

Using Intervals for Global Sensitivity and Worst Case Analyses in Multiattribute Value Trees

Jyri Mustajoki, Raimo P. Hämmäläinen* and Mats R.K. Lindstedt

Helsinki University of Technology

Systems Analysis Laboratory

P.O. Box 1100, FIN-02015 HUT, Finland

E-mails: jyri.mustajoki@hut.fi, raimo@hut.fi, mats.lindstedt@sitcom.fi

*Corresponding author: Tel. +358-9-451 3054, fax +358-9-451 3096

Abstract. Sensitivity analyses have for long been used to assess the impacts of uncertainties on outcomes of decision models. Several approaches have been suggested, but it has been problematic to get a quick overview of the total impact of all the uncertainties. Here we show how interval modeling can be used for global sensitivity analyses in multiattribute value trees, and a nuclear emergency case is used to illustrate the method. The approach is conceptually simple and computationally feasible. With intervals the decision maker can include all the possible uncertainties and quickly estimate their combined impact. This is especially useful in high-risk decisions where a worst case type of sensitivity analysis is essential. By varying the intervals one can also examine which uncertainties have the greatest impact and thus need the most consideration. Global sensitivity analysis reveals how the outcome is affected by many simultaneous variations in the model.

Keywords. Multiple criteria analysis, Sensitivity analysis, Preference Programming, Interval modeling, Worst case analysis, Nuclear emergency management

1. Introduction

Dealing with uncertainties related to data and preferential judgments is an essential part of a practical decision analysis project. How this is done is likely to strongly affect the confidence that the decision makers (DM) have in the results. If the uncertainties are not well accounted for, the credibility of the decision analysis method used can suffer. Thus an easy-to-use and transparent method for examining the effects of the uncertainties is needed to ensure the DMs' commitment to the decision.

Sensitivity analyses are commonly used to analytically assess the impacts of uncertainties on outcomes of decision models. In this paper, our focus is on multiattribute value tree analysis (MAVT) as described in Keeney and Raiffa (1976). MAVT is based on structuring the decision problem into a value hierarchy, or value tree as it is also called. The topmost objective is the overall goal that the decision maker wishes to achieve. This objective is divided into sub-objectives and on the lowest level are the measurable attributes that are important for the decision problem at hand. The overall performance of each alternative is measured by a value function, which aggregates the performance measures of the alternatives on each attribute into a single overall value measure.

Previous research has suggested that intervals and incomplete judgments can be used to incorporate uncertainties directly in the modeling phase (see e.g. Arbel 1989, Salo and Hämäläinen 1992, 1995, 2001, Weber 1985, White et al. 1984). In an interval approach, preference judgments and the outcomes of the alternatives are presented as ranges including all the possible values. The approach is also called Preference Programming (Arbel 1989, Salo and Hämäläinen 1995, 2004) which reflects the fact that one can see how the preferences over the alternatives evolve as the information increases and the preference statements become less incomplete. Interval methods have been successfully used in applications including

environmental decision making and policy analysis (see e.g. Hämäläinen and Leikola 1996). Intervals can also be used in group decision making to incorporate the preferences of all participants in a single model (see e.g. Hämäläinen et al. 1992, Hämäläinen and Pöyhönen 1996, Kim et al. 1999).

In this paper, we present a way to carry out global sensitivity analysis in MAVT by Preference Programming. That is, we allow the model parameters to vary within given intervals representing the uncertainty ranges, and study the consequent changes in the results. The aim is to quickly assess the total impact of all uncertainties. The main benefit of our approach is that it is conceptually straightforward and computationally efficient. The method suggested here captures all the uncertainties in a single analysis, and intervals used in the model are easy to understand also for non-mathematicians who might not be familiar with probability distributions. In this type of analysis, it is the extreme possibilities, i.e. the worst case results, that are important and not the probability distributions associated with the preferences and the values of the alternatives.

This work fits into the framework of sensitivity analysis proposed by Rios Insua and French (1991) and Proll et al. (2001), in which one employs constraints on the model parameters to describe uncertain information. The Preference Programming approach applied here gives a convenient and efficient way to include the constraints in the weight ratios of the attributes and in the outcomes of the alternatives also in hierarchical value trees. However, some related concepts such as potential optimality (see Hazen 1985, Rios Insua and French 1991) may not be applicable when considering the analysis from the worst case perspective, as it may eliminate alternatives that are never optimal but perform reasonably well in all situations and are thus less risky.

The interval model leads to a set of linear extremum problems. In hierarchical MAVT models, the overall problem is computationally quick to solve, as the extremum problems on each branch of the value tree can be separately solved (Salo and Hämäläinen 1992, 1995), and in a typical model there are seldom more than ten sub-objectives under any objective in the hierarchy. For example, with our WINPRE software (Workbench for INteractive PREference Programming; Hämäläinen and Helenius 1997), the DM can immediately see the changes in the results when adjusting the intervals. This requires that the calculations can be done very fast, preferably within a few tenths of a second, which is not always possible with other approaches.

Our focus is on MAVT, and thus other models such as decision trees (see Clemen 1996) are beyond the scope of this work. However, the suggested approach may also be well suited for these, but then, for example the number of parameters may set limitations for its use (e.g. the model presented in Francos et al. (2003) involves 82 input parameters). For an overview of suitable global sensitivity analysis approaches for these cases such as variance based importance measures, Bayesian networks, etc., the reader is referred to the special issue of Reliability Engineering and System Safety on sensitivity analysis (Tarantola and Saltelli 2003).

Global worst case analysis is especially relevant in high-consequence decisions with a high level of uncertainty, such as in the case of nuclear emergencies. We have successfully applied multiattribute decision analysis and decision conferences to support nuclear emergency management (Hämäläinen et al. 2000), and the example used in this paper is taken from one of these case studies (Ammann et al. 2001).

This paper is organized as follows. Section 2 describes different approaches to carry out sensitivity analysis. Section 3 describes the interval sensitivity analysis method, and how to

use it in practice. An example of using interval sensitivity analysis in a nuclear accident case is given in Section 4, and Section 5 concludes the paper.

2. Sensitivity analyses – Why and how?

Uncertainties can be grouped in several different ways. French (1995) suggested a classification into three groups depending on the step of the analysis to which they belong; modeling, interpreting the results or exploring the model (Table 1).

Table 1. Classifying uncertainties following French (1995)

When modeling the decision problem:	When interpreting the results:	When exploring the model:
- uncertainty about what might happen or what can be done	- uncertainty about the appropriateness of a descriptive/normative model	- uncertainty from physical randomness or lack of knowledge
- uncertainty about meaning or ambiguity in terminology	- uncertainty about the depth to which to conduct the analysis	- uncertainty about the evolution of future beliefs and preferences
- uncertainty about related decisions		- uncertainty about judgments - uncertainty about accuracy of calculations

Another perspective is to focus on the origins of uncertainties. For example, in the field of risk analysis Salo (2001) suggests three dimensions of technological risks; physical causation (e.g. uncertainties in causal relationships), value concerns (e.g. changes in the stakeholders' preferences), and policy response (e.g. the effectiveness of the actions taken).

