

Elicitation of criteria weights maximising the stability of pairwise outranking situations

Raymond Bisdorff^{*1}, Patrick Meyer²³, and Thomas Veneziano¹

¹ University of Luxembourg
Faculty of Sciences, Technology, and Communication
Computer Science and Communications Research Unit
Interdisciplinary Lab for Intelligent and Adaptive Systems
6, rue Richard Coudenhove-Kalergi, L-1359 Luxembourg
{raymond.bisdorff,thomas.veneziano}@uni.lu

² Institut Télécom; Télécom Bretagne
UMR CNRS 3192 Lab-STICC
Technopôle Brest Iroise CS 83818 F-29238 Brest Cedex 3, France
patrick.meyer@telecom-bretagne.eu

³ Université européenne de Bretagne

Abstract. We propose an indirect approach for assessing criteria importance weights from overall outranking statements that a decision-maker may acknowledge in a Multiple Criteria Decision Aid process. Our approach takes advantage of the bipolar characteristic valuation of the overall outranking and the stability of its associated median-cut, crisp outranking relation. Numerical criteria importance weights are inferred with the help of a mixed integer linear programming model for which we present tests results under different configurations and tunings.

keywords: multiple criteria decision analysis; disaggregation; criteria importance elicitation; stable bipolar-valued outranking.

^{*}Corresponding author

1 Introduction

We consider a decision situation in which a finite set of decision alternatives is evaluated on a finite family of performance criteria. A decision-maker is willing to pairwise compare these alternatives according to the outranking paradigm. We consider indeed that an alternative a *outranks* an alternative b when a *weighted majority* of criteria validates the fact that a is performing at least as good as b and there is no criterion where b seriously outperforms a (Roy & Bouyssou, 1993). To assess when such a weighted majority of criteria validates an outranking situation requires a more or less precise numerical knowledge of the importance of each criterion in the multiple criteria preference aggregation. Two different approaches exist to specify these values:

- either via *direct* preference information, where the criteria importance is first assessed and then the aggregated outranking situations are computed,
- or, via *indirect* preference information, where some a priori partial knowledge of the resulting aggregated outranking situations is used in order to infer plausible estimators of the criteria importance.

In this article we exclusively concentrate on the indirect preference information approach. Similar approaches, mostly in the domain of Multiple Attribute Value Theory, are generally called *disaggregation* or *ordinal regression* methods in the literature. Notice that such a process, usually called *disaggregation/aggregation* method, may be used in an iterative procedure, constructing an incremental set of preference information and capturing at each step one or several sets of weights.

The very first disaggregation approach is implemented in the seminal UTA method (Jacquet-Lagrèze & Siskos, 1982; Siskos *et al.*, 2005), using linear programming for assessing additive value functions from a partial subjective ranking of decision alternatives in order to aggregate multiple criteria into a single composite criterion. In the non-additive value theory, we may mention an overview by (Grabisch *et al.*, 2008) concerning the different methods for the identification of Choquet integral capacities.

Considering some examples of disaggregation approaches in outranking methods, different solutions for assessing the parameters of the ELECTRE TRI method have been developed: In (Mousseau & Słowiński, 1998) and (Mousseau *et al.*, 2001), the authors propose an interactive approach for assessing the criteria weights from assignments of some alternatives in defined categories; in order to decrease the computational difficulty of the resulting mixed integer linear program, (Mousseau & Dias, 2004) propose a slight adaptation of the valued outranking relation used in ELECTRE III and ELECTRE TRI. These works are complemented with the implementation of mathematical programs for inferring veto-related parameters, developed by (Dias & Mousseau, 2006), and the work of (Rocha & Dias, 2008), assessing criteria weights and cutting level parameters without having to pre-define categories. Let us also mention the IRIS software (Dias & Mousseau, 2003), an extension of the ELECTRE TRI method which allows the decision-maker to provide assignment examples and constraints on the weights and the cut levels of the valued outranking relation for sorting the alternatives. A more general approach is used in the recent theoretical work in (Meyer *et al.*, 2008), where a mixed integer linear program is presented for assessing, at the same time, criteria weights, performance discrimination thresholds as well as potential performance values directly from a median-cut outranking relation.

An important issue in all these approaches represents the computation of one particular compatible vector of weights which is, most of the time, selected among the whole compatible vectors without clear motivation. In that case, a sensitivity analysis (Triantaphyllou & Sánchez, 1997) tries to measure the impact of this algorithmic choice on the recommendation, but it considers independently each criterion, as their weights are tested separately around an “ideal” solution. Another way to deal with this issue is proposed by the UTA^{GMS} valued method (Greco *et al.*,

2008), which takes care of the robustness of the preference modeling by considering necessary preferential statements (which are valid for all value functions compatible with the given information) and possible statements (when at least one of the value function validates the statement), avoiding the use of plausibly false algorithmic assumptions. However, it requires a longer questioning of the decision-maker and the running time may become prohibitive for a real-time processing of the decision aid. Let us finally mention a sorting approach (Dias & Clímaco, 2000) where the authors compute robust assignments (*i.e.* the best and the worst assignment for each alternative) according to some given constraints by multiple decision-makers on the criteria weights.

In this article, we propose the elicitation of a unique vector of criteria importance weights, with the objective of maximising the stability of the induced median-cut outranking digraph; hence, minimising the dependency of the outranking relation to the precise numeric values of the weights. As we shall explain in the sequel, this approach tends to minimise the impact of the algorithmic choices on the decision aid recommendations. Some preliminary work has already been published in (Bisdorff *et al.*, 2009), however, the at that time proposed mathematical program was not mature and optimal.

The article is organized as follows: First, we recall the notion of stability of the outranking statements and the way of computing it. In Section 3, we present a new extended mixed integer linear program (MILP) for estimating importance weights, followed in Section 4 by experiments showing the impact of including stability constraints in the model. Notice that we compare the results for both the current extended MILP and the simpler one published in (Bisdorff *et al.*, 2009)). We finally illustrate the potential use of the current MILP in an iterative decision aid procedure.

