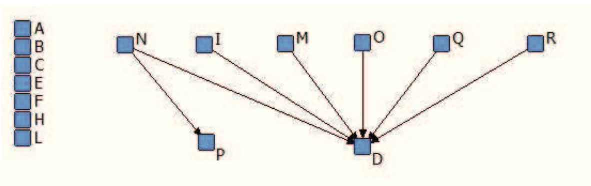


Table 3.1: Evaluations of students on the eight elementary subcriteria

student\subcriteria	$g_{(1,1,1)}$	$g_{(1,1,2)}$	$g_{(1,2,1)}$	$g_{(1,2,2)}$	$g_{(2,1,1)}$	$g_{(2,1,2)}$	$g_{(2,2,1)}$	$g_{(2,2,2)}$
A	Very Bad	Very Good	Very Bad	Good	Very Good	Very Good	Very Bad	Bad
B	Bad	Very Good	Medium	Very Good	Very Bad	Bad	Very Bad	Very Bad
C	Very Good	Medium	Medium	Very Bad	Very Good	Good	Bad	Medium
D	Medium	Very Bad	Bad	Very Bad	Very Bad	Bad	Medium	Very Bad
E	Very Good	Very Good	Medium	Medium	Bad	Very Good	Bad	Very Bad
F	Good	Bad	Bad	Medium	Very Bad	Very Bad	Very Good	Very Good
H	Medium	Very Bad	Bad	Bad	Very Good	Very Bad	Very Bad	Very Bad
I	Good	Good	Good	Medium	Medium	Bad	Good	Very Bad
L	Good	Very Bad	Bad	Good	Good	Very Bad	Very Good	Good
M	Medium	Medium	Medium	Bad	Medium	Medium	Very Good	Good
N	Good	Bad	Very Good	Medium	Bad	Very Good	Very Good	Medium
O	Good	Medium	Bad	Bad	Medium	Bad	Very Good	Very Bad
P	Bad	Very Bad	Bad	Medium	Bad	Very Good	Medium	Very Bad
Q	Very Good	Very Good	Medium	Very Bad	Bad	Medium	Medium	Bad
R	Good	Good	Bad	Very Bad	Bad	Bad	Medium	Medium

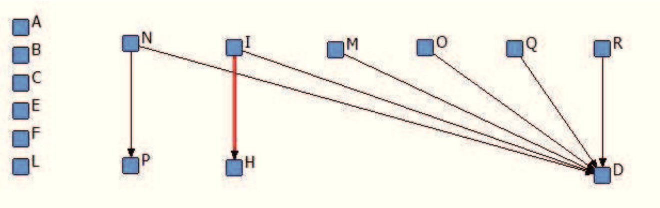
Figure 3.3: Dominance relation in the set of students



The only comprehensive relation that comes out from the problem formulation is the dominance relation in the set of students, shown in Figure 3.3. The dominance relation does not take into account the preferences of the Dean and, moreover, it leaves too many students incomparable. For this reason, the Dean decides to use the ROR approach adapted to the hierarchical structure of criteria.

The Dean provides the following preference information which is then transformed to constraints of

Figure 3.4: Necessary preference relation determined by the first piece of preference information



the ordinal regression problem:

1. On Chemistry, student **I** is preferred to student **H**. In order to take into consideration this preference information it is represented in the constraints (E^{A^R}) as follows:

$$U_2(I) > U_2(H) \Leftrightarrow U_{(2,1)}(\mathbf{I}) + U_{(2,2)}(\mathbf{I}) > U_{(2,1)}(\mathbf{H}) + U_{(2,2)}(\mathbf{H}) \quad \Leftrightarrow$$

$$\Leftrightarrow u_{(2,1,1)}(\mathbf{I}) + u_{(2,1,2)}(\mathbf{I}) + u_{(2,2,1)}(\mathbf{I}) + u_{(2,2,2)}(\mathbf{I}) > u_{(2,1,1)}(\mathbf{H}) + u_{(2,1,2)}(\mathbf{H}) + u_{(2,2,1)}(\mathbf{H}) + u_{(2,2,2)}(\mathbf{H}).$$

   Figure 3.4 shows the necessary preference relation determined by this piece of preference information. In Figure 3.4, the arrow from **I** to **H** is bold marked because it constitutes the part of necessary preference relation originating from the considered piece of preference information and, therefore, not present at the previous stage (dominance relation, see Figure 3.3). Bold marked arrows in the following figures have an analogous interpretation with respect to preference information provided in further steps.

2. On Analytical Chemistry, student **E** is preferred to student **H**. This, can be modeled using the following constraint:

$$U_{(2,1)}(\mathbf{E}) > U_{(2,1)}(\mathbf{H}) \Leftrightarrow u_{(2,1,1)}(\mathbf{E}) + u_{(2,1,2)}(\mathbf{E}) > u_{(2,1,1)}(\mathbf{H}) + u_{(2,1,2)}(\mathbf{H}).$$

   Figure 3.5 shows the necessary preference relation determined by the two pieces of preference information.

3. On Mathematics, student **N** is preferred to student **Q**. This, can be modeled using the following constraint:

$$U_1(\mathbf{N}) > U_1(\mathbf{Q}) \Leftrightarrow U_{(1,1)}(\mathbf{N}) + U_{(1,2)}(\mathbf{N}) > U_{(1,1)}(\mathbf{Q}) + U_{(1,2)}(\mathbf{Q}) \quad \Leftrightarrow$$

$$\Leftrightarrow u_{(1,1,1)}(\mathbf{N}) + u_{(1,1,2)}(\mathbf{N}) + u_{(1,2,1)}(\mathbf{N}) + u_{(1,2,2)}(\mathbf{N}) > u_{(1,1,1)}(\mathbf{Q}) + u_{(1,1,2)}(\mathbf{Q}) + u_{(1,2,1)}(\mathbf{Q}) + u_{(1,2,2)}(\mathbf{Q}).$$

Figure 3.5: Necessary preference relation determined by the two pieces of preference information

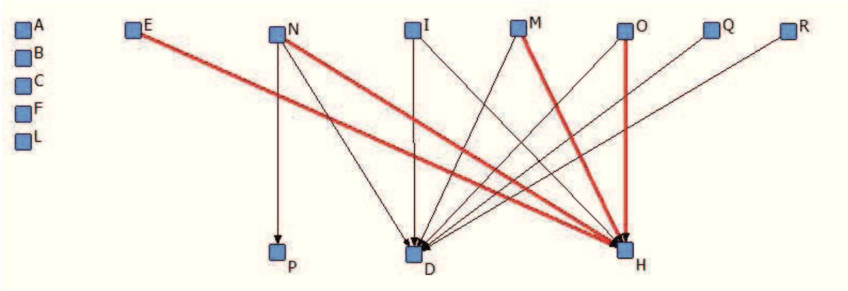


Figure 3.6: Necessary preference relation determined by the three pieces of preference information

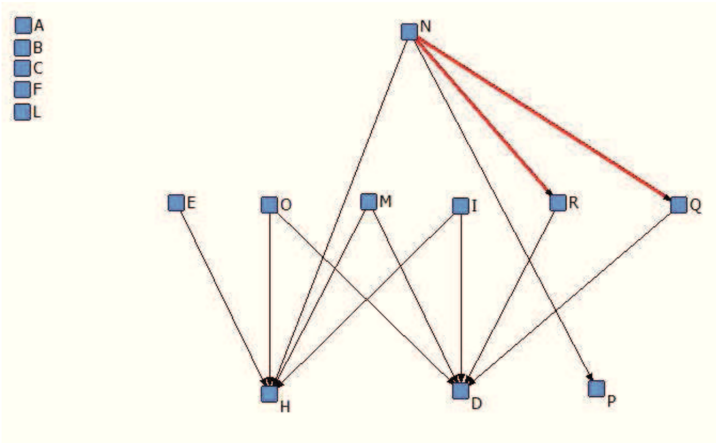


Figure 3.6 shows the necessary preference relation determined by the three pieces of preference information.

4. On Chemistry, student **L** is preferred to student **P**. This, can be modeled using the following constraint:

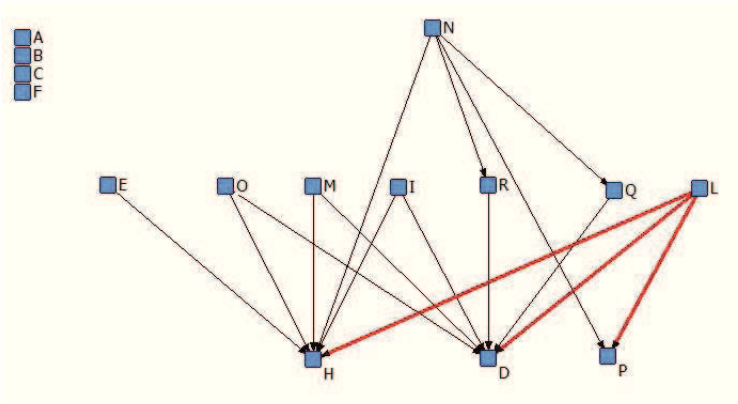
$$U_2(\mathbf{L}) > U_2(\mathbf{P}) \Leftrightarrow U_{(2,1)}(\mathbf{L}) + U_{(2,2)}(\mathbf{L}) > U_{(2,1)}(\mathbf{P}) + U_{(2,2)}(\mathbf{P}) \quad \Leftrightarrow$$

$$\Leftrightarrow u_{(2,1,1)}(\mathbf{L})+u_{(2,1,2)}(\mathbf{L})+u_{(2,2,1)}(\mathbf{L})+u_{(2,2,2)}(\mathbf{L}) > u_{(2,1,1)}(\mathbf{P})+u_{(2,1,2)}(\mathbf{P})+u_{(2,2,1)}(\mathbf{P})+u_{(2,2,2)}(\mathbf{P}).$$

Figure 3.7 shows the necessary preference relation determined by the four pieces of preference information.

In the context of the hierarchical multiple criteria evaluation, it is possible to check the necessary preference relation at intermediate levels of the hierarchy, that is we can see if student a is necessarily preferred to student b with respect to considered domain (Mathematics, Chemistry, Algebra, Analysis and so on); in Tables 3.2 and 3.3, we present the necessary preference relation with respect to macro

Figure 3.7: Necessary preference relation determined by the four pieces of preference information



subjects: Mathematics and Chemistry, respectively.

Table 3.2: Necessary preference relations for Mathematics and its subcriteria

student\subcriterion	$\succsim^{N}_{1}$	$\succsim^{N}_{(1,1)}$	$\succsim^{N}_{(1,2)}$
A			
B	A, P	A, P	A, C, D, E, F, H, L, M, O, P, Q, R
C	D	D, F, H, L, M, N, O, P	D, Q, R
D		H, P	R
E	C, D, F, H, M, O, P, Q, R	A, B, C, D, F, H, I, L, M, N, O, P, Q, R	C, D, F, H, M, O, P, Q, R
F	D, H, P	D, H, L, N, P	D, H, O, P, R
H	D	D, P	D, O, R
I	D, F, H, M, O, P, R	D, F, H, L, M, N, O, P, R	C, D, E, F, H, M, O, P, Q, R
L	D, H, P	D, H, P	A, D, F, H, O, P, R
M	D, H	D, H, P	C, D, H, O, Q, R
N	*C*,D, F, H, P, *Q*,*R*	D, F, H, L, P	C, D, E, F, H, I, M, O, P, Q, R
O	D, H	D, F, H, L, M, N, P	D, H, R
P			D, F, H, O, R
Q	C, D, R	A, B, C, D, E, F, H, I, L, M, N, O, P, R	C, D, R
R	D	D, F, H, I, L, M, N, O, P	D

Table 3.3: Necessary preference relation for Chemistry and its subcriteria

student\subcriterion	$\succsim^{N}_{2}$	$\succsim^{N}_{(2,1)}$	$\succsim^{N}_{(2,2)}$
A	B, H	B, C, D, E, F, H, I, L, M, N, O, P, Q, R	B, H
B		D, F	H
C	B, H	B, D, F, H, I, L, M, O, Q, R	A, B, E, H
D	B	B, F	B, E, H, P
E	B, H	B, D, F, H, L, N, P, Q, R	B, H
F			A, B, C, D, E, H, I, L, M, N, O, P, Q, R
H		F, L	B
I	B, D,*H*	B, D, F, O, R	B, D, E, H, P
L	*B,D,H,P*	F	A, B, C, D, E, H, I, M, N, O, P, Q, R
M	B, D, H, I, O, Q, R	B, D, F, I, O, Q, R	A, B, C, D, E, H, I, L, N, O, P, Q, R
N	B, D, E, H, P, Q, R	B, D, E, F, H, L, P, Q, R	A, B, C, D, E, H, I, O, P, Q, R
O	B, D,*H*,I	B, D, F, I, R	B, D, E, H, I, P
P	B, D, E, H	B, D, E, F, H, L, N, Q, R	B, D, E, H
Q	B, D	B, D, F, R	A, B, D, E, H, P
R	B, D	B, D, F	A, B, C, D, E, H, P, Q

In Tables 3.2 and 3.3, the alternatives in italics are those for which the necessary preference relation is true at the second level but it is not true at the level below. For example, $\mathbf{L} \succsim^{N}_{2} \mathbf{B}$ but $\mathbf{L} \not\succsim^{N}_{(2,1)} \mathbf{B}$.

As shown in subsection 3.1.6, one can compute the representative value function, taking into account a sequence of subcriteria $G_{\mathbf{r}_1}, \ldots, G_{\mathbf{r}_f} \in \mathcal{G}$ ordered with respect to the Dean's interest. Results presented in Table 3.4 show the ranking of students obtained using the representative value function in three different cases:

- the Dean considers as the most important and the second most important the criteria Mathematics (G_1) and Chemistry (G_2), respectively, and consequently, he considers the sequence of corresponding necessary preference relations $\succsim_1^N, \succsim_2^N$ (1^{st} and 2^{nd} columns),
- the Dean considers as the most important and the second most important the criteria Chemistry (G_2) and Mathematics (G_1), respectively, and consequently, he considers the sequence of corresponding necessary preference relations $\succsim_2^N, \succsim_1^N$ (3^{rd} and 4^{th} columns),
- the Dean does not discriminate criteria with respect to their importance and consequently he takes into account only the comprehensive necessary preference relation $\succsim_0^N$ (5^{th} column).

We can observe three important facts:

- student **N** is almost always the best one in the ranking obtained using different representative value functions,
- the ranking obtained by the representative value function changes between the first and the second iteration of the method,
- the ranking obtained by the representative value function changes if we consider a different order of importance between the necessary preference relations.

3.1.8 Further extensions of ROR for the hierarchy of criteria

Infeasibility

We have seen in section 3.1.3, that the first step of ROR is to check if there exists at least one value function compatible with the preference information provided by the DM. In fact, it is possible that the information provided by the DM is such that it is not possible to find a compatible additive value function. In this case, the DM, together with the analyst, can decide to continue the study while accepting to work with not fully compatible value functions, or look for sets of constraints responsible of this infeasibility (let us call them troublesome constraints), and remove them from the linear program.

Table 3.4: Ranking of students by a representative value function (in parentheses there are value of the corresponding alternatives)

$\succsim^N_{\mathbf{r_1}} = \succsim^N_1$	$\succsim^N_{\mathbf{r_2}} = \succsim^N_2$	$\succsim^N_{\mathbf{r_1}} = \succsim^N_2$	$\succsim^N_{\mathbf{r_2}} = \succsim^N_1$	$\succsim^N$
N(0.8560)	**N**(0.8586)	**N**(1)	**N**(1)	**M**(0.8808)
I(0.6635)	**I**(0.6949)	**M**(0.8752)	**M**(0.8636)	**N**(0.8622)
E(0.6250)	**E**(0.6250)	**L**(0.7663)	**L**(0.7273)	**F**(0.6690)
M(0.6023)	**M**(0.5881)	**O**(0.6934)	**O**(0.6818)	**L**(0.6690)
Q(0.5611)	**Q**(0.5453)	**F**(0.6754)	**F**(0.6364)	**A**(0.6690)
F(0.5)	**F**(0.5)	**P**(0.5844)	**P**(0.5455)	**I**(0.5426)
L(0.5)	**L**(0.5)	**I**(0.5735)	**I**(0.5)	**C**(0.4915)
C(0.4773)	**C**(0.4590)	**Q**(0.4940)	**Q**(0.4944)	**O**(0.4893)
B(0.4630)	**A**(0.4572)	**A**(0.4875)	**A**(0.4489)	**R**(0.4654)
A(0.4588)	**O**(0.4474)	**R**(0.4091)	**R**(0.4091)	**Q**(0.4617)
O(0.4559)	**B**(0.4389)	**E**(0.4026)	**E**(0.3636)	**P**(0.4190)
R(0.4087)	**R**(0.3678)	**C**(0.3621)	**C**(0.3567)	**E**(0.4190)
P(0.25)	**P**(0.25)	**D**(0.2273)	**D**(0.2273)	**B**(0.3808)
H(0.1250)	**H**(0.125)	**H**(0.1934)	**H**(0.1818)	**D**(0.2117)
D(0.0880)	**D**(0.0639)	**B**(0.1754)	**B**(0.1364)	**H**(0.1690)

In case of the hierarchy of criteria, inconsistencies can be present at different levels of the hierarchy and for this reason, differently from [87] where all constraints translate preference information concerning the same level, the DM could be interested in removing troublesome constraints regarding a particular set of criteria/subcriteria $\{G_{\mathbf{r_1}}, \ldots, G_{\mathbf{r_h}}\}$. For example, considering preference information regarding students evaluated on criteria structured according to the hierarchy shown in Section 3.1.7, the DM could be interested in removing the troublesome constraints at the lowest level possible, i.e. starting by the last but one level, that is constraints regarding Algebra, Analysis, Analytical Chemistry and Organic Chemistry. Then, if it is still not sufficient to get feasibility of the whole set of constraints E^{A^R}, one can look at the constraints of the level immediately above, that is constraints regarding Mathematics and Chemistry, and so on; in this way (s)he could examine the infeasibility going up the hierarchy of criteria. Another DM could be interested in removing troublesome constraints regarding sets of criteria from different levels, like, for example, {Mathematics, Organic Chemistry} or {Analysis, Analytical Chemistry}, or Mathematics alone, or Chemistry alone, and so on. Two important remarks concerning this procedure have to be done:

- looking for troublesome constraints among all constraints E^{A^R} translating the full preference information provided by the DM can be seen as a particular case of the above procedure; in fact, in order to get the whole set of constraints E^{A^R}, it is sufficient to consider the set $\{G_{\mathbf{r_1}}, \ldots, G_{\mathbf{r_h}}\}$ of criteria composed of all criteria from the first level of the hierarchy,

- finding a set of troublesome constraints regarding a particular set of criteria/subcriteria could be not sufficient to remove the infeasibility of the whole set of constraints E^{A^R}; if it would be the case, one should continue the search and add some criterion/subcriterion to the set of criteria $\{G_{\mathbf{r_1}}, \ldots, G_{\mathbf{r_h}}\}$ considered before in order to verify if removing troublesome constraints from the extended set is sufficient to make E^{A^R} feasible.

Knowing a few or all sets of constraints causing infeasibility, if the DM would refuse to choose the one to be removed, then the analyst could suggest a certain heuristic for ordering these sets of constraints with respect to importance of the corresponding piece of preference information. For example, given a set of constraints $S = \{C_1, \ldots, C_p\}$ coming from levels $\{h_1, \ldots, h_p\}$, respectively, one could associate to this set the number $H_S = \left(\sum_{k=1}^{p} h_k\right)/p$. H_S represents an average level of constraints belonging to set S. Supposing that a constraint from level h is more important than the one from level $h+1$, one could decide to remove S_i, such that $H_{S_i} > H_{S_j}$ for all $j \neq i$, that is the set having the greatest value of H_{S_i}. If two sets, S_i and S_j, would have the smallest score $H_{S_i} = H_{S_j}$, then we could remove the one that has less constraints coming from the lowest level. In order to find sets of troublesome constraints in a set of constraints translating preference information, one can proceed as shown in [87].

Credibility

ROR methods permit to specify incrementally the preferences of the DM, assigning them a different degree of credibility. The idea of considering a sequence of pieces of preference information ordered according to their credibility has been introduced in [55] and investigated further in [76]. More formally, the preference information given by the DM is represented as a chain of embedded preference relations $\succsim_1 \subseteq \ldots \subseteq \succsim_n$, where for each $r, s = 1, \ldots, n$, with $r < s$, the preference $\succsim_r$ is more credible than $\succsim_s$. If for any $t = 1, \ldots, n$, we denote by E_t the set of constraints obtained from $\succsim_t$, and by $\mathcal{U}_t$ the sets of value functions compatible with the preference information of $\succsim_t$, then we have $E_1 \subseteq \ldots \subseteq E_n$ and $\mathcal{U}_1 \supseteq \ldots \supseteq \mathcal{U}_n$, and consequently $\succsim_1^N \subseteq \ldots \subseteq \succsim_n^N$, and $\succsim_1^P \supseteq \ldots \supseteq \succsim_n^P$, that is the smaller the credibility of the considered preference relation $\succsim_t$, the richer the necessary preference relation $\succsim_t^N$ and the poorer the possible preference relation $\succsim_t^P$. In case of the hierarchy of criteria, for each subcriterion $G_{\mathbf{r}} \in \mathcal{G}$, we have a sequence of nested possible preference relations $\succsim_{\mathbf{r},1}^P \supseteq \ldots \supseteq \succsim_{\mathbf{r},n}^P$ and a sequence of nested necessary preference relations $\succsim_{\mathbf{r},1}^N \subseteq \ldots \subseteq \succsim_{\mathbf{r},n}^N$.

Extreme ranking

Necessary and possible preference relations give information regarding couples of alternatives. However, it could be interesting to analyse some information related to the whole set of alternatives in terms of the best and the worst ranking position assigned to each alternative by the compatible value functions. This constitutes the extreme ranking analysis introduced in [76]. In case of the hierarchy of criteria, the extreme ranking analysis can be performed for each subcriterion $G_{\mathbf{r}} \in \mathcal{G}$.

UTADIS^{GMS}

In general, MCDA considers three types of problems:

- ranking, consisting in completely or partially ordering the alternatives from the best to the worst,
- choice, consisting in selecting a subset of the best alternatives,
- sorting, consisting in assigning the alternatives to some predefined and preferentially ordered classes.

Ranking and choice problems are based on pairwise comparisons of alternatives and, therefore, they can be dealt with possible and necessary preference relations. Sorting relies instead on the intrinsic value of an alternative and not on its comparison to others. Therefore, sorting problems need specific methods. Within ROR, UTADIS^{GMS} [57] has been proposed to deal with sorting problems as follows. Given a set of pre-defined classes $C_1, C_2, \ldots, C_p$, ordered from the worst to the best, the DM gives preference information in terms of exemplary assignments of reference alternatives to some sequences of classes, such that $a^* \rightarrow [C_{L^{DM}}(a^*), C_{R^{DM}}(a^*)]$, with $L^{DM} \leq R^{DM}$, means that reference alternative a^* can be assigned to one of the classes between $C_{L^{DM}}(a^*)$ and $C_{R^{DM}}(a^*)$. Denoting by $A^R \subseteq A$ the set of reference alternatives considered by the DM, we say that a value function U is compatible if

$$\forall a^*, b^* \in A^R, L^{DM}(a^*) > R^{DM}(b^*) \Rightarrow U(a^*) > U(b^*). \tag{3.3}$$

Denoting by $\mathcal{U}$ the set of compatible value functions, we have that each $U \in \mathcal{U}$ assigns an alternative $a \in A$ to a sequence of classes $\left[L^U(a), R^U(a)\right]$, where

$$L^U(a) = \max\left(\{1\} \cup \left\{L^{DM}(a^*) : U(a^*) \leq U(a), a^* \in A^R\right\}\right),$$

$$R^U(a) = \min\left(\{p\} \cup \left\{R^{DM}(a^*) : U(a^*) \geq U(a), a^* \in A^R\right\}\right).$$

Within ROR, considering the set of all compatible value functions, for each $a \in A$ one can define the possible assignment $C^P(a)$ and the necessary assignment $C^N(a)$ as follows:

- $C^P(a) = \left[L_P^{\mathcal{U}}(a), R_P^{\mathcal{U}}(a)\right] = \cup_{U \in \mathcal{U}} \left[L^U(a), R^U(a)\right]$,

- $C^N(a) = \left[L_N^{\mathcal{U}}(a), R_N^{\mathcal{U}}(a)\right] = \cap_{U\in\mathcal{U}}\left[L^U(a), R^U(a)\right]$.

In case of the hierarchy of criteria, the DM can give exemplary assignments $a^* \rightarrow [C_{L^{DM}}(a^*), C_{R^{DM}}(a^*)]$ at a comprehensive level, but (s)he can also give assignments $a^* \rightarrow_{\mathbf{r}} \left[C^{\mathbf{r}}_{L^{DM}_{\mathbf{r}}}(a^*), C^{\mathbf{r}}_{R^{DM}_{\mathbf{r}}}(a^*)\right]$ with respect to each subcriterion $G_{\mathbf{r}}$ from the hierarchy, excluding the elementary subcriteria, i.e. $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$.

For example, suppose that a Dean has to evaluate students according to their scores in various subjects. He can say that student s_1 is assigned comprehensively to a class between "Medium" and "Very good", i.e. $s_1 \rightarrow$ [Medium,Very good], but he can also say that student s_2 (not necessarily s_2 different from s_1) is assigned to a class between "Weakly bad" and "Weakly good" with respect to Literature, i.e. $s_2 \rightarrow_{Lit}$ [Weakly bad$_{Lit}$, Weakly good$_{Lit}$]. The compatibility condition relative to the assignment with respect to subcriterion $G_{\mathbf{r}}, \mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, is as follows:

$$\forall a^*, b^* \in A^R, L^{DM}_{\mathbf{r}}(a^*) > R^{DM}_{\mathbf{r}}(b^*) \Rightarrow U_{\mathbf{r}}(a^*) > U_{\mathbf{r}}(b^*). \tag{3.4}$$

At the output, for each $a \in A$, besides the comprehensive possible assignments $C^P(a)$ and the necessary assignments $C^N(a)$, the method gives the possible assignment $C^P_{\mathbf{r}}(a)$ and the necessary assignment $C^N_{\mathbf{r}}(a)$ for each $G_{\mathbf{r}}, \mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, as follows:

- $C^P_{\mathbf{r}}(a) = \left[L^{\mathcal{U}}_{\mathbf{r},P}(a), R^{\mathcal{U}}_{\mathbf{r},P}(a)\right] = \cup_{U\in\mathcal{U}}\left[L^U_{\mathbf{r}}(a), R^U_{\mathbf{r}}(a)\right]$,
- $C^N_{\mathbf{r}}(a) = \left[L^{\mathcal{U}}_{\mathbf{r},N}(a), R^{\mathcal{U}}_{\mathbf{r},N}(a)\right] = \cap_{U\in\mathcal{U}}\left[L^U_{\mathbf{r}}(a), R^U_{\mathbf{r}}(a)\right]$,

where

$$L^U_{\mathbf{r}}(a) = \max\left(\{1\} \cup \left\{L^{DM}_{\mathbf{r}}(a^*) : U_{\mathbf{r}}(a^*) \leq U_{\mathbf{r}}(a), a^* \in A^*\right\}\right),$$

$$R^U_{\mathbf{r}}(a) = \min\left(\{p\} \cup \left\{R^{DM}_{\mathbf{r}}(a^*) : U_{\mathbf{r}}(a^*) \geq U_{\mathbf{r}}(a), a^* \in A^*\right\}\right).$$

Group decision

In many decision making situations there is a plurality of DMs. For example, in case of decision related to land development, a group of stakeholders with different perceptions of predefined criteria has to be involved. ROR ([56, 48]) has been applied to group decision as follows. Considering a set $\mathcal{D}$ of DMs, and a set of pairwise comparisons provided by the DM belonging to $\mathcal{D}' \subseteq \mathcal{D}$, for each DM $d_h \in \mathcal{D}'$ we find the necessary and possible preference relations $\succsim^N_h$ and $\succsim^P_h$. Then, we can represent consensus between decision makers from $\mathcal{D}$, defining the following preference relations for all $\mathcal{D}' \subseteq \mathcal{D}$:

- the necessary-necessary preference relation ($\succsim_{\mathcal{D}'}^{N,N}$), for which a is necessarily preferred to b for all $d_h \in \mathcal{D}'$,
- the necessary-possibly preference relation ($\succsim_{\mathcal{D}'}^{N,P}$), for which a is necessarily preferred to b for at least one $d_h \in \mathcal{D}'$,
- the possibly-necessary preference relation ($\succsim_{\mathcal{D}'}^{P,N}$), for which a is possibly preferred to b for all $d_h \in \mathcal{D}'$,
- the possibly-possibly preference relation ($\succsim_{\mathcal{D}'}^{P,P}$), for which a is possibly preferred to b for at least one $d_h \in \mathcal{D}'$.

In case of the hierarchy of criteria we can define the above four relations for each subcriterion $G_{\mathbf{r}}$ from the hierarchy, excluding the elementary subcriteria, i.e. $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$.

Interacting criteria

UTA^{GMS}, UTADIS^{GMS} and GRIP take into account an additive value function. This model is among the most popular ones because it has the advantage of being easily manageable, and, moreover, it has a very sound axiomatic basis (see, e.g., [80, 129]). However, the additive value function is not able to represent *interactions* among criteria. For example, consider evaluation of cars using such criteria as maximum speed, acceleration and price. In this case, there may exist a negative interaction (*negative synergy*) between maximum speed and acceleration because a car with a high maximum speed also has a good acceleration, so, even if each of these two criteria is very important for a DM who likes sport cars, their joint impact on reinforcement of preference of a more speedy and better accelerating car over a less speedy and worse accelerating car will be smaller than a simple addition of the impacts of the two criteria considered separately in validation of this preference relation. In the same decision problem, there may exist a positive interaction (*positive synergy*) between maximum speed and price because a car with a high maximum speed is usually expensive, and thus a car with a high maximum speed and relatively low price is very much appreciated. Thus, the comprehensive impact of these two criteria on the strength of preference of a more speedy and cheaper car over a less speedy and more expensive car is greater than the impact of the two criteria considered separately in validation of this preference relation. To handle the interactions among criteria, one can consider *non-additive integrals*, such as Choquet integral [21] and Sugeno integral [119], or an additive value function augmented by additional components reinforcing the value when there is a positive interaction for some pairs of criteria, or penalizing the value when this interaction is negative, like in UTA^{GMS}-INT [58]. In case of the hierarchy of criteria we can consider interaction among criteria at each level of the hierarchy. For example, evaluating students we can have negative synergy (redundancy) for Mathematics and Physics (because, in general, good students in Mathematics are good also in Physics) and positive synergy for Algebra and Analysis at a lower level (because Algebra and Analysis require different aptitudes, and therefore a student good in Algebra is not always good in Analysis).

3.1.9 Conclusions

In this section, in order to deal with one important issue of Multiple Criteria Decision Aiding (MCDA), that is the hierarchy of criteria, we proposed a new methodology called Multiple Criteria Hierarchy Process (MCHP). The basic idea of MCHP relies on consideration of preference relations regarding subcriteria at each level of the hierarchy of criteria, obtaining in this way a better insight into the problem at hand. MCHP can be applied to any MCDA method. In this section, we considered the case where evaluations of alternatives are aggregated by a value function, and we applied MCHP to one particular MCDA methodology that is the Robust Ordinal Regression (ROR). In this case, the preference model is the entire set of general additive value functions compatible with preference information given by the Decision Maker (DM) in terms of pairwise comparisons of some alternatives, and in terms of intensity of preference with respect to some pairs of alternatives. The advantage is twofold:

- from the point of view of preference information, the hierarchy of criteria is enriching the possibility of the DM to express his/her preferences: in fact, the DM can give preference information at a comprehensive level, e.g., student s_1 is comprehensively preferred to student s_2, as well as at an intermediate level with respect to subcriteria, e.g., student s_1 is preferred to student s_2 on a subset of criteria related to Mathematics;
- with respect to decision support, taking into account the hierarchy of criteria permits to define possible and necessary preference relations not only at a comprehensive level but also at each intermediate level of the hierarchy: in fact, as a final result, we can have not only that student s_1 is comprehensively necessarily preferred to student s_2, and student s_3 is comprehensively possibly preferred to student s_4, but also that, e.g., student s_1 is necessarily preferred to student s_2 on a subset of criteria related to Mathematics, and s_3 is possibly preferred to student s_4 on criteria related to Organic Chemistry.

Adapting ROR to the hierarchy of criteria, i.e. putting together MCHP and ROR, gives a very powerful methodology of multiple criteria decision aiding: in fact, in this way we conjugate, on one hand, the robustness concerns by taking into account the set of all value functions compatible with preference information supplied by DM, and, on the other hand, the benefits of the hierarchical decomposition of a complex multiple criteria decision problem. We have shown, moreover, that all the methodological developments proposed within the ROR can be used in the case of the hierarchy of criteria: calculation of a representative value function, consideration of different credibilities of preference information, extreme ranking analysis, application to sorting problems, group decision, handling interaction among criteria. Let us observe that we can consider preference relations referring to a subset of criteria also if there is no explicit hierarchy in the set of criteria. In fact, for any subset of criteria $\mathcal{J}$, the DM can always express preferences of the type "a is preferred to b with respect to $\mathcal{J}$", as well as we can define necessary and possible preference relations with respect to $\mathcal{J}$.

We envisage three further methodological developments of the ROR adapted to the case of the hierarchy of criteria:

- consideration of imprecise evaluations on specific criteria;
- consideration of the outranking preference models;
- consideration of a structure of criteria more complex than the hierarchy defined in this section: for example, while in this section we assume that each subcriterion descends from only one criterion located at the upper level of the hierarchy tree, we can have a real situation where one subcriterion descends from more than one criterion of the upper level; for example, in case of evaluation of students at a scientific faculty, Analytic Mechanics can descend from both Mathematics and Physics; we also plan to deal with more complex criteria structures, like those considered in Analytical Network Process (ANP) [111].

3.1.10 Appendix

Proof of Proposition 3.1.2

1. For all $a, b \in A$

$$a \succsim_{\mathbf{r}}^{N} b \Leftrightarrow \forall U \in \mathcal{U}, U_{\mathbf{r}}(a) \geq U_{\mathbf{r}}(b) \Rightarrow \exists U \in \mathcal{U} : U_{\mathbf{r}}(a) \geq U_{\mathbf{r}}(b) \Leftrightarrow a \succsim_{\mathbf{r}}^{P} b,$$

   thus we proved that $\succsim_{\mathbf{r}}^{N} \subseteq \succsim_{\mathbf{r}}^{P}$.

2. We have

$$\forall a \in A, \forall U \in \mathcal{U}, U_{\mathbf{r}}(a) \geq U_{\mathbf{r}}(a) \Leftrightarrow \forall a \in A, \;\; a \succsim_{\mathbf{r}}^{N} a,$$

and therefore $\succsim_{\mathbf{r}}^N$ is reflexive.

For all $a, b, c \in A$,

$$a \succsim_{\mathbf{r}}^N b, b \succsim_{\mathbf{r}}^N c \Leftrightarrow \forall U \in \mathcal{U}, U_{\mathbf{r}}(a) \geq U_{\mathbf{r}}(b) \geq U_{\mathbf{r}}(c) \Rightarrow \forall U \in \mathcal{U}, U_{\mathbf{r}}(a) \geq U_{\mathbf{r}}(c) \Rightarrow a \succsim_{\mathbf{r}}^N c$$

and therefore $\succsim_{\mathbf{r}}^N$ is transitive. Being reflexive and transitive $\succsim_{\mathbf{r}}^N$ is a partial preorder.

3. For all $a, b \in A$,

$$a \not\succsim_{\mathbf{r}}^P b \Leftrightarrow \forall U \in \mathcal{U}, U_{\mathbf{r}}(a) < U_{\mathbf{r}}(b) \Rightarrow \exists U \in \mathcal{U} : U_{\mathbf{r}}(b) \geq U_{\mathbf{r}}(a) \Leftrightarrow b \succsim_{\mathbf{r}}^P a$$

   and therefore $\succsim_{\mathbf{r}}^P$ is strongly complete.
   For all $a, b, c \in A$,

$$a \not\succsim_{\mathbf{r}}^P b \text{ and } b \not\succsim_{\mathbf{r}}^P c \Leftrightarrow \forall U \in \mathcal{U}, U_{\mathbf{r}}(a) < U_{\mathbf{r}}(b) < U_{\mathbf{r}}(c) \Rightarrow \forall U \in \mathcal{U}, U_{\mathbf{r}}(a) < U_{\mathbf{r}}(c) \Rightarrow a \not\succsim_{\mathbf{r}}^P c$$

   and thus we proved that $\succsim_{\mathbf{r}}^P$ is negatively transitive.

4. For all $a, b \in A$,

$$a \not\succsim_{\mathbf{r}}^N b \Leftrightarrow \exists U \in \mathcal{U} : U_{\mathbf{r}}(a) < U_{\mathbf{r}}(b) \Rightarrow \exists U \in \mathcal{U} : U_{\mathbf{r}}(a) \leq U_{\mathbf{r}}(b) \Rightarrow b \succsim_{\mathbf{r}}^P a$$

   and therefore we proved that $a \succsim_{\mathbf{r}}^N b$ or $b \succsim_{\mathbf{r}}^P a$.

5. For all $a, b, c \in A$, $a \succsim_{\mathbf{r}}^N b$ implies that $U_{\mathbf{r}}(a) \geq U_{\mathbf{r}}(b)$ for all compatible value functions; $b \succsim_{\mathbf{r}}^P c$ implies that there exist at least one compatible value function $\overline{U}$ such that $\overline{U}_{\mathbf{r}}(b) \geq \overline{U}_{\mathbf{r}}(c)$; then for this compatible value function we have $\overline{U}_{\mathbf{r}}(a) \geq \overline{U}_{\mathbf{r}}(b) \geq \overline{U}_{\mathbf{r}}(c)$, and thus $a \succsim_{\mathbf{r}}^P c$.