In the different cases, methods are needed to support the DM in dealing with the uncertainties in a constructive way. Sensitivity analyses are commonly used, and the reader is

referred to Saltelli et al. (2000a, 2000b) or Tarantola and Saltelli (2003) for a perspective on different approaches in general, or to Belton and Stewart (2001), French and Rios Insua (1999) or Rios Insua and French (1991) on different approaches specifically in MAVT.

Sensitivity analyses can be used for a wide range of purposes. Pannell (1997) grouped these into four categories:

1. Decision making (identifying critical values/parameters, testing robustness, overall riskiness of decision)
2. Communication (increasing commitment/confidence/credibility, explicitly showing critical assumptions)
3. Increased understanding (understanding relationship between input/output variables)
4. Model development (identifying needs for more accurate measurements/more information)

The single parameter test is a common sensitivity analysis method to examine how sensitive a model is to small changes in one parameter. That is, all other parameters are held fixed except for a single one that is allowed to vary. The analysis is usually visualized by graphs showing the consequential variations in the overall results. The tornado diagram is another common method providing useful diagrams. It is usually drawn for single parameter tests and employed to compare the base alternative to another option (Felli and Hazen 1999). For a detailed discussion of tornado diagrams, the reader is referred to Clemen (1996). These methods do not, however, account for parameter interactions and nor do they cover the worst case settings.

Often there is more than one parameter that the DM might be uncertain about and with the traditional approaches it can be difficult to estimate the combined overall impact. For

example, Felli and Hazen (1999) showed that single parameter tests tend to overestimate the overall sensitivity. This leads to the need for a multiparameter test.

Global sensitivity analysis incorporates the influences of the whole ranges of variation in model parameters and these variations are allowed in multiple parameters simultaneously (see e.g. Saltelli et al. 2000a). Especially in complex simulation models, for example when modeling environmental phenomena, the model can have numerous input parameters. Different types of global sensitivity analysis approaches have been developed for dealing with uncertainties in these models, and these approaches often rely on statistical or probabilistic calculations. Examples of such approaches include Monte Carlo analysis, ANOVA, FAST and Bayesian models (see e.g. Saltelli et al. 2000a, 2000b). Rank based methods can also be used to study the sensitivity of the model with respect to imprecision in the rankings of parameters (see e.g. Barron 1992 or Salo and Punkka 2004).

In MAVT, the Monte Carlo simulation technique can be used to analyze model uncertainties and to statistically rank the alternatives (see e.g. Arbel and Vargas 1993, Butler et al. 1997, Stam and Silva 1997). The main advantage of this method is that one gets a lot of information, such as mean values, variances and fractiles, about the characteristics of the decision model subject to uncertainties. However, the normalization of the weights and the reciprocity of the weight ratios make the application of distributions on the weights very difficult. In addition, the effort needed for the calculations can become substantial and one cannot always carry out what-if analyses without computational delays. To avoid this, Kirkwood (1992) has suggested a method to estimate the impact of uncertainty on the results of a multiattribute model prior to a complete probabilistic analysis. However, this would still demand the approximation of expected values, variances and covariances.

Kirkwood (1997) provides a summary of scenario analysis literature in decision making. In scenario analysis parameters are given, sometimes extreme, values so as to make the analysis favor a certain view of the actual situation. Often the set of scenarios contains the expected as well as an optimistic and a pessimistic alternative. The process of scenario generation is not, however, a straightforward one. For example, there is not a clear rule how pessimistic the pessimistic scenario should be.

The precautionary principle (see e.g. Goldstein and Carruth 2004 or Graham 2000) has recently received growing attention, and there is a similar need for global worst case analysis. The precautionary approach addresses the problems of multidimensionality, humility about knowledge, and openness to alternatives. It places the burden of proof on the advocates to prove the soundness of the suggested decision. Thus also in that approach a global sensitivity analysis or a worst case analysis would be useful.

3. Intervals in worst case sensitivity analysis

3.1 Interval methods

In MAVT the overall values of the alternatives are composed of the ratings of the alternatives in respect to each attribute, and of the weights of the attributes. If the attributes are mutually preferentially independent (see e.g. Keeney and Raiffa 1976), an additive value function can be used to derive the overall values. The overall value for the alternative x is

$$v(x) = \sum_{i=1}^n w_i v_i(x_i), \quad (1)$$

where n is the number of attributes, w_i is the weight of attribute i , x_i is the consequence of alternative x with respect to attribute i and $v_i(x_i)$ is its rating. The sum of the weights is normalized to one, and the ratings $v_i(x_i)$ are scaled onto the range $[0, 1]$. The weights denote

the relative importances of the attributes changing from their worst level to their best level compared to the changes in the other attributes.

The value tree can also be constructed hierarchically, i.e. the upper level objectives are divided into sub-objectives, and the weighting is carried out locally on each set of these. Then, w_i in (1) denotes the overall weight of the lowest level attribute i , which is calculated as a product of the local weight of this attribute and the local weights of all the preceding upper level objectives.

Preference Programming techniques can be used to model uncertainties in the DM's preference statements with intervals. PAIRS (Salo and Hämäläinen 1992) is a Preference Programming technique in which intervals are directly given to constrain both the weight ratios of any attribute pairs and the ratings of the alternatives. For example, instead of giving an exact weight ratio $w_1/w_2=2$, the DM can define that ratio $w_1/w_2 \in [1, 4]$, i.e. the ratio is at least 1 but no more than 4. The given intervals constrain the feasible region of the weights S . Figure 1 illustrates the feasible region constrained by intervals $w_1/w_2 \in [1, 4]$, $w_1/w_3 \in [1, 4]$ and $w_2/w_3 \in [1/2, 2]$. Similarly, uncertainties in decision outcomes can be modeled with ranges of possible values (e.g. $v_i(x_i) \in [0.3, 0.5]$). As a result, the overall values of the alternatives will also be intervals, which can be calculated as extremes of (1) with linear programming. The lower bound for the value of alternative x is

$$\underline{v}(x) = \min_{w \in S} \sum_{i=1}^n w_i \underline{v}_i(x_i), \quad (2)$$

where $\underline{v}_i(x_i)$ is the least allowed value for $v_i(x_i)$ and $w=(w_1, \dots, w_n) \in S$, i.e. the feasible region of the weights. The upper bound is calculated analogously. For details see Salo and Hämäläinen (1992).

