

*This is a PDF file of an unedited manuscript that has been accepted for publication in International Journal of Society Systems Science. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain. The final version will be available at: <http://dx.doi.org/10.1504/IJSS.2012.051143>*

# Comparison of decision rules to mediate negotiations with incomplete information

Paula Sarabando

Departamento de Matemática, Instituto Superior Politécnico de Viseu,  
Campus Politécnico de Viseu, 3504-510 Viseu, Portugal.  
e-mail: psarabando@mat.estv.ipv.pt

Luis C. Dias

INESC Coimbra and Faculdade de Economia da Universidade de Coimbra  
Av Dias da Silva 165, 2004-512 Coimbra, Portugal.  
e-mail: lmcdias@fe.uc.pt

## Abstract

Decisions in society resulting from negotiation among two parties are often inefficient (in terms of Pareto efficiency) as the parties fail to identify an agreement that would make both better off. A third party mediating the negotiation can increase the chances of reaching a Pareto efficient alternative. However, in most practical situations the mediator does not have a precise model of the parties' preferences, and therefore cannot identify which alternatives are efficient. This paper assesses the performance of decision rules that a mediator can apply in multi-issue bilateral Negotiation Analysis given ordinal information about the importance of the issues and the value of the potential alternatives in each issue. We assume the preferences of the parties can be modelled by an additive multiattribute value function, but without assuming the precise parameters of this model are known. We study three mediation criteria: maximizing the sum of the values, maximizing the product of the excesses, or maximizing the minimal proportion of potential. Monte-Carlo simulation is used to assess how good the alternative chosen by each decision rule is, and to provide guidelines about the use of these rules in a context of selecting a subset of promising alternatives.

## 1 Introduction

Most decisions in society result from some type of formal or informal negotiation: we are all negotiators (Fisher et al., 1992). It is possible to distinguish between two types of negotiation: distributive negotiation and integrative negotiation (see, e.g., Walton and McKersie, 1965). In distributive (or “win-lose”)

negotiations the purpose is typically the division of a single resource and the negotiators' main concern is to obtain the largest possible piece of the pie. In integrative (or "win-win") negotiations it is assumed that the integration of resources and capabilities of the parties can create more value to both. This is possible because such negotiations include multiple issues (e.g., price, quality, deadline, warranty) that are valued differently by the parties. This paper is focused on multiple-issue negotiation among two parties (integrative bilateral negotiation).

According to Thompson (2001), a major negotiation sin is "leaving money on the table", which occurs when negotiators fail to identify win-win potential. For instance, the parties have reached an agreement but there existed a different alternative that would be better for both. An alternative is called Pareto efficient if there exists no other alternative that is better for at least one of the parties without making the other party worse off.

This paper fits into Raiffa's externally prescriptive perspective (Raiffa et al., 2002), according to which the objective of the analysis is to suggest how mediators should act in order to help the parties in an impartial and balanced way. In particular we are interested in supporting a mediator who wishes to propose good potential agreements to the negotiating parties. Raiffa et al. (2002) suggest the mediator can elicit the preferences of both parties separately (modelling these preferences by means of a multiattribute value function) and then use these preferences to recommend potential agreements using a criterion that guarantees Pareto efficiency (details are provided in Section 2).

However, many authors sustain that it is difficult to obtain precise numerical values for the preference model's parameters (e.g., Lahdelma et al., 2003 and Sarabando and Dias, 2009): the decision makers do not have time to study the problem carefully enough, the decision makers have difficulties in comparing criteria, the decision makers do not want to fix their preferences because the preferences may change during the process, and in general they find that precise quantitative information is hard to provide. Even in private, there may exist also reluctance from the parties to share their complete information with the mediator.

We assume that negotiators feel more "comfortable" providing ordinal information. For instance, instead of asking for precise values for the criteria weights, a mediator can simply ask for a ranking of these weights. This is the type of information required by the SMARTER method (Edwards and Barron, 1994) and other multicriteria analysis methods. It is also possible to avoid asking for precise quantitative estimates of the value that each potential agreement has for each issue of interest. We consider in particular that the information provided by the parties to the mediator, possibly in private, is solely:

1. ordinal information about the weights of the issues (which is the issue with highest weight, the issue with second highest weight, etc.),
2. ordinal information about the value of each performance level in each

issue (which is the best level for one issue, the second best level for that issue, etc.), and, possibly,

3. ordinal information about the difference of value between consecutive levels in each issue.

One of the possibilities described in the literature to exploit ordinal information on the weights is to select a weights vector from a set of admissible weights to represent that set and then to use this vector to evaluate the alternatives. Examples of this are the use of equal weights and the use of ROC (rank order centroid) weights, which are compared in the simulation study of Barron and Barrett (1996). This study concludes that ROC weights provide a better approximation than the other weighting vectors. In posterior studies, Sarabando and Dias (2009, 2010) have extended and compared this type of rules in the context of supporting an individual decision maker.

While it is expected that eliciting ordinal information (instead of precise cardinal information) contributes to facilitate the interaction between the mediator and the parties, this implies working with less information. This raises the question of knowing whether using a simple rule to obtain results from ordinal information will lead to results that are close enough to the results that would be obtained by a fully cardinal elicitation process (assuming that would be possible). A related question is to know which rules are more interesting in terms of the compromise between the elicitation effort they require and the results they yield. Finally, it is important to know whether conclusions about the usefulness of the rules are sensitive to some characteristics of the problem.

This paper aims at providing answers to these questions in the context of supporting a mediator in bilateral multi-issue negotiation. Knowing the answers to these questions, a mediator can make informed decisions about the type of information she will elicit from each party (this will of course also depend on each party's characteristics, the available time, etc.). The rules we compare in this paper have been already compared considering the case of individual decision (e.g. Barron and Barrett, 1996; Sarabando and Dias, 2009, 2010), but, as far as we now, they have never be compared to the case of negotiation. As in these previous comparison studies, this paper provides answers that are based on a large number of Monte-Carlo simulations. We conclude that it is possible to obtain good results considering ordinal information regarding the preferences of the parties.

This paper is structured as follows. First, we present the model for the preferences of the parties and the negotiation analysis criteria that a mediator can use to identify a Pareto efficient alternative and review other types of approaches. Afterwards, we formally introduce the type of information we assume is available to the mediator and the decision rules she can follow. In Section 4 the conducted simulations are described, and results of such simulations are presented in Section 5. Section 6 presents some conclusions and some lines for future research. The tables with detailed results referred to in the text are presented in Appendix A.

## 2 Background: Negotiation Analysis and Mediation Criteria

Let us consider Raiffa's negotiation analysis setting in which two parties are negotiating over multiple issues, there is a finite number of potential alternatives for an agreement, and the preferences of each party can be modelled by a multiattribute value (or utility) function (Raiffa et al., 2002). In this Negotiation Analysis context Raiffa assumes that the preferences of each party can be modelled by a multiattribute additive value (or utility) function (Raiffa et al., 2002):

$$v^k(x^{(j)}) = \sum_{i=1}^n w_i^k v_i^k(l_i^{(j)}) = w_1^k v_1^k(l_1^{(j)}) + \dots + w_n^k v_n^k(l_n^{(j)}), j = 1, \dots, m \quad (1)$$

where  $n$  represents the number of issues in negotiation,  $m$  represents the number of alternatives (the potential agreements),  $l_i^{(j)}$  represents the performance level that alternative  $x^{(j)}$  reaches in issue  $i$ ,  $v_i^k(\cdot)$  represents the value (or utility) function of issue  $i$  ( $i = 1, \dots, n$ ) for party  $k$  ( $k = 1, 2$ ), and  $w_i^k$  represents the weight of the value function  $v_i^k(\cdot)$  (the weight of the issue  $i$ ), for party  $k$ . Let us note that there are other preference models that compute an overall global value besides the one we are presenting, such as PROMETHEE II (Brans and Vincke, 1985; Beynon and Barton, 2008) and AHP (Saaty, 1996; Hajeeh, 2008), but to our knowledge no negotiation analysis models based on these alternative models have been proposed so far.

The subjective preferences of each party are modelled by the value functions and the weights. The function  $v_i^k(l_i^{(j)})$  is such that the higher its value is, the better is the level attained by alternative  $x^{(j)}$  in issue  $i$ , for party  $k$ . The weights  $w_1^k, \dots, w_n^k$  are scaling coefficients that reflect value trade-offs between the issues, for party  $k$ . For instance, if  $w_1^k = 2w_2^k$ , then this means that one unit of value in issue 1 is valued as highly as 2 units of value in issue 2.