2 Defining and computing the stability of outranking statements

In this section, we briefly recall the notion of stability, adapted from Bisdorff *et al* (Bisdorff, 2004; Bisdorff *et al.*, 2009), and the way how to compute it.

Let $A = \{x, y, z, \dots\}$ be a finite set of $m > 1$ potential decision alternatives evaluated on a coherent finite family $F = \{1, \dots, n\}$ of $n > 1$ criteria. The alternatives are evaluated on ordinal performance scales and the performance of alternative x on criterion i is denoted x_i .

Between any two alternatives x and y of A , the marginal *at least as good as* situation $S_i(x, y)$, with each criterion i , is characterized as follows (Bisdorff, 2002; Bisdorff *et al.*, 2008):

$$S_i(x, y) = \begin{cases} 1 & \text{if } x_i \text{ is clearly at least as good as } y_i, \\ -1 & \text{if } x_i \text{ is clearly not at least as good as } y_i, \\ 0 & \text{otherwise.} \end{cases}$$

For instance, considering a real performance scale, to which an indifference $q_i \geq 0$ and a preference $p_i > q_i$ discrimination threshold (for all i in F) are associated (Roy & Bouyssou, 1993), the double threshold order $S_i(x, y)$ is given by:

$$S_i(x, y) = \begin{cases} 1 & \text{if } x_i + q_i \geq y_i, \\ -1 & \text{if } x_i + p_i \leq y_i, \\ 0 & \text{otherwise.} \end{cases}$$

We associate furthermore with each criterion $i \in F$ a rational *importance weight* w_i which represents the contribution of i to the *overall warrant or not* of the *at least as good as* preference situation between all pairs of alternatives. Let $w = (w_1, \dots, w_m)$ be the vector of relative importance weights associated with F such that $0 < w_i$ ($\forall i \in F$) and let \mathcal{W} be the set of such importance weights vectors.

In this paper, as it will be shown later, we may without loss of generality ignore the veto principle normally taken into account when dealing with a classical outranking relation (Roy & Bouyssou, 1993). The *overall valued outranking relation*, denoted \tilde{S}^w , aggregating the partial *at least as good as* situations, is then given by:

$$\tilde{S}^w(x, y) = \sum_{w_i \in W} w_i \cdot S_i(x, y), \quad \forall (x, y) \in A \times A.$$

$\tilde{S}^w(x, y)$ is thus evaluated with the following semantics (Bisdorff, 2002):

- $\tilde{S}^w(x, y) > 0$ indicates that a majority of criteria warrants the “*at least as good as*” preference situation between x and y . Then, x is said to outrank y and this statement is denoted $xS^w y$.
- $\tilde{S}^w(x, y) < 0$ indicates that a majority of criteria does not warrant the “*at least as good as*” preference situation between x and y . In this case, x does not outrank y and this statement is denoted $x\mathcal{S}^w y$.
- $\tilde{S}^w(x, y) = 0$ indicates a balanced situation, denoted $x?^w y$, where the criteria warranting the *at least as good as* preference situation between x and y are exactly as significant as those who do not warrant this situation.

Let \succsim_w be the preorder on F associated with the natural \geq relation on the elements of the vector of importance weights w . As classically done, $>_w$ denotes the asymmetric part of \succsim_w , whereas $=_w$ denotes its symmetric part that induces r ordered equivalence classes $\Pi_1^w >_w \dots >_w \Pi_r^w$ ($1 \leq r \leq m$). The criteria gathered in each equivalence class have the same importance weight in w and for $i < j$, those of Π_i^w have a higher importance weight than those of Π_j^w .

Let $w \in W$. For all $(x, y) \in A \times A$, the statement “ $xS^w y$ ” (resp. “ $x\mathcal{S}^w y$ ”) is said to be:

- *Independent* (of the weights) iff the outranking statement between x and y is warranted (resp. is not warranted) disregarding any vector of weights. For instance, this occurs when x dominates (resp. is dominated by) y .
- *Stable* (w.r.t. the weights) iff it is only depending on the preorder of the weights, *i.e.* it is validated (resp. invalidated) for all vectors of weights having the same preorder than w .
- *Unstable* (w.r.t. the weights) iff the situation between x and y is depending on the precise numeric value of the weights, *i.e.* it is not validated (resp. invalidated) for every vector having the same preorder than w .

In the context of a disaggregation approach, forcing the elicitation of importance weights that maximises the number of stable statements allows a better validation of the resulting median-cut outranking digraph with the decision-maker. Indeed, one can ensure that every stable statements will not change when considering any vector of weights in accordance with the given preorder, which is easier for a decision-maker to validate, instead of giving precise numeric values as weights. Therefore, assuming the validation of the elicited preorder by the decision-maker, he only has to focus on the validation, or invalidation, of unstable statements for a correct tuning of the weights.

We should note that if a veto situation occurs in the comparison of a pair of alternatives (x, y) , such that it invalidates the outranking situation disregarding any criteria weights, it is then considered as independent. As no preferential information from this situation can help us to capture criteria weights, we may without loss of generality ignore the veto principle in this article.

To compute the degree of stability associated with any outranking situation, let $c_k^w(x, y)$ be the sum of the local concordance relations $S_i(x, y)$ for all criteria $i \in \Pi_k^w$. Furthermore, let $C_k^w(x, y) = \sum_{i=1}^k c_i^w(x, y)$ be the cumulative sum of the local concordance relations for all criteria having importance at least equal to the one associated with Π_k^w , for all k in $\{1, \dots, r\}$.

Then, the following proposition gives us a test for evaluating the stability of any outranking statement:

Proposition 1 (Stability (Bisdorff, 2004)).

$$"xS^w y" \text{ is stable} \iff \begin{cases} \forall k \in 1, \dots, r : C_k^w(x, y) \geq 0; \\ \exists k \in 1, \dots, r : C_k^w(x, y) > 0. \end{cases} \quad (1)$$

$$"x\mathcal{S}^w y" \text{ is stable} \iff \begin{cases} \forall k \in 1, \dots, r : C_k^w(x, y) \leq 0; \\ \exists k \in 1, \dots, r : C_k^w(x, y) < 0. \end{cases} \quad (2)$$

$$"x?^w y" \text{ is stable} \iff \forall k \in 1, \dots, r : C_k^w(x, y) = 0. \quad (3)$$

The test of Proposition 1 corresponds in fact to the verification of *stochastic dominance*-like conditions (Bisdorff, 2004). Notice that this proposition is also warranted for a Pareto-dominant situation.