6. $a \succsim_{\mathbf{r}}^P b$ implies that there exist at least one compatible value function $\overline{U}$ such that $\overline{U}_{\mathbf{r}}(a) \geq \overline{U}_{\mathbf{r}}(b)$; $b \succsim_{\mathbf{r}}^N c$ implies that $U_{\mathbf{r}}(b) \geq U_{\mathbf{r}}(c), \forall U \in \mathcal{U}$; in this way for the value function $\overline{U}$ we have $\overline{U}_{\mathbf{r}}(a) \geq \overline{U}_{\mathbf{r}}(b) \geq \overline{U}_{\mathbf{r}}(c)$, and thus $a \succsim_{\mathbf{r}}^P c$;

Proof of Proposition 3.1.3

1. Remembering that $U_{\mathbf{r}}(x) = U_{(\mathbf{r},1)}(x) + \ldots + U_{(\mathbf{r},n(\mathbf{r}))}(x)$, we have

$$a \succsim_{(\mathbf{r},j)}^N b \;\; \forall j = 1, \ldots, n(\mathbf{r}) \Leftrightarrow U_{(\mathbf{r},j)}(a) \geq U_{(\mathbf{r},j)}(b) \;\; \forall U \in \mathcal{U}, \forall j = 1, \ldots, n(\mathbf{r}) \Rightarrow$$

$$\Rightarrow \forall U \in \mathcal{U},\ \sum_{j=1}^{n(\mathbf{r})} U_{(\mathbf{r},j)}(a) \geq \sum_{j=1}^{n(\mathbf{r})} U_{(\mathbf{r},j)}(b) \Leftrightarrow \forall U \in \mathcal{U},\ \ U_{\mathbf{r}}(a) \geq U_{\mathbf{r}}(b) \Leftrightarrow a \succsim_{\mathbf{r}}^{N} b.$$

2. $a \succsim_{(\mathbf{r},w)}^{P} b$ implies that there exists $\overline{U} \in \mathcal{U}$ such that $\overline{U}_{(\mathbf{r},w)}(a) \geq \overline{U}_{(\mathbf{r},w)}(b)$; considering that $a \succsim_{(\mathbf{r},j)}^{N} b$ for all $j \in \{1,\ldots,n(\mathbf{r})\} \setminus \{w\}$, we have $U_{(\mathbf{r},j)}(a) \geq U_{(\mathbf{r},j)}(b)\ \ \forall j \in \{1,\ldots,n(\mathbf{r})\} \setminus \{w\}$, and therefore also for $\overline{U} \in \mathcal{U}$ we have $\overline{U}_{(\mathbf{r},j)}(a) \geq \overline{U}_{(\mathbf{r},j)}(b)\ \ \forall j \in \{1,\ldots,n(\mathbf{r})\} \setminus \{w\}$ and thus

$$\overline{U}_{\mathbf{r}}(a) = \sum_{j=1}^{n(\mathbf{r})} \overline{U}_{(\mathbf{r},j)}(a) \geq \sum_{j=1}^{n(\mathbf{r})} \overline{U}_{(\mathbf{r},j)}(b) = \overline{U}_{\mathbf{r}}(b),$$

from which $a \succsim_{\mathbf{r}}^{P} b$.

3. Let us suppose, for contradiction, that $a \succsim_{\mathbf{r}}^{P} b$; this means that there exists a value function $\overline{U} \in \mathcal{U}$ such that $\overline{U}_{\mathbf{r}}(a) \geq \overline{U}_{\mathbf{r}}(b)$; from this we obtain that

$$\overline{U}_{\mathbf{r}}(a) \geq \overline{U}_{\mathbf{r}}(b) \Leftrightarrow \sum_{j=1}^{n(\mathbf{r})} \overline{U}_{(\mathbf{r},j)}(a) \geq \sum_{j=1}^{n(\mathbf{r})} \overline{U}_{(\mathbf{r},j)}(b) \Leftrightarrow \sum_{j=1}^{n(\mathbf{r})} \left[\overline{U}_{(\mathbf{r},j)}(a) - \overline{U}_{(\mathbf{r},j)}(b)\right] \geq 0$$

and from this, for at least one $j \in \{1,\ldots,n(\mathbf{r})\}$ we have $\overline{U}_{(\mathbf{r},j)}(a) - \overline{U}_{(\mathbf{r},j)(b)} \geq 0 \Rightarrow \overline{U}_{(\mathbf{r},j)}(a) \geq \overline{U}_{(\mathbf{r},j)}(b)$ and thus $a \succsim_{(\mathbf{r},j)}^{P} b$ which contradicts the hypothesis.

Proof of Proposition 3.1.4

Let us remember that $\forall a \in A$ we have $U_{\mathbf{r}}(a) = U_{(\mathbf{r},1)}(a) + \ldots + U_{(\mathbf{r},n(\mathbf{r}))}(a)$.

1. For any $a, b, c, d \in A$

$$(a,b) \succsim_{(\mathbf{r},j)}^{*N} (c,d)\ \ \forall j = 1,\ldots,n(\mathbf{r}) \Leftrightarrow$$

$$\Leftrightarrow U_{(\mathbf{r},j)}(a) - U_{(\mathbf{r},j)}(b) \geq U_{(\mathbf{r},j)}(c) - U_{(\mathbf{r},j)}(d), \forall U \in \mathcal{U}, \forall j = 1,\ldots,n(\mathbf{r}) \Rightarrow$$

$$\Rightarrow \forall U \in \mathcal{U},\ \sum_{j=1}^{n(\mathbf{r})} \left[U_{(\mathbf{r},j)}(a) - U_{(\mathbf{r},j)}(b)\right] \geq \sum_{j=1}^{n(\mathbf{r})} \left[U_{(\mathbf{r},j)}(c) - U_{(\mathbf{r},j)}(d)\right] \Leftrightarrow$$

$$\Leftrightarrow \forall U \in \mathcal{U}, \sum_{j=1}^{n(\mathbf{r})} U_{(\mathbf{r},j)}(a) - \sum_{j=1}^{n(\mathbf{r})} U_{(\mathbf{r},j)}(b) \geq \sum_{j=1}^{n(\mathbf{r})} U_{(\mathbf{r},j)}(c) - \sum_{j=1}^{n(\mathbf{r})} U_{(\mathbf{r},j)}(d) \Leftrightarrow$$

$$\Leftrightarrow \forall U \in \mathcal{U}, U_{\mathbf{r}}(a) - U_{\mathbf{r}}(b) \geq U_{\mathbf{r}}(c) - U_{\mathbf{r}}(d) \Leftrightarrow (a,b) \succsim_{\mathbf{r}}^{*N} (c,d).$$

2. For any $a, b, c, d \in A$, $(a,b) \succsim_{(\mathbf{r},w)}^{*P} (c,d)$ implies that there exists $\overline{U} \in \mathcal{U}$ such that $\overline{U}_{(\mathbf{r},w)}(a) - \overline{U}_{(\mathbf{r},w)}(b) \geq \overline{U}_{(\mathbf{r},w)}(c) - \overline{U}_{(\mathbf{r},w)}(d)$; considering that $(a,b) \succsim_{(\mathbf{r},j)}^{*N} (c,d)$ for all $j \in \{1,\ldots,n(\mathbf{r})\} \setminus \{w\}$

and for all compatible value functions, we have $\overline{U}_{(\mathbf{r},j)}(a) - \overline{U}_{(\mathbf{r},j)}(b) \geq \overline{U}_{(\mathbf{r},j)}(c) - \overline{U}_{(\mathbf{r},j)}(d) \;\; \forall j \in \{1,\ldots,n(\mathbf{r})\} \setminus \{w\}$, and thus

$$\sum_{j=1}^{n(\mathbf{r})} \left[\overline{U}_{(\mathbf{r},j)}(a) - \overline{U}_{(\mathbf{r},j)}(b)\right] \geq \sum_{j=1}^{n(\mathbf{r})} \left[\overline{U}_{(\mathbf{r},j)}(c) - \overline{U}_{(\mathbf{r},j)}(d)\right] \Leftrightarrow$$

$$\Leftrightarrow \sum_{j=1}^{n(\mathbf{r})} \overline{U}_{(\mathbf{r},j)}(a) - \sum_{j=1}^{n(\mathbf{r})} \overline{U}_{(\mathbf{r},j)}(b) \geq \sum_{j=1}^{n(\mathbf{r})} \overline{U}_{(\mathbf{r},j)}(c) - \sum_{j=1}^{n(\mathbf{r})} \overline{U}_{(\mathbf{r},j)}(d) \Leftrightarrow$$

$$\Leftrightarrow \overline{U}_{\mathbf{r}}(a) - \overline{U}_{\mathbf{r}}(b) \geq \overline{U}_{\mathbf{r}}(c) - \overline{U}_{\mathbf{r}}(d) \Leftrightarrow (a,b) \succsim_{\mathbf{r}}^{*^P} (c,d).$$

from which $(a,b) \succsim_{\mathbf{r}}^{*^P} (c,d)$.

3. Let us suppose, for contradiction, that for $a,b,c,d \in A$ $(a,b) \succsim_{\mathbf{r}}^{*^P} (c,d)$; this means that there exists a value function $\overline{U} \in \mathcal{U}$ such that $\overline{U}_{\mathbf{r}}(a) - \overline{U}_{\mathbf{r}}(b) \geq \overline{U}_{\mathbf{r}}(c) - \overline{U}_{\mathbf{r}}(d)$; from this we obtain that

$$\overline{U}_{\mathbf{r}}(a) - \overline{U}_{\mathbf{r}}(b) \geq \overline{U}_{\mathbf{r}}(c) - \overline{U}_{\mathbf{r}}(d) \Leftrightarrow \sum_{j=1}^{n(\mathbf{r})} \overline{U}_{(\mathbf{r},j)}(a) - \sum_{j=1}^{n(\mathbf{r})} \overline{U}_{(\mathbf{r},j)}(b) \geq \sum_{j=1}^{n(\mathbf{r})} \overline{U}_{(\mathbf{r},j)}(c) - \sum_{j=1}^{n(\mathbf{r})} \overline{U}_{(\mathbf{r},j)}(d) \Leftrightarrow$$

$$\Leftrightarrow \sum_{j=1}^{n(\mathbf{r})} \overline{U}_{(\mathbf{r},j)}(a) - \sum_{j=1}^{n(\mathbf{r})} \overline{U}_{(\mathbf{r},j)}(b) - \sum_{j=1}^{n(\mathbf{r})} \overline{U}_{(\mathbf{r},j)}(c) + \sum_{j=1}^{n(\mathbf{r})} \overline{U}_{(\mathbf{r},j)}(d) \geq 0 \Leftrightarrow$$

$$\Leftrightarrow \sum_{j=1}^{n(\mathbf{r})} \left[\overline{U}_{(\mathbf{r},j)}(a) - \overline{U}_{(\mathbf{r},j)}(b) + \overline{U}_{(\mathbf{r},j)}(c) - \overline{U}_{(\mathbf{r},j)}(d)\right] \geq 0,$$

and from this follows that, for at least one $j \in \{1,\ldots,n(\mathbf{r})\}$ we have $\overline{U}_{(\mathbf{r},j)}(a) - \overline{U}_{(\mathbf{r},j)}(b) \geq \overline{U}_{(\mathbf{r},j)}(c) - \overline{U}_{(\mathbf{r},j)}(d)$, and thus $(a,b) \succsim_{(\mathbf{r},j)}^{*^P} (c,d)$ for at least one j, which contradicts the hypothesis.

3.2 Multiple Criteria Hierarchy Process with ELECTRE and PROMETHEE

3.2.1 Introduction

Multiple Criteria Decision Aiding (MCDA) copes with three main types of decision problems: ranking, sorting and choice. Ranking problems consist in rank ordering of all alternatives from the worst to the best, looking at their evaluations on the considered criteria; sorting problems consist in assigning each alternative to a predefined and preference ordered class; choice problems consist in selecting a subset of alternatives considered as the best (for a more detailed survey, see [29, 27]). In order to handle these problems, one can use either of the two different methodologies:

- assign to each alternative a utility value, i.e. a real number reflecting the degree of desirability of a considered alternative, independently from the evaluations of other alternatives,
- compare alternatives pairwise, in order to discover if one is preferred to the other, or if they are indifferent or incomparable.

In the first case, to associate a utility value to an alternative, taking into account its evaluations on the considered criteria, multi attribute utility theory (MAUT) [80] frequently uses an additive value function defined as a sum of as many marginal value functions as there are criteria. In the second case, outranking methods [16, 103] construct a binary relation which reads: "alternative a is at least as good as alternative b", which means "a outranks b". This construction takes into account evaluation of both compared alternatives on the considered criteria, as well as some comparison thresholds and weights expressing the relative importance of the criteria.

Generally, the information provided by the dominance relation on the set of alternatives is poor, and makes many alternatives incomparable. To enrich this relation, the Decision Maker (DM) is asked to provide some preference information, so that the outranking relation giving account of it makes alternatives more comparable. As this comparability is consistent with the value system of the DM, the outranking relation can be considered as the DM's preference model.

Preference information can be direct or indirect; direct means that the DM can give information regarding values of parameters of the considered preference model, while indirect means that the DM gives information regarding some alternatives (s)he knows well, and from this information there are inferred values of parameters of the considered preference model. Generally, the indirect methodology is more realistic (see, e.g.,[72],[95],[117]), because the DM does not always understand well enough the meaning of all these parameters. Using the indirect methodology, there usually exist more than one set of parameters compatible with the preference information provided by the DM, and each of these sets of parameters could give different results to the decision problem at hand. For this reason, any choice of one specific set of parameters compatible with preference information provided by the DM could be considered as arbitrary and meaningless. In order to deal with this inconvenience, Robust Ordinal Regression (ROR) takes into account not only one set of parameters compatible with the preference information provided by the DM, but considers all these sets simultaneously defining two preference relations:

- the necessary preference relation, for which "alternative a is necessarily preferred to alternative b" if a is at least as good as b for all compatible sets of parameters,

- the possible preference relation, for which "alternative a is possibly preferred to alternative b" if a is at least as good as b for at least one compatible set of parameters.

ROR methods have been proposed for ranking and choice problems [33, 55], sorting problems [57], outranking models [47] and non additive models [7].

Remark that not all multiple criteria decision problems present evaluation criteria at the same level, but there can exist a hierarchical structure of criteria. This is the case, for example, of environmental planning in which it is possible to take into account economic, social and environmental criteria, and each of these criteria can be composed of subcriteria on which the alternatives are evaluated. In [23], we have considered the hierarchy of criteria in the context of ROR, showing the following advantages of using this procedure:

- the DM can express preference information not only in a comprehensive way but also in a partial way, that is considering preference information with respect to a subcriterion at an intermediate level of the hierarchy,

- the DM can obtain results not only with respect to the comprehensive view, but also results at intermediate levels of the hierarchy; for example, the DM can learn if a is necessarily or possibly preferred to b with respect to a subcriterion G at an intermediate level of the hierarchy.

Let us remark that the use of the hierarchy of criteria proposed by our approach is rather different from other MCDA methodologies assuming a hierarchical structure of the family of criteria. In fact, while in general the hierarchy of criteria is used to decompose and make easier the preference elicitation concerning pairwise comparisons of criteria with respect to relative importance, in our approach, a preference relation in each node of the hierarchy constitutes a base for the discussion with the DM.

Indeed, the preference relations in particular nodes of the hierarchy are presented to the DM as consequences (output) of her/his preference information provided at the input. In course of an interactive process, the DM can add, modify or remove some items of the preference information if (s)he feels that the preference relations do not reflect correctly her/his value system. This interactive process ends when the DM get convinced by the preference relations obtained in consequence of her/his preference information, and thus accepts the recommendations provided by the MCDA methodology.

Observe that consideration of preference relations at each level of the hierarchy constitutes a specific feature of our methodology, which we consider very useful in any decision process in which a hierarchy of criteria is considered. Considering preference relations in particular nodes of the hierarchy permits decomposition of arguments explaining the overall preferences. For example, in case of evaluations of students, one could say that student a is comprehensively preferred to student b, because even if a is slightly worse than b with respect to subjects related to Literature, he is much better with respect to subjects related to Mathematics and Physics. Moreover, going in depth of the hierarchy, one could add that the preference with respect to subjects related to Mathematics is based on better evaluations of student a on subjects related to Analysis rather than on subjects related to Algebra.

It is worth noting that this specific use of the hierarchy of criteria can be applied to any MCDA methodology. In this work, we are applying it to Robust Ordinal Regression approach, but it can be applied to any other MCDA methodology, even those which use the hierarchy to ask the DM for pairwise comparisons of subcriteria with respect to their importance.

In this section, we propose a generalization of outranking methods, more specifically ELECTRE and PROMETHEE methods, to the case of the hierarchy of criteria. No similar attempt is known in the literature. We extend the methodologies of ELECTRE and PROMETHEE to the case where the considered criteria are not at the same level, but they are structured into several levels. In this way, the DM can obtain information not only regarding the comprehensive outranking of an alternative a over an alternative b, but also partially, that is, considering a particular criterion/subcriterion of the hierarchy. For example, in the environmental planning problem, it will be possible to investigate if a certain location p_1 outranks another location p_2 with respect to economic criteria, or environmental criteria, or social aspects, or with respect to all criteria simultaneously. In this particular context, it is worth stressing that ELECTRE and PROMETHEE methods can be considered as particular cases of our methodology, and for this reason, it can be considered as a real generalization of these methods.

,

In the perspective of considering a constructive interaction between the DM and the analyst, we intend to use the ROR methodology to deal with outranking methods in case of the hierarchy of criteria. The application of ROR to ELECTRE and PROMETHEE methods has already been done in [47] and [76], respectively, but also in this case, our methodology can be considered as their generalization because, of course, the absence of hierarchy corresponds to the case of a hierarchy with only one level containing all the criteria.

The section is organized in the following way: in section 3.2.2, we recall the principal concept of the hierarchy of criteria and describe the ELECTRE method generalized to this case; in section 3.2.3, we extend the concept of ROR applied to ELECTRE (which constitutes ELECTRE^{GKMS} method) in case of the hierarchy of criteria, and we propose a didactic example illustrating the use of ELECTRE and ELECTRE^{GKMS} methods applied to a hierarchical structure of criteria; in section 3.2.4, we extend the PROMETHEE method to the case of the hierarchy of criteria; in section 3.2.5, we describe the application of ROR to the PROMETHEE method in case of the hierarchy of criteria, and we provide an example illustrating the PROMETHEE and PROMETHEE^{GKS} methods applied to a hierarchy of criteria; section 3.2.6 collects conclusions.

3.2.2 Hierarchical ELECTRE method

In this section, we recall the basic concepts of the hierarchy of criteria introduced in [23], and we introduce the Hierarchical ELECTRE method.
We suppose that evaluation criteria are not at the same level but they are structured into several levels (see Figure 3.8);

- $A = \{a, b, c, \ldots\}$ is the finite set of alternatives,
- l is the number of levels in the hierarchy of criteria,
- $\mathcal{G}$ is the set of all criteria at all considered levels,
- $\mathcal{I}_{\mathcal{G}}$ is the set of indices of particular criteria representing position of the criteria in the hierarchy,
- m is the number of the first level (root) criteria, $G_{\mathbf{1}}, \ldots, G_{\mathbf{m}}$,
- $G_{\mathbf{r}} \in \mathcal{G}$, with $\mathbf{r} = (i_1, \ldots, i_h) \in \mathcal{I}_{\mathcal{G}}$, denotes a subcriterion of the first level criterion G_{i_1} at level h,
- $n(\mathbf{r})$ is the number of subcriteria of $G_{\mathbf{r}}$ in the subsequent level, i.e. the direct subcriteria of $G_{\mathbf{r}}$ are $G_{(\mathbf{r},1)}, \ldots, G_{(\mathbf{r},n(\mathbf{r}))}$,

- $g_{\mathbf{t}} : A \to \mathbb{R}$, with $\mathbf{t} = (i_1, \ldots, i_l) \in \mathcal{I}_{\mathcal{G}}$, denotes an elementary subcriterion of the first level criterion G_{i_1}, i.e a subcriterion at level l,

- EL is the set of indices of all elementary subcriteria:

$$EL = \{\mathbf{t} = (i_1, \ldots, i_l) \in \mathcal{I}_{\mathcal{G}}\}, \quad \text{where} \begin{cases} i_1 = 1, \ldots, m \\ i_2 = 1, \ldots, n(i_1) \\ \ldots\ldots \\ i_l = 1, \ldots, n(i_1, \ldots, i_{l-1}) \end{cases}$$

- $E(G_{\mathbf{r}})$ is the set of indices of elementary subcriteria descending from $G_{\mathbf{r}}$, i.e.

$$E(G_{\mathbf{r}}) = \{(\mathbf{r}, i_{h+1}, \ldots, i_l) \in \mathcal{I}_{\mathcal{G}}\}, \quad \text{where} \begin{cases} i_{h+1} = 1, \ldots, n(\mathbf{r}) \\ \ldots\ldots \\ i_l = 1, \ldots, n(\mathbf{r}, i_{h+1}, \ldots, i_{l-1}) \end{cases}$$

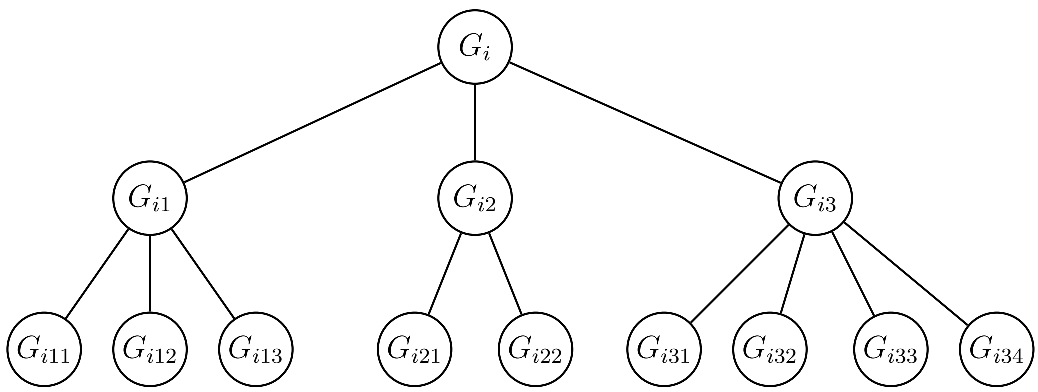
  thus $E(G_{\mathbf{r}}) \subseteq EL$ and, more precisely, $E(G_r) = EL$ if all elementary subcriteria descend from criterion $G_{\mathbf{r}}$,

- LBO is the set of indices of all subcriteria of the last but one level,

- $LB(G_{\mathbf{r}})$ is the set of indices of subcriteria of the last but one level descending from criterion/subcriterion $G_{\mathbf{r}}$,

- when $\mathbf{r} = 0$, then by $G_{\mathbf{r}} = G_{\mathbf{0}}$, we mean the entire set of criteria and not a particular criterion or subcriterion; in this particular case, we have $E(G_{\mathbf{0}}) = EL$ and $LB(G_{\mathbf{0}}) = LBO$.

Figure 3.8: Example of the hierarchy of criteria starting from the first level (root) criterion G_i



Remark that, without loss of generality, we consider a hierarchical structure where each criterion belongs to only one criterion of the level immediately above, that is a criterion $G_{\mathbf{r}}$ from the i-th level of the hierarchy is a subcriterion of only one of the criteria of the $(i-1)$-th level (we call a structure of this type a partitioned structure). Example of the hierarchy of criteria with a partitioned structure is presented in Figure 3.8. In order to understand the reason of this restriction, let us examine an example of the hierarchy of criteria with a non-partitioned structure shown in Figure 3.9. In this particular structure, criterion G_{ixx} and elementary subcriterion g_{ixxx} are subcriteria of more than one criterion of the level immediately above. In particular, criterion G_{ixx} is a subcriterion of criteria G_{i1} and G_{i2} while elementary subcriterion g_{ixxx} is a subcriterion of criteria G_{ixx} and G_{i22}. This means that both criteria influence the criteria they descend from, in a different way. That is, the evaluations of alternatives with respect to elementary subcriterion g_{ixxx} will be weighted in one way if g_{ixxx} is considered to be a subcriterion of criterion G_{ixx}, and they could be weighted in another way if g_{ixxx} is considered to be a subcriterion of criterion G_{i22}. In this way, we can distinguish the contribution of g_{ixxx} to G_{ixx}, from the contribution of g_{ixxx} to G_{i22}. In order to take into account these different types of contribution, we propose to split g_{ixxx} in two "indicators": g_{ixx2}, representing the contribution of elementary subcriterion g_{ixxx} to G_{ixx}, and g_{i221}, representing the contribution of elementary subcriterion g_{ixxx} to criterion G_{i22}. All alternatives will keep the same evaluations with respect to indicators g_{ixx2} and g_{i221}, as they had with respect to criterion g_{ixxx}, but their weights k_{ixx2} and k_{i221} could be different and, moreover, they have to satisfy the relation: $k_{ixx2} + k_{i221} = k_{ixxx}$; this means that the sum of weights of new indicators g_{i221} and g_{ixxx} has to be equal to the weight of criterion g_{ixxx}. Doing in this way, we obtain the hierarchical structure shown in Figure 3.10. At this point, we observe that G_{ixx} is a subcriterion in common of G_{i1} and G_{i2}, thus, we may proceed analogically to g_{ixxx}. So, we have to distinguish between the contribution of G_{ixx} to G_{i1} and the contribution of G_{ixx} to G_{i2}. But, in this case, subcriterion G_{ixx} influences the two above criteria via all its subcriteria (if any) and elementary subcriteria; for this reason we have to split all subcriteria and elementary subcriteria descending from it (here: g_{ixx1} and g_{ixx2}) in order to take into account the different contribution they give to the above criteria. In this way we obtain the partitioned hierarchical structure shown in Figure 3.11, where:

- subcriterion G_{ixx} is split into indicators G_{i12} and G_{i21},
- indicator g_{ixx1} is split into indicators g_{i121} and g_{i211}, and thus, $k_{ixx1} = k_{i121} + k_{i211}$,
- indicator g_{ixx2} is split into indicators g_{i122} and g_{i212}, and thus, $k_{ixx2} = k_{i122} + k_{i212}$.

Let us remark that the above splitting of criteria, which aims to take into account their contribution to different criteria at an upper level, can be applied independently of the type of the preference

Figure 3.9: Example of the hierarchy of criteria with a non-partitioned structure

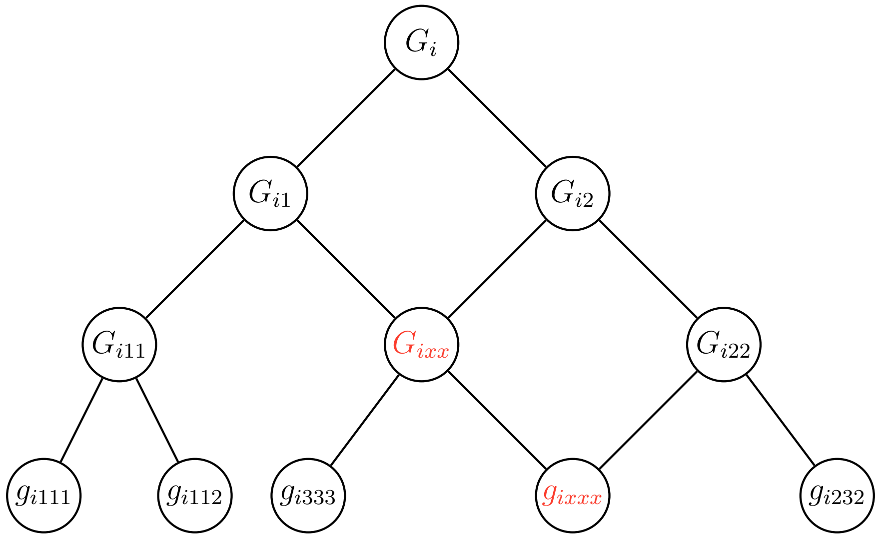
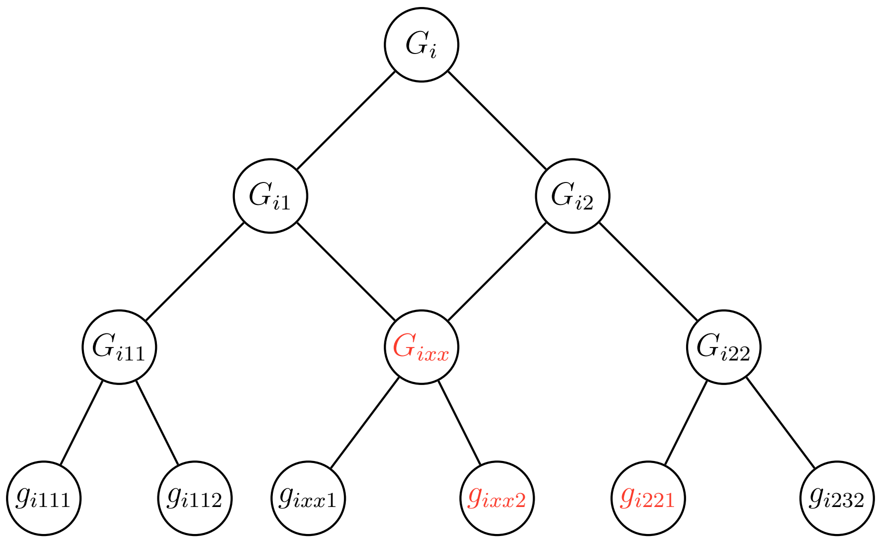


Figure 3.10: Transformation of non-partitioned structure (step 1)

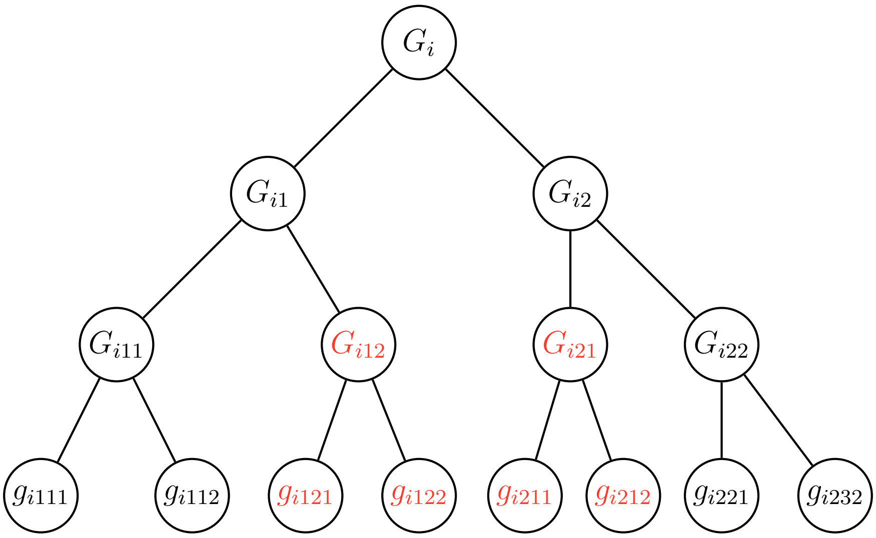


model used in the hierarchical MCDA method. Thus, it could also be used in the hierarchical method involving multiattribute utility functions presented in [23].

Handling the hierarchy of criteria in ELECTRE methods

In this sub-section, we introduce a generalization of ELECTRE methods to the case of the hierarchy of criteria. We start with the hierarchical generalization of the ELECTRE IS method, and then we

Figure 3.11: Partitioned structure resulting from the transformation of the non-partitioned one (step 2)



extend this generalization on the ELECTRE III method (see [104, 106] for description of different ELECTRE methods).

Given a criterion/subcriterion $G_{\mathbf{r}}$ with $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, an outranking relation is a binary relation $S_{\mathbf{r}} \subseteq A \times A$ (in the following $A \times A = B$), such that $aS_{\mathbf{r}}b$ means "a is at least as good as b with respect to criterion $G_{\mathbf{r}}$". Knowing if $S_{\mathbf{r}}$ is true or not for an ordered pair of alternatives $(a, b) \in B$, one is able to represent situations of *weak* ($Q_{\mathbf{r}}$) or *strict* ($P_{\mathbf{r}}$) *preference* (the two relations together called *large preference*), *indifference* ($\sim_{\mathbf{r}}$), and *incomparability* ($R_{\mathbf{r}}$) among a and b:

$$\left.\begin{array}{l} aS_{\mathbf{r}}b \text{ and } not(bS_{\mathbf{r}}a) \Leftrightarrow aQ_{\mathbf{r}}b \text{ or } aP_{\mathbf{r}}b, \\[2ex] aS_{\mathbf{r}}b \text{ and } bS_{\mathbf{r}}a \Leftrightarrow a \sim_{\mathbf{r}} b, \\[2ex] not(aS_{\mathbf{r}}b) \text{ and } not(bS_{\mathbf{r}}a) \Leftrightarrow aR_{\mathbf{r}}b. \end{array}\right\}$$

Let us denote by $k_{\mathbf{t}}$ the *weight* assigned to elementary subcriterion $g_{\mathbf{t}}, \mathbf{t} \in EL$. It is a non-negative real number representing the relative importance (strength) of elementary subcriterion $g_{\mathbf{t}}$ within the family of elementary subcriteria. The *indifference, preference*, and *veto thresholds* on elementary subcriterion $g_{\mathbf{t}}$ are denoted by $q_{\mathbf{t}}$, $p_{\mathbf{t}}$ and $v_{\mathbf{t}}$, respectively. $q_{\mathbf{t}}$ is the greatest difference between the evaluations of two alternatives, compatible with the indifference among them with respect to elementary subcriterion $g_{\mathbf{t}}$; $p_{\mathbf{t}}$ is the smallest difference between the evaluations of two alternatives, compatible with the preference of an alternative over another with respect to elementary subcriterion

$g_{\mathbf{t}}$; $v_{\mathbf{t}}$ is an upper bound beyond which the discordance about the assertion "a is at least as good as b" cannot surpass. For consistency, $v_{\mathbf{t}} > p_{\mathbf{t}} \geq q_{\mathbf{t}} \geq 0$, for all $\mathbf{t} \in EL$. The thresholds on particular elementary subcriterion can be either constant for all alternatives, or dependent on evaluation of a, $g_{\mathbf{t}}(a)$. In the sequel, we assume, for simplicity, constant thresholds, although this is not a necessary assumption for our methodology. Moreover, we assume without loss of generality that all criteria are increasing monotone with respect to the preference, i.e. the greater the evaluation, the better it is.

Construction of an outranking relation involves two concepts known as *concordance* and *non-discordance* tests. The concordance test involves calculation of concordance index $C_{\mathbf{r}}(a, b)$. It represents the strength of the coalition of elementary subcriteria $g_{\mathbf{t}}, \mathbf{t} \in E(G_{\mathbf{r}})$, being in favor of $aS_{\mathbf{r}}b$. This coalition is composed of two subsets of elementary subcriteria:

- subset of elementary subcriteria $g_{\mathbf{t}}, \mathbf{t} \in E(G_{\mathbf{r}})$, being clearly in favor of $aS_{\mathbf{r}}b$, i.e. such that, $g_{\mathbf{t}}(a) \geq g_{\mathbf{t}}(b) - q_{\mathbf{t}}$,
- subset of elementary subcriteria $g_{\mathbf{t}}, \mathbf{t} \in E(G_{\mathbf{r}})$, that do not oppose to $aS_{\mathbf{r}}b$, while being in an ambiguous position with respect to this assertion, i.e. those with $bQ_{\mathbf{r}}a$, which is equivalent to $g_{\mathbf{t}}(b) - p_{\mathbf{t}} < g_{\mathbf{t}}(a) < g_{\mathbf{t}}(b) - q_{\mathbf{t}}$.

Note that $aS_{\mathbf{r}}b$ is true not only when alternative a is preferred to alternative b on criterion/subcriterion $G_{\mathbf{r}}$ but also when a is indifferent to b on $G_{\mathbf{r}}$, and even when b dominates a on $G_{\mathbf{r}}$ by a sufficiently small amount in each elementary subcriterion descending from $G_{\mathbf{r}}$.
Consequently, the partial concordance index is defined as:

$$C_{\mathbf{r}}(a, b) = \sum_{\mathbf{t} \in E(G_{\mathbf{r}})} \phi_{\mathbf{t}}(a, b) \times k_{\mathbf{t}} = \sum_{\mathbf{t} \in E(G_{\mathbf{r}})} \psi_{\mathbf{t}}(a, b) \tag{3.5}$$

where, traditionally, for each $\mathbf{t} \in E(G_{\mathbf{r}})$,

$$\phi_{\mathbf{t}}(a, b) = \begin{cases} 1, & \text{if } \ g_{\mathbf{t}}(a) \geq g_{\mathbf{t}}(b) - q_{\mathbf{t}}, \\ \frac{g_{\mathbf{t}}(a) - [g_{\mathbf{t}}(b) - p_{\mathbf{t}}]}{p_{\mathbf{t}} - q_{\mathbf{t}}}, & \text{if } \ g_{\mathbf{t}}(b) - p_{\mathbf{t}} \leq g_{\mathbf{t}}(a) < g_{\mathbf{t}}(b) - q_{\mathbf{t}}, \\ 0, & \text{if } \ g_{\mathbf{t}}(a) < g_{\mathbf{t}}(b) - p_{\mathbf{t}}. \end{cases} \tag{3.6}$$

$\phi_{\mathbf{t}}(a, b)$ is a marginal concordance index, indicating to what extent elementary subcriterion $g_{\mathbf{t}}, \mathbf{t} \in E(G_{\mathbf{r}})$, contributes to the concordance index $C_{\mathbf{r}}(a, b)$. In order to simplify calculations, and without loss of generality, we assume that the weights of elementary subcriteria sum up to one, i.e. $\sum_{\mathbf{t} \in EL} k_{\mathbf{t}} = 1$.