Interval models require that we specify the dominance concepts to analyze the results (see e.g. Weber 1987, Salo and Hämäläinen 1992). Alternative x dominates alternative y absolutely, if the lower bound of the overall value interval of x is higher than the upper bound of the interval of y . Alternative x dominates y in a pairwise sense, if the overall value of x is higher than the overall value of y for every feasible weight and value combination, i.e. if

$$\min_{w \in S} \sum_{i=1}^n w_i [\underline{v}_i(x_i) - \bar{v}_i(y_i)] \geq 0 \quad (3)$$

and the inequality is strict at least for some $w \in S$. Thus, pairwise dominance can also exist under overlapping value intervals. Absolute dominance implies pairwise dominance, and in general, the term dominance is considered to refer to pairwise dominance.

3.2 Interval sensitivity analysis

In this paper we apply the PAIRS method to carry out interval sensitivity analysis in MAVT. That is, we extend the point estimates of weight ratios and attribute ratings into intervals to describe possible variation in these. By studying the consequential changes in the overall values and dominance relations, we can elicit how sensitive these are to combined variation in all the model parameters.

The dominance concepts can be used to study the changes in the ranking of the alternatives due to combined variations in model parameters. As long as the dominance relations between the alternatives remain unchanged, the result is not sensitive to any parameter variation within the given intervals. On the other hand, any of the non-dominated alternatives can be considered as a suitable candidate for the most preferred alternative.

In PAIRS, intervals can be simultaneously assigned to both the weight ratios between the attributes on any level of the value tree and the ratings of the alternatives. This makes it possible to study the joint effects of different types of uncertainties both in the DM's

preference assessments and in the input data on the alternatives. An extreme case is to have intervals on all the weight ratios and the ratings of the alternatives, which means that all the possible uncertainties in the model are simultaneously taken into account.

Interval sensitivity analysis can also be applied to the Analytic Hierarchy Process (AHP; Saaty 1980). In AHP, the attribute weights are computed from a matrix of all the pairwise weight ratio judgments. Thus, Preference Programming we can be used to carry out interval sensitivity analyses by setting intervals on all the pairwise judgments. However, in AHP one also has to take into account feasibility constraints due to the redundancy of comparisons.

3.2.1 Interpreting the intervals

By nature, interval sensitivity analysis is a worst case analysis, as any combinations of model parameters within the given intervals are allowed. However, in practice there are different ways to interpret the intervals. The strictest one is to set the intervals so that they indeed cover all the possible variation in the model parameters. Then, the sensitivity analysis can be seen as a way to find a true worst case solution. That is, if a dominating alternative is found, it is the best alternative for every feasible combination of model parameters, including the worst ones. However, with this interpretation, the intervals may easily become so wide that no dominance relations between the alternatives can be established.

One can also use tighter intervals that may not cover all the possible variation in parameter values. Then, the analysis should be interpreted as a “what-if” type sensitivity analysis. That is, we study *what* would be the overall value intervals, *if* we allowed the parameter values to be any values within the given intervals.

Yet another approach is to consider intervals as confidence intervals (e.g. a 95% confidence interval). However, if the distributions of parameter values on the intervals are

unknown, the confidence with which the overall values belong to the resulting intervals would also be unknown. That is, an overall value interval describes the possible variation of the overall value due to each parameter being with a certain confidence level (e.g. 95%) on its interval, but this does not imply that the overall value is on its interval with this same confidence. To get true confidence intervals for the overall values, one has to assign distributions on the parameters, but this leads to the use of a Monte Carlo simulation approach.

Attributes can have different levels of uncertainty, which can be taken into account by using intervals of different size. If the DM is not able or willing to give the intervals explicitly, he/she can, for example, assign error ratios to the point estimates (see e.g. Bryson et al. 1995, Salo and Hämäläinen 2001). They provide a quick way to model the proportional uncertainties. For example, with error ratio 2, the weight ratio $w_1/w_2=3$ extends to interval $w_1/w_2 \in [3/2, 3 \times 2] = [1.5, 6]$. Similarly, alternatives' ratings can be extended with error margins, e.g. $[v_i(x)-0.1, v_i(x)+0.1]$.

The choice of intervals could however be difficult if there is a high level of uncertainty involved. This could lead to a wide set of intervals and a non-acceptable risk region. In this case the value of the analysis would be in finding out the key factors that result in the non-acceptable risk region. That is, what uncertainties have the greatest impact and thus need the most consideration. A detailed analysis might even reveal how much the uncertainties need to be reduced to arrive at an acceptable outcome.

3.2.2 Potential optimality and the decision rules

In many cases it may be useful to first try to reduce the number of alternatives before the final decision on which alternative to choose. The dominance test eliminates clearly inferior

alternatives, but the remaining set can still be extensive. Rios Insua and French (1991) discuss the use of potential optimality (Hazen 1985) for reducing the number of alternatives. This method aims at the optimal solution and only considers alternatives that could potentially be the optimal one. In many cases this can be a useful approach, but not necessarily in the worst case analysis, as it may possibly reject non-optimal alternatives that might also be acceptable to the DM. For example, in the nuclear accident case presented here also a non-optimal alternative might be the best decision if it performs reasonably well in all the possible scenarios.

As a result of a worst case analysis there may be several non-dominated alternatives. The DM can then use what-if analyses to study with which value intervals there would be only one non-dominated alternative. Consequently, he/she can consider whether these intervals could be accepted to represent the related uncertainties. If the DM is not able or willing to modify the intervals, decision rules (see e.g. Salo and Hämäläinen 2001) can be applied to rank non-dominated alternatives, for example, according to the minimum values of the overall value intervals.

The possible loss of value estimates (Salo and Hämäläinen 2001) can also be used to further compare non-dominated alternatives. The possible loss of value for alternative x is calculated as

$$\max_{w \in S, y \neq x} \sum_{i=1}^n w_i [\bar{v}_i(y) - v_i(x)]. \quad (4)$$

This estimate indicates how much the DM can at most lose in the overall value by choosing alternative x instead of any other alternative. That is, the uncertainties might cause the chosen alternative to be overtaken by another alternative in some circumstances. However, if the possible loss of value is within acceptable limits, then the original choice might still be

acceptable. In high-risk situations, such as the one discussed in this paper, global sensitivity analyses can be used to eliminate alternatives that potentially could have disastrous outcomes in the worst case. Possible loss of value can then be used to further reduce the set of alternatives by specifying an acceptable limit for how far from optimality any alternative can be at the most.

3.2.3 *Origins of imprecision*

Depending on the origins of imprecision (see e.g. French, 1995) one can assign imprecision to the weights of the attributes or to the ratings of the alternatives. An interval assigned to a rating of an alternative describes imprecision only in this alternative, and any variation within the given interval is assumed to be independent of the allowed variations in the other alternatives. On the other hand, by definition the weight of an attribute should reflect the importance of the range of this attribute compared to the other attribute ranges (Keeney and Raiffa 1976). Consequently, any imprecision on the weight of an attribute affects in a similar way all alternatives as then the imprecision is actually assigned to the range of the attribute. Thus, for example, linear correlations in the variations of the ratings on some attribute could be modeled by assigning imprecision in the weight of this attribute and keeping the ratings constant. Then, the relative ratings of the alternatives in this attribute remain the same, but the impact of this attribute varies with respect to the other attributes.