Any outranking statement that does not validate the stability proposition is then said to be *unstable*.

Table 1. Performance table, outranking relation \tilde{S}^w and associated stability

F	1	2	3	4	5	6	\tilde{S}^w : pairwise majority (stability)		
W	3	3	2	1	1	1	<i>a</i>	<i>b</i>	<i>c</i>
<i>a</i>	7	6	5	3	5	7	<i>a</i> 1.00 (Ind.)	0.09 (Uns.)	-0.45 (Sta.)
<i>b</i>	6	8	6	3	4	5	<i>b</i> 0.09 (Sta.)	1.00 (Ind.)	-0.09 (Uns.)
<i>c</i>	8	7	4	5	8	6	<i>c</i> 0.45 (Sta.)	0.09 (Uns.)	1.00 (Ind.)

Ind.: Independent; *Sta.*: Stable; *Uns.*: Unstable.

Example 1. In order to illustrate Proposition 1, we consider a small example with 3 alternatives and 6 criteria. Simplifying without losing any specificity, we do not consider here any performance discrimination thresholds. A vector of weights w , inducing the importance ordering $\{1, 2\} > \{3\} > \{4, 5, 6\}$ on the criteria, is defined and the performance table, on which every evaluations have to be maximized, is shown in the left part of Table 1. The resulting global majority concordance relation (or the outranking relation, as we consider no veto thresholds here) is normalised between -1 and 1 and is shown in the right part of Table 1, where the stability of each majority with respect to the given preorder of criteria weights is given in parentheses.

For instance, we may notice the difference in stability of the two outranking situations $bS^w a$ and $cS^w b$. Indeed, the first one appears to be stable; it may never be invalidated, assuming the agreement of the decision-maker to the preorder of w . Its apparent weak majority of 0.09 is not a weakly-determined one, contrary to the second statement, supported with a same apparent majority, which may be lost with a different weight vector showing the same preorder. A classical sensitivity analysis that is not taking into account the preorder of w may easily consider here that an outranking situation is only validated if it is warranted by a two third majority of the criteria weights. Hence, the first outranking situation would be invalidated and the second validated. However, the first one appears to be much more robust than the second when considering the potential uncertainty inherent in the setting of precise numeric values for the weights. A similar argument was already put forward by Roy & Bouyssou (Roy & Bouyssou, 1993).

An important property for our purpose is described by (Veneziano *et al.*, 2010). Let w_1 be the weights vector for which all the criteria weights equal 1 . Then:

Proposition 2.

$$\begin{aligned} x\mathcal{S}^{w_1}y &\implies \nexists w \in \mathcal{W}, \text{ s.t. } xS^w y \text{ is stable} \\ xS^{w_1}y &\implies \nexists w \in \mathcal{W}, \text{ s.t. } x\mathcal{S}^w y \text{ is stable} \end{aligned}$$

In other words, when more than half of the criteria are invalidating (resp. validating) an outranking situation, it is impossible to find a vector of criteria weights for warranting (resp. not warranting) this situation in a stable manner. Indeed, it is simply impossible to stably warrant an outranking statement going against the one obtained with an equisignificant weights vector.

3 Inverse analysis for the elicitation of weights

We present in this section a mathematical model eliciting criteria weights from a set of overall outrankings confirmed by a decision-maker, that best enforces the stability of the expressed outranking situations, in order to improve the overall stability of the resulting complete outranking relation. First, we recall the modeling of the stability test (see Proposition 1) by linear constraints that may force the stability of some overall outranking statements, adapted from (Bisdorff *et al.*, 2009). Then, we recall the kind of preferential information on alternatives a decision-maker can provide and how it can be integrated in the model. Finally, we enrich the model by potentially adding preferential information on the relative importance of the criteria.

3.1 Constraints ensuring the stability of some given outranking situations

We may notice that no constraint can be formulated in order to force an independent statement. Indeed, we can ignore such unanimous situations, positive or negative, as they concern a trivial pairwise comparison situation between Pareto dominant (resp. dominated) alternatives. The outranking situation is then unanimously warranted (resp. unwarranted), disregarding every possible importance of the criteria. Those situations do not give us any specific information for the elicitation of the importance weights.

As criteria importance weights are supposed to be rational, we can, without any loss of generality, restrict our assessment problem to integer weights vectors. Hence, an integer weight $w_i \in [1, M]$ will be associated with each criterion i , M standing for the maximal admissible value. For the practical resolution of real decision problems, this bound may be set equal to the number m of criteria.

Let $P_{m \times M}$ be a Boolean matrix with generic term $[p_{i,u}]$, characterizing, for each line i , the number of weights units allocated to criterion i . Formally, line i^{th} represents the decomposition of the weight associated with i on M bits in a unary base, in such a way that $\sum_{u=1}^M p_{i,u} = w_i$. For example, if i is associated with an integer weight equal to 3, and if $M = 5$, then the i^{th} line of the matrix $P_{m \times 5}$ will be $(1, 1, 1, 0, 0)$. Such decomposition is necessary in order to know the variable equivalence class of each criterion.