An alternative  $x^{(j)}$  is Pareto efficient if and only if there exists no other potential agreement  $x^{(j')}$  that is better for one party without being worse for the other party:

$$\nexists x^{(j')} \neq x^{(j)} : [v^1(x^{(j')}) \geq v^1(x^{(j)}) \wedge v^2(x^{(j')}) > v^2(x^{(j)})] \vee \\ \vee [v^1(x^{(j')}) > v^1(x^{(j)}) \wedge v^2(x^{(j')}) \geq v^2(x^{(j)})]. \quad (2)$$

There exist typically multiple Pareto optimal alternatives, hence the mediator should use some criterion to propose a potential agreement to the parties. Several such criteria can be developed, depending on whether the mediator is more interested in finding an efficient solution (which maximizes the total value created) or an equitable solution (which tries to balance the interests of the parties involved). Three well-known criteria for selecting a potential compromise are the following (see Raiffa et al., 2002):

- Maximizing the sum of the values: Select the alternative  $x^{(j)}$  that maximizes

$$v^1(x^{(j)}) + v^2(x^{(j)}). \quad (3)$$

- Maximizing the product of the excesses regarding the reservation values: Select the alternative  $x^{(j)}$  that maximizes

$$(v^1(x^{(j)}) - RV^1) \times (v^2(x^{(j)}) - RV^2) \quad (4)$$

( $RV^k$  represents the reservation value for party  $k$  ( $k = 1, 2$ ), i.e., the value below which the negotiator would not accept any alternative and would rather not make a deal).

- Maximizing the minimal proportion of potential (PoP): Select the alternative  $x^{(j)}$  that maximizes

$$\min_{k=1,2} \frac{v^k(x^{(j)}) - RV^k}{v^{k-max} - RV^k} \quad (5)$$

( $v^{k-max}$  denotes the best payoff that negotiator  $k$  can reach considering the set of alternatives that are better for both parties than the reservation value).

The choice of the criteria to use depends on the preferences of the mediator and all of them present advantages and disadvantages. For example, the criterion of maximizing the sum of the values may not be attractive to many due to the inequality of results that normally provides and due to its character completely compensatory. The criterion of maximizing the sum of the values does not meet the balance and disregards the fact that a party is “rich” and the other “poor”, i.e., it assumes that each unit has the same value for both parties. However, this is the criterion that maximizes the total value created and the mediator can devise compensation mechanisms. Let us note that the objective of this paper is not to compare the criteria, but to assess and compare the decision rules we will present in the next section.

Usually it is assumed that the parameters of the multiattribute value functions are known or can be asked to each party. If the mediator knows the exact value functions of both parties, she can calculate the Pareto frontier, defined by the Pareto efficient alternatives. So, the mediator can suggest an agreement in this set, where the choice between the solutions that are Pareto efficient can be based on maximizing the sum of the values, maximizing the product of the excesses regarding the reservation values, or maximizing the minimal proportion of potential (PoP). But, in many cases, this assumption is unrealistic for the reasons mentioned in the introduction.

The importance of the information about the preferences is clearly mentioned in the negotiation literature. According to the “Dual Concern” model

(Pruitt, 1983; Thomas, 1992), solutions that improve the position of the parties relatively the current situation only can be achieved if negotiators have enough knowledge about their preferences and about the preferences of the other parties. Ehtamo et al. (1999) present a class of methods called “Constraint Proposal Methods”, which are interactive methods to find Pareto optimal solutions through common tangent hyperplanes, for negotiations between two parties about two or more continuous issues, with the help of a mediator. Heikonen (1999) proposes an interactive method to determine Pareto optimal solutions in negotiations with multiple parties about continuous issues, using a mediator. The method does not require that negotiators know the value functions of the other parties, neither that the mediator knows all the value functions.

For bilateral negotiation processes with a discrete set of alternatives and incomplete information on weights, Clímaco and Dias (2006) propose an extension of the VIP-G (Dias and Clímaco, 2005) methodology, based on the relaxation of the weights space that each negotiator accepts and defining convergence paths in this space. The methodology suits situations in which the parties agree on which issues are to maximize and which ones are to minimize. Lai et al. (2006) present a model that considers Pareto and computational efficiency, for situations where information is incomplete, the value functions are not linear and are not explicitly known. The authors refer that the major problem associated with multiattribute negotiation is the difficulty in making decisions in a multidimensional space. To overcome this difficulty, they propose a process that allows the negotiators to negotiate based in a single line, called the “negotiation base line”, with the help of a mediator. Although it is not difficult to involve a mediator in automated negotiations between software agents, there may exist situations in which the parties do not trust the mediator or where an artificial mediator is difficult to implement. Thus, Lai and Sycara (2009) focus on the development of mechanisms to search for Pareto efficiency in multicriteria negotiations without the presence of a mediator. Vetschera (2009) proposes a method to measure the amount of information that is available during the negotiation, considering that information is obtained implicitly through the offers. The method is based on the domain criterion (which was initially developed for sensitivity analysis), and can be used both when a compromise is reached and when it is not. Sarabando et al. (2009) propose to derive incomplete information about the preferences of negotiators from the statements they make and the offers they exchange during the negotiation process. The authors also present and discuss three approaches that use this information in order to help a mediator proposing a better solution than the compromise the negotiators have reached or are close to reach.

In this work we study a different and much simpler strategy, which consists of using simple decision rules based on surrogate parameter values.

### 3 Ordinal Information and Decision Rules

To identify which alternatives are Pareto efficient and which ones are the best according to a mediation criterion it is necessary that the mediator knows exactly the numerical values for the value functions and the weights. We are now going to drop the assumption of complete information and will consider instead that the mediator can only elicit information that has an ordinal character:

1. **Ordinal information about the weights of the issues.** Given two value functions, say  $v_i^k(\cdot)$  and  $v_j^k(\cdot)$ , negotiator  $k$  can indicate if one unit of value on one of the functions is worth more than one unit of value in the other function, but does not need to quantify how much. We consider that the indices of the issues are coded, in such a way that, for each negotiator  $k$  ( $k = 1, 2$ ), the weights are in decreasing order:

$$W^k = \{(w_{[1]}^k, w_{[2]}^k, \dots, w_{[n]}^k) : w_{[1]}^k \geq w_{[2]}^k \geq \dots \geq w_{[n]}^k \geq 0, \sum_{i=1}^n w_{[i]}^k = 1\} \quad (6)$$

where  $w_{[i]}^k \in \{w_1^k, \dots, w_n^k\}$  represents the weight in position  $i$  for negotiator  $k$ , for  $i = 1, \dots, n$ .

2. **Ordinal information about the value of each level in each issue.** We also consider that the negotiators provide ordinal information about the value of each performance level in each issue, stating which is the best level in that issue, the second best level, and so on (let us note that different alternatives can have the same performance level in a given issue). We assume that, for negotiator  $k$  and for issue  $i$  ( $k = 1, 2$  and  $i = 1, \dots, n$ ):

$$V_{ik} = \{(v_i^k(l_i^{([1]k)}), v_i^k(l_i^{([2]k)}), \dots, v_i^k(l_i^{([m_i]k)})) : \\ : 0 = v_i^k(l_i^{([1]k)}) \leq v_i^k(l_i^{([2]k)}) \leq \dots \leq v_i^k(l_i^{([m_i]k)}) = 1\}, \quad (7)$$

where  $l_i^{([j]k)} \in \{l_i^{(1)}, \dots, l_i^{(m_i)}\}$  represents the level in position  $m_i + 1 - j$  of issue  $i$  for negotiator  $k$ , and  $m_i$  represents the number of levels in issue  $i$ .

3. **Ordinal information about the difference of value between consecutive levels in each issue.** We consider that in some cases each negotiator can also provide ordinal information about the difference of value between consecutive levels in each issue (Sarabando and Dias, 2010). Let  $\Delta_{i(m_i-1)}^k = v_i^k(l_i^{([m_i]k)}) - v_i^k(l_i^{([m_i-1]k)})$ , ..., and  $\Delta_{i1}^k = v_i^k(l_i^{([2]k)}) - v_i^k(l_i^{([1]k)})$ . In this case, each negotiator provides ordinal information about these differences of value.



We intend to assess if a mediator can suggest good alternatives based on such information. For this purpose, we can use the Rank Order Centroid (ROC) weights rule (Barron and Barrett, 1996) to compute surrogate weights. We can also use analogous rules (introduced in Sarabando and Dias (2010) for the case of individual decisions) for the ordinal information referring to the values of the levels: when the incomplete information refers to the value of each level in each issue, we consider the ROC values rule; when the incomplete information refers to the difference of values between consecutive levels in each issue, we consider the  $\Delta$ ROC values rule.