Note 3.2.1. *When comparing two alternatives a, b on a given elementary subcriterion, the zone between $-p_{\mathbf{t}}$ and $-q_{\mathbf{t}}$ corresponds to hesitation between opting for indifference and preference. In order to take into account this ambiguity, ELECTRE methods consider $\phi_{\mathbf{t}}(a,b)$ being linear and non-decreasing functions with respect to the difference $g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b)$. The assumption of linearity of functions $\phi_{\mathbf{t}}(a,b)$ is only conventional and it is not related in any case to the concept of intensity of preference. Moreover, according to [31], slight changes of the form of $\phi_{\mathbf{t}}(a,b)$ have no impact (apart from very particular cases) on the results.*

Remark that $C_{\mathbf{r}}(a,b) \in [0, K_{\mathbf{r}}]$, where $K_{\mathbf{r}} = \sum_{\mathbf{t} \in E(G_{\mathbf{r}})} k_{\mathbf{t}}$, and $C_{\mathbf{r}}(a,b) = 0$ if $g_{\mathbf{t}}(a) \leq g_{\mathbf{t}}(b) - p_{\mathbf{t}}$, for all $\mathbf{t} \in E(G_{\mathbf{r}})$ (b is strictly preferred to a on all elementary subcriteria descending from $G_{\mathbf{r}}$), and $C_{\mathbf{r}}(a,b) = K_{\mathbf{r}}$ if $g_{\mathbf{t}}(a) \geq g_{\mathbf{t}}(b) - q_{\mathbf{t}}$, for all $\mathbf{t} \in E(G_{\mathbf{r}})$ (a outranks b on all elementary subcriteria descending from $G_{\mathbf{r}}$). When $\mathbf{r} = 0$, $C_{\mathbf{0}}(a,b) \in [0,1]$ because $E(G_{\mathbf{0}}) = EL$ and thus $K_{\mathbf{0}} = 1$.

In ELECTRE, the result of the concordance test concerning a pair of alternatives is positive when the value of the concordance index is not smaller than a fixed concordance cutting level. In the hierarchical extension of ELECTRE, we admit one concordance cutting level $\lambda_{\mathbf{r}}$ for each criterion/subcriterion $G_{\mathbf{r}}$ with $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, that is, we consider one concordance cutting level for each criterion/subcriterion except for elementary subcriteria, such that:

- $\lambda_{\mathbf{s}} \in [K_{\mathbf{s}}/2, K_{\mathbf{s}}]$, for all $\mathbf{s} \in LBO$,
- $\lambda_{\mathbf{r}} = \displaystyle\sum_{j=1}^{n(\mathbf{r})} \lambda_{(\mathbf{r},j)}$, for all $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus \{LBO \cup EL\}$.

In particular, the first condition means that each concordance cutting level $\lambda_{\mathbf{s}}, \mathbf{s} \in LBO$, is bounded between the half-sum and the sum of the weights of elementary subcriteria descending from $G_{\mathbf{s}}$; the second condition means that the concordance cutting level of a criterion/subcriterion $G_{\mathbf{r}}$, is equal to the sum of the concordance cutting levels of subcriteria $G_{(\mathbf{r},\mathbf{j})}, j = 1, \ldots, n(\mathbf{r})$, at the level immediately below; we add this condition in order to avoid the case where an alternative a outranks an alternative b with respect to all subcriteria $G_{(\mathbf{r},1)}, \ldots, G_{(\mathbf{r},n(\mathbf{r}))}$ from the level immediately below $G_{\mathbf{r}}$, but a does not outrank b with respect to criterion/subcriterion $G_{\mathbf{r}}$. For example, it is obvious that if student s_1 outranks student s_2 with respect to Algebra and Analysis, being immediate subcriteria of Mathematics, then s_1 outranks s_2 also with respect to Mathematics.

Note 3.2.2. *Remark that the two above conditions ensure that:*

$$\lambda_{\mathbf{r}} \in \left[\frac{K_{\mathbf{r}}}{2}, K_{\mathbf{r}}\right], \quad \textit{for all}\ \mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL.$$

This implies that not only the concordance cutting level of criteria from the last but one level of the hierarchy, but all concordance cutting levels $\lambda_{\mathbf{r}}$, $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, are constrained between the half-sum and the sum of weights of elementary subcriteria descending from $G_{\mathbf{r}}$.

Note 3.2.3. *In case the DM is not confident in providing a concordance cutting level for each criterion belonging to the last but one level of the hierarchy, (s)he could give another information regarding them. In fact, (s)he could state that each concordance cutting level $\lambda_{\mathbf{r}}, \mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, should be equal to a certain percentage of the sum of weights of the elementary subcriteria descending from criterion $G_{\mathbf{r}}$. For example, if the DM declared that the concordance cutting levels should be equal to 70% of the relative weights of elementary subcriteria descending from the corresponding criteria, it would give $\lambda_{\mathbf{r}} = 0.7 \times \sum_{\mathbf{t}\in E(G_{\mathbf{r}})} k_{\mathbf{t}}$. It is easy to observe that also in this case $\lambda_{\mathbf{r}} = \sum_{j=1}^{n(\mathbf{r})} \lambda_{(\mathbf{r},j)}$ for all $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus \{LBO \cup EL\}$.*

The result of the concordance test for a pair $(a, b) \in B$ is positive if $C_{\mathbf{r}}(a, b) \geq \lambda_{\mathbf{r}}$. Once the result of the concordance test has been positive, one can pass to the non-discordance test. Its result is positive for the pair $(a, b) \in B$, unless "a is significantly worse than b" on at least one elementary subcriterion descending from $G_{\mathbf{r}}$, i.e. if $g_{\mathbf{t}}(b) - g_{\mathbf{t}}(a) \geq v_{\mathbf{t}}$ for any $\mathbf{t} \in E(G_{\mathbf{r}})$.

Remark that, if we consider $\mathbf{r} = 0$ then this procedure boils down to the classical ELECTRE method in which all evaluation criteria are considered at the same level.
Summing up, for each criterion/subcriterion $G_{\mathbf{r}}$, with $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, and for each $a, b \in A$, we have:

$$aS_{\mathbf{r}}b \Leftrightarrow C_{\mathbf{r}}(a, b) \geq \lambda_{\mathbf{r}}, \text{ and } g_{\mathbf{t}}(b) - g_{\mathbf{t}}(a) < v_{\mathbf{t}}, \text{ for all } \mathbf{t} \in E(G_{\mathbf{r}}).$$

In the following Proposition we show two fundamental properties of hierarchical outranking:

Proposition 3.2.1.

1. *Given two alternatives $a, b \in A$, and $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus (LBO \cup EL)$, such that*

$$aS_{(\mathbf{r},j)}b, \quad \textit{for all } j = 1, \ldots, n(\mathbf{r}),$$

   then $aS_{\mathbf{r}}b$,

2. *Given two alternatives $a, b \in A$, and $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus (LBO \cup EL)$, such that*

$$not(aS_{(\mathbf{r},j)}b), \quad \textit{for all } j = 1, \ldots, n(\mathbf{r}),$$

   then $not(aS_{\mathbf{r}}b)$.

Proof. See Appendix A. □

Note 3.2.4. *Until now, we have applied the concepts of the hierarchy of criteria to one specific ELECTRE method that is the ELECTRE IS method. The Multiple Criteria Hierarchy Process (MCHP) can be applied also to other ELECTRE methods, including the most popular ELECTRE III method. ELECTRE III builds, for each couple of alternatives $(a,b) \in B$, the credibility index*

$$\rho(a,b) = C(a,b) \prod_{\{j: d_j(a,b) > C(a,b)\}} \frac{1 - d_j(a,b)}{1 - C(a,b)}$$

where for each criterion g_j,

$$d_j(a,b) = \begin{cases} 1, & \text{if } \; g_j(a) \leq g_j(b) - v_j, \\ \frac{g_j(a) - [g_j(b) - p_j]}{v_j - p_j}, & \text{if } \; g_j(b) - v_j < g_j(a) < g_j(b) - p_j, \\ 0, & \text{if } \; g_j(a) \geq g_j(b) - p_j. \end{cases}$$

In MCHP, for each criterion $G_{\mathbf{r}}$ with $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, we can define the following credibility index

$$\rho_{\mathbf{r}}(a,b) = C_{\mathbf{r}}(a,b) \prod_{\{\mathbf{t} \in E(G_{\mathbf{r}}) \,:\, d_{\mathbf{t}}(a,b) > C_{\mathbf{r}}(a,b)\}} \frac{1 - d_{\mathbf{t}}(a,b)}{1 - C_{\mathbf{r}}(a,b)}$$

where $d_{\mathbf{t}}(a,b)$ is defined equivalently to $d_j(a,b)$ for each elementary subcriterion $g_{\mathbf{t}}$, $\mathbf{t} \in EL$.
From the definition of $\rho_{\mathbf{r}}(a,b)$ it follows that if none of the elementary subcriteria descending from $G_{\mathbf{r}}$ opposes veto to the outranking of a over b on criterion $G_{\mathbf{r}}$ (that is $d_{\mathbf{r}}(a,b) = 0$ for all $\mathbf{t} \in E(G_{\mathbf{r}})$), then $\rho_{\mathbf{r}}(a,b) = C_{\mathbf{r}}(a,b)$; if some elementary subcriterion descending from $G_{\mathbf{r}}$ opposes the veto to the outranking of a over b with respect to criterion $G_{\mathbf{r}}$ (that is there exists at least one elementary subcriterion $g_{\mathbf{t}}$, with $\mathbf{t} \in E(G_{\mathbf{r}})$, such that $d_{\mathbf{t}}(a,b) = 1$), then $\rho_{\mathbf{r}}(a,b) = 0$ and in all other cases the credibility index $\rho_{\mathbf{r}}(a,b)$ is lower than the concordance index $C_{\mathbf{r}}(a,b)$.

3.2.3 Hierarchical ELECTRE^{GKMS}

The only information the DM can obtain from the evaluations of alternatives with respect to the considered criteria is the dominance relation. In general, information carried by the dominance relation is very poor, and thus, in order to arrive to a final decision which would be concordant with the value system of the DM, it is useful to take into account some preference information provided by the DM. This preference information can be obtained in either direct or indirect way: if the way is direct, then the DM provides precise values or interval of values for the parameters present in the model, and if the way is indirect, then the DM is invited to provide preference information from which the parameters of the model can be inferred.

In this work, we use a mix of both ways in order to infer the parameters of the model. We suppose that, considering a multiple criteria choice or ranking problem, the DM can provide preference information of two types:

- pairwise comparisons of some reference alternatives from set $A^R \subseteq A$, stating the truth or falsity of outranking relation $aS_{\mathbf{r}}b$, with $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$ and $a, b \in A^R$ (in the following $B^R = A^R \times A^R$),

- information regarding the indifference and preference thresholds $q_{\mathbf{t}}$ and $p_{\mathbf{t}}$ for each elementary subcriterion $g_{\mathbf{t}}, \mathbf{t} \in EL$ and information regarding weights $k_{\mathbf{t}}$ for some elementary subcriterion $g_{\mathbf{t}}, \mathbf{t} \in EL$.

Regarding the direct preference information, the DM can provide intervals of possible values $[q_{\mathbf{t},*}, q_{\mathbf{t}}^*]$ and $[p_{\mathbf{t},*}, p_{\mathbf{t}}^*]$ for each indifference and preference threshold $q_{\mathbf{t}}, p_{\mathbf{t}}$, $\mathbf{t} \in EL$, where $q_{\mathbf{t},*}$ and $q_{\mathbf{t}}^*$ are, respectively, the smallest and the greatest value of the indifference threshold, and $p_{\mathbf{t},*}$ and $p_{\mathbf{t}}^*$ are, respectively, the smallest and the greatest value of the preference threshold allowed by the DM. Besides, we assume that the DM could give information on the weight $k_{\mathbf{t}}$ of some elementary subcriterion providing interval of possible values $[k_{\mathbf{t},*}, k_{\mathbf{t}}^*]$, where $k_{\mathbf{t},*}$ and $k_{\mathbf{t}}^*$ are, respectively, the smallest and the greatest value of the weights allowed by the DM, or providing pairwise comparison between the elementary subcriteria of the type: "elementary subcriterion $g_{\mathbf{t_1}}$ is more important than elemenatry subcriterion $g_{\mathbf{t_2}}$" or "elementary subcriteria $g_{\mathbf{t_1}}$ and $g_{\mathbf{t_2}}$ are equally important" that are translated from the constraints $k_{\mathbf{t_1}} > k_{\mathbf{t_2}}$ and $k_{\mathbf{t_1}} = k_{\mathbf{t_2}}$ respectively.
If the DM cannot provide intervals of indifference and preference threshold values for an elementary subcriterion $g_{\mathbf{t}}$, then (s)he has to indicate at least one couple of reference alternatives $a, b \in A^R \subseteq A$ for which the difference between $g_{\mathbf{t}}(a)$ and $g_{\mathbf{t}}(b)$ is non-significant for the DM ($a \sim_{\mathbf{t}} b$), and at least one couple of reference alternatives a, b for which the difference between $g_{\mathbf{t}}(a)$ and $g_{\mathbf{t}}(b)$ is significant for the DM ($a \succ_{\mathbf{t}} b$). We denote by EL_1 and EL_2 the subsets of EL (such that $EL_1 \cup EL_2 = EL$) containing indices of elementary subcriteria for which the DM provides information about the thresholds in a direct or indirect way, respectively.
In order to ensure the consistency of the above thresholds, the following constraints need to be satisfied:

- $q_{\mathbf{t},*} \leq q_{\mathbf{t}}^*$, $p_{\mathbf{t},*} \leq p_{\mathbf{t}}^*$ and $q_{\mathbf{t}}^* \leq p_{\mathbf{t},*}$, for all $\mathbf{t} \in EL_1$,

- $|g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b)| < g_{\mathbf{t}}(c) - g_{\mathbf{t}}(d)$, if $a \sim_{\mathbf{t}} b$ and $c \succ_{\mathbf{t}} d$, for all $\mathbf{t} \in EL_2$,

- $p_{\mathbf{t}}^*$ should be not greater than $\beta_{\mathbf{t}} - \alpha_{\mathbf{t}}$, $\mathbf{t} \in EL_1$, where $\alpha_{\mathbf{t}} = \min_{a \in A} g_{\mathbf{t}}(a)$, and $\beta_{\mathbf{t}} = \max_{a \in A} g_{\mathbf{t}}(a)$.

We call *compatible model*, a set of parameters (thus variables $\psi_{\mathbf{t}}(a,b)$ for each pair of alternatives $(a,b) \in B$ and for each elementary subcriterion $g_{\mathbf{t}}$, $\mathbf{t} \in EL$, veto thresholds $v_{\mathbf{t}}$ for all $\mathbf{t} \in EL$, and concordance cutting levels $\lambda_{\mathbf{s}}$ for all $\mathbf{s} \in LBO$) which restore the preference information provided by the DM and thus satisfy the following set of constraints (see [47] for a similar formulation in a non-hierarchical case, and Appendix B for a detailed description of the constraints):

Pairwise comparison stating $aS_{\mathbf{r}}b$ or $not(aS_{\mathbf{r}}b)$:

$$C_{\mathbf{r}}(a,b) = \sum_{\mathbf{t}\in E(G_{\mathbf{r}})} \psi_{\mathbf{t}}(a,b) \geq \lambda_{\mathbf{r}} \;\; \text{and} \;\; g_{\mathbf{t}}(b) - g_{\mathbf{t}}(a) + \varepsilon \leq v_{\mathbf{t}}, \;\; \mathbf{t} \in E(G_{\mathbf{r}}),$$
$$\text{if } aS_{\mathbf{r}}b, \;\; \text{for} \;\; (a,b) \in B^R,$$

$$C_{\mathbf{r}}(a,b) = \sum_{\mathbf{t}\in E(G_{\mathbf{r}})} \psi_{\mathbf{t}}(a,b) + \varepsilon \leq \lambda_{\mathbf{r}} + M_0^{\mathbf{r}}(a,b) \;\; \text{and} \;\; g_{\mathbf{t}}(b) - g_{\mathbf{t}}(a) \geq v_{\mathbf{t}} - \delta_{\mathbf{r}} M_{\mathbf{t}}(a,b),$$
$$\text{if } \; not(aS_{\mathbf{r}}b), \;\; \text{for} \;\; (a,b) \in B^R,$$
$$M_0^{\mathbf{r}}(a,b), M_{\mathbf{t}}(a,b) \in \{0,1\}, \; \text{for all } \mathbf{t} \in E(G_{\mathbf{r}}), \;\; M_0^{\mathbf{r}}(a,b) + \sum_{\mathbf{t}\in E(G_{\mathbf{r}})} M_{\mathbf{t}}(a,b) \leq |E(G_{\mathbf{r}})|\,,$$
$$\delta_{\mathbf{r}} \geq max_{\mathbf{t}\in E(G_{\mathbf{r}})}\{\beta_{\mathbf{t}} - \alpha_{\mathbf{t}}\} \text{ where } \alpha_{\mathbf{t}} = \min_{a\in A} g_{\mathbf{t}}(a) \text{ and } \beta_{\mathbf{t}} = \max_{a\in A} g_{\mathbf{t}}(a).$$

Concordance cutting levels and values of inter-criteria parameters:

$$\lambda_{\mathbf{s}} \geq \sum_{\mathbf{t}\in E(G_{\mathbf{s}})} \frac{\psi_{\mathbf{t}}(x_{\mathbf{t}}^*, x_{\mathbf{t},*})}{2}, \;\; \text{and } \lambda_{\mathbf{s}} \leq \sum_{\mathbf{t}\in E(G_{\mathbf{s}})} \psi_{\mathbf{t}}(x_{\mathbf{t}}^*, x_{\mathbf{t},*}), \; \text{for all } \mathbf{s} \in LBO,$$
$$\lambda_{\mathbf{r}} = \sum_{j=1}^{n(\mathbf{r})} \lambda_{(\mathbf{r},j)}, \; \text{for all } \mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus \{LBO \cup EL\}\,,$$
$$\sum_{\mathbf{t}\in EL} \psi_{\mathbf{t}}(x_{\mathbf{t}}^*, x_{\mathbf{t},*}) = 1, \; \text{where } x_{\mathbf{t},*}, x_{\mathbf{t}}^* \in A \text{ for all } \mathbf{t} \in EL : g_{\mathbf{t}}(x_{\mathbf{t}}^*) = \beta_{\mathbf{t}}, \; \text{and } g_{\mathbf{t}}(x_{\mathbf{t},*}) = \alpha_{\mathbf{t}},$$
$$v_{\mathbf{t}} \geq p_{\mathbf{t}}^* + \varepsilon, \;\; \mathbf{t} \in EL,$$
$$v_{\mathbf{t}} \geq g_{\mathbf{t}}(b) - g_{\mathbf{t}}(a) + \varepsilon \;\; \text{if} \;\; a \sim_{\mathbf{t}} b, \; \text{and } g_{\mathbf{t}}(a) \leq g_{\mathbf{t}}(b), \;\; \mathbf{t} \in EL_2, \; \text{for all } (a,b) \in B,$$

Values of marginal concordance indices conditioned by intra-criterion preference information, for all $(a,b) \in B$:

$$k_{\mathbf{t},*} \leq \psi_{\mathbf{t}}(x_{\mathbf{t}}^*, x_{\mathbf{t},*}) \leq k_{\mathbf{t}}^*, \mathbf{t} \in EL,$$
$$\psi_{\mathbf{t_1}}(x_{\mathbf{t_1}}^*, x_{\mathbf{t_1},*}) \geq \psi_{\mathbf{t_2}}(x_{\mathbf{t_2}}^*, x_{\mathbf{t_2},*}) + \varepsilon, \;\; \text{if elementary subcriterion } g_{\mathbf{t_1}} \text{ is more important than elementary subcriterion } g_{\mathbf{t_2}}, \mathbf{t_1}, \mathbf{t_2} \in EL,$$
$$\psi_{\mathbf{t_1}}(x_{\mathbf{t_1}}^*, x_{\mathbf{t_1},*}) = \psi_{\mathbf{t_2}}(x_{\mathbf{t_2}}^*, x_{\mathbf{t_2},*}), \;\; \text{if elementary subcriteria } g_{\mathbf{t_1}} \text{ and } g_{\mathbf{t_2}} \text{ are equally important}, \mathbf{t_1}, \mathbf{t_2} \in EL,$$
$$\psi_{\mathbf{t}}(a,b) = 0 \;\; \text{if} \;\; g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b) \leq -p_{\mathbf{t}}^*, \mathbf{t} \in EL_1,$$
$$\psi_{\mathbf{t}}(a,b) \geq \varepsilon \;\; \text{if} \;\; g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b) > -p_{\mathbf{t},*}, \mathbf{t} \in EL_1,$$
$$\psi_{\mathbf{t}}(a,b) = \psi_{\mathbf{t}}(x_{\mathbf{t}}^*, x_{\mathbf{t},*}) \;\; \text{if} \;\; g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b) \geq -q_{\mathbf{t},*}, \mathbf{t} \in EL_1,$$
$$\psi_{\mathbf{t}}(a,b) + \varepsilon \leq \psi_{\mathbf{t}}(x_{\mathbf{t}}^*, x_{\mathbf{t},*}) \;\; \text{if} \;\; g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b) < -q_{\mathbf{t}}^*, \mathbf{t} \in EL_1,$$
$$\psi_{\mathbf{t}}(a,b) = \psi_{\mathbf{t}}(b,a) = \psi_{\mathbf{t}}(x_{\mathbf{t}}^*, x_{\mathbf{t},*}) \;\; \text{if} \;\; a \sim_{\mathbf{t}} b, \mathbf{t} \in EL_2$$
$$\psi_{\mathbf{t}}(a,b) = 0 \;\; \text{if} \;\; b \succ_{\mathbf{t}} a, \mathbf{t} \in EL_2.$$

Monotonicity of the functions of marginal concordance indices, for all $a,b,c,d \in A$, $\mathbf{t} \in EL$:

$$\psi_{\mathbf{t}}(a,b) \geq \psi_{\mathbf{t}}(c,d) \;\; \text{if} \;\; g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b) > g_{\mathbf{t}}(c) - g_{\mathbf{t}}(d),$$
$$\psi_{\mathbf{t}}(a,b) = \psi_{\mathbf{t}}(c,d) \;\; \text{if} \;\; g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b) = g_{\mathbf{t}}(c) - g_{\mathbf{t}}(d),$$

$\Bigg\}\; E^{A^R}$

The whole set of constraints E^{A^R} has the form of 0-1 Mixed Integer Linear Program (MILP), as shown above. If E^{A^R} is feasible and $\varepsilon^* = max\ \varepsilon$, subject to E^{A^R}, is greater than 0, then there exists at least one outranking model compatible with the preference information.

In general, there may exist more than one outranking model compatible with preference information provided by the DM; each one of the compatible models restores the preference information concerning the reference alternatives provided by the DM, but it can compare in a different way the other couples of alternatives not present in the preference information provided by the DM. For this reason, ROR takes into account all outranking models compatible with preference information provided by the DM simultaneously. In the ROR context, in case of the hierarchy of criteria applied to ELECTRE, and considering a criterion/subcriterion $G_{\mathbf{r}}$ of the hierarchy, with $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$ and two alternatives $a, b \in A$, we can give the following definitions:

Definition 3.2.1.

- *a necessarily outranks b with respect to* $G_{\mathbf{r}}$*, and we write* $aS_{\mathbf{r}}^{N}b$*, if a outranks b with respect to* $G_{\mathbf{r}}$*, for all compatible models,*
- *a possibly outranks b with respect to* $G_{\mathbf{r}}$*, and we write* $aS_{\mathbf{r}}^{P}b$*, if a outranks b with respect to* $G_{\mathbf{r}}$*, for at least one compatible model,*
- *a necessarily does not outrank b with respect to* $G_{\mathbf{r}}$*, and we write* $aS_{\mathbf{r}}^{CN}b$*, if a does not outrank b with respect to* $G_{\mathbf{r}}$*, for all compatible models,*
- *a possibly does not outrank b with respect to* $G_{\mathbf{r}}$*, and we write* $aS_{\mathbf{r}}^{CP}b$*, if a does not outrank b with respect to* $G_{\mathbf{r}}$*, for at least one compatible model.*

Remark that, in case of $\mathbf{r} = 0$, the necessary and possible outranking relations $S_{\mathbf{r}}^{N}$ and $S_{\mathbf{r}}^{P}$ are the same as necessary and possible outranking relations defined in [47], for a flat (non-hierarchical) structure of the set of criteria.

Given a pair of alternatives $(a, b) \in B$, and a criterion $G_{\mathbf{r}} \in \mathcal{G}$ with $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, necessary and possible outranking relations ($\succsim_{\mathbf{r}}^{N}$, $\succsim_{\mathbf{r}}^{P}$) can be computed as follows.

- To check whether $aS_{\mathbf{r}}^{N}b$, we assume that a does not outrank b with respect to criterion $G_{\mathbf{r}}$ ($not(aS_{\mathbf{r}}b)$), and we add the corresponding constraints to set E^{A^R}, getting the set of constraints $E_{\mathbf{r}}^{N}(a, b)$ shown below. Then, we verify whether $not(aS_{\mathbf{r}}b)$ is possible in the set of all outranking models compatible with the previously provided preference information.

$$\left.\begin{array}{l}E^{A^R}\\ C_{\mathbf{r}}(a,b)=\displaystyle\sum_{\mathbf{t}\in E(G_{\mathbf{r}})}\psi_{\mathbf{t}}(a,b)+\varepsilon\leq\lambda_{\mathbf{r}}+M_0^{\mathbf{r}}(a,b)\ \text{ and }\ g_{\mathbf{t}}(b)-g_{\mathbf{t}}(a)\geq v_{\mathbf{t}}-\delta_{\mathbf{r}}M_{\mathbf{t}}(a,b),\\ M_0^{\mathbf{r}}(a,b)+\displaystyle\sum_{\mathbf{t}\in E(G_{\mathbf{r}})}M_{\mathbf{t}}(a,b)\leq|E(G_{\mathbf{r}})|,\ M_0^{\mathbf{r}}(a,b),M_{\mathbf{t}}(a,b)\in\{0,1\},\ \ \mathbf{t}\in E(G_{\mathbf{r}}).\end{array}\right\}E_{\mathbf{r}}^N(a,b)$$

We say that:

$aS_{\mathbf{r}}^Nb$ if $E_{\mathbf{r}}^N(a,b)$ is infeasible or $\varepsilon_{\mathbf{r}}^N(a,b)\leq 0$ where $\varepsilon_{\mathbf{r}}^N(a,b)=max\ \varepsilon$, subject to $E_{\mathbf{r}}^N(a,b)$.

- To check whether $aS_{\mathbf{r}}^Pb$, we assume that a outranks b with respect to criterion $G_{\mathbf{r}}$ ($aS_{\mathbf{r}}b$), and we add the corresponding constraints to the set E^{A^R}, getting the set of constraints $E_{\mathbf{r}}^P(a,b)$ shown below. Then, we verify whether $aS_{\mathbf{r}}b$ is possible in the set of all outranking models compatible with the previously provided preference information.

$$\left.\begin{array}{l}E^{A^R}\\ C_{\mathbf{r}}(a,b)=\displaystyle\sum_{\mathbf{t}\in E(G_{\mathbf{r}})}\psi_{\mathbf{t}}(a,b)\geq\lambda_{\mathbf{r}}\ \text{ and }\ g_{\mathbf{t}}(b)-g_{\mathbf{t}}(a)+\varepsilon\leq v_{\mathbf{t}},\ \mathbf{t}\in E(G_{\mathbf{r}})\end{array}\right\}E_{\mathbf{r}}^P(a,b)$$

We say that:

$aS_{\mathbf{r}}^Pb$ if $E_{\mathbf{r}}^P(a,b)$ is feasible and $\varepsilon_{\mathbf{r}}^P(a,b)>0$ where $\varepsilon_{\mathbf{r}}^P(a,b)=max\ \varepsilon$, subject to $E_{\mathbf{r}}^P(a,b)$.

Note 3.2.5. *It is worth noting that the set of constraints E^{A^R} defines a set of variables $\psi_{\mathbf{t}}(a,b)$, $\mathbf{t}\in EL$, $(a,b)\in B$, being non-decreasing functions with respect to the difference $g_{\mathbf{t}}(a)-g_{\mathbf{t}}(b)$, differently from the functions $\phi_{\mathbf{t}}(a,b)$ which are non-decreasing and linear. This is due to the fact that indifference and preference thresholds, as well as the veto thresholds, are not directly provided by the DM. In this situation, taking thresholds and weights as unknown variables, makes that the optimization problems to be solved in ROR are non more linear programming ones. As there are many optimization problems to be solved in ROR, the whole approach would be practically non-tractable.*
If the DM would be able to provide all the thresholds considered in the model (indifference, preference and veto), then linear programming could be applied again within a simplified model of ROR, considering as variables only the weights $k_{\mathbf{t}},\mathbf{t}\in EL$, and the concordance cutting levels $\lambda_{\mathbf{s}}$, $\mathbf{s}\in LBO$. In this case, the feasibility constraints of the optimization problems considered in ROR can be simply modified to the following form:

$$
\left.
\begin{array}{l}
\textit{Pairwise comparison stating } aS_{\mathbf{r}}b \textit{ or } not(aS_{\mathbf{r}}b)\textit{:} \\
\quad C_{\mathbf{r}}(a,b) = \sum_{\mathbf{t} \in E(G_{\mathbf{r}})} k_{\mathbf{t}} \cdot \phi_{\mathbf{t}}(a,b) \geq \lambda_{\mathbf{r}}, \quad \textit{if } \ aS_{\mathbf{r}}b, \quad \textit{for } \ (a,b) \in B^R, \\
\quad C_{\mathbf{r}}(a,b) = \sum_{\mathbf{t} \in E(G_{\mathbf{r}})} k_{\mathbf{t}} \cdot \phi_{\mathbf{t}}(a,b) + \varepsilon \leq \lambda_{\mathbf{r}} \ \ \textit{if } \ not(aS_{\mathbf{r}}b), \quad \textit{for } \ (a,b) \in B^R, \\
\textit{Concordance cutting levels and values of inter-criteria parameters:} \\
\quad \lambda_{\mathbf{s}} \geq \sum_{\mathbf{t} \in E(G_{\mathbf{s}})} \frac{k_{\mathbf{t}}}{2}, \ \ \textit{and } \lambda_{\mathbf{s}} \leq \sum_{\mathbf{t} \in E(G_{\mathbf{s}})} k_{\mathbf{t}}, \ \textit{for all } \mathbf{s} \in LBO, \\
\quad \sum_{\mathbf{t} \in EL} k_{\mathbf{t}} = 1, \\
\quad k_{\mathbf{t},*} \leq k_{\mathbf{t}} \leq k_{\mathbf{t}}^{*}, \mathbf{t} \in EL, \\
\quad k_{\mathbf{t_1}} \geq k_{\mathbf{t_2}} + \varepsilon, \quad \textit{if elementary subcriterion } g_{\mathbf{t_1}} \textit{ is more important than} \\
\qquad\quad \textit{elementary subcriterion } g_{\mathbf{t_2}}, \mathbf{t_1}, \mathbf{t_2} \in EL, \\
\quad k_{\mathbf{t_1}} = k_{\mathbf{t_2}}, \quad \textit{if elementary subcriteria } g_{\mathbf{t_1}} \textit{ and } g_{\mathbf{t_2}} \textit{ are} \\
\qquad\quad \textit{equally important}, \mathbf{t_1}, \mathbf{t_2} \in EL,
\end{array}
\right\} E^{A^R}
$$

$$
\left.
\begin{array}{l}
E^{A^R} \\
C_{\mathbf{r}}(a,b) = \sum_{\mathbf{t} \in E(G_{\mathbf{r}})} k_{\mathbf{t}} \cdot \phi_{\mathbf{t}}(a,b) + \varepsilon \leq \lambda_{\mathbf{r}},
\end{array}
\right\} E_{\mathbf{r}}^{N}(a,b), \qquad
\left.
\begin{array}{l}
E^{A^R} \\
C_{\mathbf{r}}(a,b) = \sum_{\mathbf{t} \in E(G_{\mathbf{r}})} k_{\mathbf{t}} \cdot \phi_{\mathbf{t}}(a,b) \geq \lambda_{\mathbf{r}},
\end{array}
\right\} E_{\mathbf{r}}^{P}(a,b)
$$

The linearity and simplicity of the above formulation is concordant with reasoning of Note 3.2.1. Remark that the preference information of the type $aS_{\mathbf{r}}b$ and $not(aS_{\mathbf{r}}b)$ provided by the DM involves the concordance test only, because the veto thresholds were given before by the DM, and thus not $(aS_{\mathbf{r}}b)$ could not reasonably be caused by discordance. Indeed, it is reasonable to assume that the DM stating that $aS_{\mathbf{r}}b$ or $not(aS_{\mathbf{r}}b)$ already knows that $g_{\mathbf{t}}(b) - g_{\mathbf{t}}(a) < v_{\mathbf{t}}$ for all $\mathbf{t} \in E(G_{\mathbf{r}})$.

Properties of necessary and possible outranking relations in hierarchical ELECTRE^{GKMS}

Proposition 3.2.2.

1. *For all $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, $S_{\mathbf{r}}^{N} \subseteq S_{\mathbf{r}}^{P}$,*
2. *For all $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, $S_{\mathbf{r}}^{P}$ and $S_{\mathbf{r}}^{N}$ are reflexive,*
3. *For all $a, b \in A$, for all $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, $aS_{\mathbf{r}}^{N}b \Leftrightarrow not(aS_{\mathbf{r}}^{CP}b)$,*
4. *For all $a, b \in A$, for all $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, $aS_{\mathbf{r}}^{P}b \Leftrightarrow not(aS_{\mathbf{r}}^{CN}b)$,*
5. *$S_{\mathbf{r}}^{CN} \subseteq S_{\mathbf{r}}^{CP}$, for all $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$,*
6. *For all $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, $S_{\mathbf{r}}^{CP}$ and $S_{\mathbf{r}}^{CN}$ are irreflexive.*

Proof. See Appendix A. □

Proposition 3.2.3.

1. *Given two alternatives* $a, b \in A$ *and* $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus (EL \cup LBO)$, *such that*

$$aS^N_{(\mathbf{r},j)}b \quad \text{for all } j = 1, \ldots, n(\mathbf{r}),$$

then $aS^N_{\mathbf{r}}b$,

2. *Given two alternatives* $a, b \in A$ *and* $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus (EL \cup LBO)$, *such that:*

α) $aS^N_{(\mathbf{r},j)}b$ *for all* $j = 1, \ldots, n(\mathbf{r}), j \neq w$,

β) $aS^P_{(\mathbf{r},w)}b$,

then $aS^P_{\mathbf{r}}b$,

3. *Given two alternatives* $a, b \in A$ *and* $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus (EL \cup LBO)$, *such that*

$$aS^{CN}_{(\mathbf{r},j)}b \quad \text{for all } j = 1, \ldots, n(\mathbf{r}),$$

then $aS^{CN}_{\mathbf{r}}b$.

4. *Given two alternatives* $a, b \in A$ *and* $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus (EL \cup LBO)$, *such that:*

α) $aS^{CN}_{(\mathbf{r},j)}b$ *for all* $j = 1, \ldots, n(\mathbf{r}), j \neq w$,

β) $aS^{CP}_{(\mathbf{r},w)}b$,

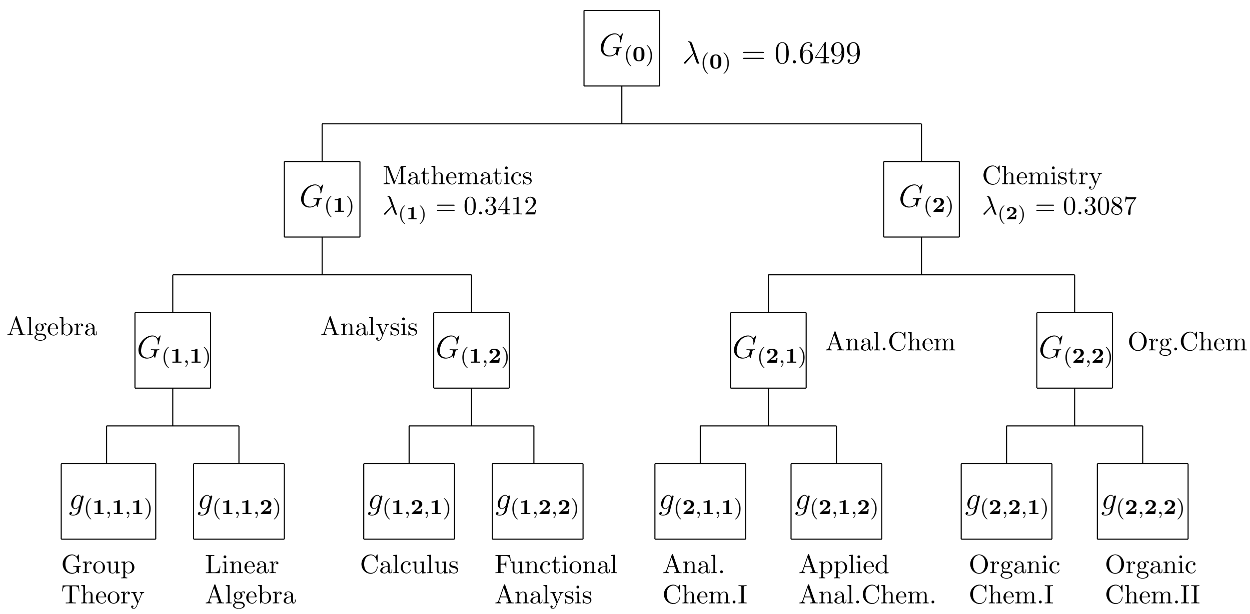
then $aS^{CP}_{\mathbf{r}}b$.

Proof. See Appendix A. □

An illustrative example

In this section, we present an illustrative example in order to show how to use the ELECTRE method and the ELECTRE^{GKMS} method in case of the hierarchical structure of criteria. At first, we describe how to use ELECTRE method in case of availability of full preference information composed of the weights, the preference, indifference and veto thresholds, and about the concordance cutting levels. Let us suppose that a university department of natural sciences, like every year, has a possibility of funding a scholarship for one of its best students. As five best students passed to the final selection, the Dean has to choose from among them one laureate. These five finalists are evaluated with respect to two macro-subjects: Mathematics and Chemistry. Both these macro-subjects present a hierarchical structure; on one hand, Mathematics has two sub-subjects: Algebra and Analysis, and each one of these has other two sub-subjects: Group Theory and Linear Algebra are sub-subjects of Algebra, while Calculus and Functional Analysis are sub-subjects of Analysis. On the other hand, Chemistry has two sub-subjects: Analytical Chemistry and Organic Chemistry, and each one of them has two sub-subjects: Analytical Chemistry I and Applied Analytical Chemistry are sub-subjects of Analytical Chemistry, while Organic Chemistry I and Organic Chemistry II are sub-subjects of Organic Chemistry. The described hierarchy of criteria is shown in Figure 3.12.