As each criterion weight must be strictly positive, we easily deduce that at least one weight unit is allocated to each criterion, *i.e.* $p_{i,1} = 1$ for all $i \in F$. We obtain the following constraint:

$$\sum_{i \in F} p_{i,1} = m. \quad (4)$$

The following constraints warrant the integrity of P , the most significant bits grouped together on the left side of the lines:

$$p_{i,u} \geq p_{i,u+1}, \quad \forall i = 1..m, \forall u = 1..M - 1. \quad (5)$$

We define \mathfrak{S}^+ (resp. \mathfrak{S}^-) as the set of ordered pairs of alternatives (x, y) on which the decision-maker expresses that x outranks (resp. does not outrank) y . In order not to write twice some similar constraints, we denote \mathfrak{S}^\pm both these sets, using symbol \pm when necessary, such that it replaces a $+$ (resp. $-$) sign when considering \mathfrak{S}^+ (resp. \mathfrak{S}^-). Now, let us introduce, for every pair $(x, y) \in \mathfrak{S}^\pm$, the following set of constraints that enforces the stability of their overall outranking situation:

$$\sum_{i \in F} (p_{i,u} \cdot \pm S_i(x, y)) \geq b_u(x, y), \quad \forall (x, y) \in \mathfrak{S}^\pm, \forall u = 1..M, \quad (6)$$

where $b_u(x, y)$ are boolean variables defined for every pair of alternatives and every equisignificance level $u = 1..M$. Those binary variables dictate at least one strict inequality case for every pairs of alternatives, as requested in Proposition 1, via the following constraints:

$$\sum_{u=1}^m b_u(x, y) \geq 1, \quad \forall (x, y) \in \mathfrak{S}^\pm. \quad (7)$$

Ensuring the stability of the outrankings validated by the decision-maker allows, the case given, to present him the overall picture of his preferences in a clearer manner (only considering the preorder), thus improving his understanding of the problem. However, as it may result in the inability of solving the problem, disregarding to the given preferential information, we have to introduce some *relaxed constraints*, obtained from the original constraints (6) by adding boolean *slack variables*, which allow a mathematical resolution by relaxing the incompatible stability constraints with the underlying problem:

$$\sum_{i \in F} (p_{i,u} \cdot \pm S_i(x, y)) + s_b(x, y) \cdot m \geq b_u(x, y), \quad \forall (x, y) \in \mathfrak{S}^\pm, \forall u = 1..M, \quad (8)$$

where $s_b(x, y)$ are boolean variables associated with the ordered pair (x, y) , multiplied by the number of criteria (as in the worst case, $C_k^w(x, y) = -m$). Notice that in (Veneziano *et al.*, 2010), we considered real slack variables (instead of a boolean one multiplied by a large enough number), and tried to minimize the sum of the slack variables. As we shall present in Section 4, this increases the number of validated stability constraints, but is not fully optimal, contrary to the minimisation of the boolean slack variables.

We may highlight that an invalidated stability constraint no longer warrants the simple majority weight support of the considered outranking situations. Consequently, in order to enforce the weighted majority for the expressed preference information, we formulate the following additional constraint, for each associated pair $(x, y) \in \mathfrak{S}^\pm$ with a relaxed stability constraint:

$$\sum_{i \in F} \left(\sum_{u=1}^M p_{i,u} \right) \cdot \pm S_i(x, y) \geq 1, \quad \forall (x, y) \in \mathfrak{S}^\pm, \quad (9)$$

Notice that we could in a similar manner enforce as well balanced situations observed for some pairs of alternatives (the right hand side of constraint (9) being modified into a 0 equality). In practice, however, such balanced situations occur when a decision-maker is unable to validate or invalidate a preference situation between two alternatives. As it is not advisable to constrain the elicitation with uncertain knowledge, we are not taking into account this kind of situations.

Notice also that we are not going to relax the simple majority weight constraints (9). In case the decision-maker would express incompatible preferential information, we will resolve this inconsistency following the approach originally proposed in (León & Liern, 2001; Roodman, 1979) and further developed and adapted to MCDA by Mousseau *et al.* (Mousseau *et al.*, 2003; Mousseau *et al.*, 2006).

3.2 Taking into account apparent preferences of the decision-maker

We propose to integrate all further preferential information that a decision-maker can provide in our elicitation model. This information may take the form of:

- a subset $E \subseteq A \times A$ of pairs of alternatives (a, b) for which a decision-maker is able to express a strict preference or an indifference;
- a partial preorder \geq_N over the weights of a subset of criteria $N \subseteq F$;
Example: criterion 1 is more valuable than criterion 4;
- some constraints over numerical values associated with some criteria weights;
Example: criterion 2 weight value is equal to 3, or is between 2 and 4;
- a partial preorder between some sets of criteria, expressing preferences about the sum of some criteria weights;
Example: the coalition of criteria 1 and 3 is more important than 2;
- some sets of criteria able to validate or invalidate an outranking statement;
Example: when an alternative x is at least as good as y over criteria 1, 2 and 3, the decision-maker considers that x outranks y .

The first type of preferential information concerns strict preferences and indifferences over a subset of ordered pairs of alternatives E . Let P be the strict preference relation over E and I the indifference relation. If the decision-maker expresses that aPb then, it necessarily results in $aS^w b$ and $b\mathcal{S}^w a$, where w is the vector of criteria weights provided by the resolution of the problem. In the same way, if the decision-maker expresses that cId then this results in $cS^w d$ and $dS^w c$.

In order to provide a solution as stable as possible to the decision-maker, we deal with the statement aPb as follows:

- We ensure the validation of the outranking statement for the ordered pair (a, b) and the invalidation of the one for the ordered pair (y, x) ;
- We try to enforce the stability of the two considered outranking relations by adding relaxed stability constraints.

Similarly, we translate an indifference judgement cId by the validation of the two associated outranking relations and try to ensure their stability.

Let us notice that, according to Proposition 2, when the decision-maker expresses the fact that $xS^w y$ (resp. $x\mathcal{S}^w y$) and otherwise we observe $x\mathcal{S}^{w_1} y$ (resp. $xS^{w_1} y$), it will not be possible to warrant these pairwise comparisons any stability. It is hence useless to keep active these stability constraints, the case given.