As a simple example, consider attribute *Costs* on two *alternatives A* and *B*. If there is imprecision e.g. in some general costs, it is likely to affect similarly to both alternatives (i.e. if the *Costs* of *A* double, so does the *Costs* of *B*). Thus, this kind of imprecision should be assigned to the weight of *Costs*, as it is related to the importance of the range of *Costs*. However, if there are some single variations in the costs of the attributes that are independent

of each other, these should be modeled by assigning the imprecision directly to the ratings of the alternatives.

3.2.4 Computer support

In practice, computer support is needed as the solutions are elicited through linear programming. In hierarchical MAVT models, the overall linear extremum problem can be decomposed into smaller problems, i.e. so that there is one extremum problem for each branch of the hierarchy (Salo and Hämäläinen 1992, 1995). Thus, when varying a model parameter, the linear programming problems needs to be re-calculated only on those branches of the value tree in which the change is made, and upwards thereof. This makes the model computationally quick to solve and update, which is especially important in *what-if* type of analyses. Computer support can also be used, for example, to visualize the results.

As noted above, there already exist software for interval techniques considered here, and sensitivity analyses can be easily run on these software. The WINPRE software provides an interactive approach by presenting the overall values of the alternatives and the dominance relations immediately to the user when making changes to the model parameters. The user can, for example, study what are the parameter values where the dominance relations change, and what are the rates of the changes in the overall values when varying the parameter intervals. The example case in Section 4 is analyzed using WINPRE and all the figures are screen captures from the software. WINPRE is freely available for academic purposes on the Decisionarium Web site (www.decisionarium.hut.fi; Hämäläinen 2000, 2003).

3.3 Comparison with other sensitivity analysis approaches

Next we shall discuss the suitability of interval sensitivity analysis in different types of situations. We compare it to one-way sensitivity analysis and related techniques such as

tornado diagrams, and the Monte Carlo simulation technique, which are commonly used sensitivity analysis approaches in MAVT framework.

3.3.1 One-way sensitivity analysis

As a result, one-way sensitivity analysis gives a graph showing the overall values of the alternatives with respect to each possible value of some single attribute. Two-way sensitivity analysis extends this analysis into two parameters, when the combined effects of variations in these are presented with a three-dimensional graph. However, with more than two parameters the visualization of the analysis becomes impossible. Yet, with tornado diagrams one can study the effects of several different parameter variations in the same graph, but also these effects are calculated by varying a single parameter at a time.

Especially in hierarchical problems, interval sensitivity analysis as a multiparameter analysis may provide useful additional information compared to the one-way analysis. In these problems, the weighting is carried out locally on each branch of the value tree, and the overall weight of a lowest level attribute is obtained as a product of the local weights of this and all the preceding upper level attributes. Thus, if one wants to take the joint effects of uncertainties in different attribute levels as well as in alternatives into account, multiparameter analysis is needed. For example, variation in the weight of attribute *Costs* with respect to any other attribute could easily affect the overall results. However, if the *Costs* are further divided into two sub-attributes, *Purchase Costs* and *Maintenance Costs*, the variation in the local weights of these is also likely to affect the results, as often an alternative with high *Purchase Costs* has low *Maintenance Costs*, and vice versa. Thus, multiparameter analysis is needed to also get this variation into account.

On the other hand, the resulting graph in one-way sensitivity analysis directly shows the overall values for each possible value for the parameter under consideration. In contrast, in

interval sensitivity analysis only the extreme values of the overall values in respect of allowed variations in model parameters are shown. Thus, if one wants to study the overall value intervals resulting from intervals of different size, he/she has to produce a separate graph for each set of intervals, or use the method interactively, i.e. by varying the parameter intervals and studying the immediate response in the overall value intervals.

3.3.2 Monte Carlo simulation

In Monte Carlo simulation one assigns distributions on the model parameters. As a result, one gets the overall value distributions for the alternatives, which reflect the variation in the model parameters. One can also estimate additional statistical measures, such as a probability for an alternative being better than some other alternative.

On the ratings of the alternatives, distributions can easily be assigned directly within their ranges. On the weights of the attributes, there are different ways to assign the distributions, e.g. directly to the weights (Butler et al. 1997), to the feasible regions of the weights (Haines 1998; Moskowitz et al. 2000), or to the set of $n-1$ weight ratios, similarly as intervals are applied in the interval SMART/SWING method (Mustajoki et al. 2004). However, in practice the use of these approaches becomes very difficult due to normalization of the weights and reciprocity of the weight ratios. For example, distributions assigned to the weights are not the same after the normalization of the weights. Also, distributions on the feasible regions and weight ratios may lead to ambiguous distributions on the weight ratios that are not explicitly defined. E.g., a uniform distribution assigned for the weight ratio interval $w_1/w_2 \in [1, 2]$ implies that the corresponding distribution on reciprocal interval $w_2/w_1 \in [1/2, 1]$ is not uniform. In addition, if the uncertainty is assumed to originate from ratio based elicitation, the interpretation of the given distributions may be problematic.

In many cases, the DM is interested in these extreme values, e.g. to get a worst case solution, or study deterministic dominance relations between the alternatives. For this purpose, our interval sensitivity analysis should be used, as it explicitly concentrates on the extremes of the intervals. Monte Carlo simulation can be used to find additional statistical information on the value distributions and relations between the alternatives. However, the assignment of the distributions and sampling under constraints could often be an overwhelming task.

4. Interval sensitivity analysis in a nuclear accident exercise

We demonstrate the use of interval sensitivity analysis with a model developed in a decision conference, i.e. facilitated training workshop exercise. The case is a hypothetical nuclear accident where protective actions on the milk production chain were to be decided (Ammann et al. 2001). The conference was one in a series of decision conferences on nuclear emergency management (see Hämäläinen et al. 2000). Although no real decisions were made in the exercise, the gained experiences are to be utilized if a real accident took place. The participants of the conference and the preparatory meetings were representatives of the safety authorities and experts on radiation, farming and the dairy industry.

4.1 Multi-attribute value tree

A MAVT approach was used to structure the problem. The value tree (Figure 2) was developed on the basis of the discussion in the decision conference and in the preliminary meetings. It describes the objective hierarchy, in which the *Overall* objective of finding the best action is composed of three main objectives: *Health* effects, *Socio-psychological* effects and *Costs*. These are further divided into sub-objectives (i.e. attributes) and the alternatives are measured with respect to these. The overall values of the alternatives are calculated from

(2), in which the weight of each attribute elicited as a product of the local weights of this and all the preceding objectives.