Figure 3.12: Hierarchical evaluation of students



The eight sub-subjects are thus the elementary subcriteria of the considered hierarchical structure and the evaluations of the students with respect to the eight sub-subjects are shown in Table 3.5. The evaluation of students on these elementary subcriteria is included between 18 and 30. Weights, indifference, preference and veto thresholds are shown in Table 3.6(a).

Table 3.5: Evaluations of students on elementary subcriteria

Student	$g_{(\mathbf{1},\mathbf{1},\mathbf{1})}$	$g_{(\mathbf{1},\mathbf{1},\mathbf{2})}$	$g_{(\mathbf{1},\mathbf{2},\mathbf{1})}$	$g_{(\mathbf{1},\mathbf{2},\mathbf{2})}$	$g_{(\mathbf{2},\mathbf{1},\mathbf{1})}$	$g_{(\mathbf{2},\mathbf{1},\mathbf{2})}$	$g_{(\mathbf{2},\mathbf{2},\mathbf{1})}$	$g_{(\mathbf{2},\mathbf{2},\mathbf{2})}$
s_1	28	22	27	21	29	21	28	20
s_2	20	23	19	22	30	20	29	19
s_3	29	21	28	20	18	24	18	23
s_4	30	20	29	19	28	22	27	21
s_5	18	24	18	23	20	23	19	22

The Dean decides to provide information regarding the concordance cutting levels in the way explained in the Note 3.2.3. Then, (s)he states that each concordance cutting level $\lambda_{\mathbf{s}}$, $\mathbf{s} \in LBO$, should be equal to 65% of the sum of the weights of elementary subcriteria descending from $G_{\mathbf{s}}$. In consequence, for each criterion of the last but one level the values of concordance cutting levels are: $\lambda_{(\mathbf{1},\mathbf{1})} = 0.1625$, $\lambda_{(\mathbf{1},\mathbf{2})} = 0.1787$, $\lambda_{(\mathbf{2},\mathbf{1})} = 0.1462$ and $\lambda_{(\mathbf{2},\mathbf{2})} = 0.1625$ as reported in Table 3.6(b).

Table 3.6: ELECTRE parameters in case of the hierarchy of criteria

(a) Weights and thresholds

Elementary subcriterion, $g_{\mathbf{t}}$	$k_{\mathbf{t}}$	$q_{\mathbf{t}}$	$p_{\mathbf{t}}$	$v_{\mathbf{t}}$
Group Theory	0.1	1	4	10
Linear Algebra	0.15	1	4	10
Calculus	0.125	1	4	10
Functional Analysis	0.15	1	4	10
Analytical Chemistry I	0.1	2	5	10
App. Anal. Chemistry	0.125	2	5	10
Organic Chemistry I	0.15	2	5	10
Organic Chemistry II	0.1	2	5	10

(b) Concordance cutting levels

Criterion, $G_{\mathbf{r}}$	$\lambda_{\mathbf{r}}$
Algebra	0.1625
Analysis	0.1787
Analytical Chemistry	0.1462
Organic Chemistry	0.1625

Following the procedure explained in section 3.2.2, we obtain the outranking relations shown in Table 3.7, where:

$$S_{(\mathbf{r})}(s_1, s_2) = \begin{cases} 1 & \text{if} \quad s_1 \text{ outranks } s_2 \text{ with respect to criterion } G_{\mathbf{r}}, \\ 0 & \text{if} \quad s_1 \text{ does not outrank } s_2 \text{ with respect to criterion } G_{\mathbf{r}}. \end{cases}$$

In Table 3.7, we obtain the "overall outranking relation", that is the outranking relation with respect to the totality of criteria, as well as "partial outranking relations", that is outranking relations with respect to a particular subcriterion at a given level of the hierarchy.

Remark that using the classical ELECTRE method, we obtain the comprehensive outranking relation only because all criteria are considered at the same level. In consequence, using the classical ELECTRE method, we could learn that student s_2 does not outrank student s_4 with respect to the totality of criteria, but we could not know that student s_2 outranks student s_4 with respect to Chemistry, Analytical Chemistry and Organic Chemistry.

According to point 1 of Proposition 3.2.1, we observe that if student s_1 outranks student s_5 with respect to Mathematics ($G_{(\mathbf{1})}$) and Chemistry ($G_{(\mathbf{2})}$), then s_1 outranks s_5 with respect to the totality of criteria ($G_{(\mathbf{0})}$), but the contrary is not true; in fact, for example, student s_2 outranks student s_3 with respect to Chemistry, but at the same time student s_2 does not outrank student s_3 with respect to Analytical Chemistry ($G_{(\mathbf{2},\mathbf{1})}$), being a sub-criterion descending from Chemistry.

Now, let us suppose that the Dean cannot provide the full preference information regarding the

Table 3.7: Outranking relations at particular levels of the hierarchy of criteria

$S_{(\mathbf{0})}$	s_1	s_2	s_3	s_4	s_5
s_1	1	1	1	1	1
s_2	1	1	1	1	1
s_3	0	0	1	0	1
s_4	1	1	1	1	1
s_5	0	0	0	0	1

$S_{(\mathbf{1})}$	s_1	s_2	s_3	s_4	s_5
s_1	1	1	1	1	1
s_2	0	1	0	0	1
s_3	1	1	1	1	0
s_4	1	0	1	1	0
s_5	0	1	0	0	1

$S_{(\mathbf{2})}$	s_1	s_2	s_3	s_4	s_5
s_1	1	1	1	1	1
s_2	1	1	1	1	1
s_3	0	0	1	0	1
s_4	1	1	1	1	1
s_5	0	0	1	0	1

$S_{(\mathbf{1,1})}$	s_1	s_2	s_3	s_4	s_5
s_1	1	1	1	1	1
s_2	0	1	0	0	1
s_3	1	1	1	1	0
s_4	1	0	1	1	0
s_5	0	1	0	0	1

$S_{(\mathbf{1,2})}$	s_1	s_2	s_3	s_4	s_5
s_1	1	1	1	1	1
s_2	0	1	0	0	1
s_3	1	1	1	1	0
s_4	1	0	1	1	0
s_5	0	1	0	0	1

$S_{(\mathbf{2,1})}$	s_1	s_2	s_3	s_4	s_5
s_1	1	1	1	1	1
s_2	1	1	0	1	1
s_3	0	0	1	0	1
s_4	1	1	1	1	1
s_5	0	0	1	0	1

$(\mathbf{2,2})$	s_1	s_2	s_3	s_4	s_5
s_1	1	1	1	1	1
s_2	1	1	1	1	1
s_3	0	0	1	0	1
s_4	1	1	1	1	1
s_5	0	0	1	0	1

parameters of the Hierarchical ELECTRE method. The only information the Dean can get from the evaluation table is the dominance relation, but in this particular case there is no student dominating another student. Thus, the Dean decides to use the Hierarchical ELECTRE^{GKMS} method. In fact, (s)he realizes that using this procedure, (s)he has two advantages: (s)he can give finer preference information, taking into account subsets of criteria at different levels of the hierarchy, and at the same time, (s)he can get more information from the partial necessary and possible outranking relations. In order to use this methodology, (s)he provides the thresholds shown in Table 3.8.

Table 3.8: Indifference and preference thresholds provided by the Dean

Elementary subcriterion, $g_{\mathbf{t}}$	$q_{\mathbf{t},*}$	$q^*_{\mathbf{t}}$	$p_{\mathbf{t},*}$	$p^*_{\mathbf{t}}$
Group Theory	1	2	3	4
Linear Algebra	1	2	3	4
Calculus	1	2	3	4
Functions Theory	1	2	3	4
Analytical Chemistry I	1	2	3	4
App. Anal. Chemistry	1	2	3	4
Organic Chemistry I	1	2	3	4
Organic Chemistry II	1	2	3	4

Looking at the evaluations of students shown in Table 3.5, the Dean specifies the following pairwise comparisons:

- student s_4 outranks student s_2 with respect to Mathematics ($s_4 S_{(\mathbf{1})} s_2$),
- student s_5 does not outrank student s_1 with respect to Organic Chemistry ($not(s_5 S_{(\mathbf{2},\mathbf{2})} s_1)$).

These two pieces of information, are translated into the following constraints regarding variables of the ordinal regression problem:

- $s_4 S_{(\mathbf{1})} s_2$ is translated into:
    1. $\psi_{(\mathbf{1},\mathbf{1},\mathbf{1})}(s_4, s_1) + \psi_{(\mathbf{1},\mathbf{1},\mathbf{2})}(s_4, s_1) + \psi_{(\mathbf{1},\mathbf{2},\mathbf{1})}(s_4, s_1) + \psi_{(\mathbf{1},\mathbf{2},\mathbf{2})}(s_4, s_1) \geq \lambda_{(\mathbf{1},\mathbf{1})} + \lambda_{(\mathbf{1},\mathbf{2})}$,
    2. $v_{(\mathbf{1},\mathbf{1},\mathbf{1})} \geq g_{(\mathbf{1},\mathbf{1},\mathbf{1})}(s_2) - g_{(\mathbf{1},\mathbf{1},\mathbf{1})}(s_4) + \varepsilon = -10 + \varepsilon$,
    3. $v_{(\mathbf{1},\mathbf{1},\mathbf{2})} \geq g_{(\mathbf{1},\mathbf{1},\mathbf{2})}(s_2) - g_{(\mathbf{1},\mathbf{1},\mathbf{2})}(s_4) + \varepsilon = 3 + \varepsilon$,
    4. $v_{(\mathbf{1},\mathbf{2},\mathbf{1})} \geq g_{(\mathbf{1},\mathbf{2},\mathbf{1})}(s_2) - g_{(\mathbf{1},\mathbf{2},\mathbf{1})}(s_4) + \varepsilon = -10 + \varepsilon$,
    5. $v_{(\mathbf{1},\mathbf{2},\mathbf{2})} \geq g_{(\mathbf{1},\mathbf{2},\mathbf{2})}(s_2) - g_{(\mathbf{1},\mathbf{2},\mathbf{2})}(s_4) + \varepsilon = 3 + \varepsilon$.
- $not(s_5 S_{(\mathbf{2},\mathbf{2})} s_1)$ is translated into:
    1. $\psi_{(\mathbf{2},\mathbf{2},\mathbf{1})}(s_5, s_1) + \psi_{(\mathbf{2},\mathbf{2},\mathbf{2})}(s_5, s_1) + \varepsilon \leq \lambda_{(\mathbf{2},\mathbf{2})} + M_0^{(\mathbf{2},\mathbf{2})}(s_5, s_1)$,
    2. $v_{(\mathbf{2},\mathbf{2},\mathbf{1})} - \delta M_{(\mathbf{2},\mathbf{2},\mathbf{1})}(5, 1) \leq g_{(\mathbf{2},\mathbf{2},\mathbf{1})}(s_1) - g_{(\mathbf{2},\mathbf{2},\mathbf{1})}(s_5) = 9$,
    3. $v_{(\mathbf{2},\mathbf{2},\mathbf{2})} - \delta M_{(\mathbf{2},\mathbf{2},\mathbf{2})}(5, 1) \leq g_{(\mathbf{2},\mathbf{2},\mathbf{2})}(s_1) - g_{(\mathbf{2},\mathbf{2},\mathbf{2})}(s_5) = -2$,
    4. $M_0^{(\mathbf{2},\mathbf{2})}(s_5, s_1) + M_{(\mathbf{2},\mathbf{2},\mathbf{1})}(s_5, s_1) + M_{(\mathbf{2},\mathbf{2},\mathbf{2})}(s_5, s_1) \leq 2$,
    5. $M_0^{(\mathbf{2},\mathbf{2})}(s_5, s_1), M_{(\mathbf{2},\mathbf{2},\mathbf{1})}(s_5, s_1), M_{(\mathbf{2},\mathbf{2},\mathbf{2})}(s_5, s_1) \in \{0, 1\}$.

The necessary outranking relation resulting from application of all sets of preference model parameters compatible with the given preference information on the set of five students is presented in Table 3.9.

Looking at Table 3.9, we can observe that with respect to the totality of criteria, the only information the Dean obtains is that student s_1 necessarily outranks student s_2 and student s_4 necessarily outranks student s_1. But, when looking at the subcriteria of the hierarchy, the Dean could observe some facts which cannot be seen when using the classic ELECTRE^{GKMS} designed for a flat structure of criteria. According to Proposition 3.2.3, from the necessary outranking of student s_4 over student s_1 with respect to Mathematics and Chemistry, follows the necessary outranking of student s_4 over student s_1 with respect to the totality of criteria, but at the same time, while student s_4 necessarily outranks student s_1 with respect to Mathematics, student s_4 does not necessarily outrank student s_1 with respect to Algebra being a subcriterion of Mathematics at the level immediately below.

Table 3.9: Necessary outranking relation obtained from application of the hierarchical version of ELECTRE^{GKMS} (1 means true, and 0 means false)

$S^N_{(\mathbf{0})}$	s_1	s_2	s_3	s_4	s_5
s_1	1	1	0	0	0
s_2	0	1	0	0	0
s_3	0	0	1	0	0
s_4	1	0	0	1	0
s_5	0	0	0	0	1

$S^N_{(\mathbf{1})}$	s_1	s_2	s_3	s_4	s_5
s_1	1	1	1	0	1
s_2	0	1	0	0	1
s_3	1	1	1	1	1
s_4	1	1	1	1	0
s_5	0	0	0	0	1

$S^N_{(\mathbf{2})}$	s_1	s_2	s_3	s_4	s_5
s_1	1	1	0	1	0
s_2	1	1	0	0	0
s_3	0	0	1	0	0
s_4	1	0	0	1	1
s_5	0	0	1	0	1

$S^N_{(\mathbf{1,1})}$	s_1	s_2	s_3	s_4	s_5
s_1	1	1	1	0	0
s_2	0	1	0	0	1
s_3	1	0	1	1	0
s_4	0	0	1	1	0
s_5	0	0	0	0	1

$S^N_{(\mathbf{1,2})}$	s_1	s_2	s_3	s_4	s_5
s_1	1	1	1	0	0
s_2	0	1	0	0	1
s_3	1	0	1	1	0
s_4	0	0	1	1	0
s_5	0	1	0	0	1

$S^N_{(\mathbf{2,1})}$	s_1	s_2	s_3	s_4	s_5
s_1	1	1	0	1	0
s_2	1	1	0	0	0
s_3	0	0	1	0	0
s_4	1	0	0	1	1
s_5	0	0	1	0	1

$S^N_{(\mathbf{2,2})}$	s_1	s_2	s_3	s_4	s_5
s_1	1	1	0	1	0
s_2	1	1	0	0	0
s_3	0	0	1	0	1
s_4	1	0	0	1	1
s_5	0	0	1	0	1

3.2.4 Handling the hierarchy of criteria in PROMETHEE methods

In this section, we describe the extension of another outranking method, called PROMETHEE, to the hierarchy of criteria (for a detailed description of PROMETHEE methods in case of a flat structure of criteria see [16]).

In the case of the hierarchy of criteria, PROMETHEE methods compare couples of alternatives with respect to criteria and subcriteria of the hierarchical family of criteria in order to construct an outranking relation in the set of alternatives. This construction involves a few parameters, that is, the weights of elementary subcriteria, as well as indifference and preference thresholds for differences of evaluations of couples of alternatives on each elementary subcriterion. The preference of the DM regarding a couple of alternatives (a, b) with respect to elementary subcriterion $g_{\mathbf{t}}$ depends on the difference between $g_{\mathbf{t}}(a)$ and $g_{\mathbf{t}}(b)$ and for this reason the preference of a over b can be represented by a function $P_{\mathbf{t}}(a, b)$, increasing with $d_{\mathbf{t}}(a, b) = g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b)$. In [16], there are given six different types of functions $P_{\mathbf{t}}(a, b)$, and each one of them involves from zero to three parameters. Let us suppose, there are m evaluation criteria and n alternatives in set A. After the DM has decided which function $P_{\mathbf{t}}$ is expressing the best her/his preferences with respect to elementary subcriterion $g_{\mathbf{t}}$, and after introducing the weights $k_{\mathbf{t}}$ for each elementary subcriterion $g_{\mathbf{t}}, \mathbf{t} \in EL$, one can calculate for each couple of alternatives (a, b) and for each criterion $G_{\mathbf{r}}$, $\mathbf{r} \in \mathcal{I}_{\mathcal{G}}$, the following indices:

- the partial aggregate preference indices:

$$\pi_{\mathbf{r}}(a,b) = \begin{cases} k_{\mathbf{r}} P_{\mathbf{r}}(a,b) & \text{if } \mathbf{r} \in EL, \\ \displaystyle\sum_{\mathbf{t} \in E(G_{\mathbf{r}})} k_{\mathbf{t}} P_{\mathbf{t}}(a,b) & \text{otherwise,} \end{cases}$$

representing, the degree of preference of a over b, with respect to criterion/subcriterion $G_{\mathbf{r}}$;

- the partial positive, negative, and net outranking flows:

$$\Phi^+_{\mathbf{r}}(a) = \frac{1}{n-1} \sum_{x \in A \setminus \{a\}} \pi_{\mathbf{r}}(a,x), \quad \Phi^-_{\mathbf{r}}(a) = \frac{1}{n-1} \sum_{x \in A \setminus \{a\}} \pi_{\mathbf{r}}(x,a), \quad \Phi_{\mathbf{r}}(a) = \Phi^+_{\mathbf{r}}(a) - \Phi^-_{\mathbf{r}}(a)$$

representing respectively how strongly alternative a outranks all other alternatives of A on $G_{\mathbf{r}}$, how strongly alternatives of A outrank a on $G_{\mathbf{r}}$, and a balance between the two previous flows.

In this case, we can build preference $P^I_{\mathbf{r}}$, indifference $I^I_{\mathbf{r}}$ and incomparability $R^I_{\mathbf{r}}$ relations of PROMETHEE I as follows:

$$\begin{cases} aP^I_{\mathbf{r}}b & \text{iff} \quad \Phi^+_{\mathbf{r}}(a) \geq \Phi^+_{\mathbf{r}}(b), \;\; \Phi^-_{\mathbf{r}}(a) \leq \Phi^-_{\mathbf{r}}(b), \text{ and at least one of the two inequalities is strict,} \\ aI^I_{\mathbf{r}}b & \text{iff} \quad \Phi^+_{\mathbf{r}}(a) = \Phi^+_{\mathbf{r}}(b) \text{ and } \Phi^-_{\mathbf{r}}(a) = \Phi^-_{\mathbf{r}}(b), \\ aR^I_{\mathbf{r}}b & \text{otherwise} \end{cases}$$

Moreover, preference $\left(P^{II}_{\mathbf{r}}\right)$ and indifference $\left(I^{II}_{\mathbf{r}}\right)$ relations of PROMETHEE II can be defined as follows:

$$aP^{II}_{\mathbf{r}}b \text{ iff } \Phi_{\mathbf{r}}(a) > \Phi_{\mathbf{r}}(b), \quad \text{while} \quad aI^{II}_{\mathbf{r}}b \text{ iff } \Phi_{\mathbf{r}}(a) = \Phi_{\mathbf{r}}(b).$$

Note 3.2.6. *Remark that in case* $\mathbf{r} = 0$*, we obtain the indices and relations of the classical PROMETHEE methods for a flat structure of criteria.*

In case of the hierarchy of criteria, we can prove the following Propositions:

Proposition 3.2.4. *For each* $a, b \in A$*, and for each* $G_{\mathbf{r}} \in \mathcal{G}$*,* $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$*, we have:*

1. $\pi_{\mathbf{r}}(a,b) = \displaystyle\sum_{j=1}^{n(\mathbf{r})} \pi_{(\mathbf{r},j)}(a,b)$

2. $\Phi^+_{\mathbf{r}}(a) = \displaystyle\sum_{j=1}^{n(\mathbf{r})} \Phi^+_{(\mathbf{r},j)}(a)$

3. $\Phi^-_{\mathbf{r}}(a) = \displaystyle\sum_{j=1}^{n(\mathbf{r})} \Phi^-_{(\mathbf{r},j)}(a)$

4. $\Phi_{\mathbf{r}}(a) = \sum_{j=1}^{n(\mathbf{r})} \Phi_{(\mathbf{r},j)}(a)$

Proof. See Appendix A. □

Proposition 3.2.5.

1. *Given two alternatives* $a, b \in A$ *and* $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$*, such that*

$$aP^I_{(\mathbf{r},j)}b \quad \textit{for all } j = 1, \ldots, n(\mathbf{r}),$$

then $aP^I_{\mathbf{r}}b$*,*

2. *Given two alternatives* $a, b \in A$ *and* $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$*, such that:*

   α) $\{C_1, C_2\}$ *is a partition of the set* $\{1, \ldots, n(\mathbf{r})\}$ *of indices of subcriteria of* $G_{\mathbf{r}}$ *in the subsequent level,*

   β) $aP^I_{(\mathbf{r},j)}b$*, for all* $j \in C_1$*,*

   γ) $aI^I_{(\mathbf{r},j)}b$*, for all* $j \in C_2$*,*

   then $aP^I_{\mathbf{r}}b$*,*

3. *Given two alternatives* $a, b \in A$ *and* $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$*, such that*

$$aI^I_{(\mathbf{r},j)}b \quad \textit{for all } j = 1, \ldots, n(\mathbf{r}),$$

then $aI^I_{\mathbf{r}}b$*,*

Proof. See Appendix A. □

Proposition 3.2.6.

1. *Given two alternatives* $a, b \in A$ *and* $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$*, such that*

$$aP^{II}_{(\mathbf{r},j)}b \quad \textit{for all } j = 1, \ldots, n(\mathbf{r}),$$

then $aP^{II}_{\mathbf{r}}b$*,*

2. *Given two alternatives* $a, b \in A$ *and* $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$*, such that:*

   α) $\{C_1, C_2\}$ *is a partition of the set* $\{1, \ldots, n(\mathbf{r})\}$ *of indices of subcriteria of* $G_{\mathbf{r}}$ *in the subsequent level,*

   β) $aP^{II}_{(\mathbf{r},j)}b$*, for all* $j \in C_1$*,*

   γ) $aI^{II}_{(\mathbf{r},j)}b$*, for all* $j \in C_2$*,*

   then $aP^{II}_{\mathbf{r}}b$*,*

3. *Given two alternatives* $a, b \in A$ *and* $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$*, such that*

$$aI^{II}_{(\mathbf{r},j)}b \quad \textit{for all } j = 1, \ldots, n(\mathbf{r}),$$

then $aI^{II}_{\mathbf{r}}b$*,*

Proof. See Appendix A. □

3.2.5 Hierarchical PROMETHEE^{GKS}

In this section we extend the principles of PROMETHEE^{GKS} to the case of the hierarchy of criteria. As stated already above, the only information the DM can obtain from the evaluation matrix is the dominance relation in the set of alternatives. In general, this information is very poor and leaves many alternatives incomparable. To enrich this information, the DM has to introduce some preference information which reveals her/his value system. In this context, we take into account both PROMETHEE I and PROMETHEE II methods, noting that the new Hierarchical PROMETHEE method, so obtained, contains PROMETHEE^{GKS} [76], as a particular case.
Given a subset A^R of A, whose elements are called reference alternatives, and a criterion $G_{\mathbf{r}}, \mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, we suppose that the DM can give two types of preference information regarding $a, b \in A^R$ (we consider $B^R = A^R \times A^R$):

- local relations (denoted by $a \succsim_{\pi_{\mathbf{r}}} b$, $a \succ_{\pi_{\mathbf{r}}} b$, and $a \sim_{\pi_{\mathbf{r}}} b$), comparing directly the performance of a and b on criterion $G_{\mathbf{r}}$, and these comparisons are translated into constraints regarding $\pi_{\mathbf{r}}(a, b)$ and $\pi_{\mathbf{r}}(b, a)$,

- global relations (denoted by $a \succsim_{\Phi_{\mathbf{r}}} b$, $a \succ_{\Phi_{\mathbf{r}}} b$, and $a \sim_{\Phi_{\mathbf{r}}} b$), comparing a and b to all other alternatives, taking into account their outranking flows, $\Phi^+_{\mathbf{r}}(a)$, $\Phi^+_{\mathbf{r}}(b)$, $\Phi^-_{\mathbf{r}}(a)$ and $\Phi^-_{\mathbf{r}}(b)$, in case of PROMETHEE I or $\Phi_{\mathbf{r}}(a)$ and $\Phi_{\mathbf{r}}(b)$ in case of PROMETHEE II.

As in the Hierarchical ELECTRE^{GKMS} method, we assume moreover that the DM can give for each elementary subcriterion information regarding indifference and preference thresholds directly, that is provide intervals of possible values, or indirectly, that is provide information on some couples of alternatives (s)he considers indifferent or not (EL_1 and EL_2 represent the sets of criteria for which the DM gives information on the indifference and preference thresholds in a direct or indirect way, respectively); besides, analogously to Hierarchical ELECTRE^{GKMS}, we assume that the DM could provide some information regarding the weights of some elementary subcriterion (for a more detailed description of these preference information and for the consistency constraints on the indifference and preference thresholds see Appendix B).

Given this preference information, a compatible outranking model is a set of preference indices $\pi_{\mathbf{t}}(a,b)$, $(a,b) \in B$, $\mathbf{t} \in EL$, restoring the preference information provided by the DM and satisfying so the following set of constraints (see [76] for a similar formulation in a non-hierarchical case and Appendix B for a detailed description of these constraints):

Pairwise comparisons (local relations), for $(a,b) \in B^R$:

$\pi_{\mathbf{r}}(a,b) \geq \pi_{\mathbf{r}}(b,a)$ if $a \succsim_{\pi_{\mathbf{r}}} b$,

$\pi_{\mathbf{r}}(a,b) \geq \pi_{\mathbf{r}}(b,a) + \varepsilon$ if $a \succ_{\pi_{\mathbf{r}}} b$,

$\pi_{\mathbf{r}}(a,b) = \pi_{\mathbf{r}}(b,a)$ if $a \sim_{\pi_{\mathbf{r}}} b$,

Pairwise comparisons (global relations), if the outranking model is exploited in the way of PROMETHEE II, for $(a,b) \in B^R$:

$\Phi_{\mathbf{r}}(a) \geq \Phi_{\mathbf{r}}(b)$ if $a \succsim_{\Phi_{\mathbf{r}}} b$,

$\Phi_{\mathbf{r}}(a) \geq \Phi_{\mathbf{r}}(b) + \varepsilon$ if $a \succ_{\Phi_{\mathbf{r}}} b$,

$\Phi_{\mathbf{r}}(a) = \Phi_{\mathbf{r}}(b)$ if $a \sim_{\Phi_{\mathbf{r}}} b$,

Pairwise comparisons (global relations), if the outranking model is exploited in the way of PROMETHEE I:

$\Phi^+_{\mathbf{r}}(a) \geq \Phi^+_{\mathbf{r}}(b)$ and $\Phi^-_{\mathbf{r}}(a) \leq \Phi^-_{\mathbf{r}}(b)$ if $a \succsim_{\Phi_{\mathbf{r}}} b$, for $(a,b) \in B^R$,

$$\left.\begin{array}{l} \Phi^+_{\mathbf{r}}(a) \geq \Phi^+_{\mathbf{r}}(b) \text{ and } \Phi^-_{\mathbf{r}}(a) \leq \Phi^-_{\mathbf{r}}(b) \text{ and} \\ \Phi^+_{\mathbf{r}}(a) - \Phi^-_{\mathbf{r}}(a) \geq \Phi^+_{\mathbf{r}}(b) - \Phi^-_{\mathbf{r}}(b) + \varepsilon \end{array}\right\} \text{ if } a \succ_{\Phi_{\mathbf{r}}} b, \text{ for } (a,b) \in B^R,$$

$\Phi^+_{\mathbf{r}}(a) = \Phi^+_{\mathbf{r}}(b)$ and $\Phi^-_{\mathbf{r}}(a) = \Phi^-_{\mathbf{r}}(b)$ if $a \sim_{\Phi_{\mathbf{r}}} b$, for $(a,b) \in B^R$.

Values of inter-criteria parameters:

$$\sum_{\mathbf{t} \in EL} \pi_{\mathbf{t}}(x^*_{\mathbf{t}}, x_{\mathbf{t},*}) = 1, \text{ where } x_{\mathbf{t},*}, x^*_{\mathbf{t}} \in A \text{ for all } \mathbf{t} \in EL : g_{\mathbf{t}}(x^*_{\mathbf{t}}) = \max_{a \in A} g_{\mathbf{t}}(a), \text{ and } g_{\mathbf{t}}(x_{\mathbf{t},*}) = \min_{a \in A} g_{\mathbf{t}}(a),$$

Values of marginal preference indices conditioned by intra-criterion preference information, for all $(a,b) \in B$:

$k_{\mathbf{t},*} \leq \pi_{\mathbf{t}}(x^*_{\mathbf{t}}, x_{\mathbf{t},*}) \leq k^*_{\mathbf{t}}, \mathbf{t} \in EL$,

$\pi_{\mathbf{t_1}}(x^*_{\mathbf{t_1}}, x_{\mathbf{t_1},*}) \geq \pi_{\mathbf{t_2}}(x^*_{\mathbf{t_2}}, x_{\mathbf{t_2},*}) + \varepsilon$, if elementary subcriterion $g_{\mathbf{t_1}}$ is more important than elementary subcriterion $g_{\mathbf{t_2}}, \mathbf{t_1}, \mathbf{t_2} \in EL$,

$\pi_{\mathbf{t_1}}(x^*_{\mathbf{t_1}}, x_{\mathbf{t_1},*}) = \pi_{\mathbf{t_2}}(x^*_{\mathbf{t_2}}, x_{\mathbf{t_2},*})$, if elementary subcriteria $g_{\mathbf{t_1}}$ and $g_{\mathbf{t_2}}$ are equally important, $\mathbf{t_1}, \mathbf{t_2} \in EL$,

$\pi_{\mathbf{t}}(a,b) = 0$ if $g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b) \leq q_{\mathbf{t},*}, \mathbf{t} \in EL_1$,

$\pi_{\mathbf{t}}(a,b) \geq \varepsilon$ if $g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b) > q^*_{\mathbf{t}}, \mathbf{t} \in EL_1$,

$\pi_{\mathbf{t}}(a,b) + \varepsilon \leq \pi_{\mathbf{t}}(x^*_{\mathbf{t}}, x_{\mathbf{t},*})$ if $g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b) < p_{\mathbf{t},*}, \mathbf{t} \in EL_1$,

$\pi_{\mathbf{t}}(a,b) = \pi_{\mathbf{t}}(x^*_{\mathbf{t}}, x_{\mathbf{t},*})$ if $g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b) \geq p^*_{\mathbf{t}}, \mathbf{t} \in EL_1$,

$\pi_{\mathbf{t}}(a,b) = 0,\ \pi_{\mathbf{t}}(b,a) = 0$ if $a \sim_{\mathbf{t}} b,\ \mathbf{t} \in EL_2$,

$\pi_{\mathbf{t}}(a,b) = \pi_{\mathbf{t}}(x^*_{\mathbf{t}}, x_{\mathbf{t},*})$ if $a \succ_{\mathbf{t}} b,\ \mathbf{t} \in EL_2$.

Monotonicity of the functions of marginal preference indices, for all $a,b,c,d \in A, \mathbf{t} \in EL$:

$\pi_{\mathbf{t}}(a,b) \geq \pi_{\mathbf{t}}(c,d)$ if $g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b) > g_{\mathbf{t}}(c) - g_{\mathbf{t}}(d)$,

$\pi_{\mathbf{t}}(a,b) = \pi_{\mathbf{t}}(c,d)$ if $g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b) = g_{\mathbf{t}}(c) - g_{\mathbf{t}}(d)$.

$\Bigg\} E^{A^R}$

If E^{A^R} is feasible and $\varepsilon^* = max\ \varepsilon$, subject to E^{A^R}, is greater than 0, then there exists at least one outranking model compatible with the preference information.

Given a criterion/subcriterion $G_{\mathbf{r}}$, $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, and two alternatives $a, b \in A$, we can give the following definitions:

Definition 3.2.2.

- *Given a compatible outranking model S exploited in the way of PROMETHEE I, we say a outranks b with respect to $G_{\mathbf{r}}$, and we write $a \succsim_{\mathbf{r}} b$, if:*

$$\Phi^{+}_{\mathbf{r}}(a) \geq \Phi^{+}_{\mathbf{r}}(b) \ \ and \ \ \Phi^{-}_{\mathbf{r}}(a) \leq \Phi^{-}_{\mathbf{r}}(b).$$

- *Given a compatible outranking model S exploited in the way of PROMETHEE II, we say that a outranks b with respect to $G_{\mathbf{r}}$, and we write $a \succsim_{\mathbf{r}} b$, if:*

$$\Phi_{\mathbf{r}}(a) \geq \Phi_{\mathbf{r}}(b).$$

Note 3.2.7. *Remark that given two alternatives $a, b \in A$, for each $G_{\mathbf{r}} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, and for each compatible outranking model S, if a outranks b with respect to criterion/subcriterion $G_{\mathbf{r}}$ in the sense of PROMETHEE I, then a outranks b with respect to criterion/subcriterion $G_{\mathbf{r}}$ in the sense of PROMETHEE II.*

In the ROR context, considering a criterion/subcriterion $G_{\mathbf{r}}$, $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, and two alternatives $a, b \in A$, we can give the following definitions:

Definition 3.2.3.

- *a necessarily outranks b with respect to $G_{\mathbf{r}}$, and we write $a \succsim^{N}_{\mathbf{r}} b$, if a outranks b with respect to $G_{\mathbf{r}}$, for all compatible outranking models,*
- *a possibly outranks b with respect to $G_{\mathbf{r}}$, and we write $a \succsim^{P}_{\mathbf{r}} b$, if a outranks b with respect to $G_{\mathbf{r}}$, for at least one compatible outranking model.*

Given a pair of alternatives $(a, b) \in B$, and a criterion/subcriterion $G_{\mathbf{r}}, \mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, necessary and possible outranking relations $(\succsim^{N}_{\mathbf{r}}, \succsim^{P}_{\mathbf{r}})$ can be computed as follows:

- To check whether $a \succsim^{N}_{\mathbf{r}} b$, we assume that a does not outrank b with respect to $G_{\mathbf{r}}$ ($not(a \succsim_{\mathbf{r}} b)$), and we add the corresponding constraints to set E^{A^R} shown below. Then, we verify whether $not(a \succsim_{\mathbf{r}} b)$ is possible in the set of all outranking models compatible with the previously provided preference information.

$$
\left.
\begin{array}{l}
E^{A^R} \\
\text{if one verifies the truth of global outranking:} \\
\quad \text{if exploited in the way of PROMETHEE II, then:} \\
\quad\quad \Phi_{\mathbf{r}}(a) + \varepsilon \leq \Phi_{\mathbf{r}}(b) \\
\quad \text{if exploited in the way of PROMETHEE I, then:} \\
\quad\quad \Phi^+_{\mathbf{r}}(a) + \varepsilon \leq \Phi^+_{\mathbf{r}}(b) + 2M_1^{\mathbf{r}} \;\; \text{and} \;\; \Phi^-_{\mathbf{r}}(a) + 2M_2^{\mathbf{r}} \geq \Phi^-_{\mathbf{r}}(b) + \varepsilon \\
\quad\quad\quad \text{where } M_i^{\mathbf{r}} \in \{0,1\}, \; i = 1,2, \; \text{and } \sum_{i=1}^{2} M_i^{\mathbf{r}} \leq 1 \\
\text{if one verifies the truth of local outranking:} \\
\quad\quad \pi_{\mathbf{r}}(a,b) + \varepsilon \leq \pi_{\mathbf{r}}(b,a)
\end{array}
\right\} E^N_{\mathbf{r}}(a,b)
$$

We say that:

$$
a \succsim^N b \;\text{ if }\; E^N_{\mathbf{r}}(a,b) \;\text{ is infeasible or }\; \varepsilon^N_{\mathbf{r}}(a,b) \leq 0, \;\; \text{where} \;\; \varepsilon^N_{\mathbf{r}}(a,b) = max\, \varepsilon, \;\; \text{subject to} \;\; E^N_{\mathbf{r}}(a,b).
$$

Observe that in $E^N_{\mathbf{r}}(a,b)$, the binary variables $M^0_{\mathbf{r}}$ and $M^1_{\mathbf{r}}$ are used in order to deny the outranking of a over b. In fact, a does not outrank b if $\Phi^+_{\mathbf{r}}(a) < \Phi^+_{\mathbf{r}}(b)$ or $\Phi^-_{\mathbf{r}}(a) > \Phi^-_{\mathbf{r}}(b)$. If $M^i_{\mathbf{r}} = 0, i = 0, 1$, then the corresponding constraint opposes a veto to the outranking of a over b (in particular, if $M^0_{\mathbf{r}} = 0$ then $\Phi^+_{\mathbf{r}}(a) < \Phi^+_{\mathbf{r}}(b)$ while if $M^1_{\mathbf{r}} = 0$ then $\Phi^-_{\mathbf{r}}(a) > \Phi^-_{\mathbf{r}}(b)$); instead, if $M^i_{\mathbf{r}} = 1, i = 0, 1$, the corresponding constraint is always verified reminding that $\Phi^+_{\mathbf{r}}(a) \in [0,1]$ and $\Phi^-_{\mathbf{r}}(a) \in [0,1]$ for all $a \in A$. Besides, the constraint $\sum_{i=0}^{1} M^i_{\mathbf{r}} \leq 1$ ensures that at least one of the two variables has to be equal to zero.