Furthermore, direct numerical information on criteria weights provided, the case given, by the decision-maker, are easy to translate into linear constraints. Thus, if a decision-maker expresses the fact that the weights of criterion i is equal to an integer value u_i , or if he wants to restrict the value of the weight of criterion j between two integers v_i and v'_i , we add some of the following constraints:

$$\sum_{c=1}^M p_{ic} = u_i, \quad \text{or} \quad \sum_{c=1}^M p_{jc} \geq v_i \quad \text{and} \quad \sum_{c=1}^M p_{jc} \leq v'_i \quad (10)$$

A decision-maker's statement "criterion i is more important than criterion j " will be taken into account by adding the below constraint:

$$\sum_{c=1}^M p_{ic} \geq \sum_{c=1}^M p_{jc} + 1 \quad (11)$$

This formula can be generalized for subsets of criteria: If a subset H of criteria is more important than a subset K , then we add the following constraint:

$$\sum_{i \in H} \left(\sum_{c=1}^M p_{ic} \right) \geq \sum_{j \in K} \left(\sum_{c=1}^M p_{jc} \right) + 1 \quad (12)$$

We can also model the fact that a subset H of criteria is, according to the decision-maker, sufficient to validate an outranking statement (the sum of its criteria weights is strictly greater than half of the sum of all criteria):

$$\sum_{i \in H} \left(\sum_{c=1}^M p_{ic} \right) \geq \frac{\sum_{j \in F} \left(\sum_{c=1}^M p_{jc} \right)}{2} + 1 \quad (13)$$

3.3 Inverse analysis mixed integer linear program

Solving the linear problem stated so far will naturally provide infinitely many admissible criteria weights. Following our goal of not exploring the whole admissible solution polytope, we try, hence, to recover a vector of criteria weights w^* that ensures every given outranking statement by maximising the number of stable statements among them. This goal is achieved with the help of the following objective function:

$$\min \quad K_1 \cdot \left(\sum_{(x,y) \in \mathfrak{S}^+ \cup \mathfrak{S}^-} s_b(x,y) \right) + K_2 \cdot \left(\sum_{i \in F} \sum_{u=1}^M p_{i,u} \right) \quad (14)$$

K_1 and K_2 are parametric constants used to order the sub-objectives. The second part of the objective function tends to minimise the sum of the weights $w_i^* \in w^*$ ($\forall i \in F$), which, in practice, means to use the smallest possible number of equi-significance classes. We show in the next section that minimising this number has a positive impact on the number of stable outranking statements. In order to prioritize the enforcement of the stability, we consider $K_1 = m \cdot M$ (which is the highest value that can be reached by the second part of the objective function) and $K_2 = 1$.

To summarize, let us present the obtained program:

MILP: STAB₂

Variables:

$$\begin{array}{ll} p_{i,u} \in \{0,1\} & \forall i \in F, \forall u = 1..M \\ b_u(x,y) \in \{0,1\} & \forall (x,y) \in \mathfrak{S}^+ \cup \mathfrak{S}^-, \forall u = 1..M \\ s_b(x,y) \geq 0 & \forall (x,y) \in \mathfrak{S}^+ \cup \mathfrak{S}^- \end{array}$$

Objective function:

$$\min \quad m \cdot M \cdot \left(\sum_{(x,y) \in \mathfrak{S}^+ \cup \mathfrak{S}^-} s_b(x,y) \right) + \left(\sum_{i \in F} \sum_{u=1}^M p_{i,u} \right)$$

Constraints:

$$\begin{array}{ll} \text{s.t.} \quad \sum_{i \in F} p_{i,1} = m & \\ p_{i,u} \geq p_{i,u+1} & \forall i \in F, \forall u = 1..M-1 \\ \sum_{u=1}^M b_u(x,y) \geq 1 & \forall (x,y) \in \mathfrak{S}^+ \cup \mathfrak{S}^- \\ \sum_{i \in F} \left(\left(\sum_{u=1}^M p_{i,u} \right) \cdot S_i(x,y) \right) \geq 1 & \forall (x,y) \in \mathfrak{S}^+ \end{array}$$

$$\begin{aligned} \sum_{i \in F} \left(\left(\sum_{u=1}^M p_{i,u} \right) \cdot S_i(x, y) \right) &\leq -1 && \forall (x, y) \in \mathfrak{S}^- \\ \sum_{i \in F} \left(p_{i,u} \cdot S_i(x, y) \right) + s_{bool}^{+2}(x, y) \cdot m &\geq b_u(x, y) && \forall (x, y) \in \mathfrak{S}^+, \forall u = 1..M \\ \sum_{i \in F} \left(p_{i,u} \cdot S_i(x, y) \right) - s_{bool}^{-2}(x, y) \cdot m &\leq -b_u(x, y) && \forall (x, y) \in \mathfrak{S}^-, \forall u = 1..M \end{aligned}$$

Direct constraints on criteria weights provided, the case given, by the decision-maker:

$$\begin{aligned} \sum_{c=1}^M p_{ic} = u_i, \quad \text{or} \quad \sum_{c=1}^M p_{ic} \geq v_i \quad \text{and} \quad \sum_{c=1}^M p_{ic} \leq v'_i &\quad \text{For some criteria } i \\ \sum_{c=1}^M p_{ic} \geq \sum_{c=1}^M p_{jc} + 1 &\quad \text{For some pairs } (i, j) \\ \sum_{i \in H} \left(\sum_{c=1}^M p_{ic} \right) \geq \frac{\sum_{j \in F} \left(\sum_{c=1}^M p_{jc} \right)}{2} + 1 &\quad \text{For some criteria subsets} \\ \sum_{i \in H} \left(\sum_{c=1}^M p_{ic} \right) \geq \sum_{j \in K} \left(\sum_{c=1}^M p_{jc} \right) + 1 &\quad \text{For some criteria subsets} \end{aligned}$$

Compared with STAB₁, the respective MILP version proposed in (Bisdorff *et al.*, 2009), the MILP STAB₂ presents the following enhancements: It specifies precise values for the constants K_i (formerly, these values were modified experimentally). In addition, as it minimises the number of violated stability constraints, not the sum of real slack variables, it is optimal. However, it is using boolean variables that may lead to a longer running time. Then, we will have to check if the additional running time is “compensated” by a significant increase in terms of stability.

Let us finally present some numerical experiments which may illustrate the usefulness of our approach.

4 Experiments on the use of stability constraints

In order to measure the behavior of STAB₂, its running time and the benefit in terms of stability of using the modeled constraints, we consider the following experiments:

1. Starting from a complete outranking relation, obtained with an unknown vector of weights w , we successfully compute another vector w^* , that reconstruct without interaction the same median-cut outranking digraph. Our objective is to first validate the model and analyse its behaviour when dealing with a full set of information (*i.e.* a maximal number of constraints).
2. In the second experiment, we iteratively construct a set of preferential information on the alternatives that allows to recapture the complete median-cut outranking relation. It highlights the reduced time for solving large instances with small sets of information, allowing the algorithm to be used in real-time interactive processes for eliciting the criteria weights.