### ROC weights

The ROC weights rule consists in using ROC weights, for each negotiator  $k$  ( $k = 1, 2$ ), defined as follows (assuming that the issue indices reflect their order, from the highest weight  $w_{[1]}^k$  to the lowest weight  $w_{[n]}^k$ ):

$$w_{[i]}^{k(ROC)} = \frac{1}{n} \sum_{l=i}^n \frac{1}{l}, \quad i = 1, \dots, n. \quad (8)$$

### ROC values

The ROC values rule consists in using ROC values. For issue  $i$  ( $i = 1, \dots, n$ ), and for each negotiator  $k$  ( $k = 1, 2$ ), the ROC values are defined as follows:

$$v_i^{k(ROC)}(l_i^{([j]k)}) = \frac{m_i - r_i^k(l_i^{([j]k)})}{m_i - 1}, \quad j = 1, \dots, m_i. \quad (9)$$

where  $r_i^k(l_i^{([j]k)})$  represents the position in the ranking of level  $l_i^{([j]k)}$  considering issue  $i$ , such that  $r_i^k(l_i^{([j]k)}) < r_i^k(l_i^{([p]k)})$  implies  $v_i^k(l_i^{([j]k)}) \geq v_i^k(l_i^{([p]k)})$ .

### $\Delta$ ROC values

Using the  $\Delta$ ROC values rule, the approximation to the values of each level in each issue can be obtained using the following algorithm:

**Step 1** Ask each negotiator  $k$  to provide a ranking of the  $m_i$  levels in each issue  $i$ . Label the resulting different levels as  $l_i^{([1]k)}, \dots, l_i^{([m_i]k)}$ , ranked from lowest to highest, with  $v_i^k(l_i^{([1]k)}) = 0$  and  $v_i^k(l_i^{([m_i]k)}) = 1$ .

**Step 2** Ask each negotiator to provide a ranking of the difference of values between consecutive levels  $\Delta_{i1}^k, \dots, \Delta_{i(m_i-1)}^k$  (where  $\Delta_{it}^k = v_i^k(l_i^{([t+1]k)}) - v_i^k(l_i^{([t]k)})$ , for  $t = 1, \dots, m_i - 1$ ). For each  $t = 1, \dots, m_i - 1$ , let  $rank^k(t)$  denote the rank of  $\Delta_{it}^k$  within the set  $\{\Delta_{i1}^k, \dots, \Delta_{i(m_i-1)}^k\}$ . This rank is an integer between 1 and  $m_i - 1$ , with rank 1 denoting the highest difference.

**Step 3** Determine a rank order centroid for  $m_i - 1$  variables:

$$\Delta_{[t]}^k = \frac{1}{m_i - 1} \sum_{l=t}^{m_i-1} \frac{1}{l}, t = 1, \dots, m_i - 1. \quad (10)$$

**Step 4** For each  $t = 1, \dots, m_i - 1$ , set the values provided by the centroid approximation:  $\Delta_{it}^{k(ROC)} = \Delta_{[rank(t)]}^k$ .

**Step 5** The approximate values for the levels in issue  $i$ , are then defined as follows:

$$v_i^k(l_i^{([1]k)}) = 0$$

$$v_i^k(l_i^{([j]k)}) = \sum_{t=1}^{j-1} \Delta_{it}^{k(ROC)}, j = 2, \dots, m_i. \quad (11)$$

**Step 6** For each negotiator  $k$  ( $k = 1, 2$ ), the approximate value in issue  $i$  for level  $l_i^{([j]k)}$  ( $i = 1, \dots, n; j = 1, \dots, m_i$ ) based on the  $\Delta ROC$  values rule,  $v_i^{k(\Delta ROC)}(l_i^{([j]k)})$ , is equal to the approximate value of the respective level, according to the values for the level correspondence created in Step 1.

The rules we test in this work are hence: ROC weights and TRUE values, ROC weights and ROC values, and ROC weights and  $\Delta ROC$  values. For a more detailed description of these rules, namely the possibility of accepting ties in the rankings provided, see Sarabando and Dias (2010). We use simulation to test if the suggested alternatives are “good” alternatives. The behavior of these rules is studied for the three mediation criteria presented in the previous section. As we have already said, our objective is not to compare criteria but to compare rules, as the choice of a criterion depends on the preferences of the mediator (namely on equity vs. efficiency considerations).

## 4 Simulations

Considering bilateral negotiation problems it is possible to vary the number of issues in negotiation and the number of levels in each issue. Since not all the issues have necessarily the same number of levels, the combination of different possibilities for these values leads to problem dimensions that can differ a lot from problem to problem. In our study, to compare the decision rules, we considered situations with 3 and 5 issues, and 3 and 5 levels in each issue. In the case with 3 issues and 3 levels in each issue there are a total of 27 alternatives, for 3 issues and 5 levels there are 125 alternatives, for 5 issues and 3 levels there are 243 alternatives, and for 5 issues and 5 levels there are 3125 alternatives. The resulting four templates will be denoted by 3\*3, 3\*5, 5\*3, and 5\*5, respectively. We deem that 5 issues is already a large number to consider

in practice (for instance, cost, quality, deadline, and warranty are typically the issues under negotiation in a commercial transaction), and 5 levels per issue is already a reasonable amount of detail.

We studied two scenarios with respect to reservation values: one scenario corresponds to the case with no reservation values, and in another scenario we considered that the reservation values of both negotiators are equal to 0.5. These are arbitrary choices, used to assess if this factor has any effect on the conclusions of the simulations. Finally, we considered, without loss of generality, that for the first negotiator all the issues are to maximize, and the opposite for the second negotiator.

To obtain random examples to the four templates presented we used Monte-Carlo Simulations. The level values were generated from a uniform distribution in the interval  $[0,1]$  and then normalized in such a way that the highest value in each issue would be equal to one and the lowest value would be equal to zero. We sorted these values to obtain monotonic value functions. The uniform distribution was also considered in Salo and Hämäläinen (2001) and Ahn and Park (2008). Barron and Barrett (1996) used both uniform and normal distributions to compare the hit rate and value loss of four rules, concluding that using the normal distribution did not significantly influence the results. The scaling weights were also generated according to a uniform distribution in  $W^k$  (acceptable set of weights for negotiator  $k$ ) using the process described in Butler et al. (1997). Note that, as sustained in Barron and Barrett (1996), if the information set on the weights is specified entirely by a complete ranking of the alternatives, then no point in  $W^k$  may be considered more likely than another, and the density of the weights is uniform over  $W^k$ .

For each random problem, defined by a value matrix and a weights vector, multiattribute value function model (1) provides the overall value of each alternative, which induces what we call the supposedly true ranking, i.e., the ranking that would be obtained if this cardinal information was known. On the other hand, each of the rules produces rankings using ordinal information about the weights vector and the values matrix. Since the decision rules use surrogate parameter values that are an approximation to the supposedly true values, the rankings in general will not coincide. We will use the notation  $x^{(real-sum)}$  to denote the supposedly best alternative according to the criterion of maximizing the sum of the values, and the notation  $x^{(rule-sum)}$  to denote the best alternative provided by the rule according to the same criterion, and similarly for the other mediation criteria.

Comparing results according to the supposedly true parameters with the results according to the decision rule used, we determined the following indicators for each simulation run:

- The position that the best alternative according to the true ranking reaches in the ranking generated by the decision rule used (this allows us to know the minimum number of alternatives that must be chosen, beginning by the top of the ranking provided by the rule, so that the true best alternative belongs to the chosen set).

- The position that the best alternative in the ranking generated by the rule reaches in the supposedly true ranking (this allows us to know how good the alternative chosen by the rule is in terms of the supposedly true ranking).
- The hit rate, i.e., the proportion of cases in which the best alternatives in the two rankings coincide. The hit rate considering the position of the supposedly best alternative in the ranking induced by the different rules is typically equal to the hit rate considering the position of the best alternative using the different rules in the supposedly true ranking, but it may differ in the presence of ties in the top of the ranking.
- The value loss, i.e., the difference between the real value of the best alternative and the real value of the alternative chosen by the rule, in the cases in which the two alternatives did not match. This allows us to know if the alternatives chosen by the rules have global value much inferior to the best alternatives in reality. For example, for the criterion of maximizing the sum of the values, the value loss is given by:

$$[v^1(x^{(real-sum)}) + v^2(x^{(real-sum)})] - [v^1(x^{(rule-sum)}) + v^2(x^{(rule-sum)})]. \quad (12)$$

## 5 Results

In this section we present the results of the simulations described in the previous section, considering 5000 random problem instances, for each problem dimension (the same number of simulations performed, for example, in Sarabando and Dias, 2009 and Sarabando and Dias, 2010). Throughout this section, “ROC TRUE” refers to the use of ROC weights and TRUE values, “ROC  $\Delta$ ROC” refers to the use of ROC weights and  $\Delta$ ROC values, and “ROC ROC” refers to the use of ROC weights and ROC values. All the tables we will refer to are presented in Appendix A.