- To check whether $a \succsim^P_{\mathbf{r}} b$, we assume that a outranks b with respect to $G_{\mathbf{r}}$ ($a \succsim_{\mathbf{r}} b$), and add corresponding constraints to the set E^{A^R} shown below. Then, we verify whether $a \succsim_{\mathbf{r}} b$ is possible in the set of all compatible outranking models.

$$\left.\begin{array}{l}
E^{A^R} \\
\text{if one verifies the truth of global outranking:} \\
\quad \text{if exploited in the way of PROMETHEE II, then:} \\
\qquad \Phi_{\mathbf{r}}(a) \geq \Phi_{\mathbf{r}}(b) \\
\quad \text{if exploited in the way of PROMETHEE I, then:} \\
\qquad \Phi^+_{\mathbf{r}}(a) \geq \Phi^+_{\mathbf{r}}(b) \;\; \text{and} \;\; \Phi^-_{\mathbf{r}}(a) \leq \Phi^-_{\mathbf{r}}(b) \\
\text{if one verifies the truth of local outranking:} \\
\qquad \pi_{\mathbf{r}}(a,b) \geq \pi_{\mathbf{r}}(b,a)
\end{array}\right\} E^P_{\mathbf{r}}(a,b)$$

We say that:

$a \succsim^P_{\mathbf{r}} b$ if $E^P_{\mathbf{r}}(a,b)$ is feasible and $\varepsilon^P_{\mathbf{r}}(a,b) > 0$, where $\varepsilon^P_{\mathbf{r}}(a,b) = max\ \varepsilon$, subject to $E^P_{\mathbf{r}}(a,b)$.

Note 3.2.8. *The same observation made for ELECTRE about application of linear programming within ROR is valid for PROMETHEE. More precisely, if the DM is able to give the marginal function $P_{\mathbf{t}}(a,b)$ and the related thresholds, the ROR optimization problems can be formulated in terms of linear programming, taking into account as variables the weights $k_{\mathbf{t}}, \mathbf{t} \in EL$, only. This amounts to substitute the set of constraints E^{A^R} with the following:*

$$
\left.
\begin{array}{l}
\textit{Pairwise comparisons (local relations), for } (a,b) \in B^R\text{:} \\
\quad \pi_{\mathbf{r}}(a,b) \geq \pi_{\mathbf{r}}(b,a) \ \textit{if } a \succsim_{\pi_{\mathbf{r}}} b, \\
\quad \pi_{\mathbf{r}}(a,b) \geq \pi_{\mathbf{r}}(b,a) + \varepsilon \ \textit{if } a \succ_{\pi_{\mathbf{r}}} b, \\
\quad \pi_{\mathbf{r}}(a,b) = \pi_{\mathbf{r}}(b,a) \ \textit{if } a \sim_{\pi_{\mathbf{r}}} b, \\
\textit{Pairwise comparisons (global relations), if the outranking model is exploited in the way} \\
\textit{of PROMETHEE II, for } (a,b) \in B^R\text{:} \\
\quad \Phi_{\mathbf{r}}(a) \geq \Phi_{\mathbf{r}}(b) \ \textit{if } a \succsim_{\Phi_{\mathbf{r}}} b, \\
\quad \Phi_{\mathbf{r}}(a) \geq \Phi_{\mathbf{r}}(b) + \varepsilon \ \textit{if } a \succ_{\Phi_{\mathbf{r}}} b, \\
\quad \Phi_{\mathbf{r}}(a) = \Phi_{\mathbf{r}}(b) \ \textit{if } a \sim_{\Phi_{\mathbf{r}}} b, \\
\textit{Pairwise comparisons (global relations), if the outranking model is exploited in the way} \\
\textit{of PROMETHEE I:} \\
\quad \Phi^+_{\mathbf{r}}(a) \geq \Phi^+_{\mathbf{r}}(b) \quad \textit{and} \quad \Phi^-_{\mathbf{r}}(a) \leq \Phi^-_{\mathbf{r}}(b) \ \textit{if } a \succsim_{\Phi_{\mathbf{r}}} b, \ \textit{for } (a,b) \in B^R, \\
\quad \left. \begin{array}{l} \Phi^+_{\mathbf{r}}(a) \geq \Phi^+_{\mathbf{r}}(b) \quad \textit{and} \quad \Phi^-_{\mathbf{r}}(a) \leq \Phi^-_{\mathbf{r}}(b) \quad \textit{and} \\ \Phi^+_{\mathbf{r}}(a) - \Phi^-_{\mathbf{r}}(a) \geq \Phi^+_{\mathbf{r}}(b) - \Phi^-_{\mathbf{r}}(b) + \varepsilon \end{array} \right\} \ \textit{if } a \succ_{\Phi_{\mathbf{r}}} b, \ \textit{for } (a,b) \in B^R, \\
\quad \Phi^+_{\mathbf{r}}(a) = \Phi^+_{\mathbf{r}}(b) \quad \textit{and} \quad \Phi^-_{\mathbf{r}}(a) = \Phi^-_{\mathbf{r}}(b) \ \textit{if } a \sim_{\Phi_{\mathbf{r}}} b, \ \textit{for } (a,b) \in B^R. \\
\textit{Values of marginal preference indices conditioned by intra-criterion preference information,} \\
\textit{for all } (a,b) \in B\text{:} \\
\quad k_{\mathbf{t},*} \leq k_{\mathbf{t}} \leq k^*_{\mathbf{t}}, \mathbf{t} \in EL, \\
\quad k_{\mathbf{t_1}} \geq k_{\mathbf{t_2}} + \varepsilon, \quad \textit{if elementary subcriterion } g_{\mathbf{t_1}} \textit{ is more important than} \\
\qquad \textit{elementary subcriterion } g_{\mathbf{t_2}}, \mathbf{t_1}, \mathbf{t_2} \in EL, \\
\quad k_{\mathbf{t_1}} = k_{\mathbf{t_2}}, \quad \textit{if elementary subcriteria } g_{\mathbf{t_1}} \textit{ and } g_{\mathbf{t_2}} \\
\qquad \textit{are equally important}, \mathbf{t_1}, \mathbf{t_2} \in EL,
\end{array}
\right\} E^{A^R}
$$

Properties of necessary and possible outranking relations in hierarchical PROMETHEE^{GKS}

Proposition 3.2.7.

1. *For all* $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, $\succsim^N_{\mathbf{r}} \subseteq \succsim^P_{\mathbf{r}}$,

2. *For all* $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, $\succsim^P_{\mathbf{r}}$ *and* $\succsim^N_{\mathbf{r}}$ *are reflexive,*

Proof. See Appendix A. □

Proposition 3.2.8.

1. *Given two alternatives* $a, b \in A$ *and* $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus (EL \cup LBO)$, *such that:*

$$a \succsim^N_{(\mathbf{r},j)} b \quad \textit{for all } j = 1, \ldots, n(\mathbf{r}),$$

then $a \succsim^N_{\mathbf{r}} b$.

2. *Given two alternatives* $a, b \in A$ *and* $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus (EL \cup LBO)$*, such that:*

   $\alpha)$ $a \succsim^{N}_{(\mathbf{r},j)} b$ *for all* $j = 1, \ldots, n(\mathbf{r}), j \neq w$,

   $\beta)$ $a \succsim^{P}_{(\mathbf{r},w)} b$,

   then $a \succsim^{P}_{\mathbf{r}} b$.

Proof. See Appendix A. □

An illustrative example

In this subsection we consider the same problem we have dealt with Hierarchical ELECTRE using both PROMETHEE and PROMETHEE^{GKS} methods extended to the case of a hierarchical family of criteria.

At first, we suppose to have the same weights, as well as the same indifference and preference thresholds as before: let us also choose for each elementary subcriterion $g_{\mathbf{t}}$, $\mathbf{t} \in EL$, the following preference function $P_{\mathbf{t}}(a,b)$, for any $a, b \in A$:

$$P_{\mathbf{t}}(a,b) = \begin{cases} 0 & \text{if } \; g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b) \leq q_{\mathbf{t}}, \\ \frac{g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b) - q_{\mathbf{t}}}{p_{\mathbf{t}} - q_{\mathbf{t}}} & \text{if } \; q_{\mathbf{t}} < g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b) < p_{\mathbf{t}}, \\ 1 & \text{if } \; g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b) \geq p_{\mathbf{t}}. \end{cases}$$

Table 3.10: Preference relations obtained using Hierarchical PROMETHEE I

Comprehensively	s_1	s_2	s_3	s_4	s_5
s_1	I	P	P	P	$\underline{P}$
s_2	P^{-1}	I	P	P^{-1}	P
s_3	P^{-1}	P^{-1}	I	P^{-1}	P
s_4	P^{-1}	P	P	I	P
s_5	P^{-1}	P	R	R	I

Maths	s_1	s_2	s_3	s_4	s_5
s_1	I	P	P	P	R
s_2	P^{-1}	I	P^{-1}	P^{-1}	R
s_3	P^{-1}	P	I	R	R
s_4	P^{-1}	P	R	I	R
s_5	R	R	R	R	I

Chemistry	s_1	s_2	s_3	s_4	s_5
s_1	I	P	P	P^{-1}	P
s_2	P^{-1}	I	P	P^{-1}	P
s_3	P^{-1}	P^{-1}	I	P^{-1}	P
s_4	P	P	P	I	P
s_5	P^{-1}	P^{-1}	P^{-1}	P^{-1}	I

Algebra	s_1	s_2	s_3	s_4	s_5
s_1	I	P	P	P	R
s_2	P^{-1}	I	P^{-1}	P^{-1}	R
s_3	P^{-1}	P	I	R	R
s_4	P^{-1}	P	R	I	R
s_5	P^{-1}	P^{-1}	P^{-1}	P^{-1}	I

Analysis	s_1	s_2	s_3	s_4	s_5
s_1	I	P	P	P	P
s_2	P^{-1}	I	P^{-1}	P^{-1}	P^{-1}
s_3	P^{-1}	P	I	R	R
s_4	P^{-1}	P	R	I	R
s_5	P^{-1}	P	R	R	I

Anal. Chem.	s_1	s_2	s_3	s_4	s_5
s_1	I	P	P	P^{-1}	P
s_2	P^{-1}	I	P	P^{-1}	P
s_3	P^{-1}	P^{-1}	I	P^{-1}	P
s_4	P	P	P	I	P
s_5	P^{-1}	P^{-1}	P^{-1}	P^{-1}	I

Org. Chem.	s_1	s_2	s_3	s_4	s_5
s_1	I	P	P	P^{-1}	P
s_2	P^{-1}	I	P	P^{-1}	P
s_3	P^{-1}	P^{-1}	I	P^{-1}	P
s_4	P	P	P	I	P
s_5	P^{-1}	P^{-1}	P^{-1}	P^{-1}	I

In Table 3.10 we present the preference relations that PROMETHEE I states for any level of the considered hierarchy of criteria. More precisely, considering $Matrix_{\mathbf{r}}$ to be one of the seven matrices presented in Table 3.10, we have:

$$Matrix_{\mathbf{r}}(s_i, s_j) = \begin{cases} P & \text{if} \quad s_i \text{ is preferred to } s_j \text{ with respect to criterion } G_{\mathbf{r}}, \\ I & \text{if} \quad s_i \text{ is indifferent to } s_j \text{ with respect to criterion } G_{\mathbf{r}}, \\ R & \text{if} \quad s_i \text{ is incomparable to } s_j \text{ with respect to criterion } G_{\mathbf{r}}, \\ P^{-1} & \text{if} \quad s_j \text{ is preferred to } s_i \text{ with respect to criterion } G_{\mathbf{r}} \end{cases}$$

In Table 3.10, $Matrix_{Comprehensively}(s_1, s_5) = P$ is underlined in order to evidence that there exists a couple of alternatives (s_i, s_j) such that s_i is preferred to s_j with respect to some criterion $G_{\mathbf{r}}$, but with respect to a subcriterion immediately descending from $G_{\mathbf{r}}$, say $G_{(\mathbf{r},\mathbf{w})}$, s_i is not preferred to s_j. In our example, s_1 is preferred to s_5 with respect to the totality of criteria, but s_1 is incomparable to s_5 with respect to Mathematics being a subcriterion of the totality of criteria. Note that the underlined couple (s_1, s_5) is not the only example of such a situation in Table 3.10.

In Table 3.11, we can see the ranking obtained by Hierarchical PROMETHEE II for each criterion/subcriterion of the hierarchy.

Table 3.11: Ranking of students at all levels of the hierarchy of criteria, obtained using Hierarchical PROMETHEE II

Position/subject	**Comprehensive**	**Maths**	**Algebra**	**Analysis**	**Chemistry**	**Analytical Chemistry**	**Organic Chemistry**
1	s_1 (0.2000)	s_1 (0.0938)	s_1 (0.0417)	s_1 (0.0521)	s_4 (0.1250)	s_4 (0.0500)	s_4 (0.0750)
2	s_4 (0.1062)	s_3 (0.0375)	s_3 (0.0125)	s_3 (0.0250)	s_1 (0.1063)	s_1 (0.0396)	s_1 (0.0667)
3	s_2 (-0.0167)	s_4 (-0.0187)	s_5 (-0.0083)	s_4 (-0.0021)	s_2 (0.0688)	s_2 (0.0188)	s_2 (0.0500)
4	s_3 (-0.0938)	s_5 (-0.0271)	s_4 (-0.0167)	s_5 (-0.0188)	s_3 (-0.1313)	s_3 (-0.0438)	s_3 (-0.0875)
5	s_5 (-0.1958)	s_2 (-0.0854)	s_2 (-0.0292)	s_2 (-0.0563)	s5 (-0.1688)	s_5 (-0.0646)	s_5 (-0.1042)

Now, let us suppose that the Dean decides to use the Hierarchical PROMETHEE^{GKS} providing some detailed outranking and non-outranking information with respect to all criteria considered together, and with respect to particular subcriteria. At the same time, (s)he wishes to obtain detailed information regarding the necessary and possible outranking relations. In order to use the methodology presented above, we suppose that the Dean can give information regarding indifference and preference thresholds on all elementary subcriteria, as shown in Table 3.12.

Let us first suppose, that the outranking relation is exploited in the way of PROMETHEE II, and that the Dean gives the following preference information:

- with respect to Mathematics, student s_4 is preferred to each other student more than student s_2 is preferred to each other student ($s_4 \succ_{\Phi_{\mathbf{1}}} s_2$),
- with respect to Organic Chemistry, student s_4 is preferred to each other student more than student s_3 is preferred to each other student ($s_4 \succ_{\Phi_{(\mathbf{2},\mathbf{2})}} s_3$).

Table 3.12: Indifference and preference thresholds provided by the DM

Elementary subcriterion, $g_{\mathbf{t}}$	$q_{\mathbf{t},*}$	$q^*_{\mathbf{t}}$	$p_{\mathbf{t},*}$	$p^*_{\mathbf{t}}$
Group Theory	1	2	4	5
Linear Algebra	1	2	4	5
Calculus	1	2	4	5
Functions Theory	1	2	4	5
Analytical Chemistry I	1	2	4	5
Applied Analytical Chemistry	1	2	4	5
Organic Chemistry I	1	2	4	5
Organic Chemistry II	1	2	4	5

These two pieces of information are translated into the following constraints regarding the variables of the ordinal regression problem:

- $s_4 \succ_{\Phi_{\mathbf{1}}} s_2$ is translated into:

$$\Phi_{\mathbf{1}}(s_4) = \sum_{x \in A \setminus \{s_4\}} \left\{ \sum_{\mathbf{t} \in E(G_{\mathbf{1}})} \frac{1}{n-1} \pi_{\mathbf{t}}(s_4, x) \right\} - \sum_{x \in A \setminus \{s_4\}} \left\{ \sum_{\mathbf{t} \in E(G_{\mathbf{1}})} \frac{1}{n-1} \pi_{\mathbf{t}}(x, s_4) \right\} \geq$$

$$\geq \Phi_{\mathbf{1}}(s_2) = \sum_{x \in A \setminus \{s_2\}} \left\{ \sum_{\mathbf{t} \in E(G_{\mathbf{1}})} \frac{1}{n-1} \pi_{\mathbf{t}}(s_2, x) \right\} - \sum_{x \in A \setminus \{s_2\}} \left\{ \sum_{\mathbf{t} \in E(G_{\mathbf{1}})} \frac{1}{n-1} \pi_{\mathbf{t}}(x, s_2) \right\} + \varepsilon$$

- $s_4 \succ_{\Phi_{(\mathbf{2},\mathbf{2})}} s_3$ is translated into:

$$\Phi_{(\mathbf{2},\mathbf{2})}(s_4) = \sum_{x \in A \setminus \{s_4\}} \left\{ \sum_{\mathbf{t} \in E(G_{(\mathbf{2},\mathbf{2})})} \frac{1}{n-1} \pi_{\mathbf{t}}(s_4, x) \right\} - \sum_{x \in A \setminus \{s_4\}} \left\{ \sum_{\mathbf{t} \in E(G_{(\mathbf{2},\mathbf{2})})} \frac{1}{n-1} \pi_{\mathbf{t}}(x, s_4) \right\} \geq$$

$$\geq \Phi_{(\mathbf{2},\mathbf{2})}(s_3) = \sum_{x \in A \setminus \{s_3\}} \left\{ \sum_{\mathbf{t} \in E(G_{(\mathbf{2},\mathbf{2})})} \frac{1}{n-1} \pi_{\mathbf{t}}(s_3, x) \right\} - \sum_{x \in A \setminus \{s_3\}} \left\{ \sum_{\mathbf{t} \in E(G_{(\mathbf{2},\mathbf{2})})} \frac{1}{n-1} \pi_{\mathbf{t}}(x, s_3) \right\} + \varepsilon$$

In Table 3.13, we show the necessary outranking relation with respect to some criteria and subcriteria of the hierarchy. For other subcriteria the necessary outranking relation is empty.

Table 3.13: Necessary outranking relations obtained using Hierarchical PROMETHEE^{GKS} and exploitation of the outranking relation in the way of PROMETHEE II

$\succsim^N_{(\mathbf{0})}$	s_1	s_2	s_3	s_4	s_5	$\succsim^N_{(\mathbf{1})}$	s_1	s_2	s_3	s_4	s_5	$\succsim^N_{(\mathbf{1},\mathbf{2})}$	s_1	s_2	s_3	s_4	s_5	$\succsim^N_{(\mathbf{2},\mathbf{2})}$	s_1	s_2	s_3	s_4	s_5
s_1	1	0	0	0	0	s_1	1	1	0	0	0	s_1	1	0	0	0	0	s_1	1	0	0	0	0
s_2	0	1	0	0	0	s_2	0	1	0	0	0	s_2	0	1	0	0	0	s_2	0	1	0	0	0
s_3	0	0	1	0	0	s_3	0	1	1	0	0	s_3	0	0	1	0	0	s_3	0	0	1	0	1
s_4	0	0	0	1	0	s_4	0	1	0	1	0	s_4	0	0	0	1	0	s_4	0	0	1	1	1
s_5	0	0	0	0	1	s_5	0	0	0	0	1	s_5	0	1	0	0	1	s_5	0	0	0	0	1

In the first matrix of Table 3.13, we observe that preference information provided by the DM does not imply any necessary outranking with respect to the totality of criteria. At the same time, we obtain partial information that cannot be obtained by PROMETHEE^{GKS} for a flat structure of the set of criteria; for example, we can see that students s_1, s_3 and s_4 are necessarily preferred to student s_2 with respect to Mathematics, so as student s_4 is necessarily preferred to student s_3 and s_5 with respect to Organic Chemistry, and so on.

Now, let us suppose that the outranking relation is exploited in the way of PROMETHHE I. We are considering the same preference information provided by the Dean. It is translated, however, in a different way than before:

- $s_4 \succ_{\Phi_{\mathbf{1}}} s_2$ is translated into constraints:

1. $\Phi^+_{(\mathbf{1})}(s_4) \geq \Phi^+_{(\mathbf{1})}(s_2) \Leftrightarrow$
$$\sum_{x \in A\setminus\{s_4\}} \left\{ \sum_{\mathbf{t} \in E(G_{\mathbf{1}})} \frac{1}{n-1} \pi_{\mathbf{t}}(s_4, x) \right\} \geq \sum_{x \in A\setminus\{s_2\}} \left\{ \sum_{\mathbf{t} \in E(G_{\mathbf{1}})} \frac{1}{n-1} \pi_{\mathbf{t}}(s_2, x) \right\},$$
2. $\Phi^-_{(\mathbf{1})}(s_4) \leq \Phi^-_{\mathbf{1}}(s_2) \Leftrightarrow$
$$\sum_{x \in A\setminus\{s_4\}} \left\{ \sum_{\mathbf{t} \in E(G_{\mathbf{1}})} \frac{1}{n-1} \pi_{\mathbf{t}}(x, s_4) \right\} \leq \sum_{x \in A\setminus\{s_2\}} \left\{ \sum_{\mathbf{t} \in E(G_{\mathbf{1}})} \frac{1}{n-1} \pi_{\mathbf{t}}(x, s_2) \right\},$$
3. $\Phi^+_{(\mathbf{1})}(s_4) - \Phi^-_{(\mathbf{1})}(s_4) \geq \Phi^+_{(\mathbf{1})}(s_2) - \Phi^-_{(\mathbf{1})}(s_2) + \varepsilon \Leftrightarrow$
$$\sum_{x \in A\setminus\{s_4\}} \left\{ \sum_{\mathbf{t} \in E(G_{\mathbf{1}})} \frac{1}{n-1} \pi_{\mathbf{t}}(s_4, x) \right\} - \sum_{x \in A\setminus\{s_4\}} \left\{ \sum_{\mathbf{t} \in E(G_{\mathbf{1}})} \frac{1}{n-1} \pi_{\mathbf{t}}(x, s_4) \right\} \geq$$
$$\geq \sum_{x \in A\setminus\{s_2\}} \left\{ \sum_{\mathbf{t} \in E(G_{\mathbf{1}})} \frac{1}{n-1} \pi_{\mathbf{t}}(s_2, x) \right\} - \sum_{x \in A\setminus\{s_2\}} \left\{ \sum_{\mathbf{t} \in E(G_{\mathbf{1}})} \frac{1}{n-1} \pi_{\mathbf{t}}(x, s_2) \right\} + \varepsilon.$$

- $s_4 \succ_{\Phi_{(\mathbf{2},\mathbf{2})}} s_3$ is translated into constraints:

1. $\Phi^+_{(\mathbf{2},\mathbf{2})}(s_4) \geq \Phi^+_{(\mathbf{2},\mathbf{2})}(s_3) \Leftrightarrow$
$$\sum_{x \in A\setminus\{s_4\}} \left\{ \sum_{\mathbf{t} \in E(G_{(\mathbf{2},\mathbf{2})})} \frac{1}{n-1} \pi_{\mathbf{t}}(s_4, x) \right\} \geq \sum_{x \in A\setminus\{s_3\}} \left\{ \sum_{\mathbf{t} \in E(G_{(\mathbf{2},\mathbf{2})})} \frac{1}{n-1} \pi_{\mathbf{t}}(s_3, x) \right\},$$
2. $\Phi^-_{(\mathbf{2},\mathbf{2})}(s_4) \leq \Phi^-_{(\mathbf{2},\mathbf{2})}(s_3) \Leftrightarrow$
$$\sum_{x \in A\setminus\{s_4\}} \left\{ \sum_{\mathbf{t} \in E(G_{(\mathbf{2},\mathbf{2})})} \frac{1}{n-1} \pi_{\mathbf{t}}(x, s_4) \right\} \leq \sum_{x \in A\setminus\{s_3\}} \left\{ \sum_{\mathbf{t} \in E(G_{(\mathbf{2},\mathbf{2})})} \frac{1}{n-1} \pi_{\mathbf{t}}(x, s_3) \right\},$$
3. $\Phi^+_{(\mathbf{2},\mathbf{2})}(s_4) - \Phi^-_{(\mathbf{2},\mathbf{2})}(s_4) \geq \Phi^+_{(\mathbf{2},\mathbf{2})}(s_3) - \Phi^-_{(\mathbf{2},\mathbf{2})}(s_3) + \varepsilon \Leftrightarrow$
$$\sum_{x \in A\setminus\{s_4\}} \left\{ \sum_{\mathbf{t} \in E(G_{(\mathbf{2},\mathbf{2})})} \frac{1}{n-1} \pi_{\mathbf{t}}(s_4, x) \right\} - \sum_{x \in A\setminus\{s_4\}} \left\{ \sum_{\mathbf{t} \in E(G_{(\mathbf{2},\mathbf{2})})} \frac{1}{n-1} \pi_{\mathbf{t}}(x, s_4) \right\} \geq$$

$$\geq \sum_{x\in A\setminus\{s_3\}} \left\{ \sum_{\mathbf{t}\in E(G_{(\mathbf{2},\mathbf{2})})} \frac{1}{n-1}\pi_{\mathbf{t}}(s_3,x) \right\} - \sum_{x\in A\setminus\{s_3\}} \left\{ \sum_{\mathbf{t}\in E(G_{(\mathbf{2},\mathbf{2})})} \frac{1}{n-1}\pi_{\mathbf{t}}(x,s_3) \right\} + \varepsilon.$$

In Table 3.14, we show the necessary outranking relation with respect to some subcriteria of the hierarchy. Also in this case, the necessary outranking relation with respect to the totality of criteria is empty, however, it is interesting to see some partial information at lower levels of the hierarchy, where the necessary outranking relation is not empty; e.g: with respect to Analysis student s_5 necessarily outranks student s_2, or with respect to Organic Chemistry, student s_3 necessarily outranks student s_5, and so on.

Table 3.14: Necessary outranking relations obtained using Hierarchical PROMETHEE^{GKS} and exploitation of the outranking relation in the way of PROMETHEE I

$\succsim^N_{(\mathbf{0})}$	s_1	s_2	s_3	s_4	s_5	$\succsim^N_{(\mathbf{1},\mathbf{2})}$	s_1	s_2	s_3	s_4	s_5	$\succsim^N_{(\mathbf{2},\mathbf{2})}$	s_1	s_2	s_3	s_4	s_5
s_1	1	0	0	0	0	s_1	1	0	0	0	0	s_1	1	0	0	0	0
s_2	0	1	0	0	0	s_2	0	1	0	0	0	s_2	0	1	0	0	0
s_3	0	0	1	0	0	s_3	0	0	1	0	0	s_3	0	0	1	0	1
s_4	0	0	0	1	0	s_4	0	0	0	1	0	s_4	0	0	0	1	0
s_5	0	0	0	0	1	s_5	0	1	0	0	1	s_5	0	0	0	0	1

3.2.6 Conclusions

In this section, we proposed a new procedure aiming at extending the outranking methods to the case of the hierarchy of criteria in the way introduced in [23]. The family of criteria is not considered at the same level, but, instead, it has a hierarchical structure. Considering the hierarchical structure of criteria, the Decision Maker (DM) can obtain not only comprehensive preference relation with respect to all criteria, but also partial preference relation with respect to subcriteria at different levels of the hierarchy. This is not possible when considering the flat structure of criteria.

Let us remark that the use of the hierarchy of criteria proposed by our approach is rather different from other MCDA methodologies [111, 28]. In fact, while in general the hierarchy of criteria is used to decompose and make easier the preference elicitation concerning pairwise comparisons of criteria with respect to relative importance, in our approach, a preference relation in each node of the hierarchy constitutes a base for discussion with the DM.

We wish to stress that this specific use of the hierarchy of criteria can be applied to any MCDA methodology. In this section we have applied it to Robust Ordinal Regression (ROR) approach, but it can be applied to any other MCDA methodology, even those which use the hierarchy to ask the DM for pairwise comparisons of subcriteria with respect to their importance.

Remark, moreover, that our hierarchical procedures boil down to the classical ELECTRE and PROMETHEE methods when criteria are considered at one level only. This proves that our hierarchical procedures generalize the classical outranking methods.

We presented the hierarchical outranking methods for two types of preference information from the part of the DM: direct, considered in classical outranking methods, and indirect, considered in Robust Ordinal Regression for outranking methods. ROR takes into account all outranking models compatible with preference information provided by the DM in terms of exemplary outranking and non-outranking relations for some pairs of reference alternatives. It is producing two binary relations: the necessary outranking relation (S^N, $\succsim^N$), for which a outranks b for all compatible outranking models, and the possible outranking relation (S^P, $\succsim^P$), for which a outranks b for at least one compatible outranking model. When ROR is applied to hierarchical outranking methods, one gets necessary ($\succsim^N_{\mathbf{r}}$) and possible ($\succsim^P_{\mathbf{r}}$) outranking relations for each criterion/subcriterion $G_{\mathbf{r}}$ belonging to the hierarchy. In this way, the DM knows the necessary and possible preference relations for given preference information, not only at the comprehensive level, for the totality of criteria, but also for any criterion/subcriterion of the hierarchy. Such finer information about preferences has an advantage over the comprehensive information because it permits to decompose the comprehensive preferences into their constituent elements. The application of ROR to ELECTRE and PROMETHEE methods was done in [47] and [76], but also in this case, our hierarchical procedures can be considered as generalizations of both ELECTRE^{GKMS} and PROMETHEE^{GKS} because the hierarchical procedures boil down to these methods when all criteria are considered at the same level.

3.2.7 Appendix A

Proof of Proposition 3.2.1

1. Let $G_{\mathbf{r}} \in \mathcal{G}$ with $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus \{LBO \cup EL\}$, and $a, b \in A$, such that $aS_{(\mathbf{r},j)}b$, for all $j = 1, \ldots, n(\mathbf{r})$. This means that:

   $\alpha)$ $C_{(\mathbf{r},j)}(a,b) \geq \lambda_{(\mathbf{r},j)}$, for all $j = 1, \ldots, n(\mathbf{r})$,

   $\beta)$ $g_{\mathbf{t}}(b) - g_{\mathbf{t}}(a) < v_{\mathbf{t}}$, for all $\mathbf{t} \in E(G_{(\mathbf{r},j)})$, for all $j = 1, \ldots, n(\mathbf{r})$.

   Noting that we are considering the case in which each criterion belongs to only one of the criteria from the upper level (see section 3.2.2), we have:

   $\gamma)$ $\cup_{j=1}^{n(\mathbf{r})} E(G_{(\mathbf{r},j)}) = E(G_{\mathbf{r}})$,

$\delta)\ C_{\mathbf{r}}(a,b) = \displaystyle\sum_{j=1}^{n(\mathbf{r})} C_{(\mathbf{r},j)}(a,b),$

$\theta)\ \lambda_{\mathbf{r}} = \displaystyle\sum_{j=1}^{n(\mathbf{r})} \lambda_{(\mathbf{r},j)},$

thus

$$C_{\mathbf{r}}(a,b) = \sum_{j=1}^{n(\mathbf{r})} C_{(\mathbf{r},j)}(a,b) \geq \sum_{j=1}^{n(\mathbf{r})} \lambda_{(\mathbf{r},j)} = \lambda_{\mathbf{r}} \qquad \text{by } \delta),\ \alpha) \text{ and } \theta)$$

and

$$g_{\mathbf{t}}(b) - g_{\mathbf{t}}(a) < v_{\mathbf{t}}, \text{ for all } \mathbf{t} \in E(G_{\mathbf{r}}) \qquad \text{by } \beta) \text{ and } \gamma).$$

This implies that $aS_{\mathbf{r}}b$.

2. Let $G_{\mathbf{r}} \in \mathcal{G}$ with $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus \{LBO \cup EL\}$, and $a,b \in A$, such that $not(aS_{(\mathbf{r},j)}b)$, for all $j = 1,\ldots,n(\mathbf{r})$. This means that for all $j = 1,\ldots,n(\mathbf{r})$ we have:

   $\alpha^{'})\ C_{(\mathbf{r},j)}(a,b) < \lambda_{(\mathbf{r},j)}$ or

   $\beta^{'})\ \exists \mathbf{t} \in E(G_{(\mathbf{r},j)}) : g_{\mathbf{t}}(b) - g_{\mathbf{t}}(a) \geq v_{\mathbf{t}}.$

   We distinguish two cases:

   - Let us suppose that $not(aS_{(\mathbf{r},j)}b)$ is satisfied because of $\beta^{'})$; thus there exists one elementary subcriterion $g_{\mathbf{t}} \in E(G_{(\mathbf{r},j)})$ such that $g_{\mathbf{t}}(b) - g_{\mathbf{t}}(a) \geq v_{\mathbf{t}}$; being $E(G_{(\mathbf{r},j)}) \subseteq E(G_{\mathbf{r}})$, $g_{\mathbf{t}}$ is an elementary subcriterion belonging also to $E(G_{\mathbf{r}})$ and so it opposes veto to the outranking of a over b with respect to criterion $G_{\mathbf{r}}$; therefore $not(aS_{\mathbf{r}}b)$.
   - Let us suppose that for all $j = 1,\ldots,n(\mathbf{r})$, $\beta^{'})$ is never satisfied, that is for all $j = 1,\ldots,n(\mathbf{r})$, for all $\mathbf{t} \in E(G_{(\mathbf{r},j)})$, $g_{\mathbf{t}}(b) - g_{\mathbf{t}}(a) < v_{\mathbf{t}}$. Thus, for all $j = 1,\ldots,n$, $not(aS_{(\mathbf{r},j)}b)$ holds because of $\alpha^{'})$, that is for all $j = 1,\ldots,n(\mathbf{r})$, $C_{(\mathbf{r},j)}(a,b) < \lambda_{(\mathbf{r},j)}$. Reminding $\gamma),\delta)$ and $\theta)$ of point 1. of this Proposition and by $\alpha^{'})$ we have:

     $$C_{\mathbf{r}}(a,b) = \sum_{j=1}^{n(\mathbf{r})} C_{(\mathbf{r},j)}(a,b) < \sum_{j=1}^{n(\mathbf{r})} \lambda_{(\mathbf{r},j)} = \lambda_{\mathbf{r}}.$$

     This opposes to outranking of a over b with respect to criterion $G_{\mathbf{r}}$, and thus $not(aS_{\mathbf{r}}b)$.

Proof of Proposition 3.2.2

1. Let $a, b \in A$, and $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$ such that $aS^N_{\mathbf{r}}b$. This means that $aS_{\mathbf{r}}b$ for all compatible outranking models, and thus there exists at least one compatible outranking model for which $aS_{\mathbf{r}}b$, thus $aS^P_{\mathbf{r}}b$.

2. Let S an outranking relation, $a \in A$ an alternative and $G_{\mathbf{r}}$, with $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$ a criterion/subcriterion. We have that:

   - for all $\mathbf{t} \in E(G_{\mathbf{r}})$, $\phi_{\mathbf{t}}(a, a) = 1$, and therefore by equation (3.5), $C_{\mathbf{r}}(a, b) = K_{\mathbf{r}}$,
   - for all $\lambda_{\mathbf{r}} \in \left[\frac{K_{\mathbf{r}}}{2}, K_{\mathbf{r}}\right]$, $C_{\mathbf{r}}(a, b) \geq \lambda_{\mathbf{r}}$, (it follows by previous point),
   - $g_{\mathbf{t}}(a) - g_{\mathbf{t}}(a) = 0 < v_{\mathbf{t}}$, for all $\mathbf{t} \in E(G_{\mathbf{r}})$.

   The last two statements bring to $aS_{\mathbf{r}}a$. Being S an arbitrary outranking relation, we obtain that $aS^N_{\mathbf{r}}a$ and by point 1. of this Proposition $aS^P_{\mathbf{r}}a$; being a an arbitrary alternative, we obtain that $S^N_{\mathbf{r}}$ and $S^P_{\mathbf{r}}$ are reflexive relations.

3. Let $G_{\mathbf{r}} \in \mathcal{G}$ with $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, and $a, b \in A$ such that $aS^N_{\mathbf{r}}b$. This means that for all compatible outranking models, a outranks b with respect to criterion $G_{\mathbf{r}}$; thus there does not exist a compatible outranking model for which a does not outrank b with respect to criterion $G_{\mathbf{r}}$, that is *not*$(aS^{CP}_{\mathbf{r}}b)$.

   Conversely, let $G_{\mathbf{r}} \in \mathcal{G}$ with $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, and $a, b \in A$ such that *not*$(aS^{CP}_{\mathbf{r}}b)$. This means that it is not true that there exists one compatible outranking model for which a does not outrank b with respect to criterion $G_{\mathbf{r}}$. Thus, for all compatible outranking models a outranks b with respect to criterion $G_{\mathbf{r}}$, that is $aS^N_{\mathbf{r}}b$.

4. Let $G_{\mathbf{r}} \in \mathcal{G}$ with $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, and $a, b \in A$ such that $aS^P_{\mathbf{r}}b$. This means that there exists at least one compatible outranking model for which a outranks b with respect to criterion $G_{\mathbf{r}}$; thus, it is not true that a does not outrank b with respect to criterion $G_{\mathbf{r}}$ for all compatible outranking models, that is *not*$(aS^{CN}_{\mathbf{r}}b)$.

   Conversely, let $G_{\mathbf{r}} \in \mathcal{G}$ with $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, and $a, b \in A$ such that *not*$(aS^{CN}_{\mathbf{r}}b)$. This means that it is not true that for all compatible outranking models a does not outrank b with respect to criterion $G_{\mathbf{r}}$. Thus, there exists at least one compatible outranking model for which a outranks b with respect to criterion $G_{\mathbf{r}}$, that is $aS^P_{\mathbf{r}}b$.

5. Let $G_{\mathbf{r}} \in \mathcal{G}$ with $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, and $a, b \in A$ such that $aS^{CN}_{\mathbf{r}}b$. This means that a does not outrank b for all compatible outranking models; thus there exists at least one compatible outranking model for which a does not outrank b, that is $aS^{CP}_{\mathbf{r}}b$.

6. For all $a \in A$, by points 2. and 3. of this Proposition, we have:

$$aS_{\mathbf{r}}^{N}a \Leftrightarrow not(aS_{\mathbf{r}}^{CP}a),$$

and thus $S_{\mathbf{r}}^{CP}$ is an irreflexive binary relation.
Analogously, for all $a \in A$, by points 2. and 4. of this Proposition, we have:

$$aS_{\mathbf{r}}^{P}a \Leftrightarrow not(aS_{\mathbf{r}}^{CN}a),$$

and thus $S_{\mathbf{r}}^{CN}$ is an irreflexive binary relation.