In the second experiment, we are considering a small set of given outranking statements. As we cannot use the stability constraint on a pairwise statement without knowing its validity, the comparison may, or not, show eventually a lack of stability. Recalling potential equi-significance of the criteria naturally provides the kind of stability of the weights we are looking for, we hence adopt as a heuristic to minimize the overall sum of the criteria weights. Indeed, this heuristic, as it tends to reduce the number of equi-importance classes, will increase the number of stable statements that are not directly imposed by the decision-maker. In order to test the adequateness of this working hypothesis, we consider in the first experiment a control algorithm A_{CON} that only takes into account the simple majority constraints, minimises the sum of the weights and drops all further stability constraints.

For the test runs, we considered 25 different sizes of problems, by varying the number of alternatives n and criteria m according to the following values: 7, 10, 13, 16 and 19. For each size, we randomly generated 500 problems: a performance table (with a gaussian distribution of the generated values on each criterion) and a vector of integer criteria weights, allowing to compute an original outranking digraph, modeling a decision-maker's set of preferences. The problems were solved using CPLEX 11.0 on a machine with two Intel Xeon X5355 2,66 GHz processors with 4 cores each.

4.1 Reconstruction of a complete median-cut outranking digraph

In the first experiment, we start from a complete given outranking digraph and successfully compute a vector of criteria weights recapturing the relation, using the three defined versions of the algorithm, namely $STAB_1$, $STAB_2$ and A_{CON} . We then measure the average percentage of stable statements in each case, showing the stability benefit of using stability constraints.

If a balanced situation occurs in the original digraph (*i.e.* $x \succ^w y$), we decided not to take it into account as a model constraint. Indeed, such cases usually express the hesitation of a decision-maker shows when trying to validate or not a pairwise preferential statement.

As a large value for the parameter M implies an exponential increase of the running time, we decided to fix it to 7 for each problem at the beginning and to increase it when a solution cannot be reached with such a low parameter (about 8% when $M = 7$, only 4% when $M = 8$, ...), until a solution is found. To compare our solutions to the optimal ones, we run again all the problems, taking the number of criteria as the value of M . No real quality improvement can be noticed: most of the time, the solutions were the same or improved by only 1 or 2 percents, for a running time ten to a hundred times longer.

Table 3. Exp₁ – Stability increase and running time

		Median percentage of stable outranking statements				Median running time (s)	
m	n	Orig	A_{CON}	$STAB_1$	$STAB_2$	$STAB_1$	$STAB_2$
7	7	0.62	0.86	0.86	0.90	0.0	0.0
7	13	0.68	0.78	0.78	0.85	0.3	0.3
7	19	0.65	0.72	0.71	0.78	1.9	1.8
13	7	0.48	0.76	0.76	0.86	0.0	0.0
13	13	0.46	0.62	0.71	0.83	4.4	4.8
13	19	0.46	0.56	0.62	0.76	32.3	25.1
16	7	0.40	0.71	0.83	0.95	0.1	0.1
16	13	0.41	0.67	0.78	0.91	35.7	44.5
16	19	0.42	0.60	0.70	0.80	106.1	412.3
19	7	0.38	0.71	0.76	0.86	0.0	0.0
19	13	0.36	0.55	0.65	0.85	10.6	11.9
19	19	0.39	0.53	0.69	0.77	151.7	459.7

On its left part, Table 3 summarizes, for some sizes of problems and each algorithm, the median percentage of stable outranking statements (according to the total number of outranking statements, $n^2 - n$). We decided not to present all of them in order to improve the reading of this article, as all tests tend to similar conclusion. On right part of Table 3, we give the median running time of the algorithms. Notice that, as A_{CON} always runs between 0.1 and 0.3 seconds, its running time is not represented.

One may remark that each method improves the average percentage of stable arcs compared to the original outranking digraph (Orig). A_{CON} , by simply minimizing the sum of the weights, tends to minimize the number of equi-significance classes, inducing an increase of the stable statements of 18 percents in average. Then, such consideration in the objective function will be useful for partially taking care of the stability when no preferential information is given on some pairs of alternatives, as we shall present in the next experiment.

As expected, the exact algorithm $STAB_2$ gives better result than $STAB_1$, with an increase of the running time. However, we can see that this increase is significant only with a large set of alternatives. Indeed, when the number of alternatives is lower or equal to 13, the difference between the two algorithms is weak and clearly compensated by a gain of at least 10% of stable arcs. Also, as M is, most of the time, constant, the running time is more correlated to the number of pairs of alternatives than the number of criteria. Consequently, in the next experiment, as we consider the construction of an incremental subset of pair of alternatives, we hope that running time will be short enough for a real-time use of the exact algorithm $STAB_2$. We shall present the positive result afterwards.

To conclude the first experiment, we retain that it is possible to reconstruct a complete outranking digraph, with a significant increase of the stability of the arcs. This can be useful when a decision-maker agrees with a digraph obtained *via* a certain outranking method, in order to reinforce the stability of the digraph before tackling the actual decision aid problematique (*e.g.* selecting the best alternatives).

4.2 An iterative construction of a limited preference information set

Getting closer to a pairwise preference elicitation protocol, we are here considering a context where a fictitious decision-maker is asked to give preferential information about selected pairs of alternatives. The eventual aim is to iteratively assess enough information in order to be able to infer criteria weights allowing to correctly model all his global pairwise preferences (modeled by the initial complete outranking digraph, or with a few percentage of changes). We are interested in trying to define an adequate number of pairs to be selected, using the $STAB_2$ algorithm for intermediate resolutions, as it previously gave us the best results in terms of stability, in an acceptable time when the considered number of alternatives was reduced.