It was assumed that due to the accident, fodder in the area becomes contaminated, and if nothing is done, the radioactivity migrates into milk products. As a precautionary action, the cattle were sheltered and provided with uncontaminated fodder and water for the first week after the accident. Our focus was on the later phase actions (from 1 to 12 weeks after the accident). Three protective action policies were considered: (i) *supplying clean fodder* ('Fod'), i.e. uncontaminated fodder is transported into the contaminated area, (ii) *production change* ('Prod'), i.e. the milk production is replaced by other dairy products, as the production processes of these can enrich, dilute or secrete radio nuclides, and (iii) *banning milk* ('Ban'), i.e. the use of contaminated milk is totally banned. In addition, an action where *nothing* is done ('---') was included in the analysis as a reference. The actions were divided into two phases. The first phase covers the actions during weeks 2 to 5 after the accident and the second phase covers weeks 6 to 12. The actual alternatives considered were combinations of these. For example, 'Ban+Fod' represents an alternative where the use of milk is banned during weeks 2 to 5 and clean fodder is supplied during weeks 6 to 12. In Figure 2, the alternatives are shown as the rightmost elements of the value tree. For further details, see Ammann et al. (2001).

The attribute weights and the alternatives' ratings used (Table 2) are based on the preferences of one of the participant groups in the conference. The estimates for the consequences of *Health* effects and *Costs* were calculated with a Real-time On-line DecisiOn Support System called RODOS (Ehrhardt and Weis 2000). The values of *Socio-psychological* attributes were directly rated by the group. Attribute weighting was carried out with the SWING method (von Winterfeldt and Edwards 1986). In SWING, the DM first identifies the

most important attribute (i.e. an attribute whose consequence he/she most preferably would change from its worst possible level to its best possible level), and gives this a hundred points. Then, he/she assigns fewer points to the other attributes to denote the relative importance of the consequence changes in these compared to the change in the most important attribute. Finally, the actual attribute weights are elicited by normalizing the sum of these points to one.

Table 2. Weights of the attributes and the ratings of the alternatives.

	Weights		Ratings					
	Local weight	Overall weight	---	Fod	Fod	Prod	Ban	Ban
			+	+	+	+	+	+
			---	---	Fod	Fod	Fod	Ban
Health	<i>0.588</i>							
Thyroid cancer	0.909	0.534	0.01	0.98	1.00	0.98	1.00	1.00
Other cancers	0.091	0.053	0.04	0.90	0.98	0.95	0.98	1.00
Socio-psychological	<i>0.294</i>							
Reassurance	0.526	0.155	0.00	0.18	0.60	0.48	0.76	0.86
Anxiety	0.053	0.016	0.00	0.21	0.87	0.75	0.49	0.35
Industry	0.158	0.046	0.10	0.46	0.67	0.34	0.19	0.14
Feasibility	0.263	0.077	1.00	0.65	0.55	0.78	0.67	0.46
Costs	<i>0.118</i>	0.118	1.00	0.88	0.82	0.70	0.26	0.00

The overall values for the alternatives are shown in Figure 3. The alternatives are listed at the bottom, and the upper and lower bounds for the overall value of each alternative are shown above its name. The bounds are also graphically shown on a [0–1] value scale between the numerical values. One should note that now the lower and upper bounds are the same, as both the attribute weights and the alternatives’ consequences are exact point estimates. The figure indicates that alternative ‘Fod+Fod’, where clean fodder is provided for both periods, is the best alternative with overall value of 0.86.

4.2 Sensitivity analysis

This example is a typical case where interval sensitivity analysis is useful. In emergency planning we often want to design precautionary actions following the worst case approach.

The alternative with the highest performance score need not be the right one to choose if a worst case analysis reveals a risk of getting an unacceptably low performance under certain circumstances. An alternative with a slightly lower scoring on an average might be preferred, if the analysis shows that it performs in an acceptable way in all possible circumstances.

To illustrate the need for multiparameter analysis, consider the phenomenon that can arise due to the hierarchical structure of the model. For example, alternative ‘----+---’ has a rating 1.00 on the attribute *Costs*, but a low rating on the *Health* effects. Thus, any variation in the weight ratio between these attributes should have a considerable effect on the overall results. However, variation between the local weights of *Reassurance* and *Feasibility* under the *Socio-psychological* effects should also be taken into account, as for example, alternative ‘----+---’ has ratings 0.00 and 1.00, respectively, on these attributes. Thus, a multiparameter analysis is needed to simultaneously take into account the possible variations in both these.

Next, we demonstrate the use of interval sensitivity analysis in this case. The objective is to study the changes in the relations between the alternatives when all the possible uncertainties in the problem are considered. The analysis consists of three phases. First, the sensitivity in the weight assessment is studied by extending the weight ratios to intervals. Next, the effects of the possible variation in the alternatives’ values are studied by giving these as intervals. Finally, uncertainties in both of these are simultaneously taken into account. The approximation of uncertainties is based on the results of the survey carried out among the DMs in the conference.

To model uncertainty in the weight assessment, the weight ratio estimates are extended into intervals. For each pairwise attribute comparison, an error ratio 2 is used to reflect this uncertainty. Figure 4 shows the resulting value intervals and dominance relations (e.g. “A →

B ” denotes that A dominates B). Attribute ‘Fod+Fod’ still dominates all the other alternatives. Thus, the choice of the most preferred alternative is not sensitive to considerable variations in the weights. One should note that alternatives ‘Fod+---’, ‘Prod+Fod’, ‘Ban+Fod’ and ‘Ban+Ban’ are all dominated, although the upper bound of these is higher than the lower bound of ‘Fod+Fod’. Thus, dominance is pairwise, i.e. there is no single feasible weight combination with which the overall value of any of these is higher than the overall value of ‘Fod+Fod’.

In the second phase, uncertainty in the DMs’ subjective estimates in the *Socio-psychological* effects is studied. This uncertainty can be taken into account by extending the original rating estimates into intervals. However, one should note that any variation within these intervals is assumed to be independent of variations in the other alternatives, and possible uncertainties in the ranges of attributes are taken into account in the variation of the weights. We assumed an uncertainty level of $\pm 10\%$ of the value interval and introduced related intervals into the model. Figure 5 shows the resulting value intervals and dominance relations. In this case, ‘Fod+Fod’ dominates all the other alternatives except ‘Prod+Fod’. Thus, the choice of the best alternative is not very sensitive to considerable variations in the participants’ value estimates either.