Proof of Proposition 3.2.3

1. Let $G_{\mathbf{r}} \in \mathcal{G}$ with $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus \{LBO \cup EL\}$, and $a, b \in A$ such that $aS_{(\mathbf{r},j)}^{N}b$, for all $j = 1, \ldots, n(\mathbf{r})$. This means that $aS_{(\mathbf{r},j)}b$ for all $j = 1, \ldots, n(\mathbf{r})$ and for all compatible outranking models. Let $\overline{M}$ one of these compatible outranking models and $\overline{S}$ the outranking relation induced by $\overline{M}$. By point 1. of Proposition 3.2.1 we obtain $a\overline{S}_{\mathbf{r}}b$. Being $\overline{M}$ an arbitrary compatible outranking model, we have that a outranks b with respect to criterion $G_{\mathbf{r}}$ for all compatible outranking models, and so $aS_{\mathbf{r}}^{N}b$.

2. Let $G_{\mathbf{r}} \in \mathcal{G}$ with $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus \{LBO \cup EL\}$, and $a, b \in A$, such that

   $\alpha)$ $aS_{(\mathbf{r},j)}^{N}b$ for all $j = 1, \ldots, n(\mathbf{r}), j \neq w$,

   $\beta)$ $aS_{(\mathbf{r},w)}^{P}b$.

   Hypothesis $\beta)$ implies that there exists at least one compatible outranking model $\overline{M}$ inducing the outranking relation $\overline{S}$ such that $a\overline{S}_{(\mathbf{r},w)}b$. But for the hypothesis $\alpha)$ we have also that $a\overline{S}_{(\mathbf{r},j)}b$, for all $j = 1, \ldots, n(\mathbf{r})$ and $j \neq w$. Together, these considerations imply that $a\overline{S}_{(\mathbf{r},j)}b$ for all $j = 1, \ldots, n(\mathbf{r})$, and thus by point 1. of Proposition 3.2.1 we obtain that a outranks b with respect to criterion $G_{\mathbf{r}}$ for outranking model $\overline{M}$ and thus $aS_{\mathbf{r}}^{P}b$.

3. Let $G_{\mathbf{r}} \in \mathcal{G}$ with $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus \{LBO \cup EL\}$, and $a, b \in A$ such that $aS_{(\mathbf{r},j)}^{CN}b$, for all $j = 1, \ldots, n(\mathbf{r})$. This means that $not(aS_{(\mathbf{r},j)}b)$, for all $j = 1, \ldots, n(\mathbf{r})$ and for all compatible outranking models. Considering $\overline{M}$ one of these compatible outranking models, and $\overline{S}$ the outranking relation induced by $\overline{M}$, by point 2. of Proposition 3.2.1 we obtain $not(a\overline{S}_{\mathbf{r}}b)$. Being $\overline{M}$ an arbitrary compatible outranking model, we have $not(aS_{\mathbf{r}}b)$ for all compatible outranking models and thus $aS_{\mathbf{r}}^{CN}b$.

4. Let $G_{\mathbf{r}} \in \mathcal{G}$ with $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus \{LBO \cup EL\}$, and $a, b \in A$ such that:

   $\alpha)$ $aS^{CN}_{(\mathbf{r},j)}b$, for all $j = 1, \ldots, n(\mathbf{r}), j \neq w$,

   $\beta)$ $aS^{CP}_{(\mathbf{r},w)}b$,

   $\beta)$ implies that there exist at least one compatible outranking model $\overline{M}$ inducing the outranking relation $\overline{S}$, such that $not(a\overline{S}_{(\mathbf{r},w)}b)$. By $\alpha)$ we have also that $not(a\overline{S}_{(\mathbf{r},j)}b)$, for all $j = 1, \ldots, n(\mathbf{r})$ and $j \neq w$. Together, these considerations imply that $not(a\overline{S}_{(\mathbf{r},j)}b)$ for all $j = 1, \ldots, n(\mathbf{r})$, and thus by point 2. of Proposition 3.2.1 we obtain $not(a\overline{S}_{\mathbf{r}}b)$. Therefore, $aS^{CP}_{\mathbf{r}}b$.

Proof of Proposition 3.2.4

1. Without loss of generality we have supposed that the hierarchy is structured in a way that each subcriterion at level l descends from only one criterion of level $l-1$ (see section 3.2.2). In this way, considering criterion $G_{\mathbf{r}}$ and its subcriteria $G_{(\mathbf{r},1)}, \ldots, G_{(\mathbf{r},n(\mathbf{r}))}$ we have:

   - $\pi_{\mathbf{r}}(a,b) = \sum_{\mathbf{t} \in E(G_{\mathbf{r}})} k_{\mathbf{t}} P_{\mathbf{t}}(a,b)$,
   - $\pi_{(\mathbf{r},j)}(a,b) = \sum_{\mathbf{t} \in E(G_{(\mathbf{r},j)})} k_{\mathbf{t}} P_{\mathbf{t}}(a,b)$, for all $j = 1, \ldots, n(\mathbf{r})$,
   - $E(G_{\mathbf{r}}) = \cup_{j=1}^{n(\mathbf{r})} E(G_{(\mathbf{r},j)})$.

   We can observe that each $\mathbf{t} \in E(G_{\mathbf{r}})$ belongs to only one of $E(G_{(\mathbf{r},j)})$, $j = 1, \ldots, n$, and thus:

$$\pi_{\mathbf{r}}(a,b) = \sum_{\mathbf{t} \in E(G_{\mathbf{r}})} k_{\mathbf{t}} P_{\mathbf{t}}(a,b) = \sum_{j=1}^{n(\mathbf{r})} \left[\sum_{\mathbf{t} \in E(G_{(\mathbf{r},j)})} k_{\mathbf{t}} P_{\mathbf{t}}(a,b) \right] = \sum_{j=1}^{n(\mathbf{r})} \pi_{(\mathbf{r},j)}(a,b).$$

2. For each criterion/subcriterion $G_{\mathbf{r}} \in \mathcal{G}$, $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, supposing that there exist n different alternatives in A, we have for all $a \in A$:

   - $\Phi^{+}_{\mathbf{r}}(a) = \frac{1}{n-1} \sum_{x \in A \setminus \{a\}} \pi_{\mathbf{r}}(a,x)$,
   - $\Phi^{+}_{(\mathbf{r},j)}(a) = \frac{1}{n-1} \sum_{x \in A \setminus \{a\}} \pi_{(\mathbf{r},j)}(a,x)$, for all $j = 1, \ldots, n(\mathbf{r})$.

   Thus, by point 1. of this Proposition and using the above expressions:

$$\Phi^{+}_{\mathbf{r}}(a) = \frac{1}{n-1} \sum_{x \in A \setminus \{a\}} \pi_{\mathbf{r}}(a,x) = \frac{1}{n-1} \sum_{x \in A \setminus \{a\}} \left[\sum_{j=1}^{n(\mathbf{r})} \pi_{(\mathbf{r},j)}(a,x) \right] =$$

$$= \sum_{j=1}^{n(\mathbf{r})} \left[\frac{1}{n-1} \sum_{x \in A \setminus \{a\}} \pi_{(\mathbf{r},j)}(a,x) \right] = \sum_{j=1}^{n(\mathbf{r})} \Phi^{+}_{(\mathbf{r},j)}(a).$$

3. Analogous to proof of point 2.

4. By points 2. and 3. of this Proposition, for each $a \in A$, and for each $G_{\mathbf{r}} \in \mathcal{G}, \mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$,

$$\Phi_{\mathbf{r}}(a) = \Phi^{+}_{\mathbf{r}}(a) - \Phi^{-}_{\mathbf{r}}(a) = \sum_{j=1}^{n(\mathbf{r})} \Phi^{+}_{(\mathbf{r},j)}(a) - \sum_{j=1}^{n(\mathbf{r})} \Phi^{-}_{(\mathbf{r},j)}(a) =$$

$$= \sum_{j=1}^{n(\mathbf{r})} \left[\Phi^{+}_{(\mathbf{r},j)}(a) - \Phi^{-}_{(\mathbf{r},j)}(a) \right] = \sum_{j=1}^{n(\mathbf{r})} \Phi_{(\mathbf{r},j)}(a).$$

Proof of Proposition 3.2.5

1. Let $a, b \in A$ and $G_{\mathbf{r}} \in \mathcal{G}, \mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, such that $aP^{I}_{(\mathbf{r},j)}b$, for all $j = 1, \ldots, n(\mathbf{r})$. By hypothesis, we have:

$$\Phi^{+}_{(\mathbf{r},j)}(a) \geq \Phi^{+}_{(\mathbf{r},j)}(b) \quad \text{and} \quad \Phi^{-}_{(\mathbf{r},j)}(a) \leq \Phi^{-}_{(\mathbf{r},j)}(b), \text{ for all } j = 1, \ldots, n(\mathbf{r}),$$

and for each j at least one of the above inequalities is strict. Then adding up with respect to j, we obtain:

$$\sum_{j=1}^{n(\mathbf{r})} \Phi^{+}_{(\mathbf{r},j)}(a) \geq \sum_{j=1}^{n(\mathbf{r})} \Phi^{+}_{(\mathbf{r},j)}(b) \quad \text{and} \quad \sum_{j=1}^{n(\mathbf{r})} \Phi^{-}_{(\mathbf{r},j)}(a) \leq \sum_{j=1}^{n(\mathbf{r})} \Phi^{-}_{(\mathbf{r},j)}(b),$$

and thus by points 2. and 3. of Proposition 3.2.4,

$$\Phi^{+}_{\mathbf{r}}(a) \geq \Phi^{+}_{\mathbf{r}}(b) \quad \text{and} \quad \Phi^{-}_{\mathbf{r}}(a) \leq \Phi^{-}_{\mathbf{r}}(b)$$

with at least one of the two inequalities being strict; therefore $aP^{I}_{\mathbf{r}}b$.

2. Let $a, b \in A$, $G_{\mathbf{r}} \in \mathcal{G}$ with $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, and $\{C_1, C_2\}$ a partition of $\{1, \ldots, n(\mathbf{r})\}$, such that $aP^{I}_{(\mathbf{r},j)}b$, for all $j \in C_1$ and $aI^{I}_{(\mathbf{r},j)}b$, for all $j \in C_2$. By the first hypothesis, we have:

$$\Phi^{+}_{(\mathbf{r},j)}(a) \geq \Phi^{+}_{(\mathbf{r},j)}(b) \quad \text{and} \quad \Phi^{-}_{(\mathbf{r},j)}(a) \leq \Phi^{-}_{(\mathbf{r},j)}(b), \text{ for all } j \in C_1 \tag{3.7}$$

with at least one of the two inequalities strict; by the second hypothesis we have:

$$\Phi^{+}_{(\mathbf{r},j)}(a) = \Phi^{+}_{(\mathbf{r},j)}(b) \quad \text{and} \quad \Phi^{-}_{(\mathbf{r},j)}(a) = \Phi^{-}_{(\mathbf{r},j)}(b), \text{ for all } j \in C_2. \tag{3.8}$$

Thus, by points 2. and 3. of Proposition 3.2.4,

$$\sum_{j\in C_1}\Phi^{+}_{(\mathbf{r},j)}(a)\geq\sum_{j\in C_1}\Phi^{+}_{(\mathbf{r},j)}(b)\quad\text{and}\quad\sum_{j\in C_1}\Phi^{-}_{(\mathbf{r},j)}(a)\leq\sum_{j\in C_1}\Phi^{-}_{(\mathbf{r},j)}(b)$$

with at least one of the two inequalities strict; by (3.7) and (3.8) we obtain:

$$\sum_{j=1}^{n(\mathbf{r})}\Phi^{+}_{(\mathbf{r},j)}(a)=\sum_{j\in C_1}\Phi^{+}_{(\mathbf{r},j)}(a)+\sum_{j\in C_2}\Phi^{+}_{(\mathbf{r},j)}(a)\geq\sum_{j\in C_1}\Phi^{+}_{(\mathbf{r},j)}(b)+\sum_{j\in C_2}\Phi^{+}_{(\mathbf{r},j)}(b)=\sum_{j=1}^{n(\mathbf{r})}\Phi^{+}_{(\mathbf{r},j)}(b)$$

and

$$\sum_{j=1}^{n(\mathbf{r})}\Phi^{-}_{(\mathbf{r},j)}(a)=\sum_{j\in C_1}\Phi^{-}_{(\mathbf{r},j)}(a)+\sum_{j\in C_2}\Phi^{-}_{(\mathbf{r},j)}(a)\leq\sum_{j\in C_1}\Phi^{-}_{(\mathbf{r},j)}(b)+\sum_{j\in C_2}\Phi^{-}_{(\mathbf{r},j)}(b)=\sum_{j=1}^{n(\mathbf{r})}\Phi^{-}_{(\mathbf{r},j)}(b);$$

therefore

$$\sum_{j=1}^{n(\mathbf{r})}\Phi^{+}_{(\mathbf{r},j)}(a)\geq\sum_{j=1}^{n(\mathbf{r})}\Phi^{+}_{(\mathbf{r},j)}(b)\quad\text{and}\quad\sum_{j=1}^{n(\mathbf{r})}\Phi^{-}_{(\mathbf{r},j)}(a)\leq\sum_{j=1}^{n(\mathbf{r})}\Phi^{-}_{(\mathbf{r},j)}(b),$$

that is, by points 2. and 3. of Proposition 3.2.4,

$$\Phi^{+}_{\mathbf{r}}(a)\geq\Phi^{+}_{\mathbf{r}}(b)\quad\text{and}\quad\Phi^{-}_{\mathbf{r}}(a)\leq\Phi^{-}_{\mathbf{r}}(b)$$

with at least one of the two inequalities being strict. From this follows that $aP^{I}_{\mathbf{r}}b$.

3. Let $a,b\in A$ and $G_{\mathbf{r}}\in\mathcal{G},\mathbf{r}\in\mathcal{I}_{\mathcal{G}}\setminus EL$ such that $aI^{I}_{(\mathbf{r},j)}b$, for all $j=1,\ldots,n(\mathbf{r})$. This means that

$$\Phi^{+}_{(\mathbf{r},j)}(a)=\Phi^{+}_{(\mathbf{r},j)}(b)\quad\text{and}\quad\Phi^{-}_{(\mathbf{r},j)}(a)=\Phi^{-}_{(\mathbf{r},j)}(b),\ \text{for all}\ j=1,\ldots,n(\mathbf{r}).$$

Adding up with respect to j we obtain:

$$\sum_{j=1}^{n(\mathbf{r})}\Phi^{+}_{(\mathbf{r},j)}(a)=\sum_{j=1}^{n(\mathbf{r})}\Phi^{+}_{(\mathbf{r},j)}(b)\quad\text{and}\quad\sum_{j=1}^{n(\mathbf{r})}\Phi^{-}_{(\mathbf{r},j)}(a)=\sum_{j=1}^{n(\mathbf{r})}\Phi^{-}_{(\mathbf{r},j)}(b),\ \text{for all}\ j=1,\ldots,n(\mathbf{r}),$$

that is, by points 2. and 3. of Proposition 3.2.4,

$$\Phi^{+}_{\mathbf{r}}(a)=\Phi^{+}_{\mathbf{r}}(b)\quad\text{and}\quad\Phi^{-}_{\mathbf{r}}(a)=\Phi^{-}_{\mathbf{r}}(b),$$

and therefore $aI^{I}_{\mathbf{r}}b$.

Proof of Proposition 3.2.6

1. Let $a,b\in A$, $G_{\mathbf{r}}\in\mathcal{G},\mathbf{r}\in\mathcal{I}_{\mathcal{G}}\setminus EL$ such that $aP^{II}_{(\mathbf{r},j)}b$, for all $j=1,\ldots,n(\mathbf{r})$. By point 4. of

Proposition 3.2.4,

$$\Phi_{(\mathbf{r},j)}(a) > \Phi_{(\mathbf{r},j)}(b), \text{ for all } j = 1, \ldots, n(\mathbf{r}) \Rightarrow \sum_{j=1}^{n(\mathbf{r})} \Phi_{(\mathbf{r},j)}(a) > \sum_{j=1}^{n(\mathbf{r})} \Phi_{(\mathbf{r},j)}(b) \Leftrightarrow \Phi_{\mathbf{r}}(a) > \Phi_{\mathbf{r}}(b),$$

and therefore $aP^{II}_{\mathbf{r}}b$.

2. Let $a, b \in A$, $G_{\mathbf{r}} \in \mathcal{G}, \mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$ and $\{C_1, C_2\}$ a partition of $\{1, \ldots, n(\mathbf{r})\}$ such that $aP^{II}_{(\mathbf{r},j)}b$ for all $j \in C_1$ and $aI^{II}_{(\mathbf{r},j)}b$ for all $j \in C_2$. By hypothesis we have:

$$\Phi_{(\mathbf{r},j)}(a) > \Phi_{(\mathbf{r},j)}(b), \quad \text{for all } j \in C_1 \quad \text{and} \quad \Phi_{(\mathbf{r},j)}(a) = \Phi_{(\mathbf{r},j)}(b), \quad \text{for all } j \in C_2.$$

Adding up with respect to j, by point 4. of Proposition 3.2.4, we obtain:

$$\sum_{j \in C_1} \Phi_{(\mathbf{r},j)}(a) > \sum_{j \in C_1} \Phi_{(\mathbf{r},j)}(b) \Rightarrow \sum_{j \in C_1} \Phi_{(\mathbf{r},j)}(a) + \sum_{j \in C_2} \Phi_{(\mathbf{r},j)}(a) > \sum_{j \in C_1} \Phi_{(\mathbf{r},j)}(b) + \sum_{j \in C_2} \Phi_{(\mathbf{r},j)}(b) \Leftrightarrow$$

$$\Leftrightarrow \sum_{j=1}^{n(\mathbf{r})} \Phi_{(\mathbf{r},j)}(a) > \sum_{j=1}^{n(\mathbf{r})} \Phi_{(\mathbf{r},j)}(b) \Leftrightarrow \Phi_{\mathbf{r}}(a) > \Phi_{\mathbf{r}}(b),$$

and therefore $aP^{II}_{\mathbf{r}}b$.

3. Let $a, b \in A$, $G_{\mathbf{r}} \in \mathcal{G}, \mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, such that $aI^{II}_{(\mathbf{r},j)}b$, for all $j = 1, \ldots, n(\mathbf{r})$. By point 4. of Proposition 3.2.4,

$$\Phi_{(\mathbf{r},j)}(a) = \Phi_{(\mathbf{r},j)}(b), \text{ for all } j = 1, \ldots, n(\mathbf{r}) \Rightarrow \sum_{j=1}^{n(\mathbf{r})} \Phi_{(\mathbf{r},j)}(a) = \sum_{j=1}^{n(\mathbf{r})} \Phi_{(\mathbf{r},j)}(b) \Leftrightarrow \Phi_{\mathbf{r}}(a) = \Phi_{\mathbf{r}}(b),$$

and therefore $aI^{II}_{\mathbf{r}}b$.

Proof of Proposition 3.2.7

We prove this Proposition in case of PROMETHEE I because the proof in case of PROMETHEE II is analogous.

1. Let be $a, b \in A$, and $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, such that $a \succsim^N_{\mathbf{r}} b$. This means that a outranks b with respect to criterion $G_{\mathbf{r}}$ for all compatible outranking models; thus, there exists at least one compatible outranking model for which a outranks b with respect to criterion $G_{\mathbf{r}}$, and therefore $a \succsim^P_{\mathbf{r}} b$.

2. For each $a \in A$, for each criterion/subcriterion $G_{\mathbf{r}}$, and for each compatible outranking model, we have:

$$\Phi^{+}_{\mathbf{r}}(a) \geq \Phi^{+}_{\mathbf{r}}(a) \;\text{ and }\; \Phi^{-}_{\mathbf{r}}(a) \leq \Phi^{-}_{\mathbf{r}}(a); \tag{3.9}$$

By equation (3.9) it follows that, for all compatible outranking models $\Phi_{\mathbf{r}}(a) = \Phi^{+}_{\mathbf{r}}(a) - \Phi^{-}_{\mathbf{r}}(a) \geq \Phi_{\mathbf{r}}(a)$ and thus $a \succsim^{N}_{\mathbf{r}} a$, for all $a \in A$ proving that $\succsim^{N}_{\mathbf{r}}$ is a reflexive binary relation. Being $\succsim^{N}_{\mathbf{r}} \subseteq \succsim^{P}_{\mathbf{r}}$, and $\succsim^{N}_{\mathbf{r}}$ a reflexive binary relation, also $\succsim^{P}_{\mathbf{r}}$ is a reflexive binary relation.

Proof of Proposition 3.2.8

We prove this Proposition in case of PROMETHEE I because the proof in case of PROMETHEE II is analogous.

1. Let $a, b \in A$ and $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus \{LBO \cup EL\}$, such that $a \succsim^{N}_{(\mathbf{r},j)} b$ for all $j = 1, \ldots, n(\mathbf{r})$. This means that a outranks b with respect to criteria/subcriteria $G_{(\mathbf{r},j)}$, for all $j = 1, \ldots, n(\mathbf{r})$, for all compatible outranking models. Thus, for all compatible outranking models we have:

$$\Phi^{+}_{(\mathbf{r},j)}(a) \geq \Phi^{+}_{(\mathbf{r},j)}(b) \text{ and } \Phi^{-}_{(\mathbf{r},j)}(a) \leq \Phi^{-}_{(\mathbf{r},j)}(b), \text{ for all } j = 1, \ldots, n(\mathbf{r}). \tag{3.10}$$

   By points 2. and 3. of Proposition 3.2.4 and equation above, for all compatible outranking models we have that:

$$\Phi^{+}_{\mathbf{r}}(a) \geq \Phi^{+}_{\mathbf{r}}(b) \text{ and } \Phi^{-}_{\mathbf{r}}(a) \leq \Phi^{-}_{\mathbf{r}}(b),$$

   implying that $a \succsim^{N}_{\mathbf{r}} b$.

2. Let $a, b \in A$ and $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus \{LBO \cup EL\}$, such that $a \succsim^{N}_{(\mathbf{r},j)} b$, for all $j = 1, \ldots, n(r), j \neq w$ and $a \succsim^{P}_{(\mathbf{r},w)} b$. This means that a outranks b with respect to criteria $G_{(\mathbf{r},j)}$, for all $j = 1, \ldots, n(\mathbf{r}), j \neq w$ for all compatible outranking models and a outranks b with respect to criterion/subcriterion $G_{(\mathbf{r},w)}$ for at least one compatible outranking model. From this we have that, for all compatible outranking models:

$$\Phi^{+}_{(\mathbf{r},j)}(a) \geq \Phi^{+}_{(\mathbf{r},j)}(b) \text{ and } \Phi^{-}_{(\mathbf{r},j)}(a) \leq \Phi^{-}_{(\mathbf{r},j)}(b), \text{ for all } j = 1, \ldots, n(\mathbf{r}), j \neq w, \tag{3.11}$$

   and for at least one compatible outranking model:

$$\Phi^{+}_{(\mathbf{r},w)}(a) \geq \Phi^{+}_{(\mathbf{r},w)}(b) \text{ and } \Phi^{-}_{(\mathbf{r},w)}(a) \leq \Phi^{-}_{(\mathbf{r},w)}(b). \tag{3.12}$$

   Let us denote by $\overline{M}$ the outranking model satisfying equation (3.12). In particular, this compatible outranking model fulfills also equation (3.11). Thus, by points 2. and 3. of Proposition 3.2.4, and considering the compatible outranking model $\overline{M}$, we have:

$$\Phi_{\mathbf{r}}^{+}(a) \geq \Phi_{\mathbf{r}}^{+}(b) \ \text{ and } \ \Phi_{\mathbf{r}}^{-}(a) \leq \Phi_{\mathbf{r}}^{-}(b),$$

and therefore a outranks b with respect to criterion/subcriterion $G_{\mathbf{r}}$ for at least one compatible outranking model, that is $a \succsim_{\mathbf{r}}^{P} b$.

3.2.8 Appendix B

Ordinal regression constraints used in the Hierarchical ELECTRE^{GKMS} method

Supposing that the DM has given some preference information of the type described in section 3.2.3, compatible outranking models are the sets of variables $\psi_{\mathbf{t}}(a,b)$ for all $(a,b) \in B$, $\mathbf{t} \in EL$, of concordance indices $C_{\mathbf{r}}(a,b)$, concordance cutting levels $\lambda_{\mathbf{s}}$, for all $\mathbf{s} \in LBO$, and veto thresholds $v_{\mathbf{t}}$ for all $\mathbf{t} \in EL$, satisfying the following set of conditions:

- Compatibility with all statements concerning the truth or falsity of the outranking relation for some reference alternatives $a, b \in A^R$:
  - For all $(a,b) \in B^R$ such that $aS_{\mathbf{r}}b$, with $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$:

$$C_{\mathbf{r}}(a,b) = \sum_{\mathbf{t} \in E(G_{\mathbf{r}})} \psi_{\mathbf{t}}(a,b) \geq \lambda_{\mathbf{r}} \ \text{ and } \ g_{\mathbf{t}}(b) - g_{\mathbf{t}}(a) < v_{\mathbf{t}}, \text{ for all } \mathbf{t} \in E(G_{\mathbf{r}}),$$

  - For all $(a,b) \in B^R$ such that $not(aS_{\mathbf{r}}b)$, with $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$:

$$C_{\mathbf{r}}(a,b) = \sum_{\mathbf{t} \in E(G_{\mathbf{r}})} \psi_{\mathbf{t}}(a,b) < \lambda_{\mathbf{r}} \ \text{ or } \ \text{there exists } \ \mathbf{t} \in E(G_{\mathbf{r}}) : g_{\mathbf{t}}(b) - g_{\mathbf{t}}(a) \geq v_{\mathbf{t}},$$

  which can be modeled as:

$$C_{\mathbf{r}}(a,b) = \sum_{\mathbf{t} \in E(G_{\mathbf{r}})} \psi_{\mathbf{t}}(a,b) + \varepsilon \leq \lambda_{\mathbf{r}} + M_0^{\mathbf{r}}(a,b) \ \text{ and } \ g_{\mathbf{t}}(b) - g_{\mathbf{t}}(a) \geq v_{\mathbf{t}} - \delta_{\mathbf{r}} M_{\mathbf{t}}(a,b),$$

  where:

  $M_0^{\mathbf{r}}(a,b), M_{\mathbf{t}}(a,b) \in \{0,1\}, \text{ for all } \mathbf{t} \in E(G_{\mathbf{r}}),$

  $M_0^{\mathbf{r}}(a,b) + \sum_{\mathbf{t} \in E(G_{\mathbf{r}})} M_{\mathbf{t}}(a,b) \leq |E(G_{\mathbf{r}})|,$

  $\delta_{\mathbf{r}}$ is an auxiliary coefficient fixed on a big positive value (i.e. $\delta_{\mathbf{r}} \geq max_{\mathbf{t} \in E(G_{\mathbf{r}})}\{\beta_{\mathbf{t}} - \alpha_{\mathbf{t}}\}$ where $\alpha_{\mathbf{t}} = \min_{a \in A} g_{\mathbf{t}}(a)$ and $\beta_{\mathbf{t}} = \max_{a \in A} g_{\mathbf{t}}(a)$).

  Differently from [47], we have one binary variable $M_0^{\mathbf{r}}(a,b)$ for each criterion $G_{\mathbf{r}}, \mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, and for each couple $(a,b) \in B^R$, because we need to distinguish the reasons for which the outranking of alternative a over alternative b is not true. In fact, let us suppose, for example, that alternative a does not outrank alternative b with respect to criteria $G_{\mathbf{r}_1}$ and $G_{\mathbf{r}_2}$, and that in the first case, a does not outrank b because there is an elementary subcriterion descending from $G_{\mathbf{r}_1}$ putting veto while the concordance test is verified. At the same time, let us suppose that a does not outrank b with respect to $G_{\mathbf{r}_2}$ because the concordance test is not verified. Then, in the first case $M_0^{\mathbf{r}_1}(a,b) = 1$, because the concordance test is verified, and in the second case $M_0^{\mathbf{r}_2}(a,b) = 0$, because the concordance test is not verified.

- Constraints on the values of $\lambda_{\mathbf{r}}$, for all $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, inter-criteria parameters and of $v_{\mathbf{t}}$ and $k_{\mathbf{t}}$, for all $\mathbf{t} \in EL$:

  - Normalization of the marginal concordance indices for all elementary subcriteria, so that the indices corresponding to the greatest difference in evaluations of two alternatives on each elementary subcriterion ($g_{\mathbf{t}}(x^*_{\mathbf{t}}) - g_{\mathbf{t}}(x_{\mathbf{t},*}) = \beta_{\mathbf{t}} - \alpha_{\mathbf{t}}$) sum up to 1:

$$\sum_{\mathbf{t}\in EL} \psi_{\mathbf{t}}(x^*_{\mathbf{t}}, x_{\mathbf{t},*}) = 1 \ \text{ with } x^*_{\mathbf{t}}, x_{\mathbf{t},*} \in A : g_{\mathbf{t}}(x^*_{\mathbf{t}}) = \beta_{\mathbf{t}} \text{ and } g_{\mathbf{t}}(x_{\mathbf{t},*}) = \alpha_{\mathbf{t}}, \text{ for all } \mathbf{t} \in EL.$$

    As we normalize weights of the elementary subcriteria so that they sum up to 1, each weight is understood as a maximal share of each elementary subcriterion in the comprehensive concordance index. Consequently, $k_{\mathbf{t}} = \psi_{\mathbf{t}}(x^*_{\mathbf{t}}, x_{\mathbf{t},*})$, for all $\mathbf{t} \in EL$.

  - Lower and upper bounds on concordance cutting level of a criterion belonging to last but one level:

$$\lambda_{\mathbf{s}} \in \left[\frac{K_{\mathbf{s}}}{2}, K_{\mathbf{s}}\right], \ \text{ where } \ K_{\mathbf{s}} = \sum_{\mathbf{t}\in E(G_{\mathbf{s}})} k_{\mathbf{t}}.$$

    In consequence of the above considerations, the concordance cutting levels of criteria belonging to the last but one level have to verify:

$$\sum_{\mathbf{t}\in E(G_{\mathbf{s}})} \frac{\psi_{\mathbf{t}}(x^*_{\mathbf{t}}, x_{\mathbf{t},*})}{2} \leq \lambda_{\mathbf{s}} \leq \sum_{\mathbf{t}\in E(G_{\mathbf{s}})} \psi_{\mathbf{t}}(x^*_{\mathbf{t}}, x_{\mathbf{t},*}).$$

  - The concordance cutting level for criterion $G_{\mathbf{r}}, \mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus \{LBO \cup EL\}$, is equal to the sum of the concordance cutting levels of subcriteria descending from it, that is $G_{(\mathbf{r},\mathbf{j})}, j = 1, \ldots, n(\mathbf{r})$:

$$\lambda_{\mathbf{r}} = \sum_{j=1}^{n(\mathbf{r})} \lambda_{(\mathbf{r},j)}.$$

– Constraints on veto thresholds $v_{\mathbf{t}}$, $\mathbf{t} \in EL$,:

$$v_{\mathbf{t}} > p^{*}_{\mathbf{t}}, \text{ for each } \mathbf{t} \in EL_1,$$

$$v_{\mathbf{t}} > g_{\mathbf{t}}(b) - g_{\mathbf{t}}(a), \text{ for each } \mathbf{t} \in EL_2 \text{ such that } a \sim_{\mathbf{t}} b.$$

- Constraints on the values of marginal concordance indices $\psi_{\mathbf{t}}(a,b)$, $\mathbf{t} \in EL$ conditioned by intra-criterion and inter-criterion preference information, for all $(a,b) \in B$:

  – $k_{\mathbf{t},*} \leq \psi_{\mathbf{t}}(x^{*}_{\mathbf{t}}, x_{\mathbf{t},*}) \leq k^{*}_{\mathbf{t}}$, $\mathbf{t} \in EL$,

  – $\psi_{\mathbf{t_1}}(x^{*}_{\mathbf{t_1}}, x_{\mathbf{t_1},*}) \geq \psi_{\mathbf{t_2}}(x^{*}_{\mathbf{t_2}}, x_{\mathbf{t_2},*}) + \varepsilon$ if elementary subcriterion $g_{\mathbf{t_1}}$ is more important than elementary subcriterion $g_{\mathbf{t_2}}$, $\mathbf{t_1}, \mathbf{t_2} \in EL$,

  – $\psi_{\mathbf{t_1}}(x^{*}_{\mathbf{t_1}}, x_{\mathbf{t_1},*}) = \psi_{\mathbf{t_2}}(x^{*}_{\mathbf{t_2}}, x_{\mathbf{t_2},*})$ if elementary subcriteria $g_{\mathbf{t_1}}$ and $g_{\mathbf{t_2}}$ are equally important, $\mathbf{t_1}, \mathbf{t_2} \in EL$,

  – $\psi_{\mathbf{t}}(a,b) = 0$ if $g_{\mathbf{t}}(b) - g_{\mathbf{t}}(a) \geq p^{*}_{\mathbf{t}}$,

  – $\psi_{\mathbf{t}}(a,b) > 0$ if $g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b) > -p_{\mathbf{t},*}$,

  – $\psi_{\mathbf{t}}(a,b) = \psi_{\mathbf{t}}(x^{*}_{\mathbf{t}}, x_{\mathbf{t},*})$ if $g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b) \geq -q_{\mathbf{t},*}$,

  – $\psi_{\mathbf{t}}(a,b) < \psi_{\mathbf{t}}(x^{*}_{\mathbf{t}}, x_{\mathbf{t},*})$ if $g_{\mathbf{t}}(b) - g_{\mathbf{t}}(a) > q^{*}_{\mathbf{t}}$,

  – $\psi_{\mathbf{t}}(a,b) = 0$ if $b \succ_{\mathbf{t}} a$,

  – $\psi_{\mathbf{t}}(a,b) = 0$ and $\psi_{\mathbf{t}}(b,a) = 0$ if $a \sim_{\mathbf{t}} b$.

- Monotonicity of the functions of marginal concordance indices $\psi_{\mathbf{t}}(a,b)$, $\mathbf{t} \in EL$:

$$\psi_{\mathbf{t}}(a,b) \geq \psi_{\mathbf{t}}(c,d) \text{ if } g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b) > g_{\mathbf{t}}(c) - g_{\mathbf{t}}(d),$$

$$\psi_{\mathbf{t}}(a,b) = \psi_{\mathbf{t}}(c,d) \text{ if } g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b) = g_{\mathbf{t}}(c) - g_{\mathbf{t}}(d),$$

Note that all strict inequalities are transformed into weak inequalities involving an auxiliary variable ε in the set of constraints E^{A^R} in the section 3.2.3.
For example, the constraint $\psi_{\mathbf{t}}(a,b) > 0$ if $g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b) > -p_{\mathbf{t},*}$, becomes $\psi_{\mathbf{t}}(a,b) \geq \varepsilon$ if $g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b) > -p_{\mathbf{t},*}$.

Ordinal regression constraints used in the Hierarchical PROMETHEE^{GKS} method

Supposing that the DM has given some preference information of the type described in section 3.2.5, compatible outranking models are the sets of preference indices $\pi_{\mathbf{t}}(a,b)$ for all $(a,b) \in B$, $\mathbf{t} \in EL$ satisfying the following conditions:

- Compatibility with local and global preference relations provided by the DM with respect to a particular criterion $G_{\mathbf{r}}$ in the hierarchy:
  - for all $a, b \in A^R$, and $G_{\mathbf{r}}$ with $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \setminus EL$, such that $a \succsim_{\pi_{\mathbf{r}}} b$,

    $$\pi_{\mathbf{r}}(a,b) = \sum_{\mathbf{t} \in E(G_{\mathbf{r}})} \pi_{\mathbf{t}}(a,b) \geq \pi_{\mathbf{r}}(b,a) = \sum_{\mathbf{t} \in E(G_{\mathbf{r}})} \pi_{\mathbf{t}}(b,a),$$

    Relations $\succ_{\pi_{\mathbf{r}}}$ and $\sim_{\pi_{\mathbf{r}}}$ are translated analogously, using strict inequality and equality, respectively.
  - Considering PROMETHEE I:

    $$\begin{array}{l} \Phi^+_{\mathbf{r}}(a) \geq \Phi^+_{\mathbf{r}}(b) \text{ and } \Phi^-_{\mathbf{r}}(a) \leq \Phi^-_{\mathbf{r}}(b), \text{ if } a \succsim_{\Phi_{\mathbf{r}}} b, \\ \left.\begin{array}{l} \Phi^+_{\mathbf{r}}(a) \geq \Phi^+_{\mathbf{r}}(b) \text{ and } \Phi^-_{\mathbf{r}}(a) \leq \Phi^-_{\mathbf{r}}(b) \text{ and} \\ \Phi^+_{\mathbf{r}}(a) - \Phi^-_{\mathbf{r}}(a) \geq \Phi^+_{\mathbf{r}}(b) - \Phi^-_{\mathbf{r}}(b) + \varepsilon \end{array}\right\} \text{ if } a \succ_{\Phi_{\mathbf{r}}} b \\ \Phi^+_{\mathbf{r}}(a) = \Phi^+_{\mathbf{r}}(b) \text{ and } \Phi^-_{\mathbf{r}}(a) = \Phi^-_{\mathbf{r}}(b) \text{ if } a \sim_{\Phi_{\mathbf{r}}} b. \end{array}$$

  - Considering PROMETHEE II:

    $$\Phi_{\mathbf{r}}(a) = \Phi^+_{\mathbf{r}}(a) - \Phi^-_{\mathbf{r}}(a) \geq \Phi_{\mathbf{r}}(b) = \Phi^+_{\mathbf{r}}(b) - \Phi^-_{\mathbf{r}}(b) \text{ if } a \succsim_{\Phi_{\mathbf{r}}} b.$$

    Relations $\succ_{\Phi_{\mathbf{r}}}$, and $\sim_{\Phi_{\mathbf{r}}}$ are treated analogously, using strict inequality and equality, respectively.

- Normalization of the marginal preference indices for all criteria, so that the indices corresponding to the greatest difference in evaluations of two alternatives on each elementary subcriterion $(g_{\mathbf{t}}(x^*_{\mathbf{t}}) - g_{\mathbf{t}}(x_{\mathbf{t},*}) = \beta_{\mathbf{t}} - \alpha_{\mathbf{t}})$ sum up to 1:

  $$\sum_{\mathbf{t} \in EL} \pi_{\mathbf{t}}(x^*_{\mathbf{t}}, x_{\mathbf{t},*}) = 1 \text{ with } x^*_{\mathbf{t}}, x_{\mathbf{t},*} \in A, \text{ for all } \mathbf{t} \in EL.$$

  We normalize weights of the criteria, so that they sum up to 1. Therefore, each weight is now understood as a maximal share of each elementary subcriterion in the aggregated preference index. Consequently, $k_{\mathbf{t}} = \pi_{\mathbf{t}}(x^*_{\mathbf{t}}, x_{\mathbf{t},*})$, for all $\mathbf{t} \in EL$.