The key steps of this process are the selection of the pairs of alternatives and a condition for stopping the iteration. For selecting a specific pair of alternatives, we test three natural heuristics:

- a *random selection* (RS),
- a selection of a pair among the *most represented class* (MRC) of pairs having the same behavior on each criterion, in order to fix the greatest number of arcs at each iteration, and
- a selection of the pair with the *worst determined outranking value* (WDV), *i.e.* arcs associated with overall values close to a balanced situation which are, in the absence of stability, anecdotic and very sensitive to criteria weights changes.

For stopping the iteration, we test a complete reconstruction, and also a 95% reconstruction (*i.e.* the algorithm stops when at least a 95% of the arcs, between the outranking digraph obtained by the current vector of weights and the original one, are similar). The second condition takes care of the fact that a decision-maker may not validate or invalidate every pairwise outranking situation. In such a case, we assume that it is better to provide the most stable outranking digraph. Indeed, validating the criteria weight preorder, will by the way, model in a stable manner some pairwise outranking statements, which the decision-maker did not explicitly considered.

Notice that, when an additional preferential information is already stably granted by the current weights vector, there is no need to compute a new vector; Consequently, the preferential

information is simply added to the set of given information and another pair of alternatives is selected.

Table 4 summarizes, for some sizes of problems, the average number of selected pairs (*i.e.* an idea on the number of questions we should ask the decision-maker before presenting him the final outranking relation, or before exploiting it) and the average number of effective resolutions. We only present the result using the WDV heuristic, as RS and MRC heuristics results are not satisfying (the number of selected pairs and the running time were two to four times higher compared to the WDV heuristic).

Table 4. Exp₂ – Median values for STAB₂ iterative construction, using the WDV heuristic

m	n	nb_pairs [†]	100% reconstruction			95% reconstruction		
			nb_select*	nb_solve*	%stable [‡]	nb_select*	nb_solve*	%stable [‡]
7	7	21	3.7 / 17.6%	1.2	83	2.0 / 9.5%	0.8	87
7	13	78	8.7 / 11.2%	2.2	76	4.5 / 5.7%	1.2	82
7	19	171	10.2 / 6.0%	2.4	76	5.0 / 2.9%	1.3	84
13	7	21	5.2 / 24.8%	2.1	79	2.9 / 13.8%	1.3	83
13	13	78	16.8 / 21.5%	4.5	70	7.3 / 9.4%	2.2	79
13	19	171	25.8 / 15.1%	6.2	63	8.0 / 4.7%	2.3	78
19	7	21	5.9 / 28.1%	2.5	77	3.8 / 18.1%	1.7	78
19	13	78	18.2 / 23.3%	5.6	70	7.8 / 10.0%	2.5	78
19	19	171	30.7 / 18.0%	8.4	59	10.3 / 6.0%	3.1	76

[†] nb_pairs : Median number of alternatives pairs that can be considered

* nb_select : Median number of alternatives pairs selected by the algorithm

* nb_solve : Median number of CPLEX executions

[‡] %stable : Median percentage of stable statements in the resulting outranking digraph

On the left part of Table 4, we detail the results for a complete reconstruction of the original median-cut outranking digraph. This experiment corresponds to an iterative version of the first one, without having to consider all pairs of alternatives. It significantly reduces the running time: even for large instances, it runs under a second for each iteration on a standard computer. In consequence, such a process can be used in a real-time decision aid process to select pairs of alternatives and to reconstruct iteratively a satisfying vector of criteria weights. Notice that the stability is slightly lower than for the first experiment, due to the fact that the algorithm only forces the stability for some selected pairs, but tries to minimize the sum of the weights and so tends to reduce the discrimination between the weights to its lowest possible level. One can run again the first experiment at the end of the iterative protocol, once the decision-maker validates the outranking relation. Notice also that the WDV heuristic considers less than 30% of the whole pairs in the worst case (a large number of criteria and a few alternatives) and only 6% in the best ones (few criteria and a large number of alternatives), helping us to find the weights parameters in a relatively fast questioning protocol.

On the right part of Table 4, we detail the results of a 95% reconstruction of the original outranking digraph, running again the previous experiment by modifying the ending condition in order to stop the iterative process when at least 95% of the original outranking digraph has been reconstructed. We can easily see the decrease of the running time, the number of selected pairs and the number of resolution. Notice that the stability is even better, as the outranking relation is less constrained.

These experiments highlight the positive use of the STAB_2 algorithm in a real-time process for eliciting the criteria weights of a given multicriteria decision problem, with an iterative and interactive construction of the set of preferential information with the decision-maker.

In concrete terms, the process briefly runs as follows: First, we question the decision-maker on his beliefs on the way some criteria are compared and add the resulting constraints to the mathematical model. Then, we question him on a few selected pairs of alternatives with the WDV heuristic, on which he expresses preference, indifference or ignorance. The resulting preorder is presented and validated, or modified if necessary, by the decision-maker. The preorder is integrated in the model and we present the decision-maker a few unstable pairs (*i.e.* at least one arc that links the two alternatives is unstable) to tune the weights. Finally the outranking digraph can be exploited in order to answer the given problematic. Notice that, if dealing with the sorting problematic, when the profiles of the categories are given, the process computes the affectations of the alternatives directly, simply asking the decision-maker's advice about the comparison of some alternatives and the profiles.

5 Conclusion and future works

In this article, we have shown that caring about stability of the majorities may result in a powerful tool for focusing the attention of the criteria weight elicitation process on the most sensitive outranking statements; thus helping to reinforce the credibility of the eventual decision aid recommendations.

We propose, first, the possibility to reconstruct a complete median-cut outranking digraph, starting from a complete set of preferences on all pairs of alternatives. Secondly, we propose to iteratively build up the set of pairwise validated outrankings, always improving the overall stability of the arcs. Furthermore, we have shown that it is possible to create an iterative process for taking interactively into account some of the decision-maker's a priori overall preferences and reconstruct an outranking digraph in best accordance with its mind, with the use of a very reduced set of selected pairs and their resulting comparisons.

This improvement on building the median-cut outranking relation induces a real gain on the credibility of eventual decision aid recommendations. Indeed, the more stable this relation is, the less critical becomes the actual choice of precise numerical criteria importance weights. As a consequence, the robustness of a solution provided by a multiple criteria method exploiting this relation may be much enhanced.