We consider that an alternative needs to be better than the reservation values, for both parties, to be acceptable. In our simulation study, we eliminated the cases in which there was no alternative better than the reservation values for both parties: to determine a realistic value loss we decided to analyze only the cases in which the alternative provided by each rule is acceptable. Indeed, no value loss would occur if the alternative proposed by the mediator was unacceptable for one of the parties, since there would be no agreement.

Table 2 presents results related to the position of the best alternative according to the different rules in the supposedly true ranking (i.e., hit rate), not considering reservation values and considering reservation values. It shows, for each rule and each criterion, the average of the position (the minimum position was always equal to 1), and the proportion of cases in which the position

reached is equal to 1,  $\leq 2$  and  $\leq 20$ . To better observe the general patterns referred below see Figure 1:

- As expected, since the total number of alternatives is high, the best alternative according to the rule is not frequently the best one.
- Increasing the number of issues decreases the hit rate; the same happens increasing the number of levels in each issue. This is a natural result, since increasing the number of issues, or increasing the number of levels in each issue, increases the number of alternatives.
- As expected, the worst results are obtained using the ROC ROC rule, since this is the case in which less information is required from the negotiators, and the best results are obtained using the ROC TRUE rule. It is noteworthy that the results of the ROC  $\Delta$ ROC rule are closer to the ones of the ROC TRUE rule than the ones of the ROC ROC rule.
- Although it is not our objective to compare the criteria, it is possible to observe that for the PoP criterion the hit rate is not as high as it is for the other criteria, for all the rules.

Note that the results are sensitive to the number of issues and to the number of levels in each issue, because as the number of issues increases and as the number of levels in each issue increases, the number of alternatives also increases. For example, for 3 issues and 3 levels in each issue, the base rate for a coincidence at the top of the ranking is equal to 3.704%. For 5 issues and 5 levels in each issue, this base rate is equal to 0.032%.

We used paired t-tests to see if a significant difference occurs in the average of the hit rate without reservation values and with reservation values for all the criteria (sum, product and PoP), after verifying that the normality assumption is not violated (the hypothesis tests are similar to those presented in Sarabando and Dias, 2009). In Table 3 we present the p-values, i.e., the lowest significance levels at which the hypothesis of equality of the rules can be rejected, and the observed value of the test statistic. Since the p-values are high, it is not possible to conclude that the use of reservation values significantly affects the quality of the approximations provided by the rules.

We also conducted some hypothesis tests to check if the difference between the rules is significant in what concerns the average of the hit rate, using ANOVA for repeated measures. This requires multivariate normality, homogeneity of covariance matrices, and sphericity. The Mauchly's test of sphericity is used to test the last assumption. However this test is not robust for small samples, and if the dimensions are small even severe violations may not be detected. So we used the Greenhouse-Geisser Epsilon correction factor (the most adequate for small dimensions). In the cases in which the equality of the rules can be rejected, we used paired t-tests to see between which rules is the difference significant. We considered significance levels equal to 1% and 5%, which are values often used in practice. In Table 4 and Table 5 we show the p-values

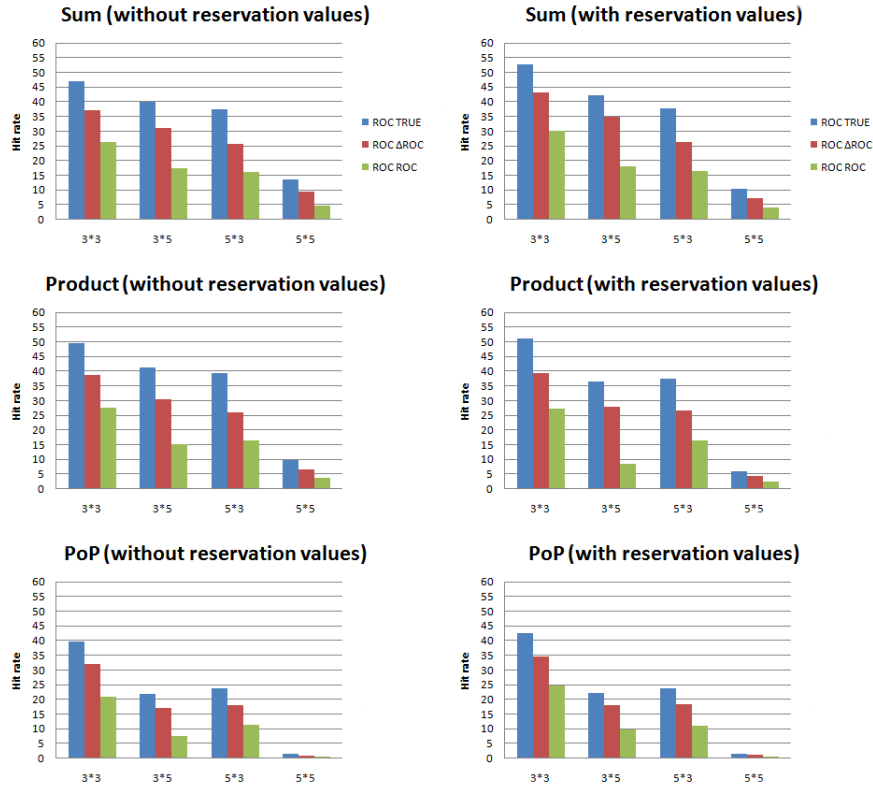


Figure 1: Comparison of the hit rate of the different rules.

related to the Greenhouse-Geisser Epsilon correction and to the paired t-tests, respectively:

- For the sum criterion, using and without using reservation values, the difference between the average of the hit rates can be considered significant for a significance level equal to 5%. Without reservation values, the difference can also be considered significant for a significance level equal to 1%. However, in this case the observed power of the test (this is, the probability of rejecting correctly the equality of the rules) is equal to 0.652, a very low value (values inferior than 0.8 should be interpreted carefully). For a significance level equal to 5%, both without and with reservation values, analyzing the p-values related to the paired t-tests it is possible to conclude that all the rules are significantly different. For a significance level equal to 1%, without reservation levels, it is not possible to conclude that the difference between the rules is significant.
- For the product criterion, using and without using reservation values, the

difference between the average of the hit rates can be considered significant for a significance level equal to 5%, but not for a significance level equal to 1%. However, with reservation values, the observed power is equal to 0.693. For a significance level equal to 5%, without reservation values, analyzing the p-values related to the paired t-tests it is possible to conclude that all the rules are significantly different. With reservation levels it is not possible to conclude that the difference between the ROC  $\Delta$ ROC rule and the ROC ROC rule is significant.

- For the PoP criterion, the difference between the average of the hit rates can not be considered significant for a significance level equal to 5%.

Results related to the position of the supposedly best alternative in the ranking induced for each rule are presented in Table 6, without reservation values and with reservation values. This table enables us to know how many alternatives should be retained to guarantee the retention of the supposedly best alternative. In Figure 2 it is possible to see how many alternatives should be chosen to guarantee the retention of the supposedly best alternative, in the case with 5 issues and 3 levels in each issue (243 alternatives), i.e., this figure shows the proportion of cases in which the best alternative is chosen depending on the number of alternatives retained. Since the total number of alternatives is high, retaining only one alternative is not enough in the majority of the cases. Obviously, the probability of retaining the supposedly best alternative increases with the number of alternatives that are retained. For the dimensions tested, it is possible to conclude:

- Without considering reservation values, retaining 20 alternatives is enough to retain the supposedly best one in more than 35% of the cases. Results not presented in the table indicate that, considering the ROC TRUE and ROC  $\Delta$ ROC rules, retaining 10 alternatives is enough to retain the supposedly best alternative in more than 42% of the cases.
- The results are worse considering reservation values (the corresponding percentages are 13% and 10%, respectively).
- Note that these values are not very high, since in the case with 5 issues and 5 levels in each issue there are 3125 alternatives. In this case, retaining 20 alternatives corresponds to retaining 0.64% of the total number of alternatives.