- Restrictions concerning the value of marginal preference indices $\pi_{\mathbf{t}}, \mathbf{t} \in EL$:

- $\pi_{\mathbf{t}}(a,b)$ needs to be equal 0 if a is not better than b on elementary subcriterion $g_{\mathbf{t}}$ by more than the least value of an indifference threshold $q_{\mathbf{t},*}$ allowed by the DM:

$$\pi_{\mathbf{t}}(a,b) = 0 \;\text{ if }\; g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b) \leq q_{\mathbf{t},*}, \;\text{ for all }\; (a,b) \in B, \mathbf{t} \in EL_1;$$

- $\pi_{\mathbf{t}}(a,b)$ needs to be greater than 0 if a is better than b on elementary subcriterion $g_{\mathbf{t}}$ by more than the greatest value of an indifference threshold $q^*_{\mathbf{t}}$ allowed by the DM:

$$\pi_{\mathbf{t}}(a,b) > 0 \;\text{ if }\; g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b) > q^*_{\mathbf{t}}, \;\text{ for all }\; (a,b) \in B, \mathbf{t} \in EL_1;$$

- $\pi_{\mathbf{t}}(a,b)$ needs to be less than the maximal value of the preference index on elementary subcriterion $g_{\mathbf{t}}$ if a is not better than b by more than the least value of a preference threshold $p_{\mathbf{t},*}$ allowed by the DM;

$$\pi_{\mathbf{t}}(a,b) < \pi_{\mathbf{t}}(x^*_{\mathbf{t}}, x_{\mathbf{t},*}), \;\text{ if }\; g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b) < p_{\mathbf{t},*}, \;\text{ for all }\; (a,b) \in B, \mathbf{t} \in EL_1;$$

- $\pi_{\mathbf{t}}(a,b)$ needs to be equal to the maximal value of preference index on elementary subcriterion $g_{\mathbf{t}}$ if a is better than b by more than the greatest value of a preference threshold $p^*_{\mathbf{t}}$ allowed by the DM:

$$\pi_{\mathbf{t}}(a,b) = \pi_{\mathbf{t}}(x^*_{\mathbf{t}}, x_{\mathbf{t},*}), \;\text{ if }\; g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b) \geq p^*_{\mathbf{t}}, \;\text{ for all }\; (a,b) \in B, \mathbf{t} \in EL_1;$$

- $\pi_{\mathbf{t}}(a,b)$ and $\pi_{\mathbf{t}}(b,a)$ need to be equal to 0 if the difference between $g_{\mathbf{t}}(a)$ and $g_{\mathbf{t}}(b)$ is not-significant for the DM:

$$\pi_{\mathbf{t}}(a,b) = 0, \; \pi_{\mathbf{t}}(b,a) = 0 \;\text{ if }\; a \sim_{\mathbf{t}} b, \mathbf{t} \in EL_2;$$

- $\pi_{\mathbf{t}}(a,b)$ needs to be equal to the maximal value of the preference index on criterion $g_{\mathbf{t}}$ if the difference between $g_{\mathbf{t}}(a)$ and $g_{\mathbf{t}}(b)$ is significant for the DM:

$$\pi_{\mathbf{t}}(a,b) = \pi_{\mathbf{t}}(x^*_{\mathbf{t}}, x_{\mathbf{t},*}), \;\text{ if }\; a \succ_{\mathbf{t}} b, \mathbf{t} \in EL_2.$$

- Monotonicity of the functions of marginal preference indices $\pi_{\mathbf{t}}(a,b)$, for all $\mathbf{t} \in EL$:

$$\pi_{\mathbf{t}}(a,b) \geq \pi_{\mathbf{t}}(c,d) \;\text{ if }\; g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b) > g_{\mathbf{t}}(c) - g_{\mathbf{t}}(d),$$

$$\pi_{\mathbf{t}}(a,b) = \pi_{\mathbf{t}}(c,d) \;\text{ if }\; g_{\mathbf{t}}(a) - g_{\mathbf{t}}(b) = g_{\mathbf{t}}(c) - g_{\mathbf{t}}(d).$$

Note that all strict inequalities are transformed into weak inequalities involving an auxiliary variable ε in the set of constraints E^{A^R} in the section 3.2.5.

3.3 Multiple Criteria Hierarchy Process for the Choquet integral

3.3.1 Introduction

In a multiple criteria decision problem (see [29] for a comprehensive state of the art), an alternative a, belonging to a finite set of m alternatives $A = \{a, b, c, \ldots\}$, is evaluated on the basis of a consistent family of n criteria $G = \{g_1, g_2, \ldots, g_n\}$. In our approach we make the assumption that each criterion $g_i \colon A \to \mathbb{R}$ is an interval scale of measurement. From here on, we will use the terms criterion g_i or criterion i interchangeably ($i = 1, 2, \ldots, n$). Without loss of generality, we assume that all the criteria have to be maximized.

The purpose of Multi-Attribute Utility Theory (MAUT) [79] is to represent the preferences of a Decision Maker (DM) on a set of alternatives A by an overall value function $U \colon \mathbb{R}^n \to \mathbb{R}$ with $U(g_1(a), \ldots, g_n(a)) = U(a)$:

- a is indifferent to $b \Leftrightarrow U(a) = U(b)$,
- a is preferred to $b \Leftrightarrow U(a) > U(b)$.

The principal aggregation model of value function is the multiple attribute additive utility [79]:

$$U(a) = u_1(g_1(a)) + u_2(g_2(a)) + \ldots + u_n(g_n(a)) \quad \text{for all} \quad a \in A, \tag{3.13}$$

where u_i are non-decreasing marginal value functions for $i = 1, 2, \ldots, n$.

As it is well-known from the literature, the underlying assumption of the preference independence of the multiple attribute additive utility is unrealistic since in real decision problems criteria often interact. In a decision problem one usually distinguishes between positive and negative interaction among criteria, corresponding to synergy and redundancy among criteria, respectively. In particular, two criteria are synergic (redundant) when the comprehensive importance of these two criteria is greater (smaller) than the sum of importances of the two criteria considered separately.

Within Multiple Criteria Decision Analysis (MCDA), the interaction of criteria has been considered in a decision model based upon a non-additive integral, *i.e.* the Choquet integral [21] (see [39, 43] for a comprehensive survey on the use of non-additive integrals in MCDA, and [44] for a state-of-the-art survey on Choquet and Sugeno integrals).

A great majority of methods designed for MCDA assume that all evaluation criteria are considered at the same level, however, it is often the case that a practical application is imposing a hierarchical structure of criteria. For example, in economic ranking, alternatives may be evaluated on indicators which aggregate evaluations on several sub-indicators, and these sub-indicators may aggregate another set of sub-indicators, etc. In this case, the marginal value functions may refer to all levels of the hierarchy, representing values of particular scores of the alternatives on indicators, sub-indicators, sub-sub-indicators, etc. Considering hierarchical, instead of flat, structure of criteria, permits decomposition of a complex decision problem into smaller problems involving less criteria. To handle the hierarchy of criteria, the Multiple Criteria Hierarchy Process (MCHP) [23] can be applied. The basic idea of the MCHP relies on consideration of preference relations at each node of the hierarchy tree of criteria. This consideration concerns both the phase of eliciting preference information, and the phase of analyzing a final recommendation by the DM. For example, in a decision problem related to evaluation of students, one can say not only that student a is comprehensively preferred to student b, i.e. $a \succ b$, but also that a is comprehensively preferred to b because a is preferred to b on the subset of subjects (subcriteria) related to Mathematics and Physics, i.e. $a \succ_{Mathematics} b$ and $a \succ_{Physics} b$, even if b is preferred to a on subjects related to Humanities, i.e. $b \succ_{Humanities} a$. Moreover, one can also say that, for example, a is preferred to b on the subset of subjects related to Mathematics because, considering Analysis and Algebra as subjects (sub-criteria) related to Mathematics, a is preferred to b on Analysis, i.e. $a \succ_{Analysis} b$, and this is enough to compensate the fact that b is preferred to a on Algebra, i.e. $b \succ_{Algebra} a$.

In this section, we apply the MCHP to the Choquet integral. Let us remark that another approach using the Choquet integral on a hierarchy of criteria has been presented in [120] (see also [128]), where the evaluation of an alternative a with respect to a certain criterion $\mathcal{G}_r$ is based on the Choquet integrals of a with respect to all subcriteria of $\mathcal{G}_r$ from the subsequent level. This means that the Choquet integral of a with respect to $\mathcal{G}_r$ is computed as the Choquet integral of other Choquet integrals, one for each subcriterion of $\mathcal{G}_r$ from the subsequent level. For example, let us consider the evaluation of student a with respect to Science and Humanities, with Mathematics and Physics as subcriteria of Science, and Literature and Philosophy as subcriteria of Humanities. In order to compute the comprehensive Choquet integral of a, one has to compute first the Choquet integral of a with respect to Science and the Choquet integral of a with respect to Humanities. Then, the comprehensive Choquet integral of a is obtained as the Choquet integral of the two Choquet integrals previously computed.

In our approach, we do not consider Choquet integrals resulting from aggregation of Choquet integrals representing evaluations at the subsequent level of the hierarchy. Instead of this, we compute the evaluation of an alternative on a certain criterion of the hierarchy as the Choquet integral of the evaluations of the alternative on all elementary criteria descending to the lowest level from that criterion, using the capacity defined on the whole set of elementary criteria only. Coming back to the above example, the comprehensive evaluation of a is calculated as the Choquet integral of the evaluations of a on all considered elementary subjects, i.e. Mathematics, Physics, Literature and Philosophy. The evaluation with respect to Sciences is obtained as the Choquet integral of the evaluations on Mathematics and Physics only, as well as, the evaluation with respect to Humanities is obtained as the Choquet integral of the evaluations on Literature and Philosophy only. In the approach of [120], the evaluations on Humanities and Sciences are also Choquet integrals, but our approach differs in two aspects: we do not need to define two different capacities to compute the two Choquet integrals, one for Science and one for Humanities; we use the two Choquet integrals on Science and Humanities to order students on the basis of Science and Humanities only, and not to aggregate them in order to get the final comprehensive evaluation.

The section is organized as follows. In Section 3.3.2, we present the basic concepts relative to interaction among criteria and to the Choquet integral. In Section 3.3.3, we describe the MCHP. In Section 3.3.4, we put together the MCHP and the Choquet integral. Section 3.3.5 contains a didactic example in which we describe the application of the new methodology, and we compare it with the approach of [120]. Some conclusions and future directions of research are presented in Section 3.3.6.

3.3.2 The Choquet integral preference model

Let 2^G be the power set of G (i.e. the set of all subsets of G); a fuzzy measure (capacity) on G is defined as a set function $\mu : 2^G \rightarrow [0,1]$ satisfying the following properties:

1a) $\mu(\emptyset) = 0$ and $\mu(G) = 1$ (boundary conditions),

2a) $\forall\, T \subseteq R \subseteq G,\ \ \mu(T) \leq \mu(R)$ (monotonicity condition).

A fuzzy measure is said to be additive if $\mu(T \cup R) = \mu(T) + \mu(R)$, for any $T, R \subseteq G$ such that $T \cap R = \emptyset$. An additive fuzzy measure is determined uniquely by $\mu(\{1\}), \mu(\{2\}) \ldots, \mu(\{n\})$. In fact, in this case, $\forall\ T \subseteq G$, $\mu(T) = \sum_{i \in T} \mu(\{i\})$. In the other cases, we have to define a value $\mu(T)$ for every subset T of G, which are as many as $2^{|G|}$. Therefore, we have to calculate the values of $2^{|G|} - 2$ coefficients, since we know that $\mu(\emptyset) = 0$ and $\mu(G) = 1$.

The Möbius representation of the fuzzy measure μ (see [99]) is defined by the function $m : 2^G \rightarrow \mathbb{R}$ (see [112]) such that:

$$\mu(R) = \sum_{T \subseteq R} m(T). \tag{3.14}$$

Let us observe that if R is a singleton, *i.e.* $R = \{i\}$ with $i = 1, \ldots, n$, then $\mu(\{i\}) = m(\{i\})$. If R is a couple (non-ordered pair) of criteria, i.e. $R = \{i, j\}$, then $\mu(\{i,j\}) = m(\{i\}) + m(\{j\}) + m(\{i,j\})$.

In general, the Möbius representation $m(R)$ is obtained by $\mu(R)$ in the following way:

$$m(R) = \sum_{T \subseteq R} (-1)^{|R \setminus T|} \mu(T). \tag{3.15}$$

In terms of Möbius representation (see [20]), properties **1a)** and **2a)** are, respectively, formulated as:

1b) $m(\emptyset) = 0, \ \sum_{T \subseteq G} m(T) = 1,$

2b) $\forall\, i \in G$ and $\forall R \subseteq G \setminus \{i\}$, $\sum_{T \subseteq R} m(T \cup \{i\}) \geq 0.$

Let us observe that in MCDA, the importance of any criterion $g_i \in G$ should be evaluated considering all its global effects in the decision problem at hand; these effects can be "decomposed" from both theoretical and operational points of view in effects of g_i as single, and in combination with all other criteria. Therefore, a criterion $i \in G$ is important with respect to a fuzzy measure μ not only when it is considered alone, i.e. for the value $\mu(\{i\})$ in itself, but also when it interacts with other criteria from G, i.e. for every value $\mu(T \cup \{i\})$, $T \subseteq G \setminus \{i\}$.

Given $a \in A$ and μ being a fuzzy measure on G, then the *Choquet integral* [21] is defined by:

$$C_\mu(a) = \sum_{i=1}^{n} \left[\left(g_{(i)}(a) \right) - \left(g_{(i-1)}\left(a\right) \right) \right] \mu\left(A_i\right), \tag{3.16}$$

where $_{(\cdot)}$ stands for a permutation of the indices of criteria such that
$g_{(1)}\left(a\right) \leq g_{(2)}\left(a\right) \leq ... \leq g_{(n)}\left(a\right)$, with $A_i = \{(i), ..., (n)\}$, $i = 1, .., n$, and $g_{(0)} = 0$.

The Choquet integral can be redefined in terms of the Möbius representation [37], without re-ordering the criteria, as:

$$C_\mu(a) = \sum_{T \subseteq G} m(T) \min_{i \in T} g_i\left(a\right). \tag{3.17}$$

One of the main drawbacks of the Choquet integral is the necessity of eliciting and giving an adequate interpretation of $2^{|G|} - 2$ parameters. In order to reduce the number of parameters to be computed and to avoid the difficult description of the interactions among criteria, which is not realistic in many applications, the concept of fuzzy k-additive measure has been considered [40].

A *fuzzy measure* is called *k-additive* if $m(T) = 0$ for $T \subseteq G$ such that $|T| > k$ and there exists at least one $T \subseteq G$, with $|T| = k$, such that $m(T) > 0$. We observe that a 1-additive measure is the common additive fuzzy measure. In many real decision problems, it suffices to consider 2-additive measures. In this case, positive and negative interactions between two criteria are modeled without considering the interaction among any n-tuples (with $n > 2$) of criteria. From the point of view of MCDA, the use of 2-additive measures is justified by observing that the information on the importance of the single criteria and the interactions between two criteria are noteworthy. Moreover, it could be not easy or not straightforward for the DM to provide information on the interactions among three or more criteria during the decision procedure. From a computational point of view, the interest in the 2-additive measures lies in the fact that any decision model needs to evaluate a number $n + \binom{n}{2}$ of parameters (in terms of Möbius representation, a value $m(\{i\})$ for every criterion i and a value $m(\{i, j\})$ for every couple of distinct criteria $\{i, j\}$). With respect to a 2-additive fuzzy measure, the inverse transformation to obtain the fuzzy measure $\mu(R)$ from the Möbius representation is defined as:

$$\mu(R) = \sum_{i \in R} m(\{i\}) + \sum_{\{i,j\} \subseteq R} m(\{i, j\}), \ \forall R \subseteq G. \tag{3.18}$$

With regard to 2-additive measures, properties **1b)** and **2b)** have, respectively, the following formulations:

1c) $m(\emptyset) = 0, \ \sum_{i \in G} m(\{i\}) + \sum_{\{i,j\} \subseteq G} m(\{i, j\}) = 1,$

2c) $\begin{cases} m(\{i\}) \geq 0, \ \forall i \in G, \\ m(\{i\}) + \sum_{j \in T} m(\{i, j\}) \geq 0, \ \forall i \in G \text{ and } \forall \, T \subseteq G \setminus \{i\}, \, T \neq \emptyset. \end{cases}$

In this case, the representation of the Choquet integral of $a \in A$ is given by:

$$C_\mu(a) \ = \ \sum_{\{i\} \subseteq G} m(\{i\}) (g_i(a)) + \sum_{\{i,j\} \subseteq G} m(\{i, j\}) \min\{g_i(a), g_j(a)\}. \tag{3.19}$$

Finally, we recall the definitions of the importance and interaction indices for couples of criteria. The Shapley value [113] expressing the importance of criterion $i \in G$, is given by:

$$\varphi(\{i\}) = \sum_{T \subseteq G \colon i \notin T} \frac{(|G \setminus T| - 1)!|T|!}{|G|!} \cdot [\mu(T \cup \{i\}) - \mu(T)], \tag{3.20}$$

while the *interaction index* [90] expressing the sign and the magnitude of the synergy in a couple of criteria $\{i,j\} \subseteq G$, is given by

$$\varphi(\{i,j\}) = \sum_{T \subseteq G:\, i,j \notin T} \frac{(|G \setminus T| - 2)!|T|!}{(|G| - 1)!} \cdot \tau(T,i,j), \tag{3.21}$$

where $\tau(T,i,j) = [\mu(T \cup \{i,j\}) - \mu(T \cup \{i\}) - \mu(T \cup \{j\}) + \mu(T)]$.

In case of 2-additive capacities, the Shapley value and the interaction index can be expressed as follows:

$$\varphi(\{i\}) = m(\{i\}) + \sum_{j \in G \setminus \{i\}} \frac{m(\{i,j\})}{2}, \;\; i \in G, \tag{3.22}$$

$$\varphi(\{i,j\}) = m(\{i,j\}). \tag{3.23}$$

3.3.3 Multiple Criteria Hierarchy Process (MCHP)

In MCHP, a set $\mathcal{G}$ of hierarchically ordered criteria is considered, i.e. all criteria are not considered at the same level, but they are distributed over l different levels (see Figure 3.13). At level 1, there are first level criteria called root criteria. Each root criterion has its own hierarchy tree. The leaves of each hierarchy tree are at the last level l and they are called elementary subcriteria. Thus, in graph theory terms, the whole hierarchy is a forest. We will use the following notation:

- l is the number of levels in the hierarchy of criteria,
- $\mathcal{G}$ is the set of all criteria at all considered levels,
- $\mathcal{I}_{\mathcal{G}}$ is the set of indices of particular criteria representing position of criteria in the hierarchy,
- m is the number of the first level criteria, $G_1, \ldots, G_m$,
- $G_{\mathbf{r}} \in \mathcal{G}$, with $\mathbf{r} = (i_1, \ldots, i_h) \in \mathcal{I}_{\mathcal{G}}$, denotes a subcriterion of the first level criterion G_{i_1} at level h; the first level criteria are denoted by G_{i_1}, $i_1 = 1, \ldots, m$,
- $n(\mathbf{r})$ is the number of subcriteria of $G_{\mathbf{r}}$ in the subsequent level, i.e. the direct subcriteria of $G_{\mathbf{r}}$ are $G_{(\mathbf{r},1)}, \ldots, G_{(\mathbf{r},n(\mathbf{r}))}$,
- $g_{\mathbf{t}} : A \rightarrow \mathbb{R}$, with $\mathbf{t} = (i_1, \ldots, i_l) \in \mathcal{I}_{\mathcal{G}}$, denotes an elementary subcriterion of the first level criterion G_{i_1}, i.e. a criterion at level l of the hierarchy tree of G_{i_1},

- EL is the set of indices of all elementary subcriteria:

$$EL = \{\mathbf{t} = (i_1, \ldots, i_l) \in \mathcal{I}_\mathcal{G}\} \quad \text{where} \begin{cases} i_1 = 1, \ldots, m \\ i_2 = 1, \ldots, n(i_1) \\ \cdots\cdots \\ i_l = 1, \ldots, n(i_1, \ldots, i_{l-1}) \end{cases}$$

- $E(G_\mathbf{r})$ is the set of indices of elementary subcriteria descending from $G_\mathbf{r}$, i.e.

$$E(G_\mathbf{r}) = \{(\mathbf{r}, i_{h+1}, \ldots, i_l) \in \mathcal{I}_\mathcal{G}\} \quad \text{where} \begin{cases} i_{h+1} = 1, \ldots, n(\mathbf{r}) \\ \cdots\cdots \\ i_l = 1, \ldots, n(\mathbf{r}, i_{h+1}, \ldots, i_{l-1}) \end{cases}$$

  thus, $E(G_\mathbf{r}) \subseteq EL$; in the case $G_\mathbf{r} \in EL$, then $E(G_\mathbf{r}) = G_\mathbf{r}$,

- when $\mathbf{r} = 0$, then by $G_\mathbf{r} = G_0$, we mean the entire set of criteria and not a particular criterion or subcriterion; in this particular case, we have $E(G_0) = EL$,

- given $\mathcal{F} \subseteq \mathcal{G}$, $E(\mathcal{F}) = \cup_{G_\mathbf{r} \in \mathcal{F}} E(G_\mathbf{r})$, that is $E(\mathcal{F})$ is composed by all elementary subcriteria descending from at least one criterion in $\mathcal{F}$,

- given $G_\mathbf{r} \in \mathcal{G}$, $\mathbf{r} \in \mathcal{I}_\mathcal{G} \cap \mathbb{N}^h$ ($G_\mathbf{r}$ is a criterion at level h), $1 \leq h < l$, and $k \in \{h+1, \ldots, l\}$, we define:

$$\mathcal{G}_\mathbf{r}^k = \left\{G_{(\mathbf{r},w)} \in \mathcal{G} : (\mathbf{r}, w) \in \mathcal{I}_\mathcal{G} \cap \mathbb{N}^k\right\}$$

  being the set of all subcriteria of criterion $G_\mathbf{r}$ at level k. (For example, in Figure 3.13, we have that

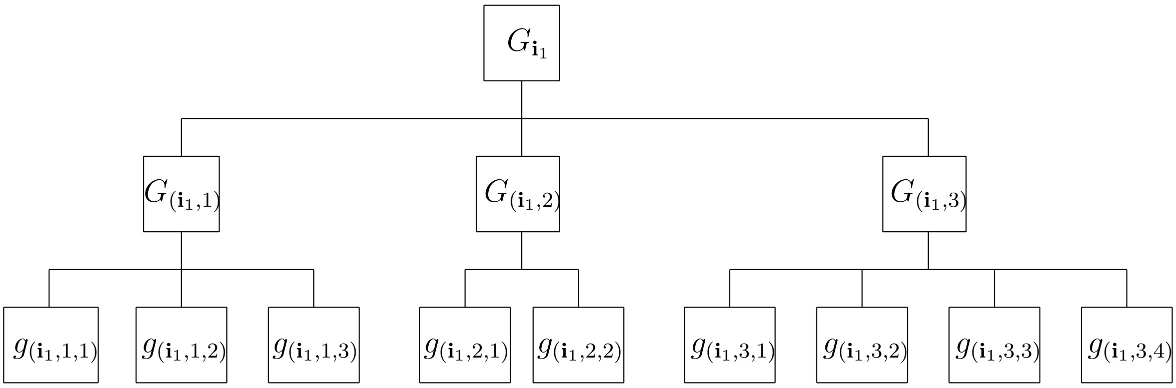
$$\mathcal{G}_{\mathbf{i_1}}^2 = \left\{G_{(\mathbf{i_1},1)}, G_{(\mathbf{i_1},2)}, G_{(\mathbf{i_1},3)}\right\} \quad \text{and} \quad \mathcal{G}_{(\mathbf{i_1},2)}^3 = \left\{g_{(\mathbf{i_1},2,1)}, g_{(\mathbf{i_1},2,2)}\right\})$$

Each alternative $a \in A$ is evaluated directly on the elementary subcriteria only, such that to each alternative $a \in A$ there corresponds a vector of evaluations:

$$(g_{\mathbf{t_1}}(a), \ldots, g_{\mathbf{t}_n}(a)), \qquad n = |EL|.$$

Within MCHP, in each node $G_\mathbf{r} \in \mathcal{G}$ of the hierarchy tree there exists a preference relation $\succsim_\mathbf{r}$ on A, such that for all $a, b \in A$, $a \succsim_\mathbf{r} b$ means "a is at least as good as b on subcriterion $G_\mathbf{r}$". In the particular case where $G_\mathbf{r} = g_\mathbf{t}, \mathbf{t} \in EL$, $a \succsim_\mathbf{t} b$ holds if $g_\mathbf{t}(a) \geq g_\mathbf{t}(b)$.

Figure 3.13: Hierarchy of criteria for the first level (root) criterion $G_{\mathbf{i}_1}$



3.3.4 Multiple Criteria Hierarchy Process for Choquet integral preference model

In this article, we will aggregate the evaluations of alternative $a \in A$ with respect to the elementary subcriteria $g_{\mathbf{t}}, \mathbf{t} \in EL$, using the Choquet integral as follows.

On the basis of a capacity μ defined on the power set of EL, for all $a, b \in A$, $a \succsim b$ if $C_\mu(a) \geq C_\mu(b)$, where $C_\mu(a)$ and $C_\mu(b)$ are the Choquet integrals with respect to μ of the vectors $[g_{\mathbf{t}}(a), t \in EL]$ and $[g_{\mathbf{t}}(b), t \in EL]$, respectively.

For all $G_{\mathbf{r}} \in \mathcal{G}$, $\mathbf{r} \in \mathcal{I}_{\mathcal{G}} \cap \mathbb{N}^h$ ($G_{\mathbf{r}}$ is a criterion at level h), $h = 1, \ldots, l-1$ and for all $k = h+1, \ldots, l$, we can define the following capacity:

$$\mu_{\mathbf{r}}^k : 2^{\mathcal{G}_{\mathbf{r}}^k} \rightarrow [0, 1]$$

such that, for all $\mathcal{F} \subseteq \mathcal{G}_{\mathbf{r}}^k$, we have that

$$\mu_{\mathbf{r}}^k(\mathcal{F}) = \frac{\mu(E(\mathcal{F}))}{\mu(E(G_{\mathbf{r}}))} \tag{3.24}$$

In this way, $\mu_{\mathbf{r}}^k$ is a capacity defined on the power set of $\mathcal{G}_{\mathbf{r}}^k$, that could be computed using the capacity μ defined on the power set of EL.

In the following, we shall write $\mu_{\mathbf{r}}$ instead of $\mu_{\mathbf{r}}^l$.

For all $a, b \in A$, $a \succsim_{\mathbf{r}} b$ if $C_{\mu_{\mathbf{r}}}(a) \geq C_{\mu_{\mathbf{r}}}(b)$, where $C_{\mu_{\mathbf{r}}}(a)$ and $C_{\mu_{\mathbf{r}}}(b)$ are the Choquet integrals with respect to $\mu_{\mathbf{r}}$ of the vectors $[g_{\mathbf{t}}(a), t \in E(G_{\mathbf{r}})]$ and $[g_{\mathbf{t}}(b), t \in E(G_{\mathbf{r}})]$, respectively. Observe that for all $a \in A$,

$$C_{\mu_{\mathbf{r}}}(a) = \frac{C_\mu(a_{\mathbf{r}})}{\mu(E(G_{\mathbf{r}}))} \tag{3.25}$$

where $a_{\mathbf{r}}$ is a fictitious alternative having the same evaluations of a on elementary criteria from $E(G_{\mathbf{r}})$ and null evaluation on criteria from outside $E(G_{\mathbf{r}})$, i.e. $g_{\mathbf{s}}(a_{\mathbf{r}}) = g_{\mathbf{s}}(a)$ if $\mathbf{s} \in E(G_{\mathbf{r}})$ and $g_{\mathbf{s}}(a_{\mathbf{r}}) = 0$ if $\mathbf{s} \notin E(G_{\mathbf{r}})$.

The Shapley value expressing the importance of criterion $G_{(\mathbf{r},w)} \in \mathcal{G}^k_{\mathbf{r}}$ being thus a subcriterion of $G_{\mathbf{r}}$ at level k is:

$$\varphi^k_{\mathbf{r}}(G_{(\mathbf{r},w)}) = \sum_{T \subseteq \mathcal{G}^k_{\mathbf{r}} \setminus \{G_{(\mathbf{r},w)}\}} \frac{(|\mathcal{G}^k_{\mathbf{r}} \setminus T| - 1)!|T|!}{|\mathcal{G}^k_{\mathbf{r}}|!} \cdot \left[\mu^k_{\mathbf{r}}(T \cup \{G_{(\mathbf{r},w)}\}) - \mu^k_{\mathbf{r}}(T)\right] \tag{3.26}$$

while the interaction index expressing the sign and the magnitude of the synergy in a couple of criteria $G_{(\mathbf{r},w_1)}, G_{(\mathbf{r},w_2)} \in \mathcal{G}^k_{\mathbf{r}}$ is given by:

$$\varphi^k_{\mathbf{r}}(G_{(\mathbf{r},w_1)}, G_{(\mathbf{r},w_2)}) = \sum_{T \subseteq \mathcal{G}^k_{\mathbf{r}} \setminus \{G_{(\mathbf{r},w_1)}, G_{(\mathbf{r},w_2)}\}} \frac{(|\mathcal{G}^k_{\mathbf{r}} \setminus T| - 2)!|T|!}{(|\mathcal{G}^k_{\mathbf{r}}| - 1)!} \cdot \tau^k_{\mathbf{r}}(T, G_{(\mathbf{r},w_1)}, G_{(\mathbf{r},w_2)}) \tag{3.27}$$

where

$$\tau^k_{\mathbf{r}}(T, A, B) = \left[\mu^k_{\mathbf{r}}(T \cup \{A, B\}) - \mu^k_{\mathbf{r}}(T \cup \{A\}) - \mu^k_{\mathbf{r}}(T \cup \{B\}) + \mu^k_{\mathbf{r}}(T)\right].$$

In case the capacity μ on $\{g_{\mathbf{t}}, \mathbf{t} \in EL\}$ is 2-additive, the Shapley value $\varphi^k_{\mathbf{r}}(G_{(\mathbf{r},w)})$ and the interaction index $\varphi^k_{\mathbf{r}}(G_{(\mathbf{r},w_1)}, G_{(\mathbf{r},w_2)})$, with $G_{(\mathbf{r},w)}, G_{(\mathbf{r},w_1)}, G_{(\mathbf{r},w_2)} \in \mathcal{G}^k_{\mathbf{r}}$, can be expressed as follows:

$$\varphi^k_{\mathbf{r}}(G_{(\mathbf{r},w)}) =$$

$$= \left\{ \sum_{\mathbf{t} \in E(G_{(\mathbf{r},w)})} m(g_{\mathbf{t}}) + \sum_{\mathbf{t_1},\mathbf{t_2} \in E(G_{(\mathbf{r},w)})} m(g_{\mathbf{t_1}}, g_{\mathbf{t_2}}) + \sum_{\substack{\mathbf{t_1} \in E(G_{(\mathbf{r},w)}) \\ \mathbf{t_2} \in E(\mathcal{G}^k_{\mathbf{r}} \setminus \{G_{(\mathbf{r},w)}\})}} \frac{m(g_{\mathbf{t_1}}, g_{\mathbf{t_2}})}{2} \right\} \cdot \frac{1}{\mu(E(G_{\mathbf{r}}))} \tag{3.28}$$

$$\varphi^k_{\mathbf{r}}(G_{(\mathbf{r},w_1)}, G_{(\mathbf{r},w_2)}) = \left\{ \sum_{\substack{\mathbf{t_1} \in E(G_{(\mathbf{r},w_1)}), \\ \mathbf{t_2} \in E(G_{(\mathbf{r},w_2)})}} m(g_{\mathbf{t_1}}, g_{\mathbf{t_2}}) \right\} \cdot \frac{1}{\mu(E(G_{\mathbf{r}}))}. \tag{3.29}$$

Taking into account the expression of the Shapley index in equation (3.26) and $G_{\mathbf{s}_1}, G_{\mathbf{s}_2} \in G^k_{\mathbf{r}_1} \cap G^k_{\mathbf{r}_2}$ (that is $G_{\mathbf{s}_1}$ and $G_{\mathbf{s}_2}$ are subcriteria of both $G_{\mathbf{r}_1}$ and $G_{\mathbf{r}_2}$ located at level k), and supposing, without loss of generality, that $\mathbf{r}_2 = (\mathbf{r}_1, w)$ (that is $G_{\mathbf{r}_2}$ is a subcriterion of $G_{\mathbf{r}_1}$), it is worth noting that the following inequalities could be verified:

$$\varphi^k_{\mathbf{r}_1}(G_{\mathbf{s}_1}) > \varphi^k_{\mathbf{r}_1}(G_{\mathbf{s}_2}) \quad \text{and} \quad \varphi^k_{\mathbf{r}_2}(G_{\mathbf{s}_1}) < \varphi^k_{\mathbf{r}_2}(G_{\mathbf{s}_2}) \quad \text{(or viceversa).}$$

This means that the importance of the criterion $G_{\mathbf{s}_1}$ is greater than the importance of the criterion $G_{\mathbf{s}_2}$ if they are considered as subcriteria of $G_{\mathbf{r}_1}$, but the importance of $G_{\mathbf{s}_2}$ is greater than importance of $G_{\mathbf{s}_1}$ if they are considered as subcriteria of $G_{\mathbf{r}_2}$. In fact, in the computation of $\varphi^k_{\mathbf{r}_1}(G_{\mathbf{s}_1})$ we take into account not only the interactions between the elementary criteria descending from $G_{\mathbf{s}_1}$ but also the interactions between elementary criteria descending from $G_{\mathbf{s}_1}$ and elementary criteria descending from $G_{\mathbf{r}_1}$. Because we have supposed that $G_{\mathbf{r}_2}$ is a subcriterion of $G_{\mathbf{r}_1}$, and consequently $E\left(G_{\mathbf{r}_2}\right) \subseteq E\left(G_{\mathbf{r}_1}\right)$, in the computation of $\varphi^k_{\mathbf{r}_1}(G_{\mathbf{s}_1})$ we take into account more interactions than those considered in the computation of $\varphi^k_{\mathbf{r}_2}(G_{\mathbf{s}_1})$. For example, evaluating students with respect to Science according to their scores on Mathematics and Physics, and with respect to Humanities according to their scores on Literature and Philosophy, one could consider Mathematics more important than Physics whithin Sciences and Literature more important than Philosophy within Humanities. However, taking into consideration that there is a great synergy between Philosophy and Physics, at the comprehensive level, Physics can be considered more important than Mathematics, as well as, at the same level, Philosophy can be considered more important than Literature.

Another interesting situation that can happen with respect to preferences represented by the Choquet integral in case of the hierarchy of criteria is the following. One can have that alternative a is evaluated better than alternative b with respect to all the subcriteria $G_{(\mathbf{r},1)}, \ldots, G_{(\mathbf{r},n(\mathbf{r}))}$ of criterion $G_{\mathbf{r}} \in \mathcal{G}$ from the subsequent level, and, nevertheless b can be evaluated better than a on criterion $G_{\mathbf{r}}$. For example, student a could be evaluated better than b on Science and Humanities but b could be evaluated better than a at the comprehensive level. This is due to the fact that when evaluating a student with respect to Science, we take into account only the interactions among subcriteria of Science as well as in the evaluation of a student with respect to Humanities we take into account only the interactions among subcriteria of Humanities. On the other hand, when evaluating a student comprehensively, we take into account also the interactions among the subcriteria of Science and subcriteria of Humanities. Thus, if there is a strong synergy between one subject from Science (for example, Physics) and another subject from Humanities (for example, Philosophy), and b is better evaluated than a in those subjects, this can result in the overall preference of b over a.

We shall show these situations in the didactic example presented in the next section.

3.3.5 A didactic example

Let us consider a set of nine students $A = \{a, b, c, d, e, f, g, h, k\}$ evaluated on the basis of two macro-subjects: Science and Humanities. Science has two sub-subjects: Mathematics and Physics, while Humanities has two sub-subjects: Literature and Philosophy. The number of levels considered is two.

In terms of notation, we have $\mathcal{G} = \left\{G_{\mathbf{1}}, G_{\mathbf{2}}, G_{(\mathbf{1},1)}, G_{(\mathbf{1},2)}, G_{(\mathbf{2},1)}, G_{(\mathbf{2},2)}\right\}$, and the elements of $\mathcal{G}$ denote respectively, Science, Humanities, Mathematics, Physics, Literature and Philosophy. The students are evaluated on the basis of the elementary criteria only; such evaluations are shown in Table 3.15(a).

Table 3.15: Matrix evaluation and Möbius measures

(a) Matrix evaluation

	Science		**Humanities**	
Student	Mathematics	Physics	Literature	Philosophy
a	18	18	12	12
b	16	16	16	16
c	14	14	18	18
d	18	12	16	16
e	15	15	18	14
f	18	14	14	18
g	15	17	18	16
h	10	20	10	20
k	14	14	14	14

(b) Möbius measures

$m(G_{(\mathbf{1},1)})$	0.29
$m(G_{(\mathbf{1},2)})$	0.19
$m(G_{(\mathbf{2},1)})$	0.29
$m(G_{(\mathbf{2},2)})$	0.19
$m(G_{(\mathbf{1},1)}, G_{(\mathbf{1},2)})$	-0.1
$m(G_{(\mathbf{1},1)}, G_{(\mathbf{2},1)})$	0
$m(G_{(\mathbf{1},1)}, G_{(\mathbf{2},2)})$	0
$m(G_{(\mathbf{1},2)}, G_{(\mathbf{2},1)})$	0
$m(G_{(\mathbf{1},2)}, G_{(\mathbf{2},2)})$	0.24
$m(G_{(\mathbf{2},1)}, G_{(\mathbf{2},2)})$	-0.1

In the following, we shall consider a 2-additive capacity determined by the Möbius measures in Table 3.15(b).