Our present work, however, assumes the a priori settings of some other outranking model parameters, like the performance discrimination thresholds or, in the particular case of the sorting problematic, the precise numerical definition of the limiting profiles of the sorting categories. Some work is in progress that will extend the introduced mathematical programming model to take into account, in fact, a simultaneous elicitation of all the outranking model parameters.

References

- [Bisdorff, 2002]Bisdorff, R. 2002. Logical Foundation of Multicriteria Preference Aggregation. *Pages 379–403 of: Bouyssou, D., & al. (eds), Aiding Decisions with Multiple Criteria*. Kluwer Academic Publishers.
- [Bisdorff, 2004]Bisdorff, R. 2004. Concordant outranking with multiple criteria of ordinal significance. *4OR*, **2**(4), 293–308.
- [Bisdorff *et al.*, 2008]Bisdorff, R., Meyer, P., & Roubens, M. 2008. Rubis: a bipolar-valued outranking method for the best choice decision problem. *4OR*, **6**(2), 143–165.
- [Bisdorff *et al.*, 2009]Bisdorff, R., Meyer, P., & Veneziano, T. 2009. Inverse analysis from a Condorcet robustness denotation of valued outranking relations. *Pages 180–191 of: Rossi, F., & Tsoukiás, A. (eds), Algorithmic Decision Theory*. Springer-Verlag.
- [Dias & Clímaco, 2000]Dias, L., & Clímaco, J. 2000. ELECTRE TRI for Groups with Imprecise Information on Parameter Values. *Group Decision and Negotiation*, **9**, 355–377.
- [Dias & Mousseau, 2003]Dias, L., & Mousseau, V. 2003. IRIS: a DSS for multiple criteria sorting problems. *Journal of Multi-Criteria Decision Analysis*, **12**(4-5), 285–298.
- [Dias & Mousseau, 2006]Dias, L., & Mousseau, V. 2006. Inferring Electre’s veto-related parameters from outranking examples. *European Journal of Operational Research*, **170**(1), 172–191.
- [Grabisch *et al.*, 2008]Grabisch, M., Kojadinovic, I., & Meyer, P. 2008. A review of capacity identification methods for Choquet integral based multi-attribute utility theory: Applications of the Kappalab R package. *European Journal of Operational Research*, **186**, 766–785.
- [Greco *et al.*, 2008]Greco, S., Mousseau, V., & Slowinski, R. 2008. Ordinal regression revisited: multiple criteria ranking using a set of additive value functions. *European Journal of Operational Research*, **191**(2), 415–435.
- [Jacquet-Lagrèze & Siskos, 1982]Jacquet-Lagrèze, E., & Siskos, Y. 1982. Assessing a set of additive utility functions for multicriteria decision making: the UTA method. *European Journal of Operational Research*, **10**, 151–164.
- [León & Liern, 2001]León, T., & Liern, V. 2001. A fuzzy method to repair infeasibility in linearly constrained problems. *Fuzzy Sets and Systems*, **122**(2), 237–243.
- [Meyer *et al.*, 2008]Meyer, P., Marichal, J.-L., & Bisdorff, R. 2008. Disaggregation of Bipolar-Valued Outranking Relations. *Pages 204–213 of: Le Thi, H.-A., Bouvry, P., & Pham Dinh, T. (eds), Modelling, Computation and Optimization in Information Systems and Management Sciences, Second International Conference*. Communications in Computer and Information Science, vol. 14. Springer.
- [Mousseau & Dias, 2004]Mousseau, V., & Dias, L. 2004. Valued outranking relations in ELECTRE providing manageable disaggregation procedures. *European Journal of Operational Research*, **156**(2), 467–482.
- [Mousseau & Slowinski, 1998]Mousseau, V., & Slowinski, R. 1998. Inferring an Electre TRI model from assignment examples. *Journal of Global Optimization*, **12**(2), 157–174.
- [Mousseau *et al.*, 2001]Mousseau, V., Figueira, J., & Naux, J.-Ph. 2001. Using assignment examples to infer weights for ELECTRE TRI method: Some experimental results. *European Journal of Operational Research*, **130**(2), 263–275.
- [Mousseau *et al.*, 2003]Mousseau, V., Dias, L.C., Figueira, J., Gomes, C., & Clímaco, J.N. 2003. Resolving inconsistencies among constraints on the parameters of an MCDA model. *European Journal of Operational Research*, **147**(1), 72–93.
- [Mousseau *et al.*, 2006]Mousseau, V., Dias, L.C., & Figueira, J. 2006. Dealing with inconsistent judgments in multiple criteria sorting models. *4OR*, **4**(2), 145–158.

- [Rocha & Dias, 2008] Rocha, C., & Dias, L. 2008. An algorithm for ordinal sorting based on ELECTRE with categories defined by examples. *Journal of Global Optimization*, **42**, 255–277.
- [Roodman, 1979] Roodman, G. 1979. Post-Infesibility Analysis in Linear Programming. *Management Science*, **25**(9), 916–922.
- [Roy & Bouyssou, 1993] Roy, B., & Bouyssou, D. 1993. *Aide Multicritère à la Décision : Méthodes et Cas*. Paris: Economica.
- [Siskos *et al.*, 2005] Siskos, Y., Grigoroudis, E., & Matsatsinis, N.F. 2005. UTA Methods. *Pages 297–344 of: Figueira, J., Greco, S., & Ehrgott, M. (eds), Multiple Criteria Decision Analysis: State of the Art Surveys*. Boston, Dordrecht, London: Springer Verlag.
- [Triantaphyllou & Sánchez, 1997] Triantaphyllou, E., & Sánchez, A. 1997. Sensitivity Analysis Approach for Some Deterministic Multi-Criteria Decision-Making Methods. *Decision Sciences*, **28**(1), 151–194.
- [Veneziano *et al.*, 2010] Veneziano, T., Meyer, P., & Bisdorff, R. 2010. Analyse inverse robuste à partir d’informations préférentielles partielles. *Pages 75–88 of: Actes Roadef 2010*.