Finally, Figure 6 shows the overall value intervals including uncertainties both in the weight assessment and in the value estimation. Although the previous analyses did not show the problem to be very sensitive on these uncertainties alone, under these joint effects all the alternatives except ‘---+---’ become non-dominated. Thus, the DM can for sure only eliminate alternative ‘---+---’ and be confident that at least some protective actions should be taken. However, of the non-dominated alternatives, ‘Fod+Fod’ still performs well. For example, it has the highest lower and upper bounds of the overall value interval (0.75 and

0.93, respectively). Thus, this analysis can also increase the confidence on the alternative 'F_{0d}+F_{0d}', as even in the worst case it is not considerably worse than any of the other alternatives.

The above analysis showed how to quickly assign uncertainties in the model. However, in practice the DM can continue the process by further adjusting the intervals interactively, and study how the overall value interval is affected by these changes. He/she can, for example, tighten the intervals and evaluate which what are the widest intervals with which 'F_{0d}+F_{0d}' is the only dominated alternative, and then consider whether to accept these intervals to represent the allowed uncertainties in the problem.

5. Conclusion

In this paper we have described how to use interval modeling in global sensitivity analyses in multilevel value trees to analyze the effects of the total impacts of all combined uncertainties on the performance of the alternatives. Our approach is concerned with the extreme values of the intervals, which are needed, e.g., if one is interested in a worst case scenario. The proposed approach is computationally fast and the interpretation of the results is conceptually straightforward also for non-mathematicians. With software, such as WINPRE, intervals can be easily given and graphical output visualizes the magnitude of the total uncertainty in an easy to understand way.

The intervals can be constructed in different ways. Strict maximum and minimum values will enable worst case analyses. The decision maker can also assign error ratios to point estimates or treat the intervals as a kind of confidence interval, although in this case one needs to be careful when interpreting the results. Yet, one can assign different levels of uncertainties in all the model parameters simultaneously. By varying the intervals the DM can also carry out what-if analyses and immediately see the results. This is useful e.g. for finding the level of

uncertainty allowed for the decision to still be the best in all circumstances. In a more detailed analysis the interval method can also show which factors affect the outcome the most. That is, by adding imprecision to different factors one by one, the analysis will reveal what information should be collected and how much that will reduce the uncertainty in the model.

As demonstrated in our example, the proposed approach is likely to be attractive in high-risk situations where a worst case analysis is needed and where the DMs might want to quickly try out different what-if analyses. Also when following the precautionary principle a global sensitivity analysis using intervals would be useful.

Acknowledgements

Jyri Mustajoki acknowledges the financial support of the Academy of Finland (project 32641) and the Finnish Cultural Foundation. We also wish to acknowledge two anonymous referees for their constructive comments.

References

- Ammann, M., Sinkko, K., Kostianen, E., Salo, A., Liskola, K., Hämäläinen, R.P., Mustajoki, J., 2001. Decision analysis of countermeasures for the milk pathway after an accidental release of radionuclides, Radiation and Nuclear Safety Authority in Finland, STUK-A186, December 2001. (Downloadable at www.stuk.fi/julkaisut/stuk-a/stuk-a186.pdf)
- Arbel, A., 1989. Approximate articulation of preference and priority derivation. *European Journal of Operational Research* 43, 317-326.
- Arbel, A., Vargas, L.G., 1993. Preference simulation and preference programming: robustness issues in priority derivation. *European Journal of Operational Research* 69, 200-209.

Barron, F.H., 1992. Selecting a best multiattribute alternative with partial information about attribute weights, *Acta Psychologica* 80, 91-103.

Belton, V., Stewart, T.J., 2001. *Multiple Criteria Decision Analysis: An Integrated Approach*. Kluwer Academic Publishers, Boston, MA.

Bryson, N., Mobolurin, A., Ngwenyama, O., 1995. Modeling pairwise comparisons on ratio scales. *European Journal of Operational Research* 83, 639-654.

Butler, J., Jia, J., Dyer, J., 1997. Simulation techniques for the sensitivity analysis of multi-criteria decision models. *European Journal of Operational Research* 103, 531-546.

Clemen, R.T., 1996. *Making Hard Decisions: An Introduction to Decision Analysis*. Duxbury Press, California.

Ehrhardt, J., Weis, A. (eds.), 2000. RODOS: Decision support system for off-site nuclear emergency management in Europe, European Commission, EUR 19144 EN. (www.rodos.fzk.de)

Felli, J., Hazen, G., 1999. Do sensitivity analyses really capture problem sensitivity? An empirical analysis based on information value. *Risk Decision and Policy* 4(2), 79-98.

Francos, A. Elorza, F.J., Bouraoui, F., Bidoglio, G., Galbiati, L., 2003. Sensitivity analysis of distributed environmental simulation models: Understanding the model behaviour in hydrological studies at the catchment scale. *Reliability Engineering and System Safety* 79, 205–218.

French, S., 1995. Uncertainty and imprecision: Modelling and analysis. *Journal of Operational Research Society* 46, 70-79.

French, S., Rios Insua, D. (eds.), 1999. *Journal of Multi-Criteria Decision Analysis*, Special Issue, Sensitivity Analysis in MCDA 8(3).

Goldstein, B., Carruth, R.S., 2004. The precautionary principle and/or risk assessment in World Trade Organization decisions: A possible role for risk perception, *Risk Analysis* 24(2), 491-499.

Graham, J.D., 2000. Perspectives on the precautionary principle, *Human and Ecological Risk Assessment* 6(3), 383-385.

Haines, L.M., 1998. A statistical approach to the analytic hierarchy process with interval judgments. (I). Distributions on feasible regions. *European Journal of Operational Research* 110, 112-125.

Hämäläinen, R.P., 2000. *Decisionarium – Global Space for Decision Support*. Systems Analysis Laboratory, Helsinki University of Technology. (www.decisionarium.hut.fi)

Hämäläinen, R.P., 2003. *Decisionarium – Aiding decisions, negotiating and collecting opinions on the Web*. *Journal of Multi-Criteria Decision Making*, 12(2-3), 101-110.

Hämäläinen, R.P., Helenius, J., 1997. *WINPRE - Workbench for Interactive Preference Programming*. Computer software. Systems Analysis Laboratory, Helsinki University of Technology. (Downloadable at www.decisionarium.hut.fi)

Hämäläinen, R.P., Leikola, O., 1996. Spontaneous decision conferencing with top-level politicians. *OR Insight* 9(1), 24-28.

Hämäläinen, R.P., Lindstedt, M.R.K., Sinkko, K., 2000. Multi-attribute risk analysis in nuclear emergency management. *Risk Analysis* 20(4), 455-468.

Hämäläinen, R.P., Pöyhönen, M., 1996. On-line group decision support by preference programming in traffic planning. *Group Decision and Negotiation* 5, 185-200.

Hämäläinen, R.P., Salo A. Pöysti, K., 1992. Observation about consensus seeking in a multiple criteria environment, Proc. of the 25th Hawaii International Conference on Systems Sciences, Hawaii, Vol. IV, 190-198.