Table 7 depicts the average, the standard deviation and the maximum of the value loss, using the different rules and considering the different criteria, not using reservation values and using reservation values. Figure 3 presents some of the results. As expected, the worst results were obtained using the ROC ROC rule. However, even using this rule, the value loss is not very high. It is not possible to compare the value loss without reservation values and with reservation values, since the results are not comparable. However we used hypothesis tests to see if the difference between the average of the hit rate of the rules is significant. The results are presented in Tables 8 and 9:

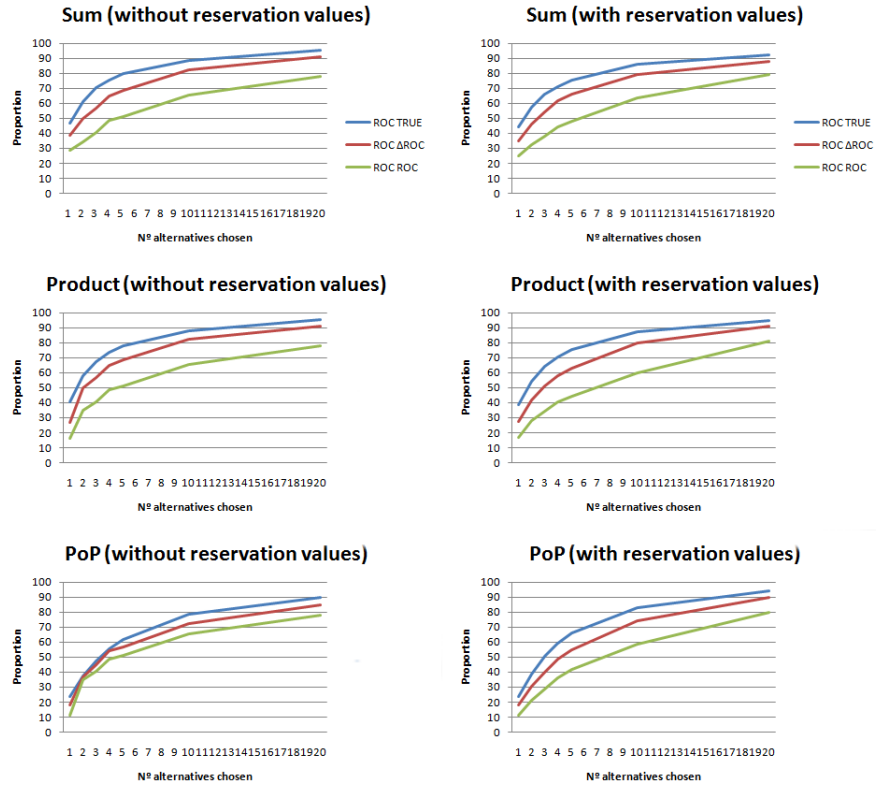


Figure 2: How many alternatives should be retained to guarantee the retention of the supposedly best alternative (5 issues and 3 levels in each issue)?

- For the sum criterion, using and without using reservation values, the difference between the average of the value losses can be considered significant for a significance level equal to 1%. However, with reservation values, for a significance level equal to 1%, the observed power of the test is equal to 0.622. Using paired t-tests it is possible to conclude that, for a significance level equal to 5%, with and without reservation values, the difference between all the rules can be considered significant. Without reservation values, for a significance level equal to 1%, the ROC TRUE rule and the ROC  $\Delta$ ROC rule can not be considered different. With reservation values, for a significance level equal to 1%, the difference between the rules can not be considered significant.
- For the product criterion, without using reservation values, the difference between the average of the value losses can be considered significant for a significance level equal to 1%. Using paired t-tests, for a significance level



equal to 5%, the difference between all the rules can be considered significant. For a significance level equal to 1%, the ROC TRUE rule and the ROC  $\Delta$ ROC rule can not be considered different. For a significance level equal to 5%, with reservation values, the difference between the rules can be considered significant. It is not possible to conclude that the difference between the ROC TRUE rule and the ROC  $\Delta$ ROC rule is significant.

- For the PoP criterion, without using reservation values, the difference between the average of the value losses can be considered significant for a significance level equal to 1% (in this case the observed power is equal to 0.774). With reservation values, the difference between the average of the value losses can be considered significant for a significance level equal to 5% (in this case the observed power is equal to 0.709), but not for a significance level equal to 1%. Without reservation levels, using paired t-tests it is possible to conclude that for a significance level equal to 5%, the difference between all the rules can be considered significant, for a significance level equal to 1%, the ROC TRUE rule and the ROC  $\Delta$ ROC rule can not be considered different. With reservation levels it is possible to conclude that all the rules are significantly different.

## 6 Conclusions

In this work we considered a setting of multiple-issue bilateral integrative negotiation, in which the preferences of both parties can be modelled by an additive value function. We compared the performance of three decision rules (ROC TRUE, ROC  $\Delta$ ROC and ROC ROC) to help a mediator suggesting an alternative under these circumstances, considering that the mediator does not have complete information (there exists ordinal information both on the scaling weights and on the levels' values), and tested them using Monte-Carlo Simulation (i.e., Monte-Carlo Simulation was used to assess the effectiveness of the three decision rules under several experimental conditions). Using four problem dimensions as templates for generating random examples we compared the rules under three mediation criteria: maximizing the sum of the values, maximizing the product of the excesses regarding the reservation values and maximizing the minimal PoP.

In the simulation study we determined the proportion of cases in which each rule chooses the truly best alternative and, whenever this did not happen, we determined the value loss. The best results were obtained using the ROC TRUE and ROC  $\Delta$ ROC rules. There are cases in which it is possible to consider that these rules are equivalent which is, in a certain way, surprising, because the ROC TRUE rule requires cardinal information about the levels' values from the negotiators. The ROC  $\Delta$ ROC rule is, almost always, better than the ROC ROC rule.

In Table 1 it is possible to see the main conclusions of the statistic tests we performed, using a significance level equal to 0.05. The conclusions in italics

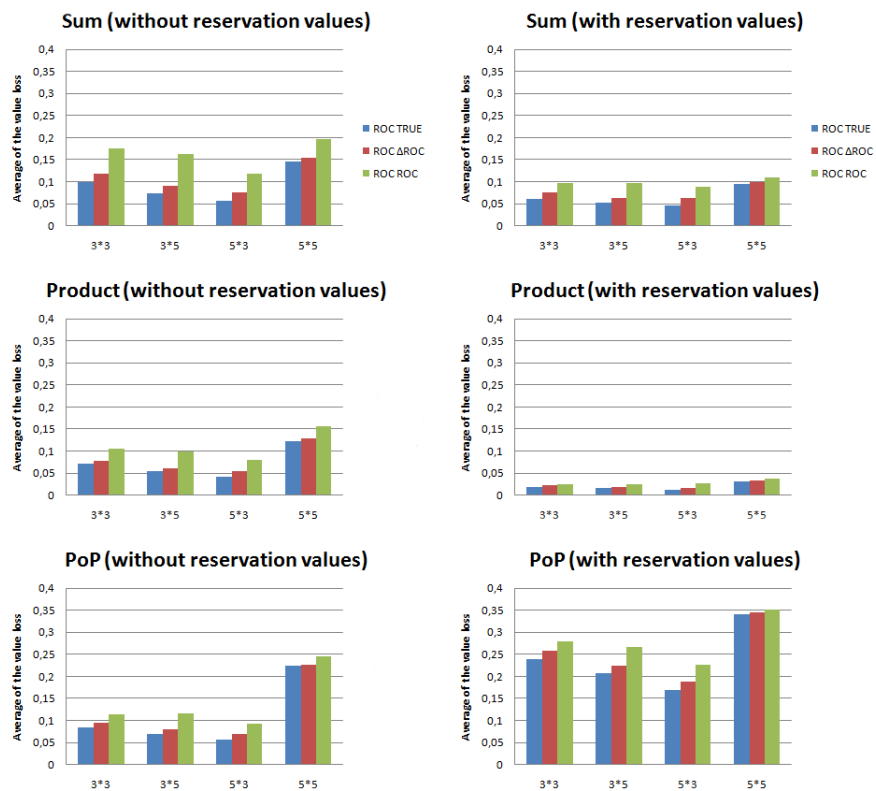


Figure 3: Comparison of the average of the value loss of the different rules.

should be read carefully since in these cases the observed power of the test is not very high.

We also compared the ranking of the alternatives according to the supposedly true parameters with the ranking of the alternatives according to the decision rule used. We considered the following results: the position that the best alternative according to the true ranking reaches in the ranking generated by the decision rule used and the position that the best alternative in the ranking generated by the rule reaches in the supposedly true ranking. The position that the best alternative according to the true ranking reaches in the ranking generated by the decision rule used allowed us to assess the strategy of retaining a small subset of alternatives instead of only one.