Applying the expression (3.25) of the hierarchical Choquet integral introduced in Section 3.3.4, we can compute the evaluation of every student with respect to macro-subjects Science ($G_{\mathbf{1}}$) and Humanities ($G_{\mathbf{2}}$), while using the expression (3.19) of the Choquet integral, we can compute the evaluation of every student with respect to the whole hierarchy of criteria (see Table 3.16).

For example, looking at the first three rows in Table 3.16, we get:

- the Choquet integral of a with respect to Science is equal to 18 and it is computed considering the fictitious alternative $a_{\mathbf{1}}$ having the same evaluations of a on the elementary criteria descending from Science, and null evaluations on all other elementary criteria,

- the Choquet integral of a with respect to Humanities is equal to 12 and it is computed considering the fictitious alternative $a_{\mathbf{2}}$ having the same evaluations of a on the elementary criteria descending from Humanities, and null evaluations on all other elementary criteria,

- the Choquet integral of a with respect to the whole hierarchy of criteria is equal to 14.28 and it is computed considering the evaluations of a on all elementary criteria.

Hereafter we underline the very interesting inversion of preference regarding alternatives h and k. In fact, looking at Table 3.16, we can observe that k is better than h with respect to both macro-subjects Science and Humanities ($C_{\mu_1}(k) > C_{\mu_1}(h)$ and $C_{\mu_2}(k) > C_{\mu_2}(h)$) but h is comprehensively better than k ($C_{\mu}(h) > C_{\mu}(k)$). The reason of this inversion of preference is explained considering that in the computation of $C_{\mu_1}(\cdot)$ and $C_{\mu_2}(\cdot)$ we take into account only the interaction between elementary criteria descending from Science and Humanities respectively, while in the computation of the comprehensive Choquet integral $C_{\mu}(\cdot)$ we take into account the possible interactions between all elementary criteria in the hierarchy.

Table 3.16: Choquet integrals with respect to the macro-subjects Science and Humanities and with respect to the whole hierarchy of criteria

	Science		**Humanities**			
	Mathematics	Physics	Literature	Philosophy		Choquet integrals
a_1	18	18	0	0	$C_{\mu_1}(a)$	18
a_2	0	0	12	12	$C_{\mu_2}(a)$	12
a	18	18	12	12	$C_{\mu}(a)$	14.28
b_1	16	16	0	0	$C_{\mu_1}(b)$	16
b_2	0	0	16	16	$C_{\mu_2}(b)$	16
b	16	16	16	16	$C_{\mu}(b)$	16
c_1	14	14	0	0	$C_{\mu_1}(c)$	14
c_2	0	0	18	18	$C_{\mu_2}(c)$	18
c	14	14	18	18	$C_{\mu}(c)$	15.52
d_1	18	12	0	0	$C_{\mu_1}(d)$	16.57
d_2	0	0	16	16	$C_{\mu_2}(d)$	16
d	18	12	16	16	$C_{\mu}(d)$	15.26
e_1	15	15	0	0	$C_{\mu_1}(e)$	15
e_2	0	0	18	14	$C_{\mu_2}(e)$	17.05
e	15	15	18	14	$C_{\mu}(e)$	15.54
f_1	18	14	0	0	$C_{\mu_1}(f)$	17.05
f_2	0	0	14	18	$C_{\mu_2}(f)$	16
f	18	14	14	18	$C_{\mu}(f)$	15.92
g_1	15	17	0	0	$C_{\mu_1}(g)$	16
g_2	0	0	18	16	$C_{\mu_2}(g)$	17.52
g	15	17	18	16	$C_{\mu}(g)$	16.58
h_1	10	20	0	0	$C_{\mu_1}(h)$	13.5
h_2	0	0	10	20	$C_{\mu_2}(h)$	13.5
h	10	20	10	20	$C_{\mu}(h)$	15.06
k_1	14	14	0	0	$C_{\mu_1}(k)$	14
k_2	0	0	14	14	$C_{\mu_2}(k)$	14
k	14	14	14	14	$C_{\mu}(k)$	14

By considering the capacities on the elementary criteria displayed in Table 3.15(b) and adopting the expression (3.28) defined in Section 3.3.4, we compute the Shapley values of the elementary criteria $G_{(\mathbf{r},i)}$ with respect to their relative parent criterion $G_{\mathbf{r}}$ (see Table 3.17(a)). Then the overall Shapley values of the elementary criteria (i.e. with respect to $G_{\mathbf{0}}$) are calculated and shown in Table 3.17(b). Finally, the Shapley values of subcriteria $G_{\mathbf{1}}$ (Science) and $G_{\mathbf{2}}$ (Humanities) and their interaction index (see the expression (3.29) introduced in Section 3.3.4) are computed and displayed in Table 3.18.

As it has been announced in Section 3.3.4, in this example, Mathematics is more important than Physics, when they are considered as subcriteria of Science (see Table 3.17(a)) and, conversely, Physics is more important than Mathematics when they are considered as subcriteria of the whole set of criteria $G_{\mathbf{0}}$ (see Table 3.17(b)).

Table 3.17: Shapley values

(a) Shapley values of every elementary criterion with respect to every macro-subject $G_{\mathbf{r}}$

	Science		**Humanities**	
	Mathematics	Physics	Literature	Philosophy
$\varphi^k_{\mathbf{r}}(G_{(\mathbf{r},w)})$	0.63	0.36	0.63	0.36

(b) Shapley values of the elementary criteria

	$\varphi^k_{\mathbf{r}}(G_{(\mathbf{r},w)})$
Mathematics	0.24
Physics	0.26
Literature	0.24
Philosophy	0.26

Table 3.18: The Shapley values and interaction index of Science (G_1) and Humanities (G_2)

	$\varphi^k_{\mathbf{r}}(G_{(\mathbf{r},w)})$
Science	0.5
Humanities	0.5
	$\varphi^k_{\mathbf{r}}(G_{(\mathbf{r},w_1)}, G_{(\mathbf{r},w_2)})$
Science and **Humanities**	0.24

In order to illustrate the difference between our approach and that of [120], in the following we shall compute the comprehensive evaluations of student g following the approach of [120]. At first, we need to define a capacity for each node of the hierarchy of criteria, which is not an elementary criterion. Because in our didactic example the hierarchy is composed of three nodes being different from the elementary criteria, we need to define three capacities, $\mu_{\{Sci\}}$, $\mu_{\{Hum\}}$, and $\mu_{\{Sci,Hum\}}$ on $\{Math, Phy\}$, $\{Lit, Phi\}$ and $\{Sci, Hum\}$, respectively. Let us suppose that the capacities are defined using the Möbius measures shown in Table 3.19.

Table 3.19: Möbius measures

Science		Humanities		Science,Humanities	
$m(\{Math\})$	0.7	$m(\{Lit\})$	0.5	$m(\{Sci\})$	0.4
$m(\{Phy\})$	0.5	$m(\{Phi\})$	0.6	$m(\{Hum\})$	0.4
$m(\{Math, Phy\})$	-0.2	$m(\{Lit, Phi\})$	-0.1	$m(\{Sci, Hum\})$	0.2

Computing the Choquet integral of $g = (15, 17, 18, 16)$ with respect to criteria Science and Humanities using the capacities $\mu_{\{Sci\}}$ and $\mu_{\{Hum\}}$, we get:

$$C_{\{Sci\}}(g) = 15m(\{Math\}) + 17m(\{Phy\}) + 15m(\{Math, Phys\}) = 16$$

$$C_{\{Hum\}}(g) = 18m(\{Lit\} + 16m(\{Phi\} + 16m(\{Lit, Phi\} = 17$$

Then, the comprehensive Choquet integral of g is obtained by aggregating the evaluations $(C_{\{Sci\}}(g), C_{\{Hum\}}(g))$ using the capacity $\mu_{\{Sci,Hum\}}$:

$$C_{\{Sci,Hum\}}(g) = C_{\{Sci\}}(g)m(\{Sci\}) + C_{\{Hum\}}(g)m(\{Hum\})+$$

$$+min(C_{\{Sci\}}(g), C_{\{Hum\}}(g))m(\{Sci, Hum\}) = 16.4.$$

The remarkable difference between the method presented in [120] and our approach, is that in the first one, one capacity has to be defined with respect to each node of the hierarchy of criteria being different from the elementary criteria (for example in the didactic example we have defined three different capacities) while in our approach one needs to define only one capacity on the set of all elementary criteria, and the capacities at higher levels are calculated according to formulas given in Section 3.3.4.

3.3.6 Conclusions

We have proposed the application of the Multiple Criteria Hierarchy Process (MCHP) to a preference model expressed in terms of Choquet integral, in order to deal with interaction among criteria. Application of the MCHP to the Choquet integral permits the handling of importance and interactions of criteria with respect to any subcriterion of the hierarchy. To apply the MCHP to the Choquet integral in real world problems, it is necessary to elicit preference model parameters, which in this case are the non-interactive weights represented by a capacity. The added value of the MCHP is that it permits the DM expressing the preference information related to any criterion of the hierarchy. When MCHP is combined with a disaggregation procedure, the DM can say, for example, that student a is globally preferred to student b, but he can also say that student c is better than student d in Humanities. DM can also say that criterion Science is more important than Humanities, or that the interaction between Physics and Philosophy is greater than the interaction between Mathematics and Literature. Many multicriteria disaggregation procedures have been proposed to infer a capacity from those types of preference information, however, without considering the hierarchy of criteria (see, for example, [86]). Recently, a new multicriteria disaggregation method has been proposed to take into account that, in general, more than one capacity is able to represent the preference expressed by the DM: Non Additive Robust Ordinal Regression (NAROR) [7]. NAROR considers all the capacities that are compatible with the preference information given by the DM, adopting the concepts of possible and necessary preference introduced in [55]. In simple words, a is necessarily or possibly preferred to b, if it is preferred for all compatible capacities or for at least one compatible capacity, respectively. In our opinion, application of NAROR to MCHP for the Choquet integral will permit to take into account interaction among hierarchically structured criteria in a very efficient way, enabling the handling of many complex real world problems.

Chapter 4

Final remarks

The objective of a Multiple Criteria Decision Aiding (MCDA) methodology is to provide a set of useful recommendations to bring the Decision Maker (DM) to make the "best" decision. Looking at the evaluations of the alternatives with respect to all considered criteria, the only information one can obtain is the dominance relation but, generally, this is very poor. For this reason, in order to gain a more insight into the problem at hand, the Multiple Attribute Utility Theory (MAUT) and the outranking methods are generally used. Two techniques, direct and indirect, are known in literature to get the parameters useful to implement the different methodologies. The direct one consists of asking the DM to directly provide all of the information on the preferential parameters, while the indirect one consists of asking the DM to provide preference information from which it is possible to elicit the preferential parameters. Generally, more than one set of parameters is compatible with the preference information provided by the DM. Robust Ordinal Regression (ROR) methodologies take into account not only one but the whole family of sets of parameters compatible with the preference information provided by the DM. On this basis ROR defines a necessary and a possible preference relation. The necessary preference relation holds between alternatives a and b, if a is at least as good as b for all sets of parameters compatible with the preferences provided by the DM while the possible preference relation holds between a and b if a is at least as good as b for at least one set of parameters. Another family of methodologies aiming to explore the whole set of parameters of a preference model is the Stochastic Multiobjective Acceptability Analysis (SMAA) that takes into account imprecision or lack of data considering probability distributions over the space of multiple criteria evaluations and over the space of preferential parameters.

The contributions given by the thesis regard Hierarchy of Criteria, Interaction of Criteria and Hierarchy of Criteria in case of interacting criteria:

- Regarding the first point, we observed that in more complex decision making problems, all evaluation criteria are not at the same level, but they are organized in a hierarchical way. This means that it is possible to define some root criteria, some subcriteria descending from each root criterion, other subsubcriteria descending from each subcriterion and so on. Basing on the hierarchy of criteria, we have introduced the Multiple Criteria Hierarchy Process (MCHP). MCHP has been applied both to MAUT and to outranking methods, in two papers: MCHP in Robust Ordinal Regression [23] and MCHP with ELECTRE and PROMETHEE [24].

- MAUT and outranking methods are based on the independence between criteria. In the case the criteria are not independent because there is a positive or a negative interaction between them, multicriteria evaluations can be aggregated using non-additive integrals or an enriched utility function, as that one used in UTA^{GMS}-INT method [58], in which, beyond the marginal utilities related to each criterion, there are components representing a bonus or a malus for synergetic or redundant criteria. In the thesis we have considered two very well known non-additive integrals: the Choquet integral and its extension to a bipolar scale that is the bipolar Choquet integral.
  Two contributions are based on non-additive integrals: the SMAA-Choquet method [3] putting together the Choquet integral and the SMAA family, and the Bipolar PROMETHEE method [22] extending the PROMETHEE I and II methods in case of interacting criteria on a bipolar scale.
  Based on a utility function taking into account interaction of criteria using bonus and malus, is instead the MUSA-int method [5] that extends the customer satisfaction analysis method MUSA.

- Putting together the two key points of the thesis, that is the hierarchy of criteria and the interaction between criteria, we developed the MCHP for the Choquet integral [4] in which we extended the Choquet integral to the case of a hierarchy of interacting criteria.

Bibliography

[1] L.S. Aiken and S.G. West. *Multiple regression: Testing and interpreting interactions.* Sage, Newbury Park, CA, 1991.

[2] J. Almeida Dias, J. Figueira, and B. Roy. A multiple criteria sorting method where each category is characterized by several reference actions: The Electre Tri-nC method. *European Journal of Operational Research*, 217(3):567–579, 2012.

[3] A. Angilella, S. Corrente, and S. Greco. SMAA-Choquet: Stochastic Multicriteria Acceptability Analysis for the Choquet Integral. In S. Greco et al., editor, *IPMU 2012, Part IV, CCIS 300*, pages 248–257. Springer-Verlag Berlin Heidelberg 2012, 2012.

[4] A. Angilella, S. Corrente, S. Greco, and R. Słowiński. Multiple criteria hierarchy process for the Choquet integral. *Accepted to the EMO 2013 conference.*

[5] A. Angilella, S. Corrente, S. Greco, and R. Słowiński. MUSA-INT: Multicriteria customer satisfaction analysis with interacting criteria. *OMEGA. Submitted*, 2012.

[6] S. Angilella, S. Greco, F. Lamantia, and B. Matarazzo. Assessing non-additive utility for multicriteria decision aid. *European Journal of Operational Research*, 158(3):734–744, 2004.

[7] S. Angilella, S. Greco, and B. Matarazzo. Non-additive robust ordinal regression: A multiple criteria decision model based on the Choquet integral. *European Journal of Operational Research*, 201(1):277–288, 2010.

[8] C.A. Bana E Costa. A multicriteria decision aid methodology to deal with conflicting situations on the weights. *European Journal of Operational Research*, 26(1):22–34, 1986.

[9] C.A. Bana E Costa. A methodology for sensitivity analysis in three-criteria problems: A case study in municipal management. *European Journal of Operational Research*, 33(2):159–173, 1988.

[10] S. Barba Romero and J.C. Pomerol. *Choix multicritère dans l'enterprise*. Heres. Collection Informatique, 1993.

[11] M. Behzadian, R.B. Kazemzadeh, A. Albadvi, and M. Aghdasi. PROMETHEE: A comprehensive literature review on methodologies and applications. *European Journal of Operational Research*, 200(1):198–215, 2010.

[12] V. Belton and T.J. Stewart. *Multiple criteria decision analysis: an integrated approach*. Springer, 2002.

[13] M. Beuthe and G. Scannella. Comparative analysis of UTA multicriteria methods. *European Journal of Operational Research*, 130(2):246–262, 2001.

[14] G. Bous, P. Fortemps, F. Glineur, and M. Pirlot. ACUTA: A novel method for eliciting additive value functions on the basis of holistic preference statements. *European Journal of Operational Research*, 206(2):435–444, 2010.

[15] J.P. Brans and B. Mareschal. PROMETHEE Methods. In J. Figueira, S. Greco, and M. Ehrgott, editors, *Multiple Criteria Decision Analysis: State of the Art Surveys*, pages 163–196. Springer, Berlin, 2005.

[16] J.P. Brans, B. Mareschal, and Ph. Vincke. PROMETHEE: a new family of outranking methods in multicriteria analysis. In J.P. Brans, editor, *Operational Research, IFORS 84*, pages 477–490. North Holland, Amsterdam, 1984.

[17] J.P. Brans and Ph. Vincke. A preference ranking organisation method: The PROMETHEE method for MCDM. *Management Science*, 31(6):647–656, 1985.

[18] J.D. Carroll and P.E. Green. Guest editorial: Psychometric methods in marketing research: Part I, conjoint analysis. *J. Marketing Res.*, 32:385–391, 1995.

[19] A. Charnes, W.W. Cooper, and R.O. Ferguson. Optimal estimation of executive compensation by linear programming. *Management Science*, 1(2):138–151, 1955.

[20] A. Chateauneuf and J.Y. Jaffray. Some characterizations of lower probabilities and other monotone capacities through the use of Möbius inversion. *Mathematical Social Sciences*, 17:263–283, 1989.

[21] G. Choquet. Theory of capacities. *Ann. Inst. Fourier*, 5(54):131–295, 1953.

[22] S. Corrente, J.R. Figueira, and S. Greco. Interaction of criteria and robust ordinal regression in bi-polar PROMETHEE methods. In S. Greco et al., editor, *IPMU 2012, Part IV, CCIS 300*, pages 469–479. Springer, Berlin, 2012.

[23] S. Corrente, S. Greco, and R. Słowiński. Multiple Criteria Hierarchy Process in Robust Ordinal Regression. *Decision Support Systems*, 53(3):660–674, 2012.

[24] S. Corrente, S. Greco, and R. Słowiński. Multiple Criteria Hierarchy Process with ELECTRE and PROMETHEE. 2012. `http://dx.doi.org/10.1016/j.omega.2012.10.009`.

[25] S. Damart, L.C. Dias, and V. Mousseau. Supporting groups in sorting decisions: Methodology and use of a multi-criteria aggregation/disaggregation DSS. *Decision Support Systems*, 43(4):1464–1475, 2007.

[26] J.S. Dyer. MAUT-Multiattribute Utility Theory. In J. Figueira, S. Greco, and M. Ehrgott, editors, *Multiple Criteria Decision Analysis: State of the Art Surveys*, pages 265–292. Springer, Berlin, 2005.

[27] M. Ehrgott, J. Figueira, and S. Greco (eds.). *Trends in Multiple Criteria Decision Analysis*. Springer, Berlin, 2010.

[28] D. Ergu, G. Kou, Y. Peng, and Y. Shi. A Simple Method to Improve the Consistency Ratio of the Pair-wise Comparison Matrix in ANP. *European Journal of Operational Research*, 213(1):246–259, 2011.

[29] J. Figueira, S. Greco, and M. Ehrgott (eds.). *Multiple Criteria Decision Analysis: State of the Art Surveys*. Springer, Berlin, 2005.

[30] J. Figueira, S. Greco, and B. Roy. ELECTRE methods with interaction between criteria: An extension of the concordance index. *European Journal of Operational Research*, 199(2):478–495, 2009.

[31] J. Figueira, S. Greco, B. Roy, and R. Słowiński. An overview of ELECTRE methods and their recent extensions. *Journal of Multicriteria Decision Analysis, To appear*, 2012, DOI: 10.1002/mcda.1482.

[32] J. Figueira, S. Greco, and R. Słowiński. Identifying the "most representative" value function among all compatible value functions in the GRIP. *Proceedings of the 68th EURO Working Group on MCDA, Chania*, October 2008.

[33] J. Figueira, S. Greco, and R. Słowiński. Building a set of additive value functions representing a reference preorder and intensities of preference: GRIP method. *European Journal of Operational Research*, 195(2):460–486, 2009.

[34] K. Fujimoto. New characterizations of k-additivity and k-monotonicity of bi-capacities. In *SCIS-ISIS 2004, 2nd Int. Conf. on Soft Computing and Intelligent Systems and 5th Int. Symp. on Advanced Intelligent Systems*, 2004.

[35] K. Fujimoto and T. Murofushi. Some characterizations of k-monotonicity through the bipolar Möbius transform in bi-capacities. *J. Adv. Comput. Intell. Intell. Informatics*, 9(5):484–495, 2005.

[36] J. Fürnkranz and E. Hüllermeier, editors. *Preference Learning.* Springer, Berlin, 2010.

[37] I. Gilboa and D. Schmeidler. Additive representations of non-additive measures and the Choquet integral. *Ann.Operational Research*, 52:43–65, 1994.

[38] A. Gofman. Emergent scenarios, synergies, and suppressions uncovered within conjoint analysis. *Journal of Sensory Studies*, 21(4):373–414, 2006.

[39] M. Grabisch. The application of fuzzy integrals in multicriteria decision making. *European Journal of Operational Research*, 89:445–456, 1996.

[40] M. Grabisch. k-order additive discrete fuzzy measures and their representation. *Fuzzy Sets and Systems*, 92:167–189, 1997.

[41] M. Grabisch and C. Labreuche. Bi-capacities-I: definition, Möbius transform and interaction. *Fuzzy sets and systems*, 151(2):211–236, 2005.

[42] M. Grabisch and C. Labreuche. Bi-capacities-II: the Choquet integral. *Fuzzy sets and systems*, 151(2):237–259, 2005.

[43] M. Grabisch and C. Labreuche. Fuzzy measures and integrals in MCDA. In J. Figueira, S. Greco, and M. Ehrgott, editors, *Multiple Criteria Decision Analysis: State of the Art Surveys*, pages 563–604. Springer, Berlin, 2005.

[44] M. Grabisch and C. Labreuche. A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid. *4OR*, 6(1):1–44, 2008.

[45] M. Grabish, S. Greco, and M. Pirlot. Bipolar and bivariate models in multicriteria decision analysis: Descriptive and constructive approaches. *International Journal of Intelligent Systems*, 23(9):930–969, 2008.

[46] S. Greco and J.R. Figueira. Dealing with interaction between bi-polar multiple criteria preferences in outranking methods. *Research Report 11-2003*, INESC-Coimbra, Portugal, 2003. `http://www.uc.pt/en/org/inescc/res_reports_docs/research_reports`.

[47] S. Greco, M. Kadziński, V. Mousseau, and R. Słowiński. ELECTRE^{GKMS}: Robust ordinal regression for outranking methods. *European Journal of Operational Research*, 214(1):118–135, 2011.

[48] S. Greco, M. Kadziński, V. Mousseau, and R. Słowiński. Robust ordinal regression for multiple criteria group decision: UTA^{GMS}-GROUP and UTADIS^{GMS}-GROUP. *Decision Support Systems*, 214(1):118–135, 2011.

[49] S. Greco, M. Kadziński, and R. Słowiński. Selection of a representative value function in robust multiple criteria sorting. *Computers & Operations Research*, 38:1620–1637, 2011.

[50] S. Greco, B. Matarazzo, and S. Giove. The Choquet integral with respect to a level dependent capacity. *Fuzzy Sets and Systems*, 175:1–35, 2011.

[51] S. Greco, B. Matarazzo, and R. Słowiński. Rough sets theory for multicriteria decision analysis. *European Journal of Operational Research*, 129(1):1–47, 2001.

[52] S. Greco, B. Matarazzo, and R. Słowiński. Axiomatic characterization of a general utility function and its particular cases in terms of conjoint measurement and rough-set decision rules. *European Journal of Operational Research*, 158:271–292, 2004.

[53] S. Greco, B. Matarazzo, and R. Słowiński. Customer satisfaction analysis based on rough set approach. *Zeitschrift für Betriebswirtschaft*, 77:325–339, 2007.

[54] S. Greco, B. Matarazzo, and R. Słowiński. Bipolar Sugeno and Choquet integrals. In *EUROFUSE Workshop on Informations Systems, Varenna, Italy*, pages 191–196, September 2002.

[55] S. Greco, V. Mousseau, and R. Słowiński. Ordinal regression revisited: multiple criteria ranking using a set of additive value functions. *European Journal of Operational Research*, 191(2):416–436, 2008.

[56] S. Greco, V. Mousseau, and R. Słowiński. The possible and the necessary for multiple criteria group decision. *Algorithmic Decision Theory*, pages 203–214, 2009.

[57] S. Greco, V. Mousseau, and R. Słowiński. Multiple criteria sorting with a set of additive value functions. *European Journal of Operational Research*, 207(3):1455–1470, 2010.

[58] S. Greco, V. Mousseau, and R. Słowiński. UTA^{GMS}–INT: robust ordinal regression of value functions handling interacting criteria. *EJOR. Submitted*, 2011.

[59] S. Greco, J. Siebert, and R. Słowiński. Modeling interactions on bipolar scales using robust ordinal regression: the UTA^{GSS} method. *Proceedings of the 74th EURO Working Group on MCDA, Yverdon, Switzerland*, 2011.

[60] S. Greco, R. Słowiński, J.R. Figueira, and V. Mousseau. Robust ordinal regression. In M. Ehrgott, J. Figueira, and S. Greco, editors, *Trends in Multiple Criteria Decision Analysis*, pages 273–320. Springer, Berlin, 2010.

[61] P.E. Green and V. Srinivasan. Conjoint analysis in consumer research: Issues and outlook. *Journal of Consumer Research*, 5:103–152, 1978.

[62] P.E. Green and Y. Wind. *Multiattribute Decision in Marketing: A Measurement Approach*. Dryden Press, Hindsale, IL, 1973.

[63] P.E. Green and Y. Wind. New way to measure consumers' judgements. *Harvard Business Review*, 53:107–117, 1975.

[64] E. Grigoroudis, Y. Politis, and Y. Siskos. Satisfaction benchmarking and customer classification: An application to the branches of a banking organization. *International Transactions in Operational Research*, 9(5):599–618, 2002.

[65] E. Grigoroudis and Y. Siskos. Preference disaggregation for measuring and analysing customer satisfaction: The MUSA method. *European Journal of Operational Research*, 143(1):148–170, 2002.

[66] E. Grigoroudis and Y. Siskos. A survey of customer satisfaction barometers: Some results from the transportation-communications sector. *European Journal of Operational Research*, 152(2):334 – 353, 2004.

[67] E. Grigoroudis and Y. Siskos. *Customer Satisfaction Evaluation*. Springer, New York, 2010.

[68] A. Gustafsson, A. Herrmann, and F. Huber. Conjoint analysis as an instrument of market research practice. In A. Gustafsson, A. Herrmann, and F. Huber, editors, *Conjoint Measurement. Methods and Applications*, pages 5–45. Springer, Berlin, 2000.

[69] L.A. Hayduk and Y.H. Wonnacott. "effect equations" or "effect coefficients": A note on the visual and verbal presentation of multiple regression interactions. *The Canadian Journal of Sociology*, 5(4):399–404, 1980.

[70] D.H. Henard and D.M. Szymanski. Customer satisfaction: A meta-analysis of the empirical evidence. *Journal of the Academy of Marketing Science*, 29(1):16–35, 2001.

[71] Y.C. Hu and C. Hsiao-Chi. Choquet integral-based hierarchical networks for evaluating customer service perceptions on fast food stores. *Expert Systems with Applications*, 37(12):7880–7887, 2010.

[72] E. Jacquet-Lagrèze and J. Siskos. Assessing a set of additive utility functions for multicriteria decision-making, the UTA method. *European journal of operational research*, 10(2):151–164, 1982.

[73] E. Jacquet-Lagrèze and Y. Siskos. Preference disaggregation: 20 years of MCDA experience. *European Journal of Operational Research*, 130(2):233–245, 2001.

[74] M. Kadziński, S. Greco, and R. Słowiński. Selection of a representative value function for robust ordinal regression in group decision making. *Group Decision and Negotiation*, pages 1–34, 2011.

[75] M. Kadziński, S. Greco, and R. Słowiński. Selection of a representative value function in robust multiple criteria ranking and choice. *European Journal of Operational Research*, 217(3):541–553, 2011.

[76] M. Kadziński, S. Greco, and R. Słowiński. Extreme ranking analysis in robust ordinal regression. *Omega*, 40(4):488–501, 2012.

[77] M. Kadziński, S. Greco, and R. Słowiński. Selection of a representative set of parameters for robust ordinal regression outranking methods. *Computers & Operations Research*, 2012.

[78] M. Kadziński and T. Tervonen. Stochastic ordinal regression for multiple criteria sorting problems. EWG-MCDA, Tarragona, 12-14 April, 2012, also EJOR, submitted, 2012.

[79] R.L. Keeney and H. Raiffa. *Decisions with multiple objectives: Preferences and value tradeoffs.* J. Wiley, New York, 1976.

[80] R.L. Keeney and H. Raiffa. *Decisions with multiple objectives: Preferences and value tradeoffs.* J. Wiley, New York, 1993.

[81] R.L. Keeney, O. Renn, and D. von Winterfeldt. Structuring West Germany's energy objectives. *Energy Policy*, 15(4):352–362, 1987.

[82] R. Lahdelma, J. Hokkanen, and P. Salminen. SMAA - stochastic multiobjective acceptability analysis. *European Journal of Operational Research*, 106(1):137–143, 1998.

[83] R. Lahdelma and P. Salminen. SMAA-2: Stochastic multicriteria acceptability analysis for group decision making. *Operations Research*, 49(3):444–454, 2001.

[84] R. Lahdelma and P. Salminen. Stochastic multicriteria acceptability analysis (SMAA). In M. Ehrgott, J. Figueira, and S. Greco, editors, *Trends in Multiple Criteria Decision Analysis*, pages 321–354. Springer, 2010.

[85] J. Lusk and F.B. Norwood. Effect of experimental design on choice-based conjoint valuation estimates. *American Journal of Agricultural Economics*, 87(3):771–785, 2005.

[86] J.L. Marichal and M. Roubens. Determination of weights of interacting criteria from a reference set. *European Journal of Operational Research*, 124(3):641–650, 2000.

[87] V. Mousseau, J. Figueira, L. Dias, C. Gomes da Silva, and J. Climaco. Resolving inconsistencies among constraints on the parameters of an MCDA model. *European Journal of Operational Research*, 147(1):72–93, 2003.

[88] V. Mousseau and R. Słowiński. Inferring an ELECTRE TRI model from assignment examples. *Journal of Global Optimization*, 12(2):157–174, 1998.

[89] V. Mousseau, R. Słowiński, and P. Zielniewicz. A user-oriented implementation of the ELECTRE-TRI method integrating preference elicitation support. *Computers & Operations Research*, 27(7-8):757–777, 2000.

[90] S. Murofushi and T. Soneda. Techniques for reading fuzzy measures (III): interaction index. *9th Fuzzy Systems Symposium, Sapporo, Japan*, pages 693–696, 1993.

[91] J. Mustajoki. Effects of imprecise weighting in hierarchical preference programming. *European Journal of Operational Research*, 218(1):193–201, 2012.

[92] K. Ogawa. Route choice behavior model using fuzzy integral utility function. *Journal of Japan Fuzzy Association*, 11(4):690–694, 1999.

[93] R.C. Overton. Moderated multiple regression for interactions involving categorical variables: A statistical control for heterogeneous variance across two groups. *Psychological Methods*, 6:218–233, 2001.

[94] Z. Pawlak. *Rough Sets - Theoretical Aspects of Reasoning about Data*. Kluwer Academic, Dordrecht, 1991.

[95] D. Pekelman and S.K. Sen. Mathematical programming models for the determination of attribute weights. *Management Science*, 20(8):1217–1229, 1974.

[96] J.R.A. Ping. A parsimonious estimating technique for interaction and quadratic latent variables. *Journal of Marketing Research*, 32:336–347, 1995.

[97] M. Poyhonen, H. Vrolijk, and R.P. Hamalainen. Behavioral and procedural consequences of structural variation in value trees. *European Journal of Operational Research*, 134(1):216–227, 2001.

[98] F.S. Roberts. *Measurement Theory with Applications to Decision-Making, Utility and the Social Sciences*. Addison-Wesley, London, 1979.

[99] G.C. Rota. On the foundations of combinatorial theory I. Theory of Möbius functions. *Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 2:340–368, 1964.

[100] B. Roy. Classement et choix en présence de points de vue multiples: la méthode ELECTRE. *Revue Francaise d'Informatique et de Recherche*, 2(8):57–75, 1968.

[101] B. Roy. La méthode ELECTRE II: une application au media-planning. In M. Ross, editor, *OR'72*, pages 291–302. North Holland, 1973.

[102] B. Roy. ELECTRE III: un algorithme de classement fondé sur une représentation floue des préférences en présence de critères multiples. *Cahiers du CERO*, 20:3–24, 1978.

[103] B. Roy. *Multicriteria Methodology for Decision Aiding*. Kluwer, Dordrecht, 1996.

[104] B. Roy. ELECTRE Methods. In J. Figueira, S. Greco, and M. Ehrgott, editors, *Multiple Criteria Decision Analysis: State of the Art Surveys*, pages 133–162. Springer, Berlin, 2005.

[105] B. Roy. Paradigm and Challenges. In J. Figueira, S. Greco, and M. Ehrgott, editors, *Multiple Criteria Decision Analysis: State of the Art Surveys*, pages 3–24. Springer, Berlin, 2005.

[106] B. Roy and D. Bouyssou. *Aide multicritère à la décision: méthodes et cas.* Economica, 1993.

[107] B. Roy and J. Hugonnard. Classement des prolongements de lignes de métro en banlieu parisienne (présentation d'une méthode multicritere originale). *Cahiers du CERO*, 24 (2-3-4):153–171, 1982.

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

[109] B. Roy and J.M. Skalka. *ELECTRE IS: Aspects méthodologiques et guide d'utilisation.* Document du LAMSADE, no. 30, Université de Paris Dauphine, 1984.

[110] B. Roy and Ph. Vincke. Pseudo-orders: definition, properties and numerical representation. *Mathematical Social Science*, 62:263–274, 1987.

[111] T.L. Saaty. The analytic hierarchy and analytic network processes for the measurement of intangible criteria and for decision-making. In J. Figueira, S. Greco, and M. Ehrgott, editors, *Multiple Criteria Decision Analysis: State of the Art Surveys*, pages 345–382. Springer, Berlin, 2005.

[112] G. Shafer. *A Mathematical Theory of Evidence.* Princeton University Press, 1976.

[113] L.S. Shapley. A value for n-person games. In Tucker A. W. Kuhn, H. W., editor, *Contributions to the Theory of Games II*, page 307. Princeton University Press, Princeton, 1953.

[114] J. Šipoš. Integral with respect to a pre-measure. *Mathematica Slovaca*, 29(2):141–155, 1979.

[115] Y. Siskos, N.F. Matsatsinis, and G. Baourakis. Multicriteria analysis in agricultural marketing: The case of French olive oil market. *European Journal of Operational Research*, 130(2):315–331, 2001.

[116] R.L Smith. Efficient Monte Carlo procedures for generating points uniformly distributed over bounded regions. *Operations Research*, 32:1296–1308, 1984.

[117] V. Srinivasan and A.D. Shocker. Estimating the weights for multiple attributes in a composite criterion using pairwise judgments. *Psychometrika*, 38(4):473–493, 1973.

[118] W.G. Stillwell, D. von Winterfeldt, and R. S.John. Comparing Hierarchical and Nonhierarchical Weighting Methods for Eliciting Multiattribute Value Models. *Management Science*, 33(4):442–450, 1987.

[119] M. Sugeno. *Theory of fuzzy integrals and its applications.* Ph.D. Thesis, Tokyo Institute of Technology, 1974.

[120] M. Sugeno, K. Fujimoto, and T. Murofushi. Hierarchical decomposition theorems for Choquet integral models. *Proceedings-IEEE International Symposium on Circuits and Systems*, 3:2075–2078, 1995.

[121] T. Tervonen and J. Figueira. A survey on stochastic multicriteria acceptability analysis methods. *Journal of Multi-Criteria Decision Analysis*, 15(1-2):1–14, 2008.

[122] T. Tervonen, G. Van Valkenhoef, N. Basturk, and D. Postmus. Efficient weight generation for simulation based multiple criteria decision analysis. EWG-MCDA, Tarragona, 12-14 April, 2012.

[123] A. Torres and J.A. Tribó. Customer satisfaction and brand equity. *Journal of Business Research*, 64(10):1089 – 1096, 2011.

[124] F.M. Tseng and Y.J. Chiu. Hierarchical fuzzy integral stated preference method for Taiwan's broadband service market. *Omega*, 33(1):55–64, 2005.

[125] F.M. Tseng and Y.C. Hu. Partitioned fuzzy integral nested logit model for the Taiwan's internet telephony market. *International Journal of Computer Mathematics*, 86(8):1283–1299, 2009.

[126] F.M. Tseng and C.Y. Yu. Partitioned fuzzy integral multinomial logit modelfor Taiwan's internet telephony market. *Omega*, 33(3):267–276, 2005.

[127] A. Tversky and D. Kahneman. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5(4):297–323, 1992.

[128] G.H. Tzeng, C.T. Ou Yang, Y.P. Lin, and C.B. Chen. Hierarchical MADM with fuzzy integral for evaluating enterprise intranet web sites. *Information Sciences*, 169(3-4):409–426, 2005.

[129] P.P. Wakker. *Additive representations of preferences: A new foundation of decision analysis*, volume 4. Springer, 1989.

[130] M. Weber, F. Eisenfuhr, and D. von Winterfeldt. The effects of splitting attributes on weights in multiattribute utility measurement. *Management Science*, 34(4):431–445, 1988.

[131] C.W. Yang. Integrating fuzzy integral with multinomial logit model to evaluate the effects of service quality on traveling airline choice. *Information Technology Journal*, 9(7):1449–1457, 2010.

[132] F.W. Young, J. De Leeuw, and Y. Takane. Regression with qualitative and quantitative variables: An alternating least squares method with optimal scaling features. *Psychometrika*, 41(4):505–529, 1976.

[133] W. Yu. Aide multicritère à la décision dans le cadre de la problématique du tri: méthodes et applications. Ph.D. thesis. *LAMSADE*, Université Paris Dauphine, Paris, 1992.