Hazen, G.B., 1985. Partial information, dominance, and potential optimality in multiattribute utility theory. *Operations Research* 34(2), 296-310.

Keeney, R.L., Raiffa, H., 1976. *Decisions with Multiple Objectives. Preferences and Value Tradeoffs*. John Wiley & Sons, Inc.

Kim, S.H., Choi, S.H., Kim, J.K., 1999. An interactive procedure for multiple attribute group decision making with incomplete information: Range-based approach. *European Journal of Operational Research* 118, 139-152.

Kirkwood, C.W., 1992. Estimating the impact of uncertainty on a deterministic multiattribute evaluation. *Management Science* 38(6), 819-826.

Kirkwood, C.W., 1997. *Strategic Decision Making: Multi-objective Decision Analysis with Spreadsheets*. Duxbury Press, Belmont, CA.

Moskowitz, H., Tang, J., Lam, P., 2000. Distribution of aggregate utility using stochastic elements of additive multiattribute utility models, *Decision Sciences* 31(2), 327-360.

Mustajoki, J., Hämäläinen, R.P., Salo, A., 2004. *Decision Support by Interval SMART/SWING – Incorporating Imprecision in the SMART and SWING Methods*, Manuscript. (Downloadable at <http://www.sal.hut.fi/Publications/pdf-files/mmus03.pdf>)

Pannell, D.J., 1997. Sensitivity analysis of normative economic models: Theoretical framework and practical strategies. *Agricultural Economics* 16, 139-152.

Proll, L.G., Salhi, A., Rios Insua, D., 2001. Improving an optimization-based framework for sensitivity analysis in multi-criteria decision-making. *Journal of Multi-Criteria Decision Analysis* 10, 1-9.

Rios Insua, D., French, S., 1991. A framework for sensitivity analysis in discrete multi-objective decision-making. *European Journal of Operational Research* 54, 176-190.

Saaty, T.L., 1980. *The Analytic Hierarchy Process*. McGraw-Hill, Inc.

Salo, A., 2001. On the role of decision analytic modelling. In Stirling A. (ed.), *On science and precaution in the management of technological risk, Vol II, Case Studies*, European Commission, Joint Research Centre, EUR 19056/EN/2, 123-141.

Salo, A., Hämäläinen, R.P., 1992. Preference assessment by imprecise ratio statements. *Operations Research* 40(6), 1053-1061.

Salo, A., Hämäläinen, R.P., 1995. Preference programming through approximate ratio comparisons. *European Journal of Operational Research* 82, 458-475.

Salo, A., Hämäläinen, R.P., 2001. Preference ratios in multiattribute evaluation (PRIME) – Elicitation and decision procedures under incomplete information. *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans* 31(6), 533-545.

Salo, A., Hämäläinen, R.P., 2004. Preference Programming. Manuscript. (Downloadable at <http://www.sal.hut.fi/Publications/pdf-files/msal03b.pdf>)

Salo, A., Punkka, A., 2004. Rank inclusion in criteria hierarchies, *European Journal of Operational Research*. (to appear)

Saltelli, A., Chan, K., Scott, E.M. (eds.), 2000a. *Sensitivity Analysis*. John Wiley & Sons, Ltd.

Saltelli, A., Tarantola, S., Campolongo, F., 2000b. Sensitivity analysis as an ingredient of modelling. *Statistical Science* 15(4), 377-395.

Stam, A., Silva, A.P.D., 1997. Stochastic judgments in the AHP: The measurement of rank reversal probabilities. *Decision Sciences* 28(3), 655-688.

Tarantola, S., Saltelli, A. (eds.), 2003. SAMO 2001: Methodological advances and innovative applications of sensitivity analysis. Special Issue on Reliability Engineering and System Safety 79(2).

Von Winterfeldt, D., Edwards, W., 1986. *Decision Analysis and Behavioral Research*. Cambridge University Press.

Weber, M., 1985. A method of multiattribute decision making with incomplete information, *Management Science* 31(11), 1365-1371.

Weber, M., 1987. Decision making with incomplete information. *European Journal of Operational Research* 28, 44-57.

White, C.C., Sage, A.P., Dozono, S., 1984. A model of multiattribute decisionmaking and trade-off weight determination under uncertainty. *IEEE Transactions on Systems, Man, and Cybernetics* 14(2), 223-229.

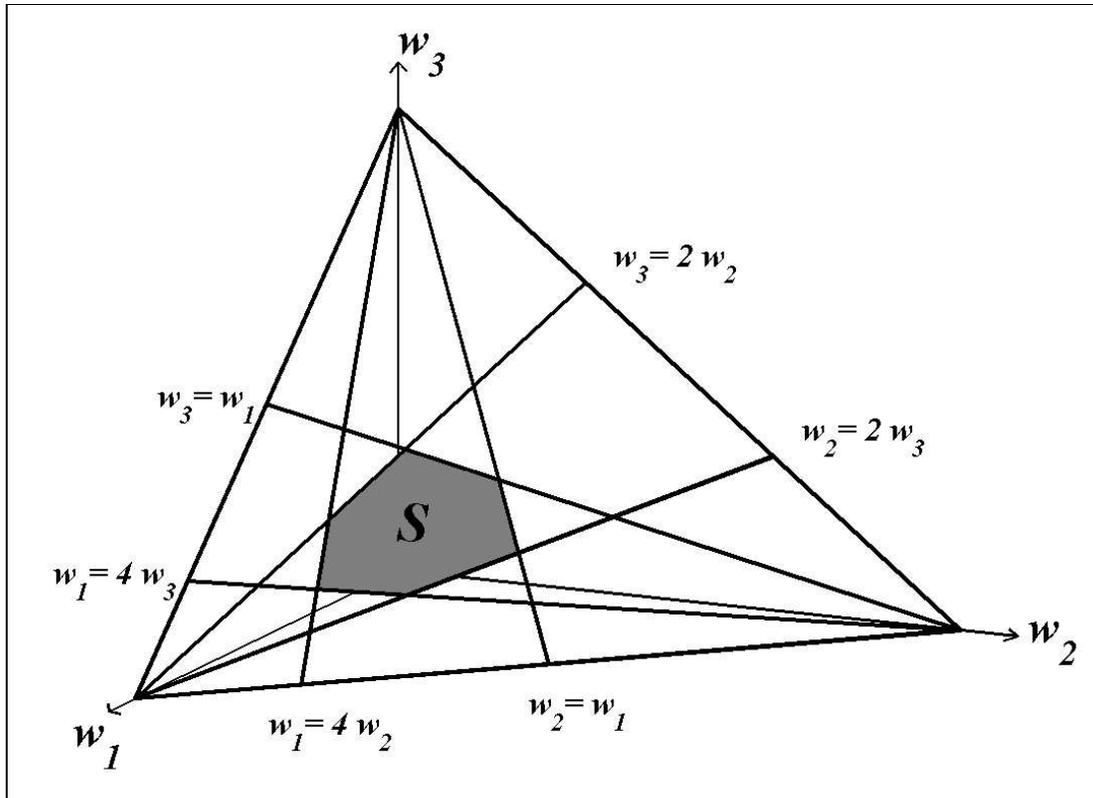


Figure 1. Example of the feasible region of the weights S on the weight plane $w_1 + w_2 + w_3 = 1$.