In our opinion the results are encouraging. Considering that the total number of alternatives is high obviously the hit rate cannot be expected to be high (the base rate for a coincidence at the top of the ranking is very low), but when the rule does not choose the real true alternative the average value loss is relatively low. The position that the best alternative according to a rule reaches

Without reservation values	With reservation values
<b>HIT RATE</b>	
<b>Sum</b>	
ROC TRUE better than ROC $\Delta$ ROC ROC TRUE better than ROC ROC ROC $\Delta$ ROC better than ROC ROC	ROC TRUE better than ROC $\Delta$ ROC ROC TRUE better than ROC ROC ROC $\Delta$ ROC better than ROC ROC
<b>Product</b>	
ROC TRUE better than ROC $\Delta$ ROC ROC TRUE better than ROC ROC ROC $\Delta$ ROC better than ROC ROC	<i>ROC TRUE better than ROC <math>\Delta</math>ROC</i> <i>ROC TRUE better than ROC ROC</i> <i>ROC <math>\Delta</math>ROC and ROC ROC can not be considered different</i>
<b>PoP</b>	
it is not possible to conclude that the average of the rules is different	it is not possible to conclude that the average of the rules is different
<b>VALUE LOSS</b>	
<b>Sum</b>	
ROC TRUE better than ROC $\Delta$ ROC ROC TRUE better than ROC ROC ROC $\Delta$ ROC better than ROC ROC	ROC TRUE better than ROC $\Delta$ ROC ROC TRUE better than ROC ROC ROC $\Delta$ ROC better than ROC ROC
<b>Product</b>	
ROC TRUE better than ROC $\Delta$ ROC ROC TRUE better than ROC ROC ROC $\Delta$ ROC better than ROC ROC	ROC TRUE better than ROC $\Delta$ ROC ROC TRUE better than ROC ROC ROC $\Delta$ ROC and ROC ROC can not be considered different
<b>PoP</b>	
ROC TRUE better than ROC $\Delta$ ROC ROC TRUE better than ROC ROC ROC $\Delta$ ROC better than ROC ROC	<i>ROC TRUE better than ROC <math>\Delta</math>ROC</i> <i>ROC TRUE better than ROC ROC</i> <i>ROC <math>\Delta</math>ROC better than ROC ROC</i>

Table 1: Main conclusions of the statistic tests (significance level equal to 5%).

in the supposedly true ranking complements the results that show if the alternative chosen by the rule is typically a good alternative. It was possible to see that, although in some cases the hit rate is relatively high, retaining one alternative is not sufficient in the majority of the cases. However, instead of one alternative, the mediator can keep, for example, a set of 20 alternatives. Note that retaining 20 alternatives does not seem to be a very encouraging result, and it is not if the initial set of alternatives is small, but, when there are hundreds or thousands of alternatives, it can be very useful for a mediator to focus on a subset of 20 alternatives.

The difference between the results related to the average of the hit rate, considering reservation values (equal to 0.5 for both negotiators) and without considering reservation values, can not be considered significant. Obviously the same does not happen with the average of the value loss, since using reservation values leads to different results (lower values losses for the sum and product criteria, and higher value losses for the PoP criterion). We also studied scenarios where we considered that the reservation value of one negotiator is equal to 0.5, and the reservation value of the other negotiator is equal to 0 (despite not showing the results in this text). It was possible to conclude that the hit rate for the product and for the PoP criteria is slightly higher in this case. For the sum criterion, the values of the hit rate are almost always between the other two cases (no reservation values and reservation values equal to 0.5 for both negotiators).

Regarding the decision rules we compared, the main results of the paper are: in almost all the cases, ROC weights and true cardinal values is better

than ROC weights and  $\Delta$ ROC values, which is better than ROC weights and ROC values. The results are somewhat sensitive to the mediation criteria used by the mediator (maximizing the sum of the values, maximizing the product of the excesses regarding the reservation values or maximizing the minimal PoP) and to the criteria of comparison (hit rate or value loss). For example, for the PoP criterion, it is not possible to conclude that the average of the rules is different, but considering the value loss it is possible to conclude that the ROC TRUE rule is better than the ROC  $\Delta$ ROC rule which is better than the ROC ROC rule. Considering the hit rate, for the PoP criterion, it is not possible to conclude that the average of the rules is different, but considering the sum criterion it is possible to conclude that the ROC TRUE rule is better than the ROC  $\Delta$ ROC rule which is better than the ROC ROC rule.

The values of the hit rate are sensitive to the number of issues and to the number of levels in each issue. As the number of issues increases and as the number of levels in each issue increases, the number of alternatives also increases, decreasing the hit rate. However, changing the problem dimension does not affect the conclusion that the ROC TRUE rule is better than the ROC  $\Delta$ ROC rule which is better than the ROC ROC rule.

When conducting the hypotheses tests there were cases in which it was not possible to conclude that the different rules lead to different results. Note that the fact of not rejecting the equality of the rules does not mean they are equal, but it means that the simulations do not provide arguments to state that the rules lead to different results.

We consider that testing the quality of approximations based on centroids to other kind of constraints is an interesting issue for future research, and we also believe that the presented type of approximations can also be extended to other methods besides the multiattribute additive value model.

While our study has led to some interesting results, it also has several limitations which indicate the need for future research. An important next step in our research will therefore consist in creating a larger empirical basis, to apply the rules to real data from (experimental) negotiations.

## References

- B. Ahn, K. Park (2008) ‘Comparing methods for multiattribute decision making with ordinal weights’, *Computers and Operational Research*, Vol. 35 (5), pp.1660–1670.
- F. Barron, B. Barrett (1996) ‘Decision quality using ranked attribute weights’, *Management Science*, Vol. 42 (11), pp.1515–1523.
- M. Beynon, H. Barton (2008) ‘A PROMETHEE based uncertainty analysis of UK police force performance rank improvement’, *Int. J. Society Systems Science*, Vol. 1 (2), pp.176–193.
- J. Brans, P. Vincke (1985) ‘A preference ranking organisation method (the

- PROMETHEE method for multiple-criteria decision making)', *Management Science*, Vol. 42, pp.647–656.
- J. Butler, J. Jia, J. Dyer (1997) 'Simulation techniques for the sensitivity analysis of multi-criteria decision models', *European Journal of Operational Research*, Vol. 103 (3), pp.531–546.
- J. Clímaco, L. Dias (2006) 'An approach to support negotiation processes with imprecise information multicriteria additive models', *Group Decision and Negotiation*, Vol. 15 (2), pp.171–184.
- L. Dias, J. Clímaco (2005) 'Dealing with imprecise information in group multi-criteria decisions: A methodology and a GDSS architecture', *European Journal of Operational Research*, Vol. 160, pp.291–307.
- W. Edwards, F. Barron (1994) 'SMARTS and SMARTER: Improved simple methods for multiattribute utility measurement', *Organizational Behavioral Human Decision*, Vol. 60 (3), pp.306–325.
- H. Ehtamo, R. Hämäläinen, P. Heiskanen, J. Teich, M. Verkama, S. Zionts (1999) 'Generating pareto solutions in a two-party setting: Constraint proposal methods', *Management Science*, Vol. 45 (12), pp.1697–1709.
- R. Fisher, W. Ury, B. Patton (1992) *Getting to yes: negotiation an agreement without giving in*, Century Business, London.
- M. Hajeesh (2008) 'Water conservation in the Gulf Cooperation Council countries: a decision support system approach', *Int. J Society Systems Science*, Vol. 1 (1), pp.84–99.
- P. Heiskanen (1999) 'Decentralized method for computing pareto solutions in multiparty negotiation', *European Journal of Operational Research*, Vol. 117, pp.578–590.
- R. Lahdelma, K. Miettinen, P. Salminen (2003) 'Ordinal criteria in stochastic multicriteria acceptability analysis (SMAA)', *European Journal of Operational Research*, Vol. 147 (1), pp.117–127.
- G. Lai, C. Li, K. Sycara (2006) 'Efficient multi-attribute negotiation with incomplete information', *Group Decision and Negotiation*, Vol. 15, pp.511–528.
- G. Lai, K. Sycara (2009) 'A generic framework for automated multi-attribute negotiation', *Group Decision and Negotiation*, Vol. 18, pp.169–187.
- D. Pruitt (1983) 'Strategic choice in negotiation', *The American Behavioral Scientist*, Vol. 27 (2), pp.167–194.
- H. Raiffa, J. Richardson, D. Metcalfe (2002) *Negotiation analysis: the science and art of collaborative decision making*, Belknap Press of Harvard, University Press, Cambridge (Ma).

- A. Salo, R. Hämäläinen (2001) ‘Preference ratios in multiattribute evaluation (PRIME) Elicitation and decision procedures under incomplete information’, *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, Vol. 31 (6), pp.533–545.
- P. Sarabando, L. Dias (2009) ‘Multi-attribute choice with ordinal information: a comparison of different decision rules’, *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, Vol. 39 (3), pp.545–554.
- P. Sarabando, L. Dias (2010) ‘Simple procedures of choice in multicriteria problems without precise information about the alternatives’ values’, *Computers and Operations Research*, Vol. 37, pp.2239–2247.
- P. Sarabando, L. Dias, R. Vetschera (2009) ‘Approaches to suggest potential agreements: Perspectives of mediation with incomplete information’, Tech. rep., Research Reports of INESC Coimbra, No. 11.
- T. Saaty (1996) *Decision making with dependence and feedback: the analytic network process*, RWS Publications, Pittsburg.
- K. Thomas (1992) ‘Conflict and conflict management: Reflections and update’, *Journal of Organizational Behavior*, Vol. 13, pp.265–274.
- L. Thompson (2001) *The mind and heart of the negotiator (2nd ed)*, Upper Saddle River, Prentice-Hall.
- R. Vetschera (2009) ‘Learning about preferences in electronic negotiations - a volume based measurement method’, *European Journal of Operational Research*, Vol. 194, pp.452–463.
- R. Walton, R. McKersie (1965), *A Behavioral Theory of Labor Negotiations*, McGraw-Hill, New York.

## A TABLES

	Without reservation values				With reservation values			
	Sum							
3 * 3	Average	% 1	% < 2	% < 20	Average	% 1	% < 2	% < 20
ROC TRUE	2.88	47.14	68.18	99.78	2.02	52.60	75.59	100
ROC ΔROC	3.74	37.10	56.88	98.86	2.34	43.19	67.33	100
ROC ROC	5.52	26.34	43.46	95.78	2.80	30.07	56.17	100
3 * 5	Average	% 1	% < 2	% < 20	Average	% 1	% < 2	% < 20
ROC TRUE	6.37	40.14	40.14	93.36	4.86	42.20	42.20	96.47
ROC ΔROC	9.04	31.14	31.14	88.56	6.16	34.89	34.89	94.75
ROC ROC	21.67	17.36	17.36	70.20	9.85	18.04	18.04	86.86
5 * 3	Average	% 1	% < 2	% < 20	Average	% 1	% < 2	% < 20
ROC TRUE	6.44	37.50	55.08	93.34	5.76	37.71	54.37	93.21
ROC ΔROC	10.19	25.60	49.64	91.02	7.95	26.40	41.08	89.65
ROC ROC	20.35	16.04	34.62	77.74	12.13	16.44	27.83	80.79
5 * 5	Average	% 1	% < 2	% < 20	Average	% 1	% < 2	% < 20
ROC TRUE	214.82	13.48	40.98	67.16	186.31	10.45	10.45	37.80
ROC ΔROC	246.80	9.32	31.24	55.30	211.9441	7.12	7.12	32.56
ROC ROC	407.48	4.66	23.18	35.42	260.88	4.09	4.09	21.28
	Product							
3 * 3	Average	% 1	% < 2	% < 20	Average	% 1	% < 2	% < 20
ROC TRUE	2.49	49.58	71.32	99.98	1.93	51.23	76.81	100
ROC ΔROC	2.94	38.56	62.34	99.98	2.28	39.26	68.86	100
ROC ROC	4.28	27.44	46.46	99.86	2.73	27.17	57.60	100
3 * 5	Average	% 1	% < 2	% < 20	Average	% 1	% < 2	% < 20
ROC TRUE	5.90	41.20	41.20	94.36	4.84	36.54	36.54	97.68
ROC ΔROC	7.47	30.50	30.50	91.60	5.84	27.99	27.99	96.70
ROC ROC	15.85	15.06	15.06	73.84	9.54	8.59	8.59	89.43
5 * 3	Average	% 1	% < 2	% < 20	Average	% 1	% < 2	% < 20
ROC TRUE	5.76	39.36	56.72	94.36	4.75	37.49	53.36	96.57
ROC ΔROC	8.66	26.08	49.64	91.02	6.65	26.56	41.12	92.90
ROC ROC	16.32	16.40	34.62	77.74	10.27	16.48	27.47	85.71
5 * 5	Average	% 1	% < 2	% < 20	Average	% 1	% < 2	% < 20
ROC TRUE	279.69	9.74	40.98	67.16	229.51	6.02	6.02	24.69
ROC ΔROC	302.33	6.72	31.24	55.30	252.65	4.32	4.32	21.87
ROC ROC	453.63	3.82	23.18	35.42	291.82	2.35	2.35	14.08
	PoP							
3 * 3	Average	% 1	% < 2	% < 20	Average	% 1	% < 2	% < 20
ROC TRUE	2.48	39.54	66.86	100	2.17	42.67	68.62	100
ROC ΔROC	2.90	32.00	64.76	100	2.48	34.72	61.17	100
ROC ROC	3.76	20.88	48.40	99.96	2.72	24.84	55.40	100
3 * 5	Average	% 1	% < 2	% < 20	Average	% 1	% < 2	% < 20
ROC TRUE	6.47	21.96	35.80	94.36	5.63	22.24	35.05	97.36
ROC ΔROC	7.98	17.24	29.30	90.54	6.54	18.13	30.53	95.68
ROC ROC	13.67	7.64	14.92	77.80	8.80	9.66	18.69	90.84
5 * 3	Average	% 1	% < 2	% < 20	Average	% 1	% < 2	% < 20
ROC TRUE	7.36	23.82	39.80	92.78	6.52	23.85	38.68	94.06
ROC ΔROC	9.74	18.12	49.64	91.02	8.31	18.27	30.85	90.16
ROC ROC	15.78	11.30	21.24	75.68	11.62	10.91	20.47	83.08
5 * 5	Average	% 1	% < 2	% < 20	Average	% 1	% < 2	% < 20
ROC TRUE	390.62	1.42	40.98	67.16	325.83	1.39	1.99	9.29
ROC ΔROC	704.70	0.74	31.24	55.30	341.45	1.08	1.64	8.97
ROC ROC	810.22	0.66	23.18	35.42	353.05	0.64	1.09	6.23

Table 2: Position of the best alternative according to the different rules in the supposedly true ranking (in  $n * m$ ,  $n$  represents the number of issues and  $m$  the number of levels in each issue).

<b>ROC TRUE</b>	$t_{obs}$	p-value
Sum - Sum	-0.662	0.555
Product - Product	1.543	0.220
PoP - PoP	-1.119	0.345
<b>ROC <math>\Delta</math>ROC</b>	$t_{obs}$	p-value
Sum - Sum	-1.173	0.325
Product - Product	1.059	0.367
PoP - PoP	-1.748	0.179
<b>ROC ROC</b>	$t_{obs}$	p-value
Sum - Sum	-1.140	0.337
Product - Product	1.346	0.271
PoP - PoP	-1.383	0.261

Table 3: Comparison of the hit rate without using reservation values and using reservation values.

	<b>Without reservation values</b>	<b>With reservation values</b>
<b>Sum</b>	0.008	0.017
<b>Product</b>	0.021	0.034
<b>PoP</b>	0.056	0.055

Table 4: Comparison of the hit rate for the different rules (p-values - Greenhouse-Geisser Epsilon correction).

	<b>Without reservation values</b>		<b>With reservation values</b>	
<b>Sum</b>	$t_{obs}$	p-value	$t_{obs}$	p-value
ROC TRUE - ROC $\Delta$ ROC	5.315	0.013	4.582	0.020
ROC TRUE - ROC ROC	5.697	0.011	4.514	0.020
ROC $\Delta$ ROC - ROC ROC	5.108	0.015	3.665	0.035
<b>Product</b>	$t_{obs}$	p-value	$t_{obs}$	p-value
ROC TRUE - ROC $\Delta$ ROC	4.250	0.024	3.588	0.037
ROC TRUE - ROC ROC	4.249	0.024	3.582	0.037
ROC $\Delta$ ROC - ROC ROC	3.762	0.033	3.048	0.056

Table 5: Comparison of the hit rate for the different rules (p-values and observed values of the test statistic - paired t-tests).