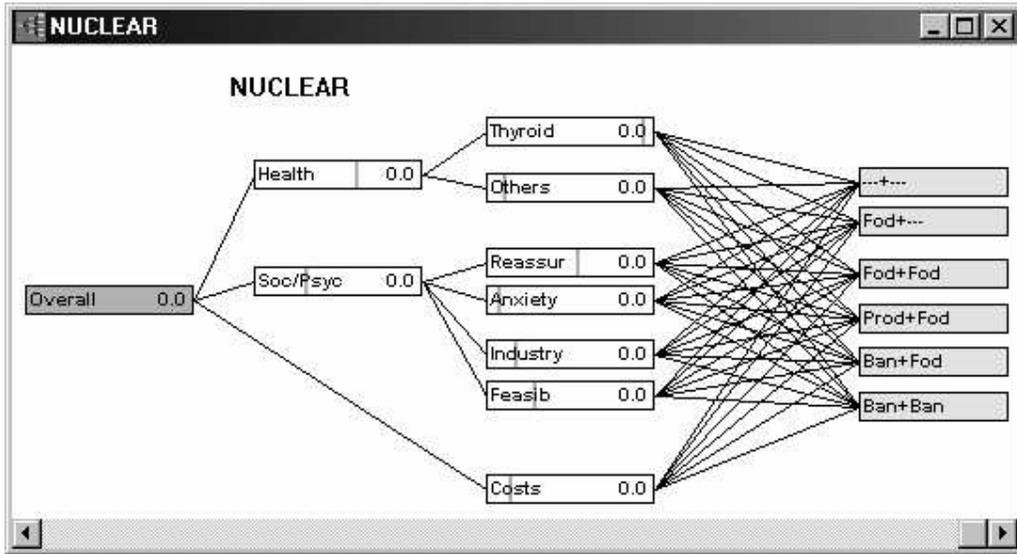


Figure 2. Value tree for the protective actions in the case.

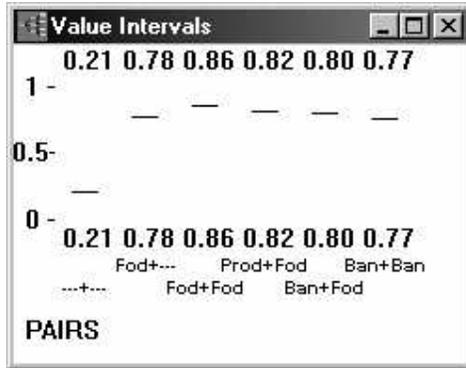


Figure 3. Overall values for the alternatives.

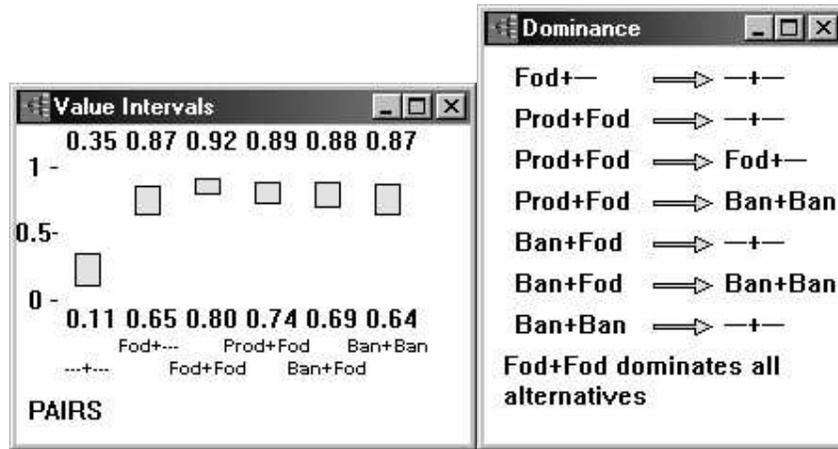


Figure 4. Overall intervals and dominance relations calculated with uncertainty in the weight assessment.

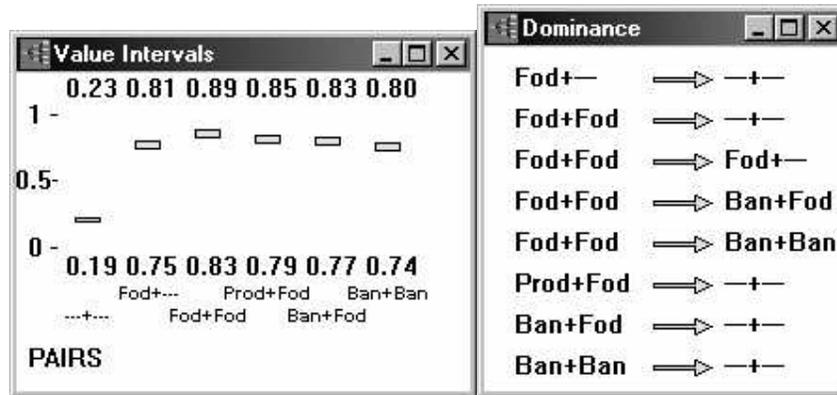


Figure 5. Overall intervals and dominance relations calculated with uncertainty in value estimation.

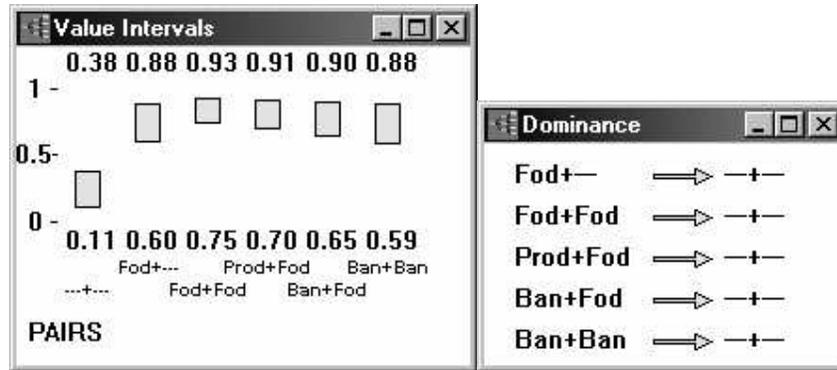


Figure 6. Overall intervals and dominance relations when all the possible uncertainty in the problem is taken into account.