	Without reservation values				With reservation values			
	Sum							
	Average	% 1	% < 2	% < 20	Average	% 1	% < 2	% < 20
3 * 3	2.34	59.94	76.28	99.80	1.76	61.08	82.43	100
ROC TRUE	2.76	54.94	69.22	99.70	2.16	55.34	76.82	100
ROC ΔROC	3.88	40.72	52.28	98.30	2.40	35.29	62.35	100
ROC ROC								
3 * 5	5.91	45.30	45.30	93.76	4.19	45.73	45.73	97.47
ROC TRUE	7.18	41.98	41.98	91.66	5.34	42.18	42.18	95.23
ROC ΔROC	14.84	34.58	34.58	78.92	6.74	27.13	27.13	93.12
ROC ROC								
5 * 3	4.93	46.78	61.18	95.44	5.97	43.93	57.60	92.16
ROC TRUE	7.71	38.40	49.64	91.02	8.33	34.81	46.19	88.15
ROC ΔROC	15.34	28.58	34.56	77.74	11.93	24.82	32.59	78.87
ROC ROC								
5 * 5	51.36	40.98	40.98	67.16	101.44	32.99	33.99	56.47
ROC TRUE	79.63	31.24	31.24	55.30	124.95	23.84	23.84	48.11
ROC ΔROC	203.47	23.18	23.18	35.42	183.91	16.33	16.33	32.22
ROC ROC								
	<b>Product</b>							
3 * 3	2.13	56.18	76.20	100	1.91	55.36	78.33	100
ROC TRUE	2.67	43.56	66.62	100	2.28	43.08	72.28	100
ROC ΔROC	4.06	27.44	53.60	100	2.50	27.17	65.11	100
ROC ROC								
3 * 5	5.37	42.98	42.98	95.16	4.90	37.46	37.46	97.30
ROC TRUE	7.10	32.06	32.06	91.94	5.97	28.72	28.72	95.84
ROC ΔROC	15.52	15.06	15.06	73.80	7.19	8.59	8.59	94.66
ROC ROC								
5 * 3	5.09	40.60	57.96	95.24	5.37	38.44	54.31	94.47
ROC TRUE	7.87	26.86	49.64	91.02	7.44	27.37	41.78	91.02
ROC ΔROC	15.58	16.40	34.62	77.74	11.81	16.48	27.85	81.03
ROC ROC								
5 * 5	149.90	29.24	40.98	67.16	124.57	18.28	18.28	39.73
ROC TRUE	193.95	21.04	31.24	55.30	143.64	12.47	12.47	34.41
ROC ΔROC	360.83	12.04	23.18	35.42	192.63	7.54	7.54	21.78
ROC ROC								
	<b>PoP</b>							
3 * 3	2.49	39.54	60.08	100	2.26	42.67	64.47	100
ROC TRUE	2.85	32.16	51.48	99.96	2.53	34.85	58.74	100
ROC ΔROC	4.22	20.88	31.04	100	2.66	24.84	53.32	100
ROC ROC								
3 * 5	6.90	21.96	34.02	93.82	5.58	22.24	36.46	96.70
ROC TRUE	8.49	17.26	28.18	90.36	6.51	18.13	31.21	94.93
ROC ΔROC	15.43	7.82	14.28	74.04	6.95	9.66	30.68	95.25
ROC ROC								
5 * 3	8.03	23.82	37.56	90.04	6.31	23.85	38.90	94.38
ROC TRUE	11.05	18.12	49.64	91.02	8.53	18.27	30.24	90.00
ROC ΔROC	19.63	11.30	34.62	77.74	12.26	10.91	20.93	79.62
ROC ROC								
5 * 5	517.89	4.30	40.98	67.16	184.25	4.59	4.59	18.25
ROC TRUE	564.49	3.26	31.24	55.30	194.37	3.64	3.64	17.43
ROC ΔROC	732.49	2.20	23.18	35.42	228.25	2.50	2.50	13.29
ROC ROC								

Table 6: Position of the supposedly best alternative in the ranking induced by the different rules.

	Without reservation values			With reservation values		
<b>Sum</b>						
3*3	Average	Std. Dev.	Maximum	Average	Std. Dev.	Maximum
ROC TRUE	0.0999	0.1069	0.8571	0.0610	0.0603	0.3825
ROC ΔROC	0.1189	0.1108	0.7266	0.0753	0.0693	0.4267
ROC ROC	0.1750	0.1609	1.2622	0.0967	0.0852	0.4722
3*5	Average	Std. Dev.	Maximum	Average	Std. Dev.	Maximum
ROC TRUE	0.0736	0.0855	0.7956	0.0519	0.0577	0.5851
ROC ΔROC	0.0907	0.0974	0.9864	0.0633	0.0642	0.4605
ROC ROC	0.1636	0.1517	1.1498	0.0970	0.0849	0.5164
5*3	Average	Std. Dev.	Maximum	Average	Std. Dev.	Maximum
ROC TRUE	0.0575	0.0671	0.5649	0.0469	0.0487	0.3391
ROC ΔROC	0.0754	0.0762	0.6106	0.0631	0.0599	0.3811
ROC ROC	0.1184	0.1116	0.7978	0.0874	0.0774	0.4799
5*5	Average	Std. Dev.	Maximum	Average	Std. Dev.	Maximum
ROC TRUE	0.1467	0.1453	0.9588	0.0940	0.0926	0.7352
ROC ΔROC	0.1539	0.1460	0.8653	0.0987	0.0917	0.6603
ROC ROC	0.1959	0.1573	0.9073	0.1106	0.0954	0.5963
<b>Product</b>						
3*3	Average	Std. Dev.	Maximum	Average	Std. Dev.	Maximum
ROC TRUE	0.0718	0.0802	0.5308	0.0190	0.0209	0.1773
ROC ΔROC	0.0779	0.0752	0.5050	0.0234	0.0250	0.1984
ROC ROC	0.1061	0.0867	0.5064	0.0257	0.0258	0.1855
3*5	Average	Std. Dev.	Maximum	Average	Std. Dev.	Maximum
ROC TRUE	0.0538	0.0627	0.4471	0.0154	0.0179	0.1855
ROC ΔROC	0.0608	0.0654	0.4932	0.0175	0.0194	0.2119
ROC ROC	0.0988	0.0805	0.5043	0.0256	0.0236	0.1724
5*3	Average	Std. Dev.	Maximum	Average	Std. Dev.	Maximum
ROC TRUE	0.0424	0.0498	0.0508	0.0129	0.0134	0.0110
ROC ΔROC	0.0536	0.0547	0.5029	0.0161	0.0157	0.1130
ROC ROC	0.0807	0.0703	0.5169	0.0268	0.0203	0.1297
5*5	Average	Std. Dev.	Maximum	Average	Std. Dev.	Maximum
ROC TRUE	0.1232	0.1185	0.7963	0.0316	0.0273	0.2071
ROC ΔROC	0.1285	0.1158	0.8010	0.0331	0.0277	0.1706
ROC ROC	0.1559	0.1214	0.7496	0.0365	0.0273	0.1816
<b>PoP</b>						
3*3	Average	Std. Dev.	Maximum	Average	Std. Dev.	Maximum
ROC TRUE	0.0842	0.0792	0.4695	0.2385	0.1987	0.9706
ROC ΔROC	0.0942	0.0875	0.5318	0.2576	0.2035	0.9528
ROC ROC	0.1148	0.1015	0.6690	0.2799	0.2134	0.9600
3*5	Average	Std. Dev.	Maximum	Average	Std. Dev.	Maximum
ROC TRUE	0.0701	0.0622	0.4489	0.2080	0.1738	0.9236
ROC ΔROC	0.0796	0.0686	0.5330	0.2247	0.1775	0.9512
ROC ROC	0.1166	0.0904	0.5776	0.2659	0.1938	0.9663
5*3	Average	Std. Dev.	Maximum	Average	Std. Dev.	Maximum
ROC TRUE	0.0575	0.0546	0.5653	0.1682	0.1177	0.9227
ROC ΔROC	0.0689	0.0599	0.5108	0.1888	0.1506	0.9340
ROC ROC	0.0927	0.0747	0.5682	0.2259	0.1637	0.8956
5*5	Average	Std. Dev.	Maximum	Average	Std. Dev.	Maximum
ROC TRUE	0.2235	0.1344	0.7737	0.3409	0.1796	0.8930
ROC ΔROC	0.2270	0.1342	0.8474	0.3460	0.1842	0.8740
ROC ROC	0.2455	0.1401	0.7871	0.3516	0.1758	0.8516

Table 7: Value loss.

	Without reservation values	With reservation values
<b>Sum</b>	0.003	0.009
<b>Product</b>	0.000	0.022
<b>PoP</b>	0.005	0.032

Table 8: Comparison of the value loss of the different rules (p-values - Greenhouse-Geisser Epsilon correction).

	Without reservation values		With reservation values	
	$t_{obs}$	p-value	$t_{obs}$	p-value
<b>Sum</b>				
ROC TRUE - ROC $\Delta$ ROC	-5.609	0.011	-4.627	0.019
ROC TRUE - ROC ROC	-7.791	0.004	-5.507	0.012
ROC $\Delta$ ROC - ROC ROC	-7.410	0.005	-5.085	0.015
<b>Product</b>				
ROC TRUE - ROC $\Delta$ ROC	-5.634	0.011	-4.382	0.022
ROC TRUE - ROC ROC	-13.709	0.001	-4.485	0.021
ROC $\Delta$ ROC - ROC ROC	-11.523	0.001	-3.099	0.053
<b>PoP</b>				
ROC TRUE - ROC $\Delta$ ROC	-4.923	0.016	-4.371	0.022
ROC TRUE - ROC ROC	-6.579	0.007	-3.776	0.033
ROC $\Delta$ ROC - ROC ROC	-6.013	0.009	-3.287	0.046

Table 9: Comparison of the value loss of the different rules (p-values and observed values of the test statistic - paired t-